

Isolation of a mouse cDNA coding for a developmentally regulated, testis-specific transcript containing homeo box homology

D.J.Wolgemuth^{1,2}, E.Engelmyer², R.N.Duggal²,
E.Gizang-Ginsberg¹, G.L.Mutter^{2,3}, C.Ponzetto²,
C.Viviano¹ and Z.F.Zakeri²

¹Department of Genetics and Development, ²The Center for Reproductive Sciences, and ³Department of Pathology, Columbia University College of Physicians and Surgeons, 630 West 168th Street, New York, NY 10032, USA

Communicated by W.J.Gehring

A clone, pHBT-1, containing sequences homologous to *Drosophila* homeo boxes has been isolated from a mouse testis cDNA library. The sequence is 80% homologous at the DNA level and 88% homologous at the amino acid level to the homeo box sequence of the *Antennapedia* gene of *Drosophila*. Sequences flanking the 3' end of the homeo box are highly diverged from other murine homeo box-containing genes characterized to date. RNA blot hybridization analysis of mouse testis poly(A)⁺ RNA revealed transcripts of ~1.4 kb in length. Within the limits of sensitivity of detection of Northern blot analysis, no transcripts were seen in any of the adult somatic tissues examined. Other tissues that contain stem cells, namely those of the hemopoietic system, also lacked detectable amounts of HBT-1 transcripts. HBT-1 transcripts were limited to male germ cell-containing tissues, since RNAs from juvenile and adult ovaries did not contain detectable amounts of the 1.4-kb transcripts. Expression of the HBT-1 gene was not detected in embryonic testes, nor in testes of neonatal animals which contain germ cells up to the Type B stage of spermatogonial development. A role for the expression of the HBT-1 gene in the meiotic stages of male germ cell differentiation is postulated.

Key words: homeo box sequences/mouse testis mRNAs/spermatogenesis

Introduction

Germ cells undergo one of the most complex differentiation processes of any cell within the body. The production of mature gametes is achieved through unique developmental programs in the male and female, resulting in strikingly dissimilar cells. Genes involved in this process would be predicted to exhibit tissue- and developmental-specificity of expression. These developmentally regulated genes can be divided into two categories: those that encode various enzymes and structural components of germ cells, and those whose products play a key regulatory role in the commitment of cells to a given pathway of differentiation, perhaps by activating genes responsible for the differentiated phenotype. The identification of regulatory genes has been elusive, not only in germ cell differentiation but also in other developing systems.

The homeotic genes expressed during larval development of *Drosophila* are an example of regulatory genes. Homeotic genes affect the morphogenetic fate of cells within segments of the body (Lewis, 1978). Recent molecular analysis of the *Antennapedia* and *Bithorax* complexes (McGinnis *et al.*, 1984a, b; Scott and

Weiner, 1984) revealed that some homeotic genes contain a region of homology of ~180 bp, termed the 'homeo box' (reviewed in Gehring, 1985). Homeo box sequences have also been identified in the homeotic-like segmentation genes *fushi tarazu* (Kuroiwa *et al.*, 1984; Laughon and Scott, 1984; Hafen *et al.*, 1984) and *engrailed* (Fjose *et al.*, 1985; Poole *et al.*, 1985) and the maternal-effect gene *caudal* (Mlodzik *et al.*, 1985).

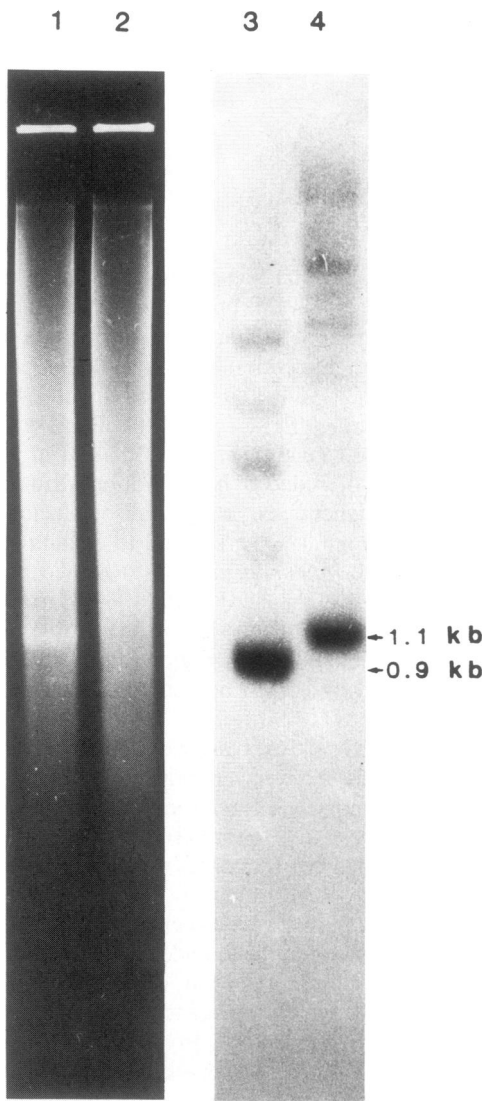
Under conditions of reduced hybridization stringency, the genomes of diverse higher eukaryotic organisms, including frog, chicken, mouse, and human, were found to contain sequences with homology to the *Antennapedia* (*Antp*) homeo box region (McGinnis *et al.*, 1984b; Shepherd *et al.*, 1984; Carrasco *et al.*, 1984; Levine *et al.*, 1984). The significance of the high level of conservation of the homeo box region, and the functions of the products encoded by homeo box-containing genes, remain to be determined.

Recent observations of the expression of homeo box-containing genes during oogenesis in lower vertebrates (Muller *et al.*, 1984) and in mammalian embryonal carcinoma cell lines of germ cell origin (Colberg-Poley *et al.*, 1985a) prompted us to consider if genes containing homeo box sequences would be expressed during mammalian germ cell development. To test this hypothesis, a mouse testis cDNA library was screened with cloned homeo box sequences from the *Antp* gene under conditions of reduced stringency (McGinnis *et al.*, 1984a, b, c). We isolated a cDNA clone, named pHBT-1, which appears to represent a new homeo box-containing gene expressed only in the testis. The pattern of expression of the HBT-1 gene during testicular development suggests a role in the meiotic stages of male germ cell differentiation.

Results

Isolation of a homeo box-containing cDNA clone from a testis cDNA library

Approximately 5×10^5 plaques from an amplified mouse testis cDNA library in λ gt10 were screened using a cloned insert containing the homeo box region of *Antp* (McGinnis *et al.*, 1984b) and pMo10 (a murine homeo box-containing genomic clone, McGinnis *et al.*, 1984c) as probes, under conditions of reduced stringency as described by McGinnis *et al.* (1984a, b, c). Two of the initial six positive plaques were shown through subsequent rounds of purification to hybridize to both *Antp* and pMo10. Further analysis showed that these two clones were identical by restriction mapping. They are designated pHBT-1. Rescreening this library has not revealed any additional clones, although HBT-1 has been isolated 12 times. Upon digestion with *EcoRI*, pHBT-1 yielded inserts of 229 bp and ~500 bp in length. Only the 229-bp fragment was recognized by the homeo box-containing probes. In addition, the ~500-bp fragment did not recognize any transcripts in testicular RNA by Northern blot analysis, exhibited a genomic Southern pattern quite different from that shown for the 229-bp fragment, and did not contain homeo box sequences as revealed by hybridization and sequencing analysis (data not shown). Therefore, this fragment was not considered further in



HBT-1

Fig. 1. Southern blot of mouse genomic DNA hybridized with HBT-1. High molecular weight DNA isolated from adult mouse testes was digested with *EcoRI* (lanes 1 and 3) or *HindIII* (lanes 2 and 4). Fragments were separated on a 1% agarose gel (10 µg DNA per lane), transferred to nitrocellulose, and hybridized with ³²P-labeled HBT-1 insert. Major fragments of ~0.9 and ~1.1 kb in length were detected for *EcoRI* and *HindIII*, respectively. Several weakly hybridizing bands were also visible, probably representing other mouse genes containing homeo box sequences with homology to the HBT-1 homeo box. Exposure was 4 days.

the present study. HBT-1 is used to designate only the 229-bp fragment.

HBT-1 was also recognized by homeo box probes isolated from the *fushi tarazu* gene (Kuroiwa *et al.*, 1984) but not by homeo box sequences from the *engrailed* gene (Fjose *et al.*, 1985) (data not shown), even under conditions of reduced stringency, suggesting that the homeo box of HBT-1 is more divergent from the *engrailed* homeo box than from the homeo boxes of *Antp* and *fushi tarazu*. This prediction was confirmed by DNA sequence analysis (see below).

Genomic organization of HBT-1

To determine if HBT-1 represents a unique locus and to provide a preliminary characterization of its genomic organization with

						10		20			30
HBT-1	---	---	---	---	---	---	---	---	---	---	---
<i>Antp</i>	CGC	AAA	CGC	GGA	AGG	CAG	ACA	TAC	ACC	CGG	
m6	A*G	**G	**G	**G	**G	C*C	**G	**C	**G	**C	**C
m5	G*G	CGC	A*A	**C	C*C	**G	**C	**G	**G	**C	**C
Mo10	TC*	**G	**G	**C	C*C	AC*	G*G	**G	**G	A**	
											EcoRI 60
HBT-1	---	---	---	---	---	---	---	---	---	---	GAA_TTC
<i>Antp</i>	TAC	CAG	ACT	CTA	GAG	CTA	GAG	AAG	**G	**T	
m6	**G	**G	**G	**G	**A	**G	**G	**G	**G	**G	**G
m5	**G	**G	**A	**G	**G	**T	**G	**G	**G	**G	**G
Mo10	CCG	**G	CTG	G**	**G	**G	**G	**G	**G	**G	**G
											90
HBT-1	CAC	TTT	AAC	CGC	TAC	CTG	ACC	CGG	CGG	CGC	
<i>Antp</i>	**G	**C	**T	**G	**G	T**	**G	**T	**G	**A	**A
m6	**T	**C	**G	**G	**G	**G	**G	**G	**C	**G	**G
m5	**G	**C	**G	**G	**G	**G	**T	**G	**G	**G	**G
Mo10	**G	**C	**G	**G	**G	**A	*TG	**G	*C*	**G	**G
											120
HBT-1	CGC	ATC	GAG	ATC	GCC	CAC	ACG	CTC	TGC	TTG	
<i>Antp</i>	A*G	**G	**G	**G	**G	**G	G*C	**G	**G	C*C	C*C
m6	ACG	C**	**G	**G	**T	**G	G**	**G	**G	C*C	C*C
m5	**G	**G	**G	**G	**T	A**	G**	**T	**G	C*C	C*C
Mo10	**G	G*G	**G	**G	**G	A**	CT*	**G	AA*	C*C	C*C
											150
HBT-1	TCG	GAG	CGC	CAG	GTC	AAG	ATC	TGG	TTT	CAG	
<i>Antp</i>	A**	**G	**G	**G	A*A	**G	**T	**G	**C	**G	**G
m6	A*T	**G	**G	**G	A**	**G	**G	**G	**C	**G	**G
m5	A*C	**G	**G	**G	A**	**G	**G	**G	**C	**G	**G
Mo10	A*C	**G	**G	**G	A**	**G	**G	**G	**C	**G	**G
											180
HBT-1	AAC	CGG	AGA	ATG	AAG	TGG	<u>AAG</u>	<u>AAA</u>	GAC	CAC	
<i>Antp</i>	**T	**G	C*C	**G	**G	**G	**G	**G	**G	A**	A**
m6	**T	**G	C*C	**G	**G	**G	**G	**G	**G	**T	**T
m5	**T	**G	C*C	**G	**G	**G	**A	**G	**A	A*T	A*T
Mo10	**G	**G	C*C	**G	**G	*AC	**G	**G	**G	**G	**G
											210
HBT-1	AAA	CTT	CCC	AAC	ACC	AAG	ATG	CGA	TCT	TCC	
<i>Antp</i>	**G	ACG	AAG	GG*	GAG	CC*	GAT				
m6	**G	GA*	GAG	*G*	CAG	GCT	CCC	ACT	G*A	G**	
m5	**G	**C	AT*	**G	T**	*C*	CA*	GCC	AGC	GGG	
Mo10	**G	GGC	AAA	GG*	*TG	CT*	*CC				
											240
HBT-1	AAC	ACT	GCC	TCG	GCC	CCT	GCC	GGC	CCG	CCT	
m6	CCG	GAA	*A*	G**	*TG	**C	T**	*TT	T*C	A*A	
m5	G*A	GAC	T*T	GA*	**G	AAA	*G	**G	GA*	TAG	
											270
HBT-1	GGG	AAA	GCA	CAA	ACT	CAC	AGC	CCA	CAC	CAC	
m6	*CT	GCT	*AC	AAG	G*G	G**	GAG	GAG	G*A	G*G	
											EcoRI
HBT-1	CAT	CCC	CCG	<u>GAA_TTC</u>							
m6	G*G	GAA	GA*								

Fig. 2. Nucleotide sequence of HBT-1, compared to homeo box regions of the *Antp* gene of *Drosophila* and the murine genes m6 (Colberg-Poley *et al.*, 1985a), m5 (Colberg-Poley *et al.*, 1985b), and Mo10 (McGinnis *et al.*, 1984b). The *EcoRI* site of HBT-1 at nucleotides 54–60 and the *MboII* site at nucleotides 168–173 are underlined (*MboII* cuts downstream of its recognition site). Asterisks indicate identical nucleotides to those found in the corresponding positions of HBT-1 (after nucleotide 54); before position 54, the asterisk indicates homology to the *Antp* sequence.

respect to other mouse homeo box-containing genes, Southern blot analysis was performed with restriction enzymes which had been used in studies on other homeo box-containing genes containing *EcoRI* sites within the homeo box region (Colberg-Poley *et al.*, 1985a, b). The results of digesting high molecular weight DNA with *EcoRI* and *HindIII* and probing with the HBT-1 insert are shown in Figure 1. Strongly hybridizing bands of 0.9

	1	2	3	4	5	6	7	8	9	10
HBT-1	---	---	---	---	---	---	---	---	---	---
Antp	arg	lys	arg	gly	arg	gln	thr	tyr	thr	arg
m6	*	*	*	*	*	*	*	*	*	*
m5	gly	arg	*	*	*	*	*	*	*	*
Mo10	ser	*	*	*	*	thr	ala	*	*	*
HBT-1	---	---	---	---	---	---	---	---	glu	phe
Antp	tyr	gln	thr	leu	glu	leu	glu	lys	*	*
m6	*	*	*	*	*	*	*	*	*	*
m5	*	*	*	*	*	*	*	*	*	*
Mo10	pro	*	leu	val	*	*	*	*	*	*
HBT-1	---	---	---	---	---	---	---	---	---	---
Antp	his	phe	asn	arg	tyr	leu	thr	arg	arg	arg
m6	*	*	*	*	*	*	*	*	*	*
m5	*	*	*	*	*	*	*	*	*	*
Mo10	*	*	*	*	*	*	met	*	pro	*
HBT-1	---	---	---	---	---	---	---	---	---	---
Antp	arg	ile	glu	ile	ala	his	thr	leu	cys	leu
m6	*	*	*	*	*	*	*	*	*	*
m5	thr	leu	*	*	*	*	ala	*	*	*
Mo10	*	val	*	met	*	asp	ala	*	*	*
HBT-1	---	---	---	---	---	---	---	---	---	---
Antp	ser	glu	arg	gln	val	lys	ile	trp	phe	gln
m6	thr	*	*	*	ile	*	*	*	*	*
m5	thr	*	*	*	ile	*	*	*	*	*
Mo10	thr	*	*	*	ile	*	*	*	*	*
HBT-1	---	---	---	---	---	---	---	---	---	---
Antp	asn	arg	arg	met	lys	trp	lys	lys	asp	his
m6	*	*	*	*	*	*	*	*	glu	asn
m5	*	*	*	*	*	*	*	*	glu	asn
Mo10	*	*	*	*	*	tyr	*	*	*	gln
HBT-1	lys	leu	pro	asn	thr	lys	met	arg	ser	ser
	asn	thr	ala	ser	ala	pro	ala	gly	pro	pro
	gly	lys	ala	gln	thr	his	ser	pro	his	his
	his	pro								

Fig. 3. Conceptual translation of HBT-1 (homeo box region and 3' flanking region), compared with the predicted amino acid sequences of the homeo boxes of Mo10 (McGinnis *et al.*, 1984c), m6 (Colberg-Poley *et al.*, 1985a), m5 (Colberg-Poley *et al.*, 1985b), and *Antp* (McGinnis *et al.*, 1984b). Asterisks indicate amino acid homology to the corresponding position in HBT-1 (after amino acid 18); before amino acid 18, the asterisk reflects homology with the *Antp* amino acids.

and 1.1 kb were observed with *EcoRI* and *HindIII* digests, respectively (Figure 1). The patterns were the same with DNAs isolated from different strains and from several mouse tissues, including testis, ovary, liver, and brain (data not shown). The pattern observed with HBT-1 is distinct from the patterns characteristic for other mouse homeo box-containing genes reported to date. Weakly hybridizing fragments were also detected (Figure 1), suggesting that within the mouse genome there are additional genes containing sequences with homeo box homology – an observation consistent with data obtained by others (McGinnis *et al.*, 1984b, c; Colberg-Poley *et al.*, 1985a). Isolation of genomic clones for the HBT-1 gene is in progress; to date, ~30 kb of genomic DNA flanking the homeo box region have been isolated in overlapping lambda clones (data not shown).

Sequence of the homeo box-containing testis cDNA clone (pHBT-1) and homology with other homeo box sequences

To determine the degree of homology of the HBT-1 homeo box with other homeo box regions, the DNA sequence of the 229-bp *EcoRI* fragment was obtained by subcloning into M13 vectors

(Messing *et al.*, 1981) and sequencing in both directions by the dideoxy-nucleotide chain termination method of Sanger *et al.* (1977), and by sequencing in a pBR322 subclone using the chemical modification method of Maxam and Gilbert (1977). The sequence is given in Figure 2. Comparisons of homology were made to the *Antp* homeo box, since *Antp* was used to select HBT-1. In addition, the DNA sequence of HBT-1 was compared with the homeo box regions of the mouse genomic clones of Mo10 (McGinnis *et al.*, 1984c) and m6 and m5 (Colberg-Poley *et al.*, 1985a, b). M6 and m5 were included because they represent other mouse homeo box-containing genes that have an *EcoRI* site within the homeo box region and because both have been reported to be transcribed in the testis, among other tissues (Colberg-Poley *et al.*, 1985b). Sequence analysis of the 229-bp insert of HBT-1 revealed that it contains only 126 bp of the 180-bp homeo box sequence and ~100 bp of 3' flanking sequence. The *EcoRI* sites of HBT-1, m6 and m5 are in identical positions but the genes differ in 21 of the 126 (HBT-1 versus m6) and 20 of the 126 (HBT-1 versus m5) nucleotides that make up the homeo box region (83% and 84% homology, respective-

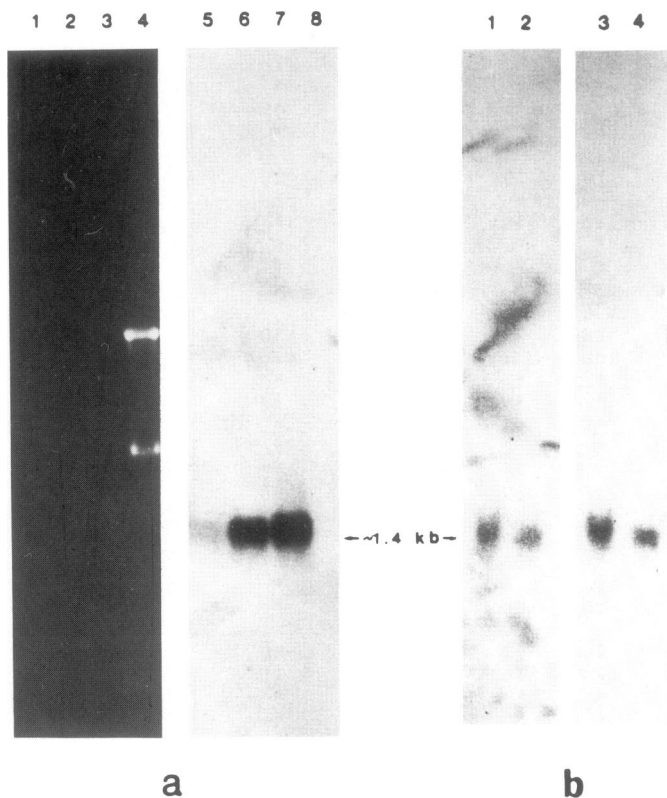


Fig. 4. HBT-1 transcripts in mouse testicular RNA. (Panel a) RNA was isolated from mature mouse testes and analyzed by Northern blot analysis using the pHBT-1 cDNA probe. Lanes 1–4 are the ethidium bromide staining patterns of 1, 4 and 8 μg of poly(A)⁺ RNA and 25 μg of flow-through RNA, respectively. Lanes 5–8 represent the autoradiographic patterns of lanes 1–4 after hybridization with ³²P-labeled HBT-1 insert. Exposure was 3 days. (Panel B) Poly(A)⁺ (4–5 μg), lanes 1 and 3, and total RNA (>25 μg), lanes 2 and 4, were hybridized with either the nick-translated homeo box fragment (lanes 1 and 2) or the 3' flanking region fragment (lanes 3 and 4). Exposure was 4 days.

ly). There is only slightly less homology with *Antp* (80%). The sequence 3' of the homeo box region in HBT-1 is highly diverged from the 3' regions of other mouse homeo box-containing genes described to date (e.g. Figure 2).

A conceptual translation of HBT-1 is presented in Figure 3 and is compared to the amino acid sequences of the homeo box regions of *Antp*, m6, m5 and Mo10. At the amino acid level, the degree of homology of the homeo box region of HBT-1 to that of the *Antp* protein is 88%. The homology of the HBT-1 homeo box with the homeo box sequences of three other mouse genes, m6 (Colberg-Poley *et al.*, 1985a), m5 (Colberg-Poley *et al.*, 1985b) and Mo10 (McGinnis *et al.*, 1984c), is 86%, 86% and 74%, respectively.

Transcripts recognized by HBT-1 and tissue specificity of expression

Total, poly(A)⁺ and poly(A)⁻ (flow-through) RNAs were isolated from mouse testes and analyzed by Northern blot hybridization with HBT-1 at high stringency. HBT-1 recognized testicular transcripts of ~1.4 kb in length (Figure 4, panel a). No hybridization was noted in the poly(A)⁻ sample. To confirm that the transcripts being recognised were the result of hybridization to the HBT-1 gene sequences, and not by sequences of highly homologous homeo box-containing genes (e.g. m6 or m5, Colberg-Poley *et al.*, 1985b), the 229-bp HBT-1 fragment

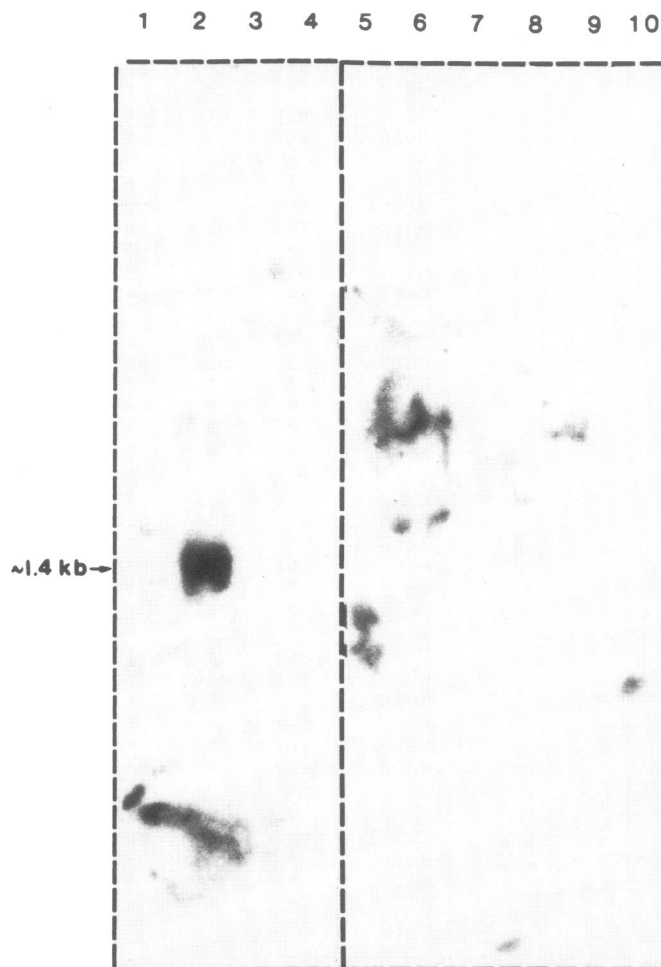


Fig. 5. HBT-1 transcripts in mouse stem cell-containing tissues. Poly(A)⁺ and flow-through RNAs (A)⁻ were analyzed by Northern blot hybridization for the presence of HBT-1 transcripts. Lanes 1 and 2, 30 μg (A)⁻ and 4 μg (A)⁺ RNAs from mature mouse testes; lanes 3 and 4, 30 μg (A)⁻ and 4 μg (A)⁺ RNAs from liver; lanes 5 and 6, 30 μg (A)⁻ and 5 μg (A)⁺ RNAs from mature thymus; lanes 7 and 8, 30 μg (A)⁻ and 4.5 μg (A)⁺ RNAs from immature thymus (day 11 mice); lanes 9 and 10, 30 μg (A)⁻ and 5 μg (A)⁺ RNAs from bone marrow of adult mice. Exposure was 3 days; no additional transcripts were visualized after exposure for 12 days (not shown).

was cut by *Mbo*II into homeo box-containing and 3'-flanking segments (Figure 2). These fragments were purified, nick translated, and used for Northern blot analysis. The ~1.4-kb testicular transcripts were recognized by both fragments (Figure 4, panel b).

To examine the tissue specificity of distribution of the transcripts recognized by HBT-1, RNAs from several tissues were examined. No hybridization to HBT-1 was seen with poly(A)⁺ or flow-through RNAs from liver, brain, uterus, or kidney when probed at high stringency (Figure 5 and data not shown), although homeo box-containing sequences have been detected in some of these tissues by other mouse probes (Colberg-Poley *et al.*, 1985b). Since the testis contains the full range of spermatogenic cells, including stem cells, we then considered that expression of HBT-1 might be characteristic of stem cell-containing tissue in general. To this end, poly(A)⁺ and flow-through RNAs from bone marrow, immature thymus, and mature thymus were examined. Although the question of whether thymus contains true stem cells is unresolved, the thymus clearly contains cells which

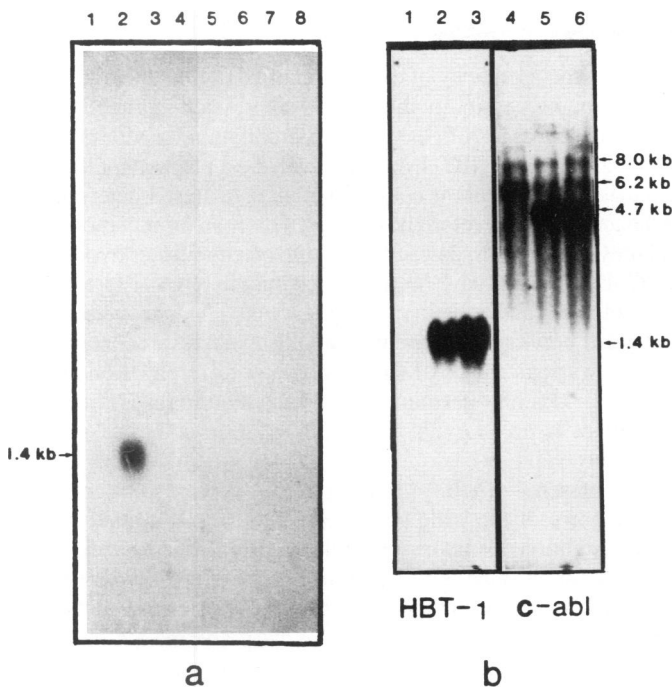


Fig. 6. HBT-1 transcripts in RNAs from germ cell-containing tissues. RNAs were isolated from mature testes (**panel a, lanes 1 and 2; panel b, lanes 2, 3, 5 and 6**), mature ovaries (**panel a, lanes 3 and 4**), and juvenile ovaries (day 11–18) (**panel b, lanes 1 and 4**) and analyzed by Northern blot hybridization with pHBT-1 (**panels a and b**) or pabl sub9 as a control (**panel b, right portion**). The ~1.4-kb transcripts are readily apparent in the testis poly(A)⁺ RNAs (**panel a, lanes 1 and 2; 25 μ g (A)⁻ and 4 μ g (A)⁺**) but not in the RNA of mature ovaries [**panel a, lanes 3 and 4, 25 μ g (A)⁻ and 4 μ g (A)⁺**, respectively]. Two somatic tissues, liver and brain (**panel a, lanes 5–8**) also failed to reveal HBT-1 transcripts (**lanes 5 and 7** each contain 25 μ g (A)⁻ RNA and **lanes 6 and 8** contain 4 μ g of (A)⁺ RNA. Similarly, RNAs from juvenile ovaries also lacked detectable amounts of HBT-1 transcripts (**panel b, lane 1, 5 μ g (A)⁺ RNA**). As a control for integrity of RNAs analyzed in these studies, all blots were hybridized with other probes, an example of which is shown in **panel b, lanes 4–6**, where pabl sub9 easily detected the somatic *c-abl* transcripts of ~8.0 and ~6.2 kb in all three RNA samples, as well as the testis-specific *c-abl* 4.7-kb variant in testicular RNA. **Lanes 1 and 4, 5 μ g of juvenile ovary (A)⁺ RNA; lanes 2, 5, 3 and 6, 5 μ g and 10 μ g (A)⁺ mature testis RNA. Panel a, exposure 9 days; panel b, 6 days.**

are very early in the T-cell lineage (Rothenberg and Lugo, 1985). Within the limits of sensitivity of detection of Northern blot analysis, no HBT-1 transcripts were detected (Figure 5), although control hybridizations with EC1 gene probes (Stacey and Evans, 1983) readily revealed EC1 transcripts (G.L.Mutter, A.J.Stacey and D.J.Wolgemuth, unpublished observations).

Of additional interest to us was an examination of the other germ cell-containing tissue, namely the ovary. No hybridization was observed to RNAs obtained from ovaries of mature or juvenile animals (Figure 6, panel a, lanes 3 and 4; panel b, lanes 1 and 4). In all Northern blot analyses, the quality of the RNAs examined was verified by the ability to detect discrete transcripts homologous to other genes known by concurrent studies in our laboratory to be expressed in these tissues [e.g. the proto-oncogene *c-abl* (Ponzetto and Wolgemuth, 1985; Wolgemuth and Ponzetto, 1985), (Figure 6, panel b); the gene homologous to pEC1 (Figure 5, data not shown); or β -actin (Figure 7, panel b)].

HBT-1 expression during testicular development

To determine whether the expression of HBT-1 is correlated with

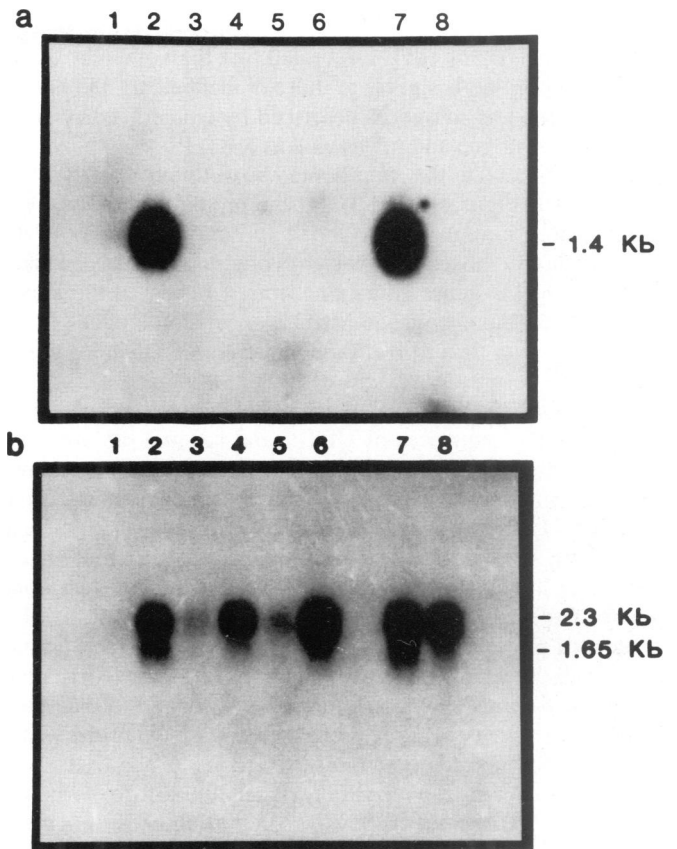


Fig. 7. Expression of HBT-1 during development of the testis. Poly(A)⁺ and flow-through RNAs from testes of different developmental stages were analyzed by Northern blot hybridization for the presence of HBT-1 transcripts. No transcripts were detected in embryonic or neonatal testes, but transcripts were readily apparent in RNAs from mature testes (**panel a**). Control hybridizations with β -actin are shown in **panel b**. The RNAs were clearly intact at all developmental stages as demonstrated by the typical pattern of β -actin transcripts for these stages (C.P.Ponzetto and D.J.Wolgemuth, unpublished observations). **Lanes 2, 4 and 6** each contain 3.8 μ g of (A)⁺ RNA and **lanes 1, 3 and 5** contain 20 μ g of flow-through RNAs of mature testis (**lanes 1 and 2**), day 7–8 neonatal testis (**lanes 3 and 4**), and day 17–20 embryonic testis (**lanes 5 and 6**). **Lanes 7 and 8** each contain 5 μ g of (A)⁺ RNA from mature testis and day 7–8 neonatal testis, respectively. Probe was the 3' flanking region of the HBT-1 homeo box region. Exposures were 8 days (**panel a**) and 3 days (**panel b**).

discrete developmental stages of testicular differentiation, or rather, is a property of the male germ cell-containing tissue throughout development of the testis, RNAs were isolated from testes of day 17–20 embryos, day 7–8 neonatal animals, and adult animals. With respect to the types of germ cells present in these tissues, the embryonic testis contains only germinal stem cells (gonocytes), whereas testes from day 7–8 neonates contain primitive Type A spermatogonia as well as the more advanced Type A and Type B spermatogonia (Nebel *et al.*, 1961; Bellve *et al.*, 1977). Before day 9 of neonatal development, none of the germ cells have entered meiosis. From about day 9–10 onward, cells enter pre-meiotic DNA synthesis and meiotic prophase ensues. Testes at all stages examined in the present study contain the typical somatic cellular components, including Leydig and Sertoli cells. As depicted in Figure 7a, no HBT-1 transcripts were detected in embryonic or neonatal testes, whereas transcripts were readily apparent in RNAs isolated from mature testes. The control hybridization for this Northern blot was β -actin, the results of which are shown in Figure 7b.

Discussion

Sequence analysis of HBT-1 revealed that the homeo box contained in this clone is similar to, but not identical to, the homeo box of the m6 and m5 genes described by Colberg-Poley *et al.* (1985a, b). Both m6 and m5 as well as HBT-1 homeo boxes are quite homologous to the *Antp* homeo box at both the DNA and protein level (Figures 2 and 3). At the protein level, HBT-1 is actually more homologous to the *Antp* homeo box than to the m6 or m5 homeo boxes. An analogous observation has been noted for two *Xenopus* genes containing homeo boxes: at the amino acid level, the *Xenopus* gene MM3 is more homologous to the *Drosophila Antp* than to the *Xenopus* AC1 gene (Muller *et al.*, 1984).

In contrast to the high level of conservation of the homeo boxes, the 3' flanking sequences of HBT-1 and m6 and m5 are completely divergent (Figure 2). This is of particular interest since it has been reported that m6 and m5 also recognize testicular transcripts (Colberg-Poley *et al.*, 1985b). However, both m5 and m6 recognize multiple testicular transcripts and also detect transcripts in several other adult tissues, including ovary and kidney (Colberg-Poley *et al.*, 1985a, b). In contrast, HBT-1 recognized transcripts only in the 1.4-kb size range and only in testis of all the adult tissues examined. To our knowledge, none of the other homeo box-containing genes isolated to date have the tissue-specificity of expression, as assayed by high-stringency Northern blot analysis, exhibited by HBT-1 in the testis.

The testis contains a wide variety of cell types, from hormone-producing and hormone-responsive supporting cells of somatic origin to virtually a complete germ cell lineage, from spermatogenic stem cells through to mature spermatozoa (reviewed in Bellve, 1979). It was the presence of the stem cell population in the testis that prompted our investigation as to whether other tissues containing stem cells — those of the hemopoietic lineage — would also contain HBT-1 transcripts. The fact that none were observed lends support to our postulate that the HBT-1 gene product is involved primarily in male germ cell differentiation.

However, the question of HBT-1 expression in the female germ cell-containing tissue was also addressed. RNAs were isolated from adult ovaries, which contain oocytes in stages of development from primordial follicles through to Graafian follicles (Peters, 1969). Since homeo box-containing transcripts have been observed in growing *X. laevis* oocytes (Muller *et al.*, 1984), RNA was also isolated from ovaries of juvenile mice. Juvenile ovaries contain a higher proportion of growing oocytes, relative to somatic tissues, than is found in adult ovaries (Peters, 1969). No transcripts were detected in any samples. We conclude that HBT-1 transcripts are not abundant in the somatic components of the ovary, because the characteristic somatic *c-abl* transcripts were readily detected in these samples (Figure 6, panel b). Germ cells represent >90% of the cells in the adult testis (Meistrich *et al.*, 1973) but only at most several thousand cells in the adult ovary (Peters, 1969). Considering that HBT-1 transcripts are detected in, for example, 5 µg of poly(A)⁺ testicular RNA with an exposure time of ~6 h, but cannot be detected in comparable amounts of ovary RNAs after one week or longer exposure, if there are HBT-1 transcripts in the female germ cells, they are probably not very abundant. However, certain stages of oocyte development are very rare or missing; therefore if HBT-1 were expressed at only one stage of oogenesis, transcripts might have been missed in this analysis. Definitive resolution of the presence or absence of HBT-1 transcripts in the female germ line will await the results of *in situ* hybridization of sections of ovarian tissues from fetal through adult stages of mouse development and Nor-

thern blot analysis of RNA prepared from isolated oocytes (in progress).

The tissue-specificity of expression of HBT-1 in the testis raised the intriguing possibility that HBT-1 expression might be critical in particular stages of germ cell differentiation, or alternatively, that expression of HBT-1 might actually be a property of the testis as a tissue, i.e., that it could function as a 'testis-determining' gene and might therefore be expressed early in testicular development. Our results to date strongly support the former hypothesis. HBT-1 transcripts have been detected only in RNAs isolated from mature testes and none in embryonic or day 7–8 neonatal testes. As noted previously, embryonic and neonatal testes contain stem cells in various stages of spermatogonial differentiation as well as the various non-germinal cells of the mature testis, including Sertoli cells and Leydig cells. No meiotic cells are present, however.

The presence of HBT-1 transcripts thus appears to be correlated to the entry of the male germ cells into the committed meiotic state. Additional support for a role of HBT-1 during meiotic differentiation may be inferred from observations indicating first, that HBT-1 transcripts can be detected in sterility mutants in which meiosis progresses as far as early spermatids but not in germ cell deficient mutant strains; and second, that enriched populations of pachytene spermatocytes indeed contain HBT-1 transcripts (D.J. Wolgemuth, C. Viviano, E. Gizang-Ginsberg and E. Engelmyer, unpublished observations). It will be important to determine the precise temporal transcriptional and translational expression of HBT-1 throughout meiosis. In addition, it will be of interest to compare the expression of HBT-1 with the pattern of expression of the m6 and m5 genes, which contain homeo boxes highly homologous to HBT-1 and which are also expressed in the testis.

Materials and methods

Sources of tissues

Swiss Webster mice (Camm, Inc.) were used as the source of tissues in all aspects of this study. Adult males refers to mice aged 60 days or older. Adult females refers to animals aged 30 days or older; juvenile females were 11–18 days of age. Embryonic testes were dissected from fetuses of pregnant females at day 17 to day 20 of gestation (the day of vaginal plug was designated day 0 of gestation). Neonatal testes were recovered from animals sacrificed at day 7 or day 8 of post-natal development. Testes, ovaries, and other tissues were removed from animals sacrificed by cervical dislocation. Tissues were either used immediately for RNA isolation or were frozen in liquid nitrogen for subsequent use as sources of RNA or DNA.

Sources of probes

The following cloned probes were used in these studies. *Antp* — p903G, a 600-bp subclone derived from a 3' cDNA of the *Drosophila Antennapedia* gene (McGinnis *et al.*, 1984b); *fushi tarazu* — pFS2, a subclone of the 3' region of the *Drosophila fushi tarazu* gene (Kuriowa *et al.*, 1984); *engrailed* — pFS7036, a 920-bp subclone from the *Drosophila engrailed* gene (Fjose *et al.*, 1985); pMo10, a 1.4-kb mouse genomic homeo box-containing clone (McGinnis *et al.*, 1984b). All the homeo box-containing clones were generously provided by W. Gehring. Control probes for ascertaining efficiency of hybridization and integrity of RNAs were (i) psub 9, a permuted *v-abl* clone which hybridizes to *c-abl* transcripts (Goff *et al.*, 1980, note Ponsetto and Wolgemuth, 1985; and Wolgemuth and Ponsetto, 1985); (ii) pEC1, a cDNA clone isolated from a mouse teratocarcinoma cDNA library (Stacey and Evans, 1985; the gift of A. Stacey); and (iii) pA1, a chicken β-actin cDNA clone (Cleveland *et al.*, 1983; the gift of D. Cleveland).

RNA isolation and Northern blot analysis

RNA was isolated using the LiCl precipitation method described by Cathala *et al.* (1983). Poly(A)⁺ RNA was selected by oligo-dT cellulose chromatography according to Aviv and Leder (1972). RNA samples were electrophoresed on denaturing 1% agarose/2.2 M formaldehyde gels, blotted onto nitrocellulose for 8–12 h according to Maniatis *et al.* (1982), and baked under vacuum for 3 h at 80°C. Prehybridization used sonicated salmon sperm DNA (100 µg/ml); hybridizations in the presence of dextran sulfate were performed essentially as

described by Wahl *et al.* (1979). Probes for hybridization were labeled with [³²P]-dXTPs by nick translation (Weinstock *et al.*, 1978). Following hybridization, filters were sequentially washed for high-stringency hybridization: 20 min each in 2 × SSC, 0.1% SDS at 25°C and at 65°C; 2 × in 1 × SSC, 0.1% SDS at 65°C; and 2 × 0.1 × SSC, 0.1% SDS at 65°C. Filters were air dried and exposed to autoradiographic film at -70°C for 6 h to 21 days.

cDNA library preparation and screening

RNA was isolated from total testis using the LiCl precipitation method, poly(A)-selected as described above, and used to prepare a cDNA library in lambda gt10 according to procedures of Huynh *et al.* (1984). Oligo-dT was used to prime the first strand synthesis using reverse transcriptase. Second strand synthesis was obtained with the Klenow fragment of DNA PolI, using natural hairpin primers. The double stranded cDNAs were trimmed with S1 nuclease, methylated with DNA methylase, and EcoRI linkers were added. cDNAs of greater than ~300 bp in length were cloned into the EcoRI site of the vector. Titration on *E. coli* Hfl strain revealed a library of ~4 × 10⁶ independent clones, which was then amplified. Recombinant phages were plated, grown to near confluency, and plaques were lifted as described by Maniatis *et al.* (1982). Duplicate filters were prepared and hybridized to nick translated inserts of the *Antp* and Mo10 homeo box genes.

Southern blot analysis

High molecular weight DNA was isolated from testis, ovary, liver and brain of mature mice. The methods used were essentially those described by Wigler *et al.* (1979) with minor modifications described by Ponzetto-Zimmerman and Wolgemuth (1984). DNA digestions were performed with various restriction enzymes purchased from commercial sources and performed according to the manufacturers' recommended conditions. Digested DNAs were run on 1% agarose gels and transferred to nitrocellulose filters according to Maniatis *et al.* (1982). High-stringency Southern blot analysis was performed as described by McGinnis *et al.* (1984b), using [³²P]HBT-1 insert, nick translated according to Weinstock *et al.* (1978).

DNA sequencing

The 229-bp HBT-1 insert was subcloned in both orientations into the EcoRI site of M13 derivatives (M13mp11, Messing *et al.*, 1981) and used as template for DNA sequencing by the enzymatic dideoxynucleotide chain termination sequencing method of Sanger *et al.* (1977). The majority of the sequence was determined by reading both strands. In addition, regions of ambiguous sequence were confirmed by sequencing both strands using the chemical modification method of Maxam and Gilbert (1977).

Acknowledgements

We wish to thank Dr W.J.Gehring for generously providing plasmids containing the *Drosophila* homeo box probes and the mouse probe Mo10, and for helpful discussions throughout the course of this study, Dr Eric Schon for help in setting up dideoxynucleotide DNA sequencing protocols, Dr M.-C.Nguyen-Huu and M.Rubin for help in preparing the mouse testis cDNA library, George Grills for photographic assistance, and Drs J.L.Roberts and R.D.Palmiter for critical review of this manuscript. This work was supported in part by grants NIH SO7RR 0535-24, NIH R01 18122, and a Research Scientist Career Development Award from the I.T.Hirschl Trust to D.J.W.

References

Aviv,H. and Leder,P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408-1412.
 Bellve,A.R. (1979) In Finn,C.A. (ed.), *Oxford Reviews of Reproductive Biology*, Oxford University Press, London, Vol. **1**, pp. 159-261.
 Bellve,A.R., Cavicchia,J.-C., Millette,C.F., O'Brien,D.A., Bhatnagar,Y.M. and Dym,M. (1977) *J. Cell. Biol.*, **74**, 68-85.
 Carrasco,A.E., McGinnis,W., Gehring,W.J. and DeRobertis,E.M. (1984) *Cell*, **37**, 409-414.
 Cathala,G., Savouret,J.-F., Mendez,B., West,B.L., Karin,M., Martial,J.A. and Baxter,J.D. (1983) *DNA*, **2**, 329-335.
 Cleveland,D.W., Lopata,M.A., MacDonald,R.J., Rutter,W.J. and Kirschner,M.W. (1983) *Cell*, **20**, 95-105.
 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.
 Fjose,A., McGinnis,W. and Gehring,W.J. (1985) *Nature*, **313**, 284-289.
 Gehring,W. (1985) *Cell*, **40**, 3-5.
 Gizang-Ginsberg,E. and Wolgemuth,D.J. (1985) *Dev. Biol.*, **111**, 293-305.
 Goff,S.P., Gilboa,E., Witte,O.N. and Baltimore,D. (1980) *Cell*, **22**, 777-785.
 Hafen,E., Kuroiwa,A. and Gehring, W. (1984) *Cell*, **37**, 833-841.

Huynh,T.V., Young,R.A. and Davis,R.W. (1984) In Glover,D. (ed.), *DNA Cloning: A Practical Approach*, IRL Press, Oxford, UK.
 Kuroiwa,A., Hafen,E. and Gehring,W.J. (1984) *Cell*, **37**, 825-831.
 Laughon,A. and Scott,M.P. (1984) *Nature*, **310**, 25-31.
 Levine,M., Rubin,G.M. and Tjian,R. (1984) *Cell*, **38**, 667-673.
 Lewis,E.B. (1978) *Nature*, **276**, 565-570.
 Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning, A Laboratory Manual*. Cold Spring Laboratory, Cold Spring Harbor, NY.
 Maxam,A.M. and Gilbert,W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560-654.
 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.J. (1984c) *Cell*, **38**, 675-680.
 Meistrich,M.L., Bruce,W.R. and Clermont,Y. (1973) *Exp. Cell Res.*, **79**, 213-227.
 Messing,J., Crea,R. and Seeberg,P.H. (1981) *Nucleic Acids Res.*, **9**, 309-321.
 Mlodzik,M., Fjose,A. and Gehring,W.J. (1985) *EMBO J.*, **4**, 2961-2969.
 Muller,M.M., Carrasco,A.E. and DeRobertis,E.M. (1984) *Cell*, **39**, 157-162.
 Nebel,B., Amarose,A. and Hackett,E. (1961) *Science*, **13**, 832-833.
 Peters,H. (1969) *Acta Endocr.*, **62**, 98-116.
 Poole,S.J., Kauer,L.M., Drees,B. and Kornberg,T. (1985) *Cell*, **40**, 37-43.
 Ponzetto,C.P. and Wolgemuth,D.J. (1985) *Mol. Cell. Biol.*, **5**, 1791-1794.
 Ponzetto-Zimmerman,C. and Wolgemuth,D.J. (1984) *Nucleic Acids Res.*, **12**, 2807-2822.
 Rothenberg,E. and Lugo,J.P. (1985) *Dev. Biol.*, **1121**, 1-17.
 Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.
 Scott,M.P. and Weiner,A.J. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 4115-4119.
 Shepherd,J.C.W., McGinnis,W., Carrasco,A.E., DeRobertis,E.M. and Gehring,W.J. (1984) *Nature*, **310**, 70-71.
 Southern,E.M. (1975) *J. Mol. Biol.*, **98**, 503-517.
 Stacey,A.J. and Evans,M.J. (1984) *EMBO J.*, **3**, 2279-2285.
 Wahl,G.M., Stern,M. and Stark,G.R. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 3683-3687.
 Weinstock,R., Sweet,R., Weiss,M., Cedar,H. and Axel,R. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 1299-1303.
 Wigler,M., Sweet,R., Sim,G.K., Wold,B., Pellicer,A., Lacy,E., Maniatis,T., Silverstein,S. and Axel,R. (1979) *Cell*, **16**, 777-785.
 Wolgemuth,D.J. and Ponzetto,C.P. (1985) In Slavkin,H. (ed.), *New Technologies and New Discoveries*. Alan R.Liss, Inc., New York, NY, in press.

Received on 25 February 1986; accepted on 20 March 1986.