The EMBO Journal vol.4 no.1 pp.127–131, 1985

Inbred and wild mice carry identical deletions in their E_α MHC genes

Zlatko Dembič, Mohamed Ayane[1], Jan Klein, Michael Steinmetz[2], Christophe O.Benoist[1] and Diane J.Mathis[1]

Max-Planck-Institut für Biologie, Abteilung Immungenetik, D-7400 Tübingen, FRG, [1]Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Institut de Chimie Biologique, Faculté de Médecine, 67085 Strasbourg Cédex, France, and [2]Basel Institute for Immunology, Basel, Switzerland

Communicated by J.Klein

Several inbred strains and a certain percentage of wild mice bear a deletion in the E_α gene of the mouse major histocompatibility complex (*H-2*). This mutation prevents transcription of the E_α gene and hence functional expression of the $E_\alpha E_\beta$ dimer on the cell surface. Two strains were selected for a more precise localization of this deletion. BALB.B is a congenic line carrying the *H-2*^{b} haplotype on the BALB/c background. CRO435 is an outbred stock derived from a wild mouse captured near Cairo, Egypt; it carries the *H-2*^{w37} haplotype including a null E_α^{w28} allele, as well as semi-lethal mutations in the *H-2* linked *t* complex (t^{Tuw7}). From these two strains, we have isolated genomic clones that contain fragments spanning the E_α deletion, and have sequenced the breakpoint region. The deletions in the two strains are identical, spanning 627 bp which include the promoter region and the signal peptide exon of the E_α gene. Limited sequence comparison suggests that the E_α^b allele of BALB.B is more closely related to the E_α^{w28} allele of CRO435 than both of these are to an E_α-expressor allele, E_α^d. It is therefore likely that the E^o deletions in the various inbred strains and wild mice are of the same origin, and we propose that they have been disseminated throughout the mouse population because of linkage to the *t* complex.

Key words: major histocompatibility complex/deletion/*H-2*/*t* complex

Introduction

In the initial phase of the immune response, an antigen is presented to T lymphocytes together with molecules controlled by the individual's major histocompatibility complex (MHC; for review seen Klein *et al.*, 1983; Kaufmann *et al.*, 1984). Antigens integrated in the host-cell membranes are recognized in the context of class I MHC molecules, whereas soluble antigens are offered to regulatory T lymphocytes together with class II molecules. In the mouse, the functionally active class II molecules are controlled by four loci, A_α, A_β, E_α and E_β . The polypeptide chains controlled by these loci assemble as two heterodimers, the $A_\alpha A_\beta$ and the $E_\alpha E_\beta$ molecules. Both the A and E molecules seem to be important in regulating the immune response since some antigens are recognized in the context of the former and others in the context of the latter (Baxevanis *et al.*, 1980). Thus, it was somewhat surprising to find that all mouse strains so far tested express the A complex but certain mice do not assemble the E complex at the cell surface (Jones *et al.*, 1981).

Inbred laboratory strains that lack a cell-surface E complex (E^o strains) have a deficiency in intracellular protein levels of either E_α (haplotypes *H-2*^{b} and *H-2*^{s}) or both E_α and E_β (haplotypes *H-2*^{f} and *H-2*^{q}; Jones *et al.*, 1981). By analyzing the mRNA of mice that do not express E_α it has been established that at least three E^o phenotypes exist: mice of the *H-2*^{f} haplotype synthesize predominantly an aberrant E_α mRNA because of defective splicing (G.Garcin, C.O.Benoist and D.J.Mathis, in preparation); *H-2*^{q} mice have very little E_α mRNA, none of which is of the correct size; and mice of haplotypes *H-2*^{b} and *H-2*^{s} express no E_α mRNA (Mathis *et al.*, 1983a). In *H-2*^{b} and *H-2*^{s} mice, there is no transcription of the E_α gene because of a 650-bp deletion encompassing the promoter region and the first exon (Mathis *et al.*, 1983a; Hyldig-Nielsen *et al.*, 1983). Interestingly, the deletion appears to be identical for the two strains examined, at least at the level of resolution afforded by Southern blotting.

The failure to express an E complex is not limited to inbred laboratory strains of mice (Götze *et al.*, 1980). In particular, 18 *t* haplotype-carrying strains of wild mice were found to have an E^o mutation (Dembič *et al.*, 1984); these represent more than half of the known *t*-bearing strains. Among the E^o strains, 15 exhibit a deletion at the 5′ end of the E_α gene, and strikingly, the location and length of the deletion appears to be the same as in the *H-2*^{b} and *H-2*^{s} laboratory strains. This observation leads to the hypothesis that this E_α deletion is an old mutation that has been disseminated throughout mouse populations by the *t*-bearing chromosome (Dembič *et al.*, 1984).

To test this hypothesis, we have cloned the E_α gene from two strains bearing the deletion and sequenced the deletion end-points. We have chosen a laboratory strain BALB.B (*H-2*^{b}), and a wild strain, CRO435 (*H-2*^{w37}), for which there is no evidence of genetic relatedness. We find that the E_α deletion in these two strains is identical.

Results

Endpoints of the E^o deletion in H-2^b mice

A lambda genomic library was constructed from partially *Eco*RI digested BALB.B liver DNA, and was screened with an E_α probe, the 2.9-kb *Sst*I fragment depicted in Figure 1a. A clone containing the *H-2*^{b} haplotype E_α gene (Bb 2.9.1) was thus identified and isolated.

Figure 1b shows a partial restriction enzyme map of a 2.5-kb *Sal*I fragment at the 5′ end of the *H-2*^{k} and *H-2*^{d} E_α genes, both of which are transcribed normally. Previous results indicated that the *Sst*I and the two *Xmn*I sites were absent in mice bearing the 5′ E^o mutation (Mathis *et al.*, 1983a). To facilitate more precise localization of the deletion, we subcloned into pBR322 the corresponding 1.9-kb *Sal*I fragment from our BALB.B genomic clone, thus producing pSLEAB1.9. Restriction enzyme analysis on the purified *Sal*I fragment confirmed the absence of the *Xmn*I and *Sst*I sites and allowed us to make the tentative conclusion that the in-

© IRL Press Limited, Oxford, England.

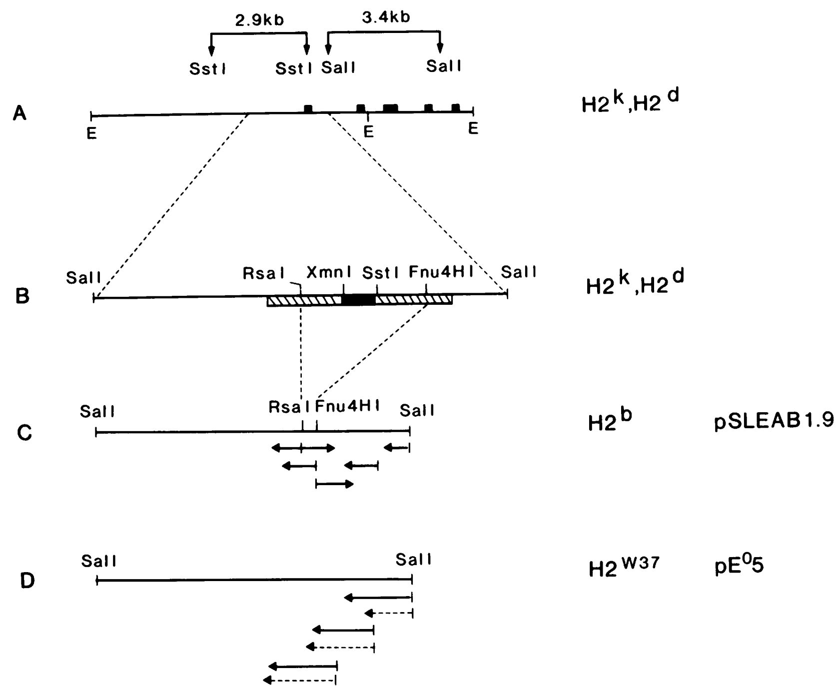

Fig. 1. E_α gene maps. **(A)** Configuration of the E_α genes in $H\text{-}2^k$ and $H\text{-}2^d$ mice (as per Hyldig-Nielsen *et al.*, 1983; Mathis *et al.*, 1983b). Thicker regions represent exons. E = *Eco*RI. **(B)** Partial restriction enzyme map of the 5′ 2.5-kb *Sal*I fragment in $H\text{-}2^k$ and $H\text{-}2^d$ mice (taken from Hyldig-Nielsen *et al.*, 1983; Mathis *et al.*, 1983a; and Ayane *et al.*, unpublished results). For any given enzyme, not all sites are shown, only those relevant to the discussion. Below the maps are the maximum possible limits of the 650-bp E^o deletion as described in Mathis *et al.* (1983b). The solid area represents the known deleted region of ~200 bp; the hatched areas demarcate an additional 450 bp on each side. **(C)** Sequencing strategy used for pSLEAB1.9. **(D)** Sequencing strategy used for pE°5. Solid lines, 5′ labeling; hatched lines, 3′ labeling.

dicated *Rsa*I and *Fnu*4HI sites were present in $H\text{-}2^b$ mice. Hence, these sites were used as end-labeled start-points for Maxam and Gilbert sequencing, as illustrated in Figure 1c.

The endpoints of the $H\text{-}2^b$ deletion are situated along the E_α sequence in Figure 2. The mutation boundaries are nucleotides 162–164 and nucleotides 787–789, excising 627 bases, in good agreement with our previous estimate of 650 ± 50 bases. Since the deletion starts and ends with an AT, it is not possible to determine the precise boundaries.

Endpoints of the E^o deletion in $H\text{-}2^{w37}$ mice

A cosmid genomic library was constructed from partially *Sau*3AI digested CRO435 liver DNA, and was screened with an E_α probe, the 3.4-kb *Sal*I fragment depicted in Figure 1a. Four positive clones, encompassing a region of ~70 kb, were isolated. Restriction enzyme mapping revealed considerable homology between this stretch of DNA and the corresponding regions of $H\text{-}2^k$ and $H\text{-}2^d$ mice (data not shown). The 1.9-kb *Sal*I fragment spanning the E^o deletion in cosmid clone 2.1 was subcloned into pUC8, and portions of the resulting clone, pE°5, were sequenced according to the strategy illustrated in Figure 1d. The sequence data (Figure 2) revealed exactly the same deletion end-points as described above for $H\text{-}2^b$ mice.

$H\text{-}2^d$, $H\text{-}2^b$ and $H\text{-}2^{w37}$ sequence comparison

The sequence shown in Figure 2 includes a stretch of 603 nucleotides which has been sequenced in all three haplotypes. This region is part of the intron separating exons 1 and 2. The $H\text{-}2^d$ sequence differs from that of $H\text{-}2^{w37}$ in 12 positions and from that of $H\text{-}2^b$ in 10 positions. The $H\text{-}2^{w37}$ and $H\text{-}2^b$ sequences differ at only four residues. Hence, $H\text{-}2^b$ and $H\text{-}2^{w37}$ mice appear to be more closely related to each other than either haplotype is to $H\text{-}2^d$ mice.

Discussion

A deletion at the 5′ end of the E_α gene has thus far been demonstrated in 16 *H-2* haplotypes *via* restriction enzyme analysis (Mathis *et al.*, 1983a; Dembić *et al.*, 1984). Five of these haplotypes are present in six strains of laboratory mice (*b, s, dx, sg4* and *w67*), nine occur in wild mouse strains carrying *t* haplotypes (*w28, w31, w36, w37, w38, w56, w59, w60* and *w63*) and two are present in other strains derived from wild mice (*w17* and *w53*). Restriction enzyme analysis suggested that the deletions may be the same in all 16 haplotypes, because their size and location within the E_α gene was indistinguishable. The demonstration here, that the $H\text{-}2^b$ and $H\text{-}2^{w37}$ haplotypes carry an identical E_α deletion renders it likely that in the other 14 haplotypes the deletions are also the same. If so, how did this mutation disperse throughout the various haplotypes?

One possibility is that the same deletion occurred repeatedly in different haplotypes. It has been reported that large deletions in the globin gene cluster may occur by exchange between locally homologous sequences at two different sites (Baglioni, 1963; Flavell *et al.*, 1978), while shorter deletions seem often to be bordered by direct repeats (Marotta *et al.*, 1977; Efstratiadis *et al.*, 1980). However, there is no striking homology between the sequences bordering the deletion, and the short AT repeat (Figure 2) seems a very unlikely cause for the repeated occurrence of precisely identical deletions.

Recently, Vanin and his co-workers (1983) have suggested that similar independent deletions in approximately the same regions of a gene may arise because the 5′ and 3′ breakpoints happen to be physically close in the nucleus (e.g., at the stem of a chromatin loop), although far apart on the linear DNA. Following this line of reasoning, the E_α deletions, too, might

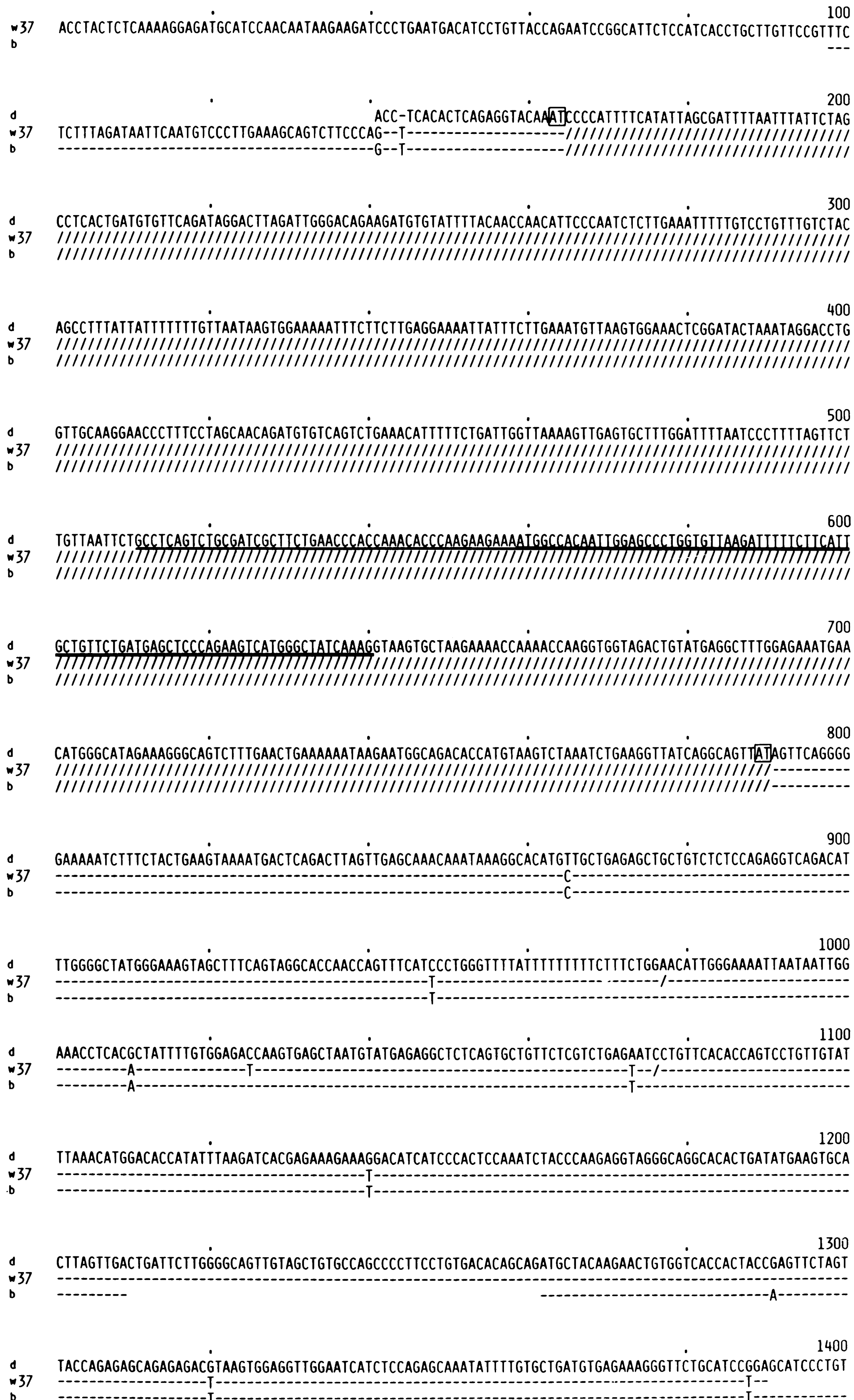

Fig. 2. Comparison of the E_α sequences in $H\text{-}2^b$, $H\text{-}2^{w37}$ and $H\text{-}2^k$ haplotype mice. The first exon is underlined with the thicker area representing the protein coding region. The directly repeated dinucleotides at the deletion boundaries are boxed. '—', indicates identity with top sequence; '/', indicates deletion.

be a consequence of DNA secondary structure or of chromatin superhelix formation. However, such a mechanism seems unlikely because one would not expect exactly the same breakpoints in different haplotypes. There is thus no obvious mechanism that could explain the repeated occurrence of the same deletion in the E_α gene, although it is, of course, possible that the borders of this deletion contain some as yet unrecognized feature that make them a 'hot spot' for repeated excision of an identical DNA stretch by some as yet unidentified mechanism.

An alternative possibility is that the 16 *H-2* haplotypes carrying the E_α deletion, although different, are all related to one another and that the deletion arose only once and was then disseminated throughout the mouse population. We favor this possibility for two reasons. The first is that, in fact, evidence for relatedness of the 16 *H-2* haplotypes is accumulating. In the present study we could show that the short determined sequence of the E_α^{0b} gene is clearly more closely related to that of the E_α^{w28} allele of the CRO435 strain than both of these sequences are to the sequence of an E_α gene that does not carry a deletion (e.g., E_α^{d}). This relatedness of E_α^{0b} and E_α^{w28} occurs in spite of the fact that the non-functional, non-expressed E^0 genes seem to diversify more rapidly than the expressed E_α genes, as evidenced by a higher incidence of restriction enzyme fragment length polymorphism (Dembič *et al.*, 1984). Relatedness of the 16 *H-2* haplotypes is also suggested by serological analysis of class I and class II molecules (Nižetič *et al.*, 1984). For example, $H\text{-}2^b$ and several of the *t*-bearing strains share the determinant H-2.33 which is, otherwise, very rare among wild mice. Also, the limited knowledge of the genealogy of the different inbred strains suggests several possible connections between the different E^0 haplotypes.

The second reason for believing that the E_α deletion occurred just once is that the *t* haplotypes provide a logical explanation for the spreading of a single mutation. Most of the naturally occurring haplotypes carry a complex of genes that distorts gene transmission through the sperm in favor of the *t* chromosome (reviewed in Klein and Hammerberg, 1977). In some matings, males transmit the *t* chromosome to nearly 100% of the progeny. Since the E^0 deletion is closely linked to the distorter gene by a strong suppressor of recombination acting on the entire centromeric arm of the *t-H-2* chromosome, it can be easily imagined how the original E_α deletion could have been spread throughout the world together with the *t* chromosomes. The presence of this deletion in non-*t* mice could then be explained by rare crossing over between *t* and *H-2*. According to this hypothesis, the E_α deletion arose a long time ago, in the original chromosome from which most if not all of the present-day *t* chromosomes derive. The *t* chromosomes spread the deletion into different mouse populations where it occasionally separates from the *t* haplotype. Once relieved of the crossing-over suppressor effect, the E^0 gene is free to recombine with other *H-2* genes and so it ends up in different *H-2* haplotypes.

This hypothesis may seem to be contradicted by the recent reports of an inversion associated with the *t* haplotypes and encompassing the entire *H-2* complex (Artzt *et al.*, 1982). An intra-*H-2* recombination should then represent crossing over within the inverted region and as such should lead to the formation of nullicentric and dicentric chromosomes. However, there is evidence that the *H-2* complex can separate itself by rare recombination from the *t* chromosome without any of the impairments which are predicted in the inversion hypothesis. An example of such a separation is the t^{h18} chromosome which carries the *H-2* complex of the original t^6 chromosome (Bechtol and Lyon, 1978) and yet recombines normally with wild-type chromosomes (Lyon and Meredith, 1964). These observations indicate that not all the *t* chromosomes carry an *H-2*-encompassing inversion and suggest that the spreading of the E_α^0 gene, as put forward here, is theoretically possible.

Whether the E_α deletion can persist stably in a population without *t* is not clear. Since some antigenic determinants are recognized exclusively in the context of E molecules, the absence of these molecules is probably a deleterious trait. However, since virtually all wild mice are *H-2* heterozygotes (Duncan *et al.*, 1979), mice that lack E molecules are probably rare outside of the laboratory. Nevertheless, a heterozygous mouse carrying the E_α deletion can express only one allelic form of the E molecule whereas a mouse with functional E_α genes expresses at least two E allomorphs. Because a particular antigen may be recognized in the context of only some E allomorphs, a mouse heterozygous for the E_α deletion might still be at a selective disadvantage. One could, therefore, imagine that the *t*-free E^0 haplotypes are slowly being eliminated from mouse populations, only to be replaced by deletions newly released from the *t* complex. Because *t* haplotypes are present in virtually all mouse populations, there is probably never a shortage of opportunities for the replenishment of *t*-free E_α^0 genes.

Materials and methods

Mice

CRO435 mice were produced in the breeding colony at the Max-Planck Institute for Biology, Tübingen. They were derived from a mating between a wild mouse captured near Cairo, Egypt and a laboratory strain. They carry the $H\text{-}2^{w37}$ haplotype and a semi-lethal t^{Tuw7} haplotype (Nižetič *et al.*, 1984). BALB.B ($H\text{-}2^b$) mice were bred in the laboratory of Dr. H.O.McDevitt, Stanford University, Stanford, CA.

DNA cloning and analysis

A BALB.B liver DNA genomic library was produced in Charon 4 exactly as described (Mathis *et al.*, 1983b). Six hundred thousand clones were screened with an E_α probe (The 2.9-kb *SstI* fragment from $H\text{-}2^k$ mice – see Figure 1a and Mathis *et al.*, 1983a). A positive clone, Bb2.9.1, was isolated and the 1.9-kb *SalI* fragment from this recombinant was subcloned into pBR322 to produce clone pSLEA1.9.

A CRO435 liver DNA genomic library was produced in cosmid pNNL by described methods (M.Steinmetz *et al.*, 1985). DNA was partially digested with *Sau*3AI, the digest size-fractionated on a sucrose gradient, 30–50 kb fractions isolated and ligated to the purified *Bam*HI-digested arms of pNNL, and the ligated mix packaged and used to infect *Escherichia coli* 490 A. One million clones were screened with an E_α probe (the 3.4-kb *SalI* fragment from cos $H\text{-}2^d$ 32.1, see Figure 1a). Cosmid clone 2.1 was used for further analysis. The 1.9-kb *SalI* fragment was subcloned into pUC8 to produce clone pE°5.

DNA sequencing

Sequences were determined according to the method of Maxam and Gilbert (1980). Deletion subcloning of pE°5 was performed as described by Frischauf *et al.* (1980). Relevant regions of pSLEAB1.9 were sequenced after 5′ end-labeling of the *RsaI*, *Fnu*4HI or *HinfI* sites.

Acknowledgements

We gratefully acknowledge the technical assistance of Mr. P.Gerber, the secretarial assistance of C.Aron and the encouragement and support of Professor P.Chambon. This work was supported in part by a grant from the National Cancer Institute, National Institutes of Health, Bethesda, MD and in part by funds from the Foundation Simone and Cino del Duca. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche, Limited Company, Basel, Switzerland.

References

Artzt,K., Shin,H.-S. and Bennett,D. (1982) *Cell,* **28**, 471-476.
Baglioni,C. (1963) *Proc. Natl. Acad. Sci. USA,* **48,** 1880-1886.
Baxevanis,C.N., Wernet,D., Nagy,Z.A., Maurer,P.H. and Klein,J. (1980) *Immunogenetics,* **11**, 617-628.
Bechtol,K.B. and Lyon,M.F. (1978) *Immunogenetics,* **6**, 571-583.
Dembič,Z., Singer,P.A. and Klein,J. (1984) *EMBO J.,* **3**, 1647-1654.
Duncan,W.R., Wakeland,E.K. and Klein,J. (1979) *Nature,* **281**, 603-605.
Efstratiadis,A., Posakony,J.W., Maniatis,T., Lawn,R.M., O'Connell,C., Spritz,R.A., DeRiel,J.K., Forget,B.G., Weissman,S.M., Slightom,J.L., Blechl,A.E., Smithies,O., Baralle,F.E., Shoulders,C.C. and Proudfoot, N.J. (1980) *Cell,* **21**, 653-688.
Flavell,R.A., Kooher,J.M., De Boer,E., Little,P.F.R. and Williamson,R. (1978) *Cell,* **15**, 25-41.
Frischauf,A.M., Garoff,H. and Lehrach,H. (1980) *Nucleic Acids Res.,* **8**, 5541-5549.
Götze,D., Nadeau,J., Wakeland,E.K., Berry,R.J., Bonhomme,F., Egorov, I.K., Hjorth,J.P., Hoogstraal,H., Vives,J., Winking,J. and Klein,J. (1980) *J. Immunol.,* **124**, 2675-2681.
Hyldig-Nielsen,J.J., Schenning,L., Hämmerling,U., Widmark,E., Helding, E., Lind,P., Servenius,B., Lund,T., Flavell,R., Lee,J.S., Trowsdale,J., Schreier,P.H., Zablitzky,F., Larhammer,D., Peterson,P.A. and Rask,L. (1983) *Nucleic Acids Res.,* **11**, 5055-5071.
Jones,P., Murphy,D. and McDevitt,H. (1981) *Immunogenetics,* **12**, 321-337.
Kaufman,J.F., Auffray,C., Korman,A.J., Shackelford,D.A. and Strominger,J. (1984) *Cell,* **36**, 1-13.
Klein,J., Figueroa,F. and Nagy,Z.A. (1983) *Annu. Rev. Immunol.,* **1**, 219-242.
Klein,J. and Hammerberg,C. (1977) *Immunol. Rev.,* **33**, 71-104.
Lyon,M.F. and Meredith,R. (1964) *Heredity,* **19**, 301-312.
Marotta,C.A., Wilson,J.T., Forget,B.G. and Weissman,S.M. (1977) *J. Biol. Chem.,* **252**, 5040-5053.
Mathis,D.J., Benoist,C.O., Williams,V.E., II, Kanter,M.R. and McDevitt, H.O. (1983a) *Proc. Natl. Acad. Sci. USA,* **80**, 273-277.
Mathis,D.J., Benoist,C.O., Williams,V.E.,II, Kanter,M.R. and McDevitt, H.O. (1983b) *Cell,* **32**, 745-754.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.,* **65**, 499-560.
Nižetič,D., Figueroa,F. and Klein,J. (1984) *Immunogenetics,* **19**, 311-320.
Steinmetz,M., Stephan,D., Dastoornikoo,G.R., Gibb,E. and Romanink,R. (1985) in Lefvovits,I. and Pernis,B. (eds.), *Immunological Methods*, Vol. **III**, Academic Press, in press.
Vanin,E.F., Henthorn,P.S., Kioussis,D., Grosveld,F. and Smithies,O. (1983) *Cell,* **35**, 701-709.

Received on 25 September 1984; revised on 29 October 1984