The EMBO Journal vol.6 no.1 pp.83–90, 1987

The molecular organization of the *H-2K* region of two *t*-haplotypes: implications for the evolution of genetic diversity

Hiroshi Uehara, Kuniya Abe, Chong-Hiyo T.Park, Hee-Sup Shin[1], Dorothea Bennett and Karen Artzt

Memorial Sloan-Kettering Cancer Center and Sloan-Kettering Division, Cornell University Graduate School of Medical Sciences, New York, NY 10021, USA

[1]Present address: Whitehead Institute, Cambridge, MA 02142, USA

Communicated by A.Frischauf

The genetic diversity between the t^{12} and t^{w5} haplotype chromosomes was studied by analyzing the molecular organization of the *H-2K* region. Twenty-one cosmid clones spanning over 150 kb of the *H-2K* region of both *t*-haplotypes were defined, and high resolution restriction maps were determined. Detailed comparison of the t^{12} and t^{w5} restriction maps revealed the following. (i) The *H-2K* regions of both *t*-haplotypes retain a very similar molecular organization to that reported for B10, BALB/c and AKR. The nucleotide sequence diversity estimated from restriction site polymorphism is 0.68% between the t^{12} and t^{w5} haplotypes; these two *t*-haplotypes are no more similar to one another than BALB/c is to AKR. (ii) Genetic recombination is strongly implicated in generating *H-2* polymorphism. (iii) Genetic polymorphisms, defined as small restriction fragment size differences, are observed at multiple sites along the *H-2K* region. An Alu-like B2 sequence and BAM5-R homologous sequence were identified as the inserted/deleted DNA segments of two of these sites, suggesting that insertion/deletion of mobile elements is a general mechanism for generating genetic diversity.

Key words: genetic polymorphism/insertion elements/MHC/restriction endonuclease mapping/*t*-haplotypes

Introduction

Mouse *t*-haplotypes are variant forms of chromosome 17 which are frequently found in wild mouse populations (Klein, 1975). Each *t*-haplotype contains at least one recessive lethal or semilethal mutation in a region where recombination versus normal chromosomes is suppressed (see reviews: Bennett, 1975; Lyon *et al.*, 1981; Silver, 1985). This region, known as the *t*-complex, spans 15–20 centimorgans (cM) of the proximal part of the chromosome, starting ~5–10 cM distal to the centromere (Forejt *et al.*, 1980; K.Artzt, unpublished data). In addition to affecting embryonic development, mutant genes found in the *t*-complex produce a wide variety of effects such as male sterility in t^x/t^y males, transmission ratio distortion in *t*/+ males, and interaction with the Brachyury (*T*) mutation, which results in tailessness in *T/t* mice. The high transmission of *t*-haplotypes through male sperm (Bennett *et al.*, 1983; Lyon, 1984) is clearly responsible for their maintenance in wild populations. The suppression of recombination between *t*-haplotypes and wild-type chromosome 17's is presumably due to two chromosomal inversions, one around the tail interaction factor (Herrmann *et al.*, 1986), and one involving the major histocompatibility complex (MHC) and tufted (*tf*) (Artzt *et al.*, 1982; Shin *et al.*, 1983). Although any two complementing *t*-haplotypes are capable of crossing over with one another, recombination in the wild is very limited, not only because it requires geographical contiguity between populations with complementing lethal genes, but also because of embryonic mortality and male sterility in the double heterozygotes. It has been estimated that genes in *t*-haplotypes have thus been effectively locked together for at least 1 000 000 years (Figueroa *et al.*, 1985).

Serological studies (Hammerberg and Klein, 1975; Sturm *et al.*, 1982; Nižetić *et al.*, 1984; Artzt *et al.*, 1985) and analyses of restriction fragment length polymorphisms (RFLPs) with DNA probes for H-2 class I and class II genes (Shin *et al.*, 1982; Silver, 1982; Figueroa *et al.*, 1985; Artzt *et al.*, 1985) have revealed strong linkage disequilibrium between specific lethal genes and *H-2* haplotypes. The *H-2* class I antigens encoded by *H-2K*, *D* and *L* loci of the MHC are integral membrane glycoproteins (Nathenson *et al.*, 1981) involved in recognition of foreign antigens such as viral proteins by cytotoxic T-cells (Zinkernagel and Doherty, 1979). Studies of these class I genes and their products have revealed an extreme degree of polymorphism (Klein and Figueroa, 1981; Hood *et al.*, 1983).

Since the MHC in *t*-haplotypes has been 'locked in' by recombination suppression over a long evolutionary period, comparative studies of the genetic organization of *H-2* class I genes of different *t*-haplotypes in this polymorphic region may give insights for understanding the evolutionary relationships between *t*-haplotypes. In addition, comparisons of gene organization between *t*-haplotypes and wild-type chromosomes will shed light on the genetic mechanism(s) that produce diversity of class I genes and their close and remote flanking regions, since opportunities for homogenization and the exchange of DNA sequence information between chromosomes is greatly reduced in *t*-haplotypes.

In this paper, we report precise restriction mapping and information on the molecular organization of the *H-2K* region of the t^{12} and t^{w5} haplotypes. These studies provide significant new information showing that the overall molecular organization in the *H-2K* region of *t*-haplotypes does not differ from that of wild-type chromosomes. Furthermore, insertion of repetitive DNA elements is suggested as a novel mechanism for the generation of diversity.

Results

Isolation of cosmid clones containing class I genes and the H-2K region

Five probes isolated from the MHC were used for screening a cosmid library constructed from liver DNA of t^{12}/t^{w5} compound mice. Probe pH-2IIa (Steinmetz *et al.*, 1981) hybridizes to most class I genes. Three other probes (LS 1/1, Bm1-18 and Bm1-R 3/1) were unique copy sequences that flank the $H\text{-}2K^b$ gene in C57BL/10 (B10) mice (Weiss *et al.*, 1984), and span about 130 kb of the *H-2K* region. Probe 8D was a unique copy specific to the *H-2K* gene (Kress *et al.*, 1983). These probes enabled us to select cosmids that presumably contain most, if not all, of the MHC, and more specifically to identify a subset that spanned the *H-2K* region from both chromosomes.

Cosmid clones (6.4×10^5) were screened and 121 positive clones were isolated. DNA prepared from these cosmids was dot blotted and hybridized separately to each of the five probes initially used, and also to five additional probes which specifically recognize sequences in the *D*, *Qa* and *T1a* regions of the MHC (see Materials and methods). On this basis 53 clones were assigned to *H-2K*, *H-2D*, *Qa* or *T1a* regions; an additional 68 clones were positive with class I probes but negative with all of the *H-2* subregion probes and await more precise assignment. In this paper we describe analyses of 21 clones that define the *H-2K* region of the MHC of *t*-haplotypes.

Table I. RFLPs between t^{12} and t^{w5} chromosomes with *H-2K* probes and classification of cosmid clones

Probe	Restriction enzyme	t^{12} specific fragment (kb)	No. of cosmids	t^{w5} specific fragment (kb)	No. of cosmids
LS 1/1	*Pst*I	2.5	5[a]	1.9	3
Bm1-18	*Eco*RI	7.0	4[a]	5.9	1
8D	*Apa*I	3.7	4	3.3	5
	*Ban*I	2.3		2.0	
Bm1-R 3/1	*Taq*I	2.1	1	2.0	1
Total			11		10

[a]Three cosmids hybridized with both probes.

Assignment of K-region clones to t^{12} or t^{w5} chromosomes

Genomic DNA from t^{12}/t^{w5} compound mice and t^{w5}/t^{w5} embryonal carcinoma (EC) cells was digested with various restriction enzymes, Southern blotted and restriction fragment length polymorphisms (RFLPs) between the two DNAs identified with the four *K*-end specific probes. RFLPs between t^{12} and t^{w5} chromosomes were detected as extra bands in t^{12}/t^{w5} DNA compared with t^{w5}/t^{w5} DNA. The *K*-end cosmid clone DNAs were then digested with the same enzymes and the sizes of the fragments hybridized with the appropriate probe were scored. In this way 22 *K*-end clones were assigned to either the t^{12} or the t^{w5} chromosome (Table I). One clone containing the Bm1-18 region was subsequently found by restriction mapping to be a cloning artifact carrying a segment of unlinked DNA and was omitted from further study.

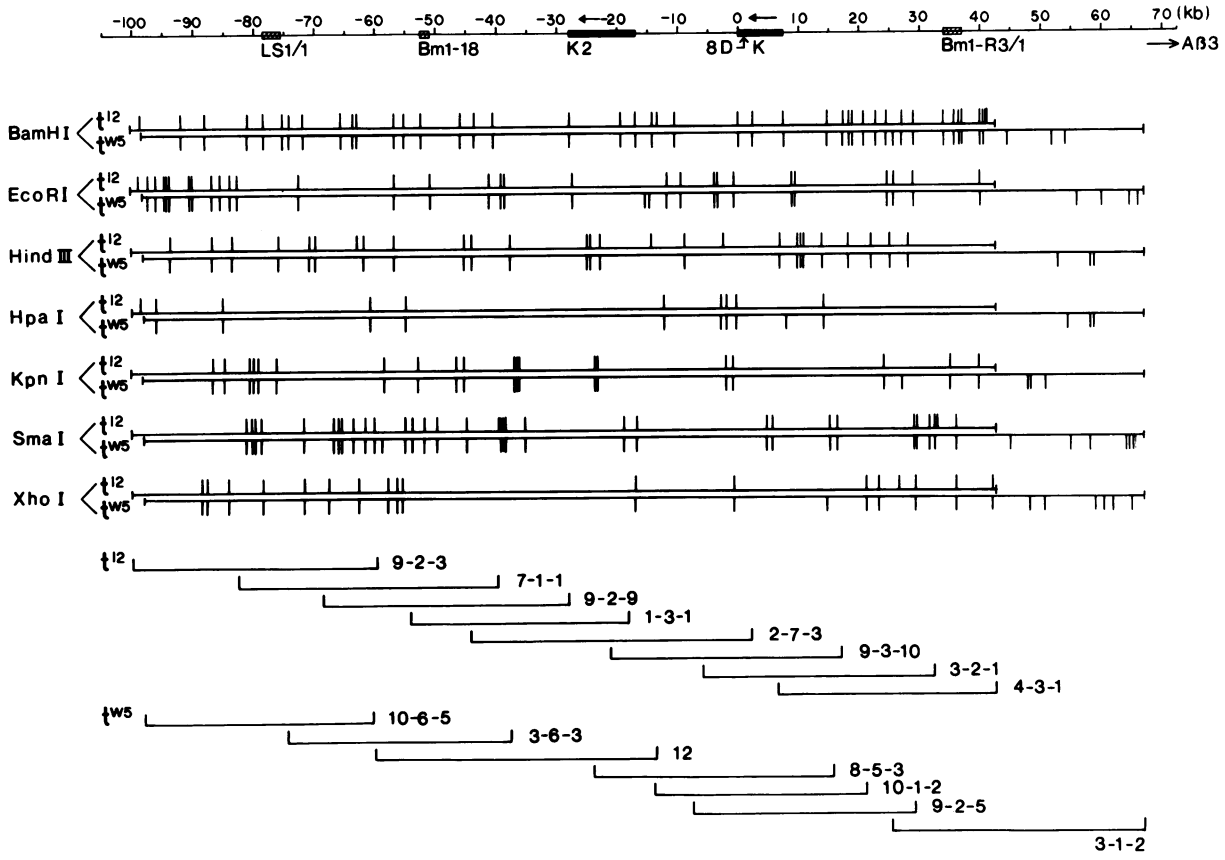

Fig. 1. Restriction map of the *H-2K* region of the t^{12} and t^{w5} haplotypes. Restriction sites were determined for *Bam*HI, *Eco*RI, *Hin*dIII, *Hpa*I, *Kpn*I, *Sma*I and *Xho*I. For every enzyme, the restriction sites are indicated above the horizontal line for the t^{12} chromosome, and below the line for t^{w5}. The map was drawn so that the left-hand side is centromeric (since the MHC is inverted in *t*-haplotypes) to facilitate comparison with wild-type chromosomes. The DNA segments homologous to the *H-2* region probes are shown as *Bam*HI fragments. The *H-2K* and *K2* class I genes are shown as *Bam*HI fragments hybridized with pH-2IIa and pH-2III probes (*H-2K*, 5.15-kb and 2.3-kb *Bam*HI fragments; *K2*, 2.3-kb and 8.0-kb fragments). The regions homologous to 3 *K*-end probes were shown as LS 1/1: 3.0-kb *Bam*HI fragment; Bm1-18: 1.6-kb *Bam*HI−*Eco*RI fragment; Bm1-R 3/1 1.8-kb, 0.8-kb and 0.56-kb *Bam*HI fragments. The *Bam*HI site of the 2.3-kb *Bam*HI fragment homologous with probe 8D was used as the starting point (0). Left of this point is shown as a minus scale; toward the right (*I* region) is the plus scale. All locations of fragments mentioned in the text refer to this scale. The distance is based on the t^{12} haplotype map and small adjustments were made to align the map of the t^{w5} haplotype. (As described in the text, small scattered length differences (20−700 bp) were detected between the t^{12} and t^{w5} chromosomes throughout the region, and the length of the DNA in these areas were adjusted to the t^{12} map in order to obtain maximum matches of restriction sites between the two chromosomes.) Only the clones which were necessary to define the area are shown, although other overlapping clones were mapped. The 14-kb stretch proximal to the *K2* pseudogene on t^{w5} was filled in and mapped using a clone (clone 12) isolated separately from cosmid library of t^{12}/t^{w5} EC cell DNA. The restriction sites where clone 12 overlaps the two flanking clones (clone 3-6-3 and clone 8-5-3) from the t^{12}/t^{w5} library were identical.

Gene organization of the H-2K region of t^{12} and t^{w5} haplotypes

The restriction map of the cosmid clones was determined by the terminase–oligomer method developed by Lehrach and co-workers (Rackwitz *et al.*, 1984, 1985). All clones were aligned along the *H-2K* region via overlapping restriction sites (Figure 1).

Class I genes and their orientation were determined using three *H-2* cDNA probes: pH-2III, a cDNA clone homologous to the third exon of class I genes, pH-2IIa (see above) and 8D (see above). Two class I genes were identified in the *H-2K* region, and the one homologous with 8D was assigned as the *H-2K* gene. These two genes defined by *Bam*HI sites are shown at the top of Figure 1, together with maps of chromosomal segments homologous to probes LS 1/1, Bm1-18 and Bm1-R 3/1.

The molecular organization of the *H-2K* region of these two *t*-haplotypes is very similar to *H-2K* of B10 mice (Weiss *et al.*, 1984; R.Flavell, personal communication). The location of all the unique DNA probes specific for the *H-2K* region (LS 1/1, Bm1-18 and Bm1-R 3/1) relative to the *H-2K* gene is very similar

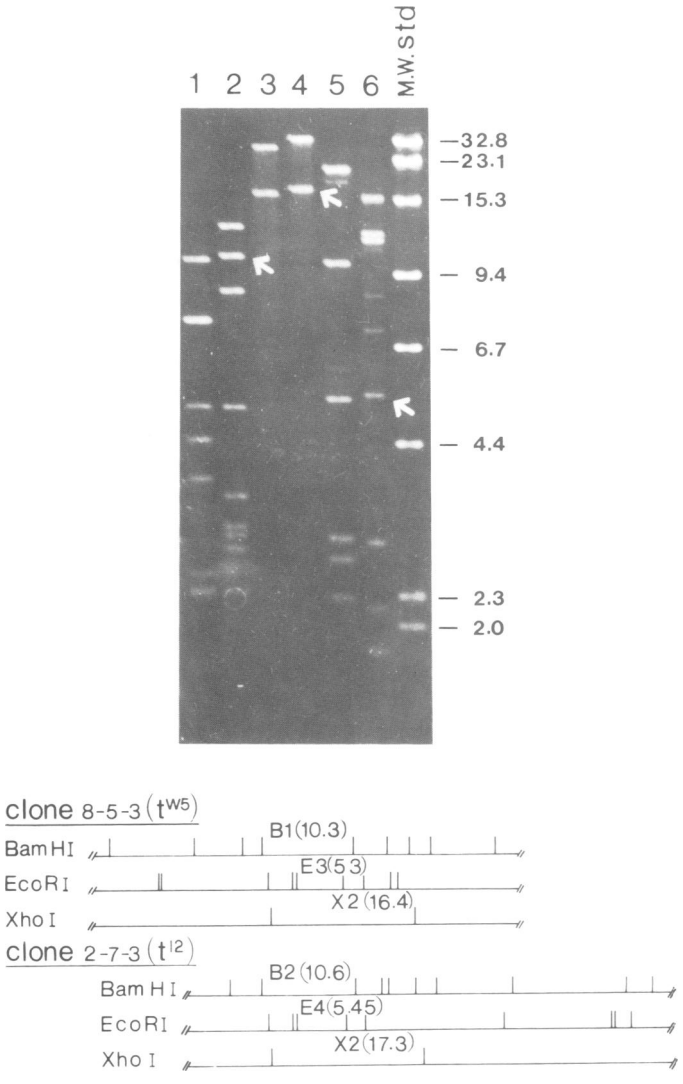

Fig. 2. Detection of small restriction fragment size difference. Ethidium bromide stained 0.7% agarose gel of t^{12} DNA (clone 2-7-3) and t^{w5} DNA (clone 8-5-3) digested wtih *Bam*HI, *Xho*I and *Eco*RI, showing size differences between restriction fragments from the two chromosomes. The bands indicated by arrows derive from corresponding regions of t^{12} and t^{w5} chromosomes. **Lanes 1** and **2**: *Bam*HI; **lanes 3** and 4, *Xho*I; **lanes 5** and **6**, *Eco*RI. Odd lanes are t^{w5}; even lanes are t^{12}. The faint bands seen in **lanes 5** and **6** result from incomplete *Eco*RI digestion.

although not identical in the two *t*-haplotypes and wild-type chromosomes. Thus, the distances of probes Bm1-18, LS 1/1 and Bm1-R 3/1 from the *Kpn*I site at −1 kb in Figure 1 are 74.5, 50 and 34 kb, respectively, while the same DNA segments in B10 DNA are located at 68.5, 47 and 34 kb from the same *Kpn*I site (R.Flavell, personal communication). The major reason for the difference in distance of LS 1/1 and Bm1-18 can be attributed to the longer interval between the *H-2K* gene and the *K2* gene in the t^{12} and t^{w5} haplotypes compared with B10 DNA. In addition, the orientation of *H-2K* and the *K2* gene (if it is transcribed) relative to other markers are also the same as those on wild-type chromosomes such as B10 (Weiss *et al.*, 1984), BALB/c and AKR (Steinmetz *et al.*, 1984) except that the whole region of the MHC is inverted with respect to wild type.

Structural differences between the H-2K region of t^{12} and t^{w5} chromosomes

Detailed analysis of the restriction map of t^{12} and t^{w5} cosmid clones revealed three distinct classes of DNA polymorphism between the two *t*-haplotypes. The first represented small differences in the size of many restriction fragments in t^{12} and t^{w5}. In order to obtain more accurate sizes of the restriction fragments as defined by partial digests, the size of DNA fragments obtained by complete digestion of cosmid DNA was determined. Each fragment was then assigned precisely to the regions in the cosmid determined by partial digests. In these analyses we found that in general the order of specific restriction enzyme sites was conserved between the $t^{12}$and t^{w5} chromosomes; however, the absolute distances between them were often significantly different in the two haplotypes (Figure 2). These observations were reproducibly confirmed with several different restriction enzymes and the regions responsible for the size differences were localized to as small a fragment as possible. As shown in Figure 3 (line A) fragment size differences are observed throughout the *H-2K* region. The size difference ranges from 20 bp (which was the lower limit of detection sensitivity) to ~700 bp.

A second kind of polymorphism between the t^{12} and t^{w5} chromosomes consists of the presence or absence of restriction sites for a given enzyme. Of 183 restriction sites available for comparison between the two chromosomes (Figure 1), 169 sites are identical while 14 sites were scored as present on only one of the two chromosomes (Figure 3, line C).

A third diversity involves the apparent replacement of a DNA segment. Between the two *K* genes, in the region 5′ to the *K2* pseudogene a DNA segment about 3 kb in the t^{12} chromosome and about 2.5 kb in the t^{w5} chromosome, appears to be totally different in the two chromosomes (Figure 3, lines A, B and C. The size difference of this region is also shown in Figure 3, line A). As a result, multiple restriction enzyme site differences are evident.

Analyses of polymorphisms generated by insertion/deletion of DNA segments

Families of short interspersed (Alu-like) repetitive sequences (SINEs) and long interspersed sequences (LINEs) have been identified (see reviews: Jelinek and Schmid, 1982; Rogers, 1985). Their properties suggest that they have dispersed throughout the genome by a mechanism similar to mobile DNA elements. The restriction fragment size differences identified between the t^{12} and t^{w5} chromosomes are small and distributed throughout the *H-2K* region, implying that at least some of these size differences might be caused by the insertion or deletion of repetitive sequences.

In order to examine this possibility, we first defined the regions

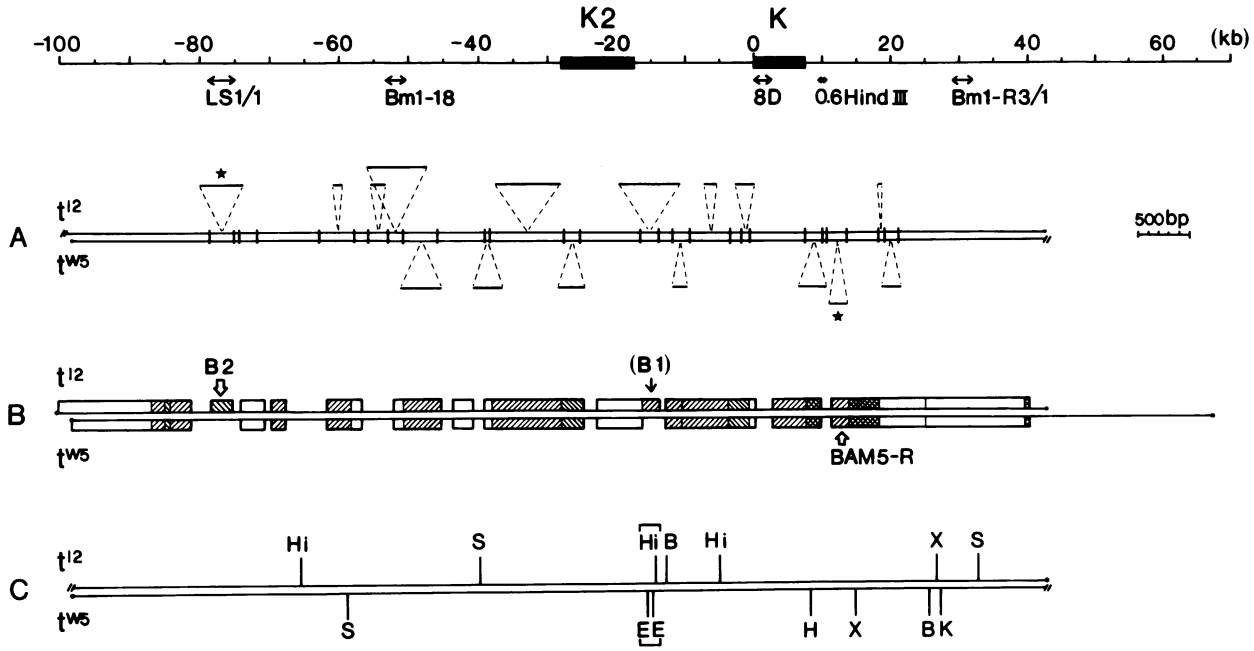

Fig. 3. Structural differences between the t^{12} and t^{w5} chromosomes. The two class I genes and three *K*-end probes are shown at the top of the figure as described in Figure 1. The size of the region is indicated at the top of the figure as described in Figure 1. **Line A**: chromosomal regions showing size differences localized between restriction enzyme sites are indicated by vertical lines. The length of each bar indicates the size difference in that region; the scale is shown on the right-hand side in base pairs. A bar above the base line indicates that the t^{12} fragment is longer than the t^{w5} fragment, while a bar under the line indicates that the t^{w5} fragment is longer than that in t^{12}. **Line B**: distribution of repetitive sequences. Regions containing undefined highly repetitive sequences (visualized by hybridization with ^{32}P-labeled total mouse DNA) are shown as open blocks. Regions where sequences homologous to B1, B2 and BAM5-R are located are shown as: ▨ B1, ▧ B2, ▩ B1 and B2. The limits of resolution are governed by the smallest restriction fragments identified. The open arrows indicate that the indicated sequence (B2 or BAM5-R) was detected on only one of the chromosomes. The B1 sequence (B1 and solid arrow) is in the region where one of the two chromosomes appears to have a small interstitial translocation. **Line C**: restriction sites present only on one chromosome are shown. B, *Bam*HI; E, *Eco*RI; Hi, *Hin*dIII; H, *Hpa*I; K, *Kpn*I; S, *Sma*I; X, *Xho*I. The polymorphisms within the brackets are those caused apparently by the replacement of the DNA segment.

in the cosmid DNAs that contain high copy number repetitive sequences (Figure 3, lane B). The same blots were hybridized with cloned DNA homologous to three families of mouse repetitive sequences, B1, B2 and BAM5-R. Significant differences in hybridization patterns between the t^{12} and t^{w5} chromosomes were found at two of the 15 sites defined as insertions or deletions. As shown in Figure 4B, total mouse DNA hybridizes strongly to a 3.0-kb *Bam*HI/3.05-kb *Kpn*I fragment of t^{12} clone 9-2-3, while no hybridization to the corresponding 2.55-kb *Bam*HI/2.60-kb *Kpn*I fragment of the t^{w5} clone 10-6-5 was observed. The *Bam*HI/*Kpn*I fragments of t^{12} but not t^{w5} also specifically hybridized with a cloned B2 probe (Figure 4C). Hybridization pattern differences were also seen 5′ of the *H-2K* gene. Although both the 7.1-kb *Bam*HI/2.95-kb *Hin*dIII fragment of the t^{12} clone 3-2-1, and the 7.6-kb *Bam*HI/3.15-kb *Hin*dIII piece of the t^{w5} clone 9-2-5 hybridize with total mouse DNA, only fragments from 9-2-5 hybridize with the BAM5-R sequence (data not shown). These results strongly suggest that the size differences at these two sites are generated by the insertion of the B2 and a fragment homologous to BAM5-R, respectively. We could not identify any specific insertions or deletions of repetitive sequences in the other 13 sites, since the smallest restriction fragments that we could obtain by the enzymes used for mapping contained repetitive sequences in both the t^{12} and t^{w5} chromosomes and we could not distinguish newly inserted repetitive elements from pre-existing sequences. The 3-kb *Bam*HI fragment derived from the −14.5 to −17.5 region of t^{12} also hybridizes with total genomic DNA, whereas the corresponding 2.5-kb fragment of t^{w5} does not. This is the region immediately 5′ of the *K2* gene, where t^{12} and t^{w5} may contain different DNA segments, as already described. A cloned B1 sequence shows weak hybridization with the t^{12} fragment.

In order to determine whether the DNA insertions/deletions observed in the t^{12} and t^{w5} chromosomes are diversities generated during the evolution of mice in general or whether they are more common in *t*-haplotypes, we further analyzed genomic DNA from eight *t*-haplotypes (including another member of the t^{12} complementation group t^{w32}), and four inbred strains by hybridization with three *K*-end probes (Figure 5). All of these probes detect polymorphisms between t^{12} and t^{w5}. These polymorphisms are most likely caused by small DNA insertions/deletions (see Figure 3, line A). When *Bam*HI digested DNA was probed wtih LS 1/1, only chromosomes of the t^{12} complementation group have a 3.0-kb *Bam*HI fragment, while all other mouse chromosomes, both *t*-haplotypes and inbred strains, show a 2.55-kb fragment, suggesting that the insertion of a 0.45-kb B2 sequence has occurred only in chromosomes of the t^{12} complementation group (Figure 5A). Probe Bm1-18 also detected a 6.8-kb *Eco*RI fragment specific for the t^{12} and t^{w32} haplotypes, whereas a 6.0-kb fragment is seen in all other genotypes. On the other hand, the hybridization pattern obtained with a 0.6-kb *Hin*dIII probe isolated from clone 9-2-5 is rather complicated (Figure 5B). Hybridization to a 7.6-kb *Bam*HI fragment is observed not only with the t^{w5} chromosome but also with $H\text{-}2^k$ haplotype chromosomes (i.e. AKR and C3H). A 7.1-kb *Bam*HI band is seen in t^{12} (t^{w32}), t^{w1} and t^{w73}, whereas

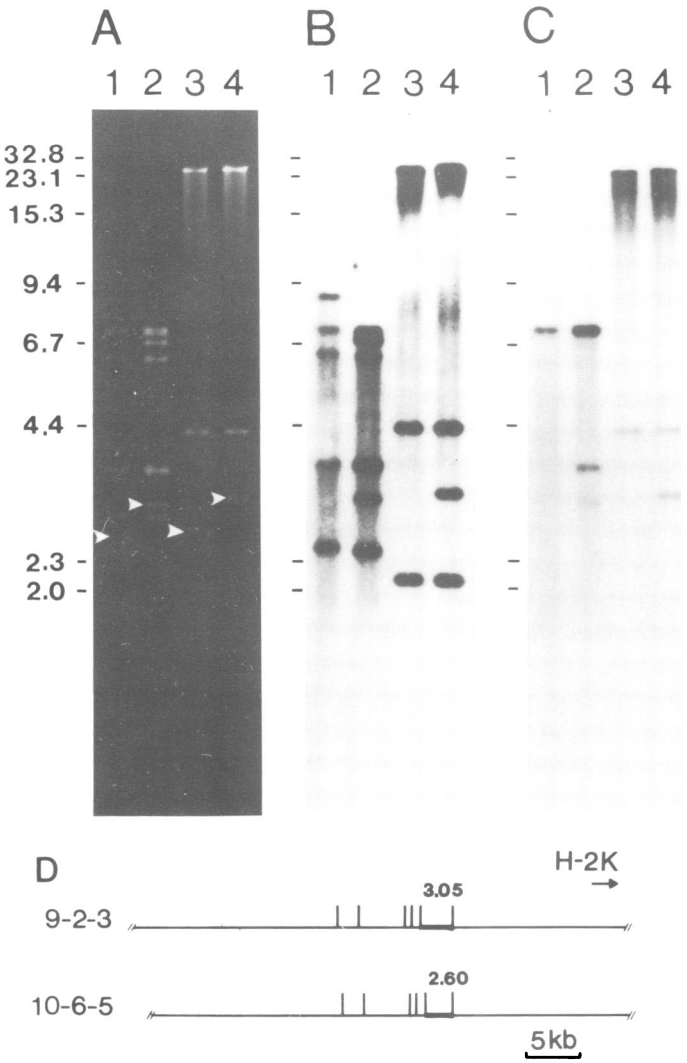

Fig. 4. Characterization of insertional elements in the t^{12} chromosome. DNAs from t^{w5} (clone 10-6-5) and t^{12} (clone 9-2-3) were digested with *Bam*HI (**lanes 1** and **2**) and *Kpn*I (**lanes 3** and **4**), the fragments were separated on an agarose gel and Southern blotted. Odd lanes are t^{w5}; even lanes are t^{12}. (**A**) Ethidium bromide staining of the gel. The bands indicated by arrows are the 3.0-kb *Bam*HI/3.05-kb *Kpn*I of clone 9-2-3, and the 2.55-kb *Bam*HI/2.6-kb *Kpn*I of clone 10-6-5. (**B**) Autoradiograph of the blot hybridized with nick translated total mouse DNA to show repetitive sequences. (**C**) Hybridization with a nick translated clone containing a B2 repetitive sequence. (**D**) Restriction map of clones 9-2-3 and 10-6-5. Only sites for *Kpn*I are shown. The 3.05-kb *Kpn*I fragment of 9-2-3 and 2.60-kb *Kpn*I fragment of 10-6-5 are indicated by the thick line. These fragments contain DNA sequences homologous to probe LS 1/1.

the BALB/c and t^0 chromosomes both show a 6.7-kb band, and B6 shows a 6.3-kb band.

Discussion

Evolution and origin of mouse t-haplotypes

We have isolated 22 clones from the *H-2K* region of either the t^{12} or t^{w5} chromosome. High resolution restriction mapping and alignment of 21 of these clones have allowed detailed analysis of this region of *t*-haplotypes for the first time.

As shown in Figure 1, the *H-2K* region of both *t*-haplotypes encodes two class I genes which are about 17 kb apart and oriented in the same direction as in wild-type chromosomes when the inversion in *t*-haplotypes is taken into consideration. All three DNA segments, represented by the unique K-end probes used, are detected in *t*-haplotypes and their relative positions on the chromosomes are very similar to that in B10 mice (Weiss *et al.*, 1984; R.A.Flavell, personal communication). Furthermore, the restriction maps of this region among t^{12}, t^{w5}, BALB/c and AKR are very similar (Figure 6) except with respect to the *H-2K* gene and its immediate vicinity. The high degree of homology between *t*-haplotypes and wild-type chromosomes in a relatively large region (~170 kb) strongly suggests that they arose from a common mouse ancestral chromosome and that *t*-haplotypes are not derivatives of an alien chromosome as was suggested by Silver (1982).

A large chromosome inversion that includes the whole MHC and several centimorgans on either side of it has been detected by classical genetic mapping in mice heterozygous for different *t*-haplotypes (Artzt *et al.*, 1982). The data presented here indicate this inversion occurred on a large scale and that local chromosome structure around the MHC was not significantly rearranged, if at all, during that event.

Nei and Li (1979) have developed a mathematical model for estimating nucleotide sequence diversity from restriction map comparisons. Using this method, we calculated diversities in the *H-2K* region among the two *t*-haplotypes and two inbred strains (BALB/c and AKR) (Figures 3 and 6). If we ignored the region immediately adjacent to the *H-2K* gene (−5 to +5 kb), where many restriction site differences are observed, we obtained the following diversities for the *H-2K* flanking region: t^{w5} versus t^{12}, BALB/c or AKR is 0.68% ± 0.03, 0.97% ± 0.2 or 0.65% ± 0.14 respectively; t^{12} versus BALB/c or AKR is 2.4% ± 0.75 or 3.2% ± 0.95; BALB/c versus AKR is 1.3% ± 0.19. Thus, although the data available for comparison are rather limited (especially those of t^{12} versus BALB/c or AKR), we can conclude that diversity in the flanking region of *H-2K* is about the same among all four chromosomes compared, and there are no special features distinguishing any chromosome or set of chromosomes. This diversity between *t*-haplotypes and wild-type chromosomes is compatible with that reported for other regions of the *t*-complex. Willison *et al.* (1986) calculated a 1.4% nucleotide sequence diversity by comparing the nucleotide sequence of *Tcp-1*^{a} and *Tcp-1*^{b} cDNAs and their introns. Similarly, by sequence analysis Labeit and Lehrach calculated a 1.0% diversity in 3.5 kb of random sequences isolated from a *t*-haplotype and a wild-type chromosome (Frischauf, 1985). The major source of the restriction site polymorphisms observed by these comparisons is probably point mutations. Nucleotide changes which occur in non-coding regions, in introns, or which result in changes at wobble positions in the third base of a codon are expected to be free of, or under very low, natural selective pressure. If these neutral changes reflect evolutionary relationships among chromosomes (Hayashida and Miyata, 1983), in this respect t^{12} and t^{w5} haplotypes are evolutionarily no closer to each other than to the BALB/c and AKR inbred strains. (If we use the maturation rate of $\sim 3 \times 10^{-9}$ per site per year for the MHC region estimated statistically by Hayashida and Miyata (1983), t^{12} and t^{w5} diverged from each other about 2 000 000 years ago.)

On the other hand, serological studies of *H-2K* and *D* antigens in *t*-haplotypes revealed only limited polymorphism among those molecules (Artzt *et al.*, 1985; Nižetić *et al.*, 1984), clearly different from the high degree of polymorphism in wild-type mice (reviewed by Klein and Figueroa, 1981; Kimball and Coligan, 1983). Thus, although our data suggest a long evolutionary history for *t*-haplotypes, the serological similarities of the *H-2* antigens compared to wild type imply that genetic recombination must play a very important role in generating polymorphism of *H-2* class I antigens.

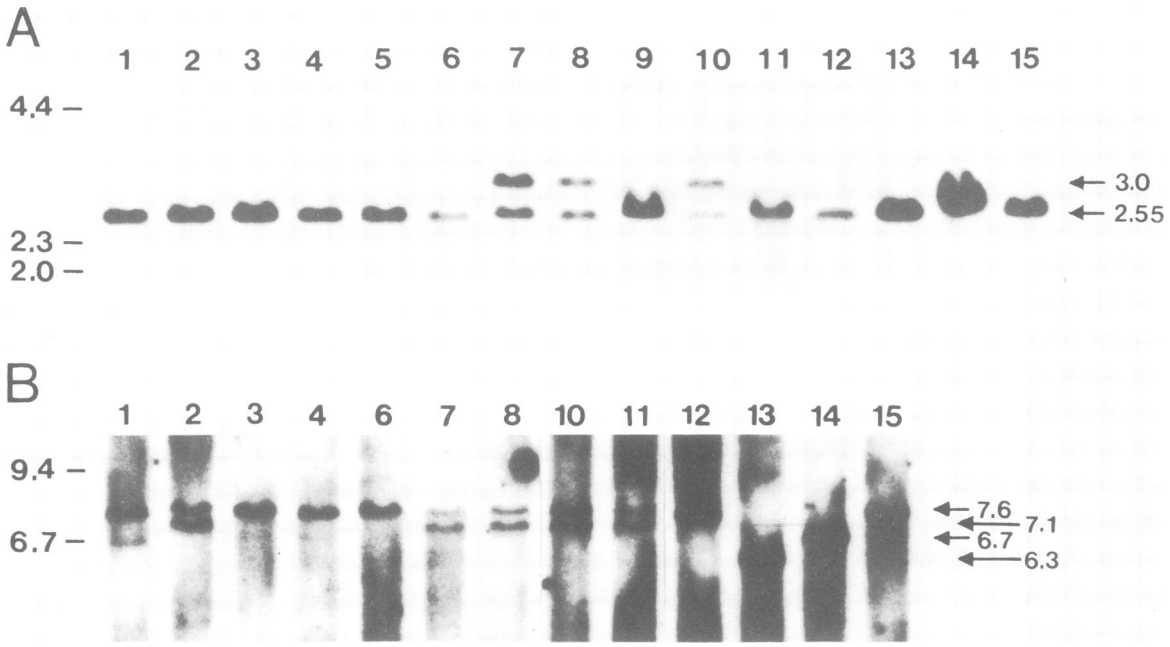

Fig. 5. Southern blot hybridization of *Bam*HI digested genomic DNA from various *t*-haplotypes and inbred strains. DNAs: **lane 1,** C3H.t^0/+; **lane 2,** C3H.t^{w1}/+; **lane 3,** C3H.t^{w5}/+; **lane 4,** t^{w5}/t^{w5} EC cell line; **lane 5,** C3H.t^{w5G}/+; **lane 6** , C3H.t^{w5G}/t^{w5G}; **lane 7,** C3H.t^{12}/+; **lane 8,** C3H.t^{12}/t^{w5}, **lane 9,** C3H.t^{w18}/+; **lane 10,** C3H.t^{w32}/+; **lane 11,** C3H.t^{w73}/+; **lane 12,** C3H; **lane 13,** BL/6; **lane 14,** BALB/c; **lane 15,** AKR. (**A**) Hybridization with LS 1/1. (**B**) Hybridization with the 0.6-kb *Hin*dIII fragment located at map position 10 kb in Figure 3 (indicated by arrow) isolated from clone 9-2-5.

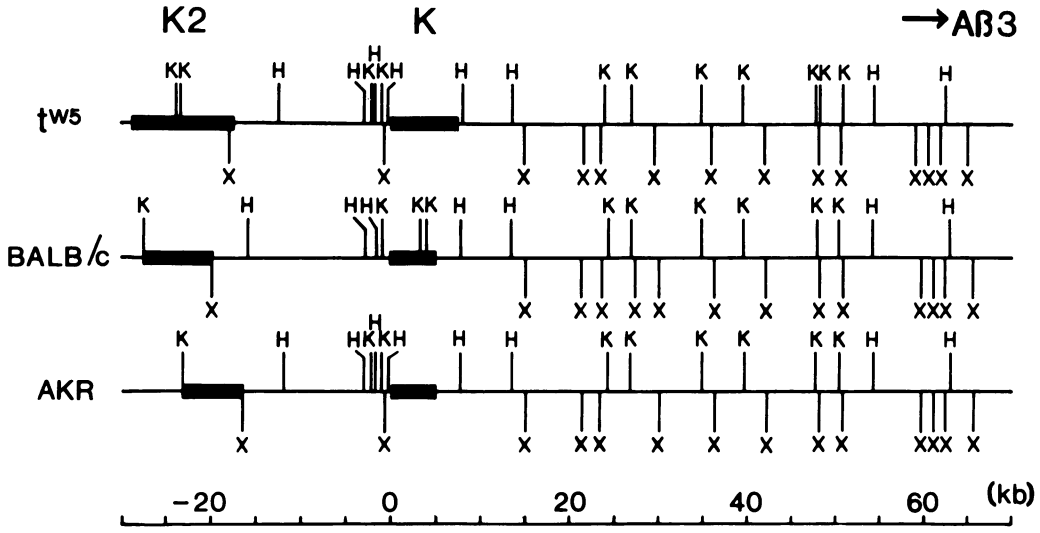

Fig. 6. Comparison of the restriction map in the *H-2K* region between t^{w5} and two inbred strains. Only the region available for comparison is shown. The map for BALB/c and AKR is taken from Steinmetz *et al.* (1984). The map was drawn as described in Figure 1. The *Kpn*I site at map position −1 was used as the starting point. H, *Hpa*I; K, *Kpn*I; X, *Xho*I.

Insertion of DNA elements as an important mechanism for generating genetic diversity

Surprisingly, the t^{12} and t^{w5} chromosomes show size differences at multiple sites along the *H-2K* region examined. These sites were localized by precisely determining the fragment sizes produced by restriction enzyme digests of cosmid DNA. While many studies have focused on the structure of *H-2* genes and their extensive flanking regions and provided information relevant to understanding genetic mechanisms that may produce *H-2* polymorphisms, little is known about the role of the insertion or deletion of short DNA segments in generating diversity in the flanking regions.

The following observations suggest that most, if not all, of the restriction fragment size differences are the result of insertion of DNA elements rather than deletion. First, we analyzed genomic DNA from nine *t*-haplotypes and four inbred mouse strains with three *K*-end probes (LS 1/1, Bm1-18 and 0.6-kb *Hin*dIII), which detect RFLPs between the t^{12} and t^{w5} chromosomes that appear to be caused by insertion/deletion of DNA segments. As shown in Figure 5A, in two cases (probes LS 1/1 and Bm1-18) the larger fragments are specific to those particular chromosomes, and all the other chromosomes examined retain the short form of the fragment. This result can most readily be explained by the insertion of DNA elements into the t^{12} chromosome.

A more complicated hybridization pattern was obtained with the 0.6-kb *Hin*dIII probe (Figure 5B). The result cannot be explained by single insertion/deletion of DNA element. This degree of polymorphism may imply frequent events of insertion/deletion in this region. Second, repetitive sequences (B2 and BAM5-R) were identified as the inserted element in at least two of 15 insertion/deletion sites. B2 is an Alu-like short interspersed repetitive sequence and BAM5-R is a portion of a major long interspersed repetitive sequence MIF-1 of mouse (Jelinek and Schmid, 1982; Rogers, 1985). Their properties, such as a functional RNA polymerase III promoter site, an A-rich 3′ end, direct terminal repeats, nonspecific insertion sites suggest, although they do not prove, that they are mobile elements. Thus, we believe that the chromosome size differences observed at multiple sites between the t^{12} and t^{w5} chromosomes were most likely caused by the insertion of repetitive DNA elements. Some of these are Alu-like B1 or B2 sequences, or BAM5-R; others may be mobile elements which have not yet been defined (Rogers, 1985).

The 15 sites we identified as insertion/deletion of DNA segments are a minimum estimate for two reasons. (i) The sensitivity of the detection of size differences depends on the sizes of the restriction fragments compared between the two genotypes, since when fragments are over 5 kb long, differences under 50 bp cannot be detected on agarose gels. (ii) The possibility remains that two or more independent insertion/deletion events were involved in any fragments defined as different in length. Thus, more insertions or deletions may be discovered by differing approaches, including DNA sequencing.

The question is whether these phenomena are limited to *t*-haplotypes or, more specifically to the MHC, or whether they represent general mechanisms for diversification of the genome. *t*-Haplotype chromosomes exist under special circumstances in that homozygotes for *t*-haplotypes are lethal. However, Paterniti *et al.* (1983) suggest that *t*-haplotypes do not appear to accumulate other 'parasitic' mutations at particularly high rates. It has been found that some class I genes also contain B1 or B2 sequences in introns or 3′ untranslated regions (Kress *et al.*, 1984; Pontarotti *et al.*, 1986), and Steinmetz *et al.* (1984) have reported restriction fragment size differences between BALB/c and AKR chromosomes in the I region in BALB/c and AKR chromosomes. Therefore, there is no indication that *t*-haplotypes are more prone to insertion/deletion of DNA segments than wild-type chromosomes. Nor is there any reason to think that insertional events are especially characteristic of the MHC, since retroposons have also been found in the introns and flanking regions of many genes (Rogers, 1985). DNA polymorphisms generated by Alu-like SINE insertions (Schuler *et al.*, 1983) and LINE insertions (Economou-Pachnis *et al.*, 1985) have also been reported. Sequence properties of SINE families and LINE families among different species suggest their recent amplification throughout the genome (Deininger and Daniels, 1986) and thus provide support for the notion of selfish or parasitic behavior of repetitive sequences (Orgel and Crick, 1980; Doolittle and Sapienza, 1980). Our evidence strongly suggests that insertion of mobile elements such as SINEs and LINEs into chromosomes is an ongoing general evolutionary mechanism for generating chromosome diversity.

Materials and methods

Mice

C3H t^{12}/t^{w5} compound mice were obtained as normal tailed progeny from crosses of tailess C3H·T/t^{12} × C3H·T/t^{w5} parents. The presence of two *t*-haplotype chromosomes in these mice was confirmed serologically, biochemically and by Southern blot hybridization with class II probes. C3H/DiSn and all C3H mice bearing *t*-haplotypes were obtained from our laboratory stock. BALB/c, AKR and C57BL/6 mice were the gift of E.A.Boyse. The t^{w5}/t^{w5} EC cell line was a gift of Terry Magnuson (Magnuson *et al.*, 1982).

Materials

All restriction enzymes were obtained from Bethesda Research Laboratory. Other enzymes were purchased from the following manufacturers: pronase, Calbiochem; ribonuclease, Sigma; T4 DNA ligase, Bethesda Research Laboratory; T4 polynucleotide kinase, New England Biolabs; calf intestinal phosphatase, Boehringer. [α-^{32}P]dCTP, [γ-^{32}P]ATP, and [^{35}S]methionine were obtained from New England Nuclear. Nitrocellulose filters used for genomic DNA hybridization and screening of the cosmid library were purchased from Millipore (HATF 0.45 μm pore size). Hybond-N was from Amersham. Packaging extract was purchased from Vector Cloning Systems.

Probes

Class I probes. pH-2IIa is a cDNA probe homologous to the fourth exon, transmembrane segment and first internal domain of class I genes (Steinmetz *et al.*, 1981). pH-2III is a cDNA probe homologous to the second and third exons of class I genes (Steinmetz *et al.*, 1981).

H-2K region probes. Probes Bm1-18 and LS 1/1 were described previously (Shin *et al.*, 1984). Bm1-R 3/1 is a genomic clone ~30 kb telomeric to the K^b gene in B10 mice: the 4.4-kb *Kpn*I fragment from cosmid clone Bm1-R 3/1 (R.A.Flavell, personal communication). 8D is a cDNA probe obtained from the 3′ untranslated region of the $H\text{-}2K^b$ gene and is homologous to all *K* genes: the 200-bp *Pst*I fragment from pBR322 clone 8D (Kress *et al.*, 1983).

H-2D region probes. 1.6-kb *D* and 0.3-kb *D* were described previously (Shin *et al.*, 1983).

Tla region probe. H11-TL is a 2.05-kb *Eco*RI fragment from cosmid H-11, a *Tla* clone isolated from the $H\text{-}2^b$ haplotype (Weiss *et al.*, 1984).

Qa region probe. Qa1432 is a genomic DNA probe obtained from the *Q4* gene, and a gift of L.Flaherty.

B1 probe. B1 was isolated from an intron of the α-fetoprotein gene, a gift of S.Tilghman (Adeniyi-Jones and Zasloff, 1985).

B2 probe. The B2 probe subcloned into SP6 from clone pMR142 was obtained from K.Bennett (Bennett *et al.*, 1984).

BAM5-R probe. The BAM5-R clone was obtained from L.Leinwand (Heller *et al.*, 1984).

0.6-kb HindIII probe. This fragment was obtained from cosmid clone 9-2-5. It maps at position 10 in Figure 1.

Construction of the cosmid library

A cosmid library was constructed as described by Steinmetz *et al.* (1982) in the vector pTL5 (Lund *et al.*, 1982) using liver DNA from C3H t^{12}/t^{w5} compound mice.

Screening of the cosmid library

We have screened ~6.4 × 10^5 cosmid clones using five probes. All probes were nick translated as described (Shin *et al.*, 1983) to a specific activity of 1−10 × 10^8 c.p.m./μg DNA. Filters were prepared as described by Grosveld *et al.* (1981) and hybridization was performed as described by Shin *et al.* (1983) with the five nick translated probes each used at 2−5 × 10^5 c.p.m. per ml hybridization solution.

Southern blot hybridization

Preparation of DNA, agarose gel electrophoresis, Southern transfer, hybridization with nick translated probes and washing filters were as described (Shin *et al.*, 1982, 1983). The hybridization with cosmid DNA was done using Hybond-N according to the method recommended by the supplier. For determining restriction fragments containing repetitive sequences, cosmid DNA (400 ng) was digested, electrophoresed, transferred to Hybond-N and hybridized with nick translated mouse liver DNA (specific activity 5−10 × 10^8 c.p.m./μg DNA) at a concentration of 2−3 × 10^6 c.p.m./ml of hybridization solution in the presence of 20 μg/ml denatured salmon sperm DNA. Blots were washed at low stringency (2 × SSC at 55°C for 15 min and 2 × SSC/0.1% SDS at 55°C for 15 min). In order to assure seeing faint bands, the films were overexposed.

Dot hybridization

Approximately 100 ng of cosmid DNA purified by small scale alkaline extraction methods (Maniatis *et al.*, 1982) was spotted on Hybond-N. All procedures followed the protocol provided by the supplier.

Restriction mapping of cosmid clones

All procedures are essentially the same as described by Rackwitz *et al.* (1984, 1985) with minor modifications. Cosmid DNAs were isolated from 50 ml over-

night cultures (Maniatis *et al.*, 1982). Terminase-treated cosmid DNA (100–300 ng) was mixed with carrier mouse DNA and partially digested with restriction enzymes. Hybridization was performed with 25 times molar excess of end-labeled oligomers and fragments were separated as described by Rackwitz *et al.* (1984) on 0.5% Tris-phosphate agarose gels at 1.5 V/cm for 15–17 h.

Estimation of nucleotide sequence diversity

The nucleotide sequence diversity among different chromosomes was estimated using equations 8 and 10 described by Nei and Li (1979). Variance was obtained according to equation 12. Only differences in restriction site patterns were scored for this calculation. Thus, neither the deletion/insertion events seen between t^{12} and t^{w5} nor the insertion 5′ to the *K2* gene in BALB/c were included in this value.

Acknowledgements

We thank Drs K.Bennett, R.A.Flavell, L.Flaherty, L.Hood, G.Jay, L.Leinwand, M.Steinmetz and S.Tilghman for their generous gifts of probes used in this study. We also thank Dr H.Lehrach for sending us bacterial strain AZ 1069. We are grateful to S.G.Nathenson for synthesis of the oligonucleotides used for cosmid mapping. We appreciate the comments and suggestions of S.Carleton, G.Gummere and E.Lader. The expert technical assistance of L.Fitton, B.Clark and A.DiMeo-Talento were greatly appreciated. We acknowledge A.Kennard for patient assistance in preparing the manuscript. This study was supported by NIH grants HD 10668 and CA 21651 and NCI Core Grant CA 08748.

References

Adeniyi-Jones,S. and Zasloff,M. (1985) *Nature*, **317**, 81–85.
Artzt,K., Shin,H.-S. and Bennett,D. (1982) *Cell*, **28**, 471–476.
Artzt,K., Shin,H.-S., Bennett,D. and DiMeo-Talento,A. (1985) *J. Exp. Med.*, **162**, 93–104.
Bennett,D. (1975) *Cell*, **6**, 441–454.
Bennett,D., Alton,A.K. and Artzt,K. (1983) *Genet. Res.*, **41**, 29–45.
Bennett,K.L., Hill,R.E., Pietras,D.F., Woodworth-Gutai,M., Kane-Haas,C., Houston,J.M., Heath,J.K. and Hastie,N.D. (1984) *Mol. Cell. Biol.*, **4**, 1561–1571.
Deininger,P.L. and Daniels,G.R. (1986) *Trends Genet.*, **2**, 76–80.
Doolittle,W.F. and Sapienza,C. (1980) *Nature*, **283**, 601–603.
Economou-Pachnis,A., Lohse,M.A., Furano,A.V. and Tsichlis,P.N. (1985) *Proc. Natl. Acad. Sci. USA*, **77**, 2856–2861.
Figueroa,F., Golubic,M., Nižetić,D. and Klein,J. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 2819–2823.
Forejt,J., Capkova,J. and Gregorova,S. (1980) *Genet. Res.*, **35**, 165–177.
Frischauf,A.-M. (1985) *Trends Genet.*, **1**, 100–103.
Grosveld,F.G., Dahl,H.-H., deBoer,E. and Flavell,R.A. (1981) *Gene*, **13**, 227–237.
Hammerberg,C. and Klein,J. (1975) *Nature*, **258**, 296–299.
Hayashida,H. and Miyata,T. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 2671–2675.
Heller,D., Jackson,M. and Leinwand,L. (1984) *J. Mol. Biol.*, **173**, 419–436.
Herrmann,B., Bucan,M., Mains,P.E., Frischauf,A.-M. and Lehrach,H. (1986) *Cell*, **44**, 469–476.
Hood,L., Steinmetz,M. and Malissen,B. (1983) *Annu. Rev. Immunol.*, **1**, 529–568.
Jelinek,W.R. and Schmid,C.W. (1982) *Annu. Rev. Biochem.*, **51**, 813–844.
Kimball,E.S. and Coligan,J.E. (1983) In Inman,F.P. and Kindt,T.J. (eds), *Contemporary Topics in Molecular Immunology*. Plenum Press, New York, Vol. 9, pp. 1–63.
Klein,J. (1975) *Biology of the Mouse Histocompatibility-2 Complex*. Springer Verlag, Berlin.
Klein,J. and Figueroa,F. (1981) *Immunol. Rev.*, **60**, 23–57.
Klein,J. and Hammerberg,C. (1977) *Immunol. Rev.*, **33**, 79–104.
Kress,M., Liu,W.-Y., Jay,E., Khouly,G. and Jay,G. (1983) *J. Biol. Chem.*, **258**, 13929–13936.
Kress,M., Barra,Y., Seidman,J.G., Khoury,G. and Jay,G. (1984) *Science*, **226**, 974–977.
Lund,T., Grosveld,F.G. and Flavell,R.A. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 520–524.
Lyon,M.F. (1981) In Berry,R.J. (ed.), *Biology of the House Mouse*. Academic Press, London, pp. 455–477.
Lyon,M.F. (1984) *Cell*, **37**, 621–628.
Magnuson,T., Epstein,C.J., Silver,L.M. and Martin,G.R. (1982) *Nature*, **298**, 750–753.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.
Nathenson,S.G., Uehara,H., Ewenstein,B.M., Kindt,T.J. and Coligan,J.E. (1981) *Annu. Rev. Biochem.*, **50**, 1025–1052.
Nei,M. and Li,W.-H. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 5269–5273.
Nižetić,D., Figueroa,F. and Klein,J. (1984) *Immunogenetics*, **19**, 311–320.
Orgel,L.E. and Crick,F.H. (1980) *Nature*, **284**, 604–607.
Paterniti,J.R., Brown,W.V., Ginsberg,H.N. and Artzt,K. (1983) *Science*, **221**, 167–169.
Pontarotti,P.A., Mashimo,H., Zeff,R.A., Fisher,D.A., Hood,L., Mellor,A., Flavell,R.A. and Nathenson,S.G. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 1782–1786.
Rackwitz,H.R., Zehetner,G., Frischauf,A.-M. and Lehrach,H. (1984) *Gene*, **30**, 195–200.
Rackwitz,H.R., Zehetner,G., Murialdo,H., Delius,H., Chai,J.H., Poustka,A., Frischauf,A. and Lehrach,H. (1985) *Gene*, **40**, 259–266.
Rogers,J.H. (1985) *Int. Rev. Cytol.*, **93**, 187–279.
Schuler,L.A., Weber,J.L. and Gorski,J. (1983) *Nature*, **305**, 159–160.
Shin,H.-S., Stavnezer,J., Artzt,K. and Bennett,D. (1982) *Cell*, **29**, 969–976.
Shin,H.-S., Flaherty,L., Artzt,K., Bennett,D. and Ravetch,J. (1983) *Nature*, **306**, 380–383.
Shin,H.-S., Bennett,D. and Artzt,K. (1984) *cell*, **39**, 573–578.
Silver,L.M. (1982) *Cell*, **29**, 961–968.
Silver,L.M. (1985) *Annu. Rev. Genet.*, **19**, 179–208.
Steinmetz,M., Frelinger,J.G., Fisher,D., Hunkapiller,T., Pereira,D., Weissman,S.M., Uehara,H., Nathenson,S.G. and Hood,L. (1981) *Cell*, **24**, 125–134.
Steinmetz,M., Winoto,A., Minard,K. and Hood,L. (1982) *Cell*, **28**, 489–498.
Steinmetz,M., Malissen,M., Hood,L., Orn,A., Maki,R.A., Dastoornikoo,G.R., Stephan,D., Gibb,E. and Romaniuk,R. (1984) *EMBO J.*, **3**, 2995–3003.
Sturm,S., Figueroa,F. and Klein,J. (1982) *Genet. Res.*, **40**, 73–88.
Weiss,E.H., Golden,L., Fahrner,K., Mellor,A.L., Devlin,J.J., Bullman,H., Tiddens,H., Bud,H. and Flavell,R.A. (1984) *Nature*, **310**, 650–655.
Willison,K.R., Dudley,K. and Potter,J. (1986) *Cell*, **44**, 727–738.
Zinkernagel,R.M. and Doherty,P.C. (1979) *Adv. Immunol.*, **27**, 51–177.

Received on 12 August 1986; revised on 14 October 1986