

The nucleotide sequence of the mouse immunoglobulin epsilon gene: comparison with the human epsilon gene sequence

Norio Ishida, Shintaro Ueda, Hidenori Hayashida¹, Takashi Miyata¹, and Tasuku Honjo*

Department of Genetics, Osaka University, Medical School, Osaka 530, and ¹Department of Biology, Faculty of Science, Kyushu University, Fukuoka 812, Japan

Communicated by K. Rajewsky
Received on 11 August 1982

We have determined the nucleotide sequence of the immunoglobulin epsilon gene cloned from newborn mouse DNA. The epsilon gene sequence allows prediction of the amino acid sequence of the constant region of the epsilon chain and comparison of it with sequences of the human epsilon and other mouse immunoglobulin genes. The epsilon gene was shown to be under the weakest selection pressure at the protein level among the immunoglobulin genes although the divergence at the synonymous position is similar. Our results suggest that the epsilon gene may be dispensable, which is in accord with the fact that IgE has only obscure roles in the immune defense system but has an undesirable role as a mediator of hypersensitivity. The sequence data suggest that the human and murine epsilon genes were derived from different ancestors duplicated a long time ago. The amino acid sequence of the epsilon chain is more homologous to those of the gamma chains than the other mouse heavy chains. Two membrane exons, separated by an 80-base intron, were identified 1.7 kb 3' to the CH4 domain of the epsilon gene and shown to conserve a hydrophobic portion similar to those of other heavy chain genes. RNA blot hybridization showed that the epsilon membrane exons are transcribed into two species of mRNA in an IgE hybridoma.

Key words: mouse epsilon gene/nucleotide sequence/membrane exons/high evolutionary rate

Introduction

Immunoglobulins are classified into five major classes, IgM, IgG, IgA, IgD, and IgE, by the primary structure of the heavy (H) chain constant (C) regions mu, gamma, alpha, delta, and epsilon, respectively. Each class has a different physiological function. IgE is unique in that it mediates an undesirable reaction, reaginic hypersensitivity, although it is not yet clear whether this immunoglobulin plays an essential role in the immune defense system mechanism. Structural analyses of the epsilon gene and the protein it codes are essential to understand the molecular mechanism which triggers the series of events ultimately leading to the release of vasoactive amines and other pharmacologically active substances responsible for the clinical manifestations of hypersensitivity (Ishizaka and Ishizaka, 1975).

We have already reported the cloning of the mouse epsilon gene, which showed that the epsilon gene is located between the gamma 2a and the alpha genes (Nishida *et al.*, 1981). We have now determined the nucleotide sequence of the gene coding for the secreted and membrane-bound forms of the mouse epsilon chain and compared it with that of the human

epsilon gene (Nishida *et al.*, 1982; Ueda *et al.*, in preparation). Such comparison revealed that the epsilon gene diverged much more rapidly than others, suggesting that the epsilon gene might be dispensable.

Results

Sequencing strategy

The clone (IgE-7) bearing the mouse genomic epsilon gene was isolated from a Charon 28 phage library containing partial *Sau3A* digests of DNA from newborn BALB/c mice (Nishida *et al.*, 1981). The 4.6-kb *Bam*HI fragment containing the structural sequence of IgE-7 was subcloned into pBR322. Restriction site mapping was carried out by cleavage of the clone with several restriction enzymes and by the partial cleavage of terminally ³²P-phosphorylated fragments. The restriction sites used to determine the nucleotide sequence are shown in Figure 1.

Nucleotide sequence of the mouse immunoglobulin epsilon gene

The complete nucleotide sequence of the epsilon gene is presented in Figure 2. The sequence predicts the entire amino acid sequence of the C region of the secreted form of the mouse IgE H chain. The epsilon gene has four exons, each encoding a domain and being interrupted by the intervening sequence (IVS). Each domain was defined on the basis of homology to the amino acid sequence of the human epsilon chain (Bennich *et al.*, 1974; Ueda *et al.*, in preparation). The borders of each domain were determined by the consensus sequences of RNA splicing (Breathnach *et al.*, 1978; Konkel *et al.*, 1978) and by the distance from two cysteine residues responsible for the intra-chain disulfide bond. The structure of the mouse epsilon gene is summarized as C_{H1} exon (273 bp)-IVS1 (542 bp)-C_{H2} exon (324 bp)-IVS2 (81 bp)-C_{H3} exon (324 bp)-IVS3 (78 bp)-C_{H4} exon (350 bp)-3' untranslated sequence (107 bp).

In comparison with other classes of immunoglobulin genes the epsilon gene has the longest IVS between the C_{H1} and C_{H2} exons (Honjo *et al.*, 1979; Kawakami *et al.*, 1980; Tucker *et al.*, 1980, 1981; Yamawaki-Kataoka *et al.*, 1980, 1981). At the 5' portion of the IVS1 there is an extremely AT-rich region (217 bp) surrounded by inverted repeat sequences (13 bp). Such structure suggests that this 217-bp segment might have been inserted into the IVS1 of the epsilon gene, thus increasing the length of IVS1 from the usual 300–350 bp to 542 bp.

The amino acid sequence predicted has six putative sites of carbohydrate attachment, Asn-X-Ser(Thr) (Figure 2), the same number as in the human epsilon chain. Four of these six sites are in fact identical to those of the human epsilon chain. There is one intra-domain disulfide bond in each domain, one inter-chain (light-H) bond in the C_{H1} domain and there are two inter-chain (H-H) bonds in the C_{H2} domain. An extra intra-domain disulfide bond in the C_{H1} domain of the human epsilon chain is not present in the mouse epsilon chain.

Identification of membrane exons of the epsilon gene

As all the C_H genes so far examined have two extra exons

*To whom reprint requests should be sent.

	20	40	60	80	100
GAATCAACCATCTGTCTCTAGACTGGAATGGGGTCCCAGAGCCCTGCTCTGTGCACAGTGCCTTAAACGAGTCCCCACGCTGCAGGTATGATGTG	100				
ATATAAATAAGTCTAACCTAGGTCTCTCTTCTACAGCCCTTATCAGGAACCCCTACGCTCTACCCTTAAAGCCCTGTAAAGGCACCTGCTTCATAGAC	200				
SerLEuArgAsnProGlnLeuTyrProLeuLysProCysLysGlyIThrAlaSerMetTh CCTAGGCTGCCCTAGTAAAGGACTACTTCCCTAATCCTGTGACTGTGACCTGGTATTCACAGCTTGCACCTGAACATGAGCAGCTTGAACCTCCCTGCCCTCGGT	300				
RLeuGLYCYsLEuVALLysAspTyrPheProAsnProValThrValThrTrpTyrSerAspSerLeuAsnMetSerThrValAsnPheProAlaLeuGLY TCTGAACTCAAGGTACACCCAGGCAAGTGACCACTGGGCGCAAGTCAGCCAAAGACTTCACATGCCAGCTGACACATCTCTCATCTTCAACGAAAGTA	400				
SERGLUeULysVALThrThrSERGLNValThrSerTrpGlyLysSERAlaLysAsnPheThrLysHisVALThrHisProProSerPheAsnGLUeUSerA GGACTATCTAGGTAAGTAGGGATGGGTGACAGTTACACTGTGATTCTCCCTTGGAGATGGAACAGATTCTTGCTATATCAATCAGGAACATGTCACAAATTC	500				
RGThrLEuLeuV					
CTTTCATAGGAGCTTCTAAGAGATTTTTTCTACTTCTATCATGTTTAGTGATCCAAATAGATTTTAAAAACTGGTGGATGATATTACTTTTAGC	600				
CTCAGAAGACATCAGGCACATTTAAGAGGCATTTAACTATTGTAAATTATTCGTGACTTAAAAATGCTCAGGACTGAGTCGTATATTTTAAATAAA	700				
TTTTATTAGTTTAGTTAAAAAAGAAAAGAAATATTAAATTTATTA	800				
AAAAAATCTCTATAJTATTA TAAAAAAGAGGAGAAAAGGCAGAGCTGGGCTGAG	900				
CTACGGTTACCAAGGACAATGGTCAGGAGGAGGAGGACTCTTATACAACCTATGACAGGAGAAGGGAGACCAACATCACTCGGGGGCCTACCTTCG	1000				
AGAGAACAAGGCGACGGGCAATACACAGAGCAGGTGTTTCGCCCTGCCCTAGTTCGACCTGTACACATCTCAGTGCACCACTTGGAGCTACTTCACT ALArgProValThrHisSerLeuSerProProTrpSerTyrSerIL	1100				
TCTCGCTGCGACCCCAATGCATTCCTTCCACCATCCAGCTGTACTGCTTCAATGATGCCACATCTAAATGATGTCCTCGGCTGAGTAAATGGAC EHisARGCysAspProAsnAlaPheHisSerThrILEGnLLeuTyrCYsPHEILETyrGLYHisILELeuAsnAspValSerLeuTrpLeuMetASP	1200				
GATCGGGAGATAACTGATACACTTGCACAACTGTTCTAATCAGGAGGAAAGGCAAACTAGCCCTCTACCTCAGTAACTCAACATCACTGAGCAGCAAT ASPArgGLUILEThrASPThrLeuVALAGnLThrVALLeuILELysGLUUGLUGLYLysLEuVALSerThrCYsSerLysLEuASnILEThrGLUGnLGNL	1300				
GGATGCTGAAGACCTTCACTGCAAGGTCACCTCCCAAGGCTAGACTAATTTGGCCACACTCGGAGATGCCAGGTAGGTTCTACACTCGCCTGATG RMetSERLeuSerThrPheThrLysARGValThrSERGLNGLYValASPTryrLEuAlaHisThrARGArgCYsProA	1400				
CCCAGACCTCAGAGTCCCTGAGGGAAGGCGAGGCTACACAGCCCTTCTCCCGCAGACATGATGAGCCAGGGGGCGGATTACTTACTGATCCACCC SPHisGLUProArgGLYValILEThrTrpLEuLeuProPro	1500				
AGCCCCCTGGACCTGTATCAAAACGGTGCTCCAAGCTTACTGCTGCTGGTGGTGGACCTGGAAAGCGAGAAAGTATCAATGTGACCTGGAAACAGAGA SerProLeuASPLeuTyrGLNAsnGLYAlaProLysLEuThrCYsLEuValValASPLeuGLUeUSerGLULysAsnValAsnValThrTrpAsnGLNGLU	1600				
AGAAGACTTCAGTCTCAGACTCCAGTGATACCTAAGCACAACAAAGCCACAAGTATCACCCTCATCTGCTGTAGTGTGCAAGGACTGAT LysTyrHisSERValAlaSerALEuSerILEThrTyrThrLYsHisHisAsnAsnAlaThrHisSERLeuILEuProValValASPLeuTrpTrpIL	1700				
TGAAGGCTACGGCTATCAGTGTGCTGATGGACCGCCCTGATTTTCCCAAGGCCAATTTGGCTTCCATCACCTTCCCCAGGTTAGTACAGGAGGTTGAGAG EGLUGLYLyrGLYTrpGLNLYsValValASPArgProASPheProLysProILEValARGSerILEThrLeuProGLNVALS	1800				
TGGCCAGCCCTTCTCATGTTACAGAGAATCGTTAACTGGTTAAGTCATGCTGCTCCACAGGCCAGCGCTCAGCCCCAGGTTATGTGTGCCACCA ERGLNARGValAlaProGLUeULyrValPheProProP	1900				
CAGAGGAGGAGAGCGAGGACAAACGCACATCTACCTGTTTGATCCAGAACTTCTCCCTGAGGATATCTTGTGCGAGTGGCTGGGGGATGGCAACTGAT ROGLUUGLUGLUSerGLUASPLeuArgThrLeuThrCYsLEuILEGLNAsnPheProGLUeUASPLeuSERVALGnLTrpLeuGLYAspGLYLYsLEuIL	2000				
CTCAAACAGCCGACAGACTACCAACACCCCTGAATCAATGGCTTCCCAATCAAGGCTTCTTCTCATCTCAGTGGCTCAGGAGTGGCAAGACATCTGG ESerAsnSERGLNHisSerThrThrThrTrpProLeuLysSerAsnGLYSerAsnGLNGLYPhePHEILEPheSerARGLeuGLUValALALysTThrLeuTrp	2100				
ACACAGAGAAAAGGTTCACTTGCCAAGTGATCATGAGGCATGCAAGAAACCGAGAAACTGAGAAACATATCCCAAGACGCTTGTATCAACACTCCC ThrGLNArgGLYsGLNpheThrLYsGLNLeuILEHisGLUALeUeUGnLysProARGLysLEUGLUYsThrILESerThrSERLeuGLYAsnThrSERIL	2200				
TCGTCCCTCTAGGCTCCATGTAGCTGTGGTGGGGAAGTGGATGACAGACATCCGCTCAUGCTTGTGAACACAGGAAGCTACCCGATAAACTCA EUArgProSer	2300				
GTGGCTGATTAGAGCCCTGGGCGCCGTTCTTCGGGAGAGCAGGGGTTACTGGGCGAGAAATATCTCGGCTTGAAGAAAGGGACACCCAGAGAAGGACAG	2400				
GAGTGAAGCAGCTGGATCCCATCTGTCTGTGTGTAATATTTAAACATAAGCGTCAAGGACTCAGGACTGAGTGGTCTTCACGCTCCCTGATCTTCTG					
CTGCTCTCAGCTGGATATCATGCGGCTGATCTAGA					

1118

[illegible]

M2 exon was shown to be transcribed into the mRNA of the membrane form. As the sum of the length of all exons is ~1.8 kb, the 3' untranslated sequence of the mRNA of the membrane form is at least 1.35 kb. The primary structural difference of the two membrane-form mRNAs could be due to the presence of a third membrane exon, to RNA splicing within the 3' untranslated region (Cheng *et al.*, 1982), to aberrant splicing in the hybridoma, or to the presence of the cytoplasmic precursor molecule to the mature mRNA.

Discussion

Comparison of nucleotide sequences of the murine and human epsilon genes

The nucleotide sequence of the mouse epsilon gene was compared with that of the human epsilon gene determined in our laboratory (Nishida *et al.*, 1982; Ueda *et al.*, in preparation; Max *et al.*, 1982; Flanagan and Rabbitts, 1982). The two sequences were aligned to maximize the homology with a computer program described previously (Miyata and Yasunaga, 1980). The amino acid sequences predicted from the nucleotide sequences of the epsilon genes were also aligned simultaneously (Figure 5).

The divergence of the amino acid sequences (k_{aa}) was calculated for each domain as well as for the whole C region. The rate of evolution, defined as per residue per year (v_{aa}), was evaluated assuming that the time of divergence between man and mouse is 75×10^6 years. The v_{aa} values of the epsilon chain are extremely high compared with those of the other immunoglobulin chains as shown in Table I. The v_{aa} value (5.2×10^{-9}) of the epsilon chain is the highest among known values except for the value (9.0×10^{-9}) of fibrinopeptides (Wilson *et al.*, 1977). We have also found that the delta chain has a very high rate of evolution ($v_{aa} = 5.9 \times 10^{-9}$). In contrast, the rest of the immunoglobulin chains including the mu, gamma, alpha, lambda, and kappa chains have v_{aa} values ranging from 3.5×10^{-9} to 2.3×10^{-9} . These results indicate that the epsilon and delta chains are under weaker constraint than the other immunoglobulin chains. The last domain of the epsilon chain is the most conserved, probably because the C_{H4} domain is the most important part of the molecule for binding to the surface receptor of the mast cells (Ishizaka *et al.*, 1970).

The divergence of the nucleotide sequences of the coding region was calculated separately for synonymous (K_s) and amino acid-substituting (K_a) sites. Nucleotide divergence at the synonymous sites of the coding regions is free from the selective constraint operating at the protein level. The divergence of the non-coding region (K_n) was calculated by comparing every nucleotide position. It is known that K_s and K_n values are constant among different genes, whereas K_a values vary among different genes (Miyata *et al.*, 1980). The K_s and K_n values of the epsilon gene are about the same as those of the gamma gene as shown in Table II. In contrast, the K_a value of the epsilon gene is almost 50% higher than that of the gamma gene. The K_s (0.484) of the mu gene, based on a partial sequence, is similar to that of the gamma gene but the K_a value (0.171) is slightly lower. These results, as expected, agree with the comparison of the amino acid sequences and clearly indicate that the epsilon gene is relatively

free from the selective constraint operating at the protein level, although the mutation rate at the nucleotide positions irrelevant to the protein structure is equivalent to those of the gamma and mu genes. The K_a value of the C_{H4} exon is close to those of the gamma exons, again in agreement with the amino acid sequence comparison. Note that the hinge exon of the gamma gene seems to diverge most rapidly.

Why does the epsilon chain drift faster than the other H chains? Since the difference is only in the K_a value but not in the K_s or K_n value, this phenomenon is most likely to be related to the property of the protein – either structural or functional characteristics of the epsilon chain. Although the epsilon chain has a structure similar to that of the mu chain, i.e., four C domains, the epsilon chain or IgE has a unique physiological function. IgE mediates an undesirable and harmful reaction, hypersensitivity, whereas it is not clear whether IgE is essential to the immune defense system. Although IgE may help to protect the animal from parasite infections to a limited extent, the epsilon gene may be dispensable. In general, functionally less important molecules evolve faster (Dickerson, 1971; Kimura and Ohta, 1974; Wilson *et al.*, 1977). From such a view point it might be reasonable to assume that the epsilon chain is not constrained strongly by the selection pressure to conserve the protein structure but rather allowed to drift relatively faster to evolve a less harmful H chain molecule. Note that the delta chain, which does not have any clear physiological role in the defense mechanisms, because of its extreme lability in the serum, also seems to be less constrained by selection pressure at the protein level. It is also interesting that both epsilon and delta chains are least abundant in the serum.

Another explanation which we have to take into account is that the human and murine epsilon genes may have different ancestors. In this case we assume that the duplication of the epsilon gene took place before mammalian radiation (paralogous). The ancestor of the murine epsilon gene might have turned into a pseudogene or been lost from the human genome. On the other hand, the ancestor of the human epsilon gene must have been lost from the mouse genome, which has only one epsilon gene but no epsilon pseudogene. This model can be easily tested by analyzing the organisation of the epsilon gene in many mammals. The presence of two epsilon genes in the rat genome (Y. Yaoita and T. Honjo, unpublished data) is consistent with this model.

Comparison of amino acid sequences of mouse immunoglobulin epsilon and other H chains

The amino acid sequence of the murine epsilon chain was compared with those of other murine H chains. As shown in Table III, the epsilon chain is more homologous to the gamma chains than to the mu, delta, and alpha chains. The homology between the epsilon and gamma chains is most significant in the M1 exon, followed by the C_{H1} and C_{H4} exons. The amino acid sequence of the hydrophobic segment of the epsilon gene is 74% and 69% homologous to those of the gamma 2a and the mu genes, respectively, while the homology of the other coding region is <41% (Figure 3B). The results indicate that the transmembrane segment is under a stronger constraint from selective pressure than the other parts of the H chain. Note that the lengths of the M2 exons of

Fig. 5. Aligned nucleotide and amino acid sequences of mouse and human epsilon genes. The human (upper two rows) and mouse (lower two rows) sequences are aligned to maximize homology. Amino acid sequences are shown by one letter code. The matched nucleotides and amino acids are shown by : and *, respectively. Double dashes indicate deletions to maximize the homology. Boxed regions were omitted from calculation of divergence because frame-shift mutations completely changed amino acid residues.

Table I. Comparison of amino acid sequences of human and murine immunoglobulin chains

Immunoglobulin chains	C _{H1}		C _{H2} (hinge)		C _{H3} (C _{H2})		C _{H4} (C _{H3})		Total chain	
	k_{aa}^C	v_{aa}^*	k_{aa}^C	v_{aa}^*	k_{aa}^C	v_{aa}^*	k_{aa}^C	v_{aa}^*	k_{aa}^C	v_{aa}^*
mu									0.43	2.9
delta									0.89	5.9
gamma ^a	0.45	3.0	1.03	6.9	0.42	2.8	0.52	3.5	0.48	3.2
epsilon	0.83	5.5	0.96	6.4	0.83	5.5	0.57	3.8	0.78	5.2
alpha									0.53	3.5
lambda ^b									0.35	2.3
kappa									0.51	3.4

Sources of sequences are as follows; human mu (Rabbitts *et al.*, 1981; Dayhoff, 1978), mouse mu (Kawakami *et al.*, 1980), human gamma (Ellison *et al.*, 1981), mouse gamma (Honjo *et al.*, 1979; Yamawaki-Kataoka *et al.*, 1981, 1982), human alpha (Tsuzukida *et al.*, 1979), mouse alpha (Tucker *et al.*, 1980), human and mouse kappa (Hieter *et al.*, 1980), human lambda (Hieter *et al.*, 1981), mouse lambda (Bothwell *et al.*, 1981), and human and mouse epsilon (this paper). v_{aa}^* is equal to $v_{aa} \times 10^6$. Notation of domains of the gamma chain are shown in parenthesis.

^aThree mouse gamma sequences, gamma 1, gamma 2a and gamma 2b were separately compared with human gamma 4 sequence. The k_{aa}^C values of three pairs were averaged.

^bMouse lambda 1, sequence was separately compared with three human lambda chains, mcg, KernOZ⁺ and OZ⁻. The k_{aa}^C values of three pairs were averaged.

Table II. Comparison of nucleotide sequences of the immunoglobulin gamma and epsilon genes between man and mouse

Segment of gene	Gamma gene ^a			Epsilon gene		
	K_a	$K_s(K_n)$	K_a/K_s	K_a	$K_s(K_n)$	K_a/K_s
5' flanking		0.380			0.353	
IVS1		0.388			0.419	
IVS2		0.371			0.349	
IVS3		0.360			0.375	
total non-coding		0.375			0.374	
C _{H1}	0.220	0.492	0.44	0.299	0.484	0.617
C _{H2} (hinge)	0.429	0.664	0.646	0.339	0.546	0.621
C _{H3} (C _{H2})	0.157	0.500	0.314	0.336	0.525	0.64
C _{H4} (C _{H3})	0.214	0.486	0.44	0.239	0.532	0.449
total coding	0.204	0.502	0.406	0.301	0.521	0.578

Sources of sequences are shown in Table I. Notations of domains of the gamma chain are shown in parenthesis.

^aHuman gamma₄ and mouse gamma₁ gene sequences were compared. Similar values were obtained with other combinations.

the epsilon and gamma genes are identical, whereas those of the mu and delta genes are much shorter (Figure 3). Such homology might be due in part to gene conversion between gamma and epsilon genes. Nikaido *et al.* (1982), however, have shown that the switch (S) region sequence of the mu gene, which is responsible for class switch recombination, has greatest homology to that of the epsilon gene. Such paradoxical dissociation of the segmental homology may suggest that recombination took place among mouse immunoglobulin H chain genes as suggested previously (Miyata *et al.*, 1980; Ollo and Rougeon, 1982; Schreier *et al.*, 1981; Yamawaki-Kataoka *et al.*, 1981, 1982).

Materials and methods

Materials

Restriction endonucleases, *Bam*HI, *Eco*RI, *Hae*III, *Hap*II, and T₄ polynucleotide kinase were purchased from Takara Shuzo Co. (Kyoto, Japan). Sources of other restriction enzymes were described (Honjo *et al.*, 1979). [α -³²P]dCTP and [γ -³²P]ATP were obtained from Radiochemical Centre (Amersham, UK).

Subcloning of the epsilon gene segments

The clone Igε-7 bearing the epsilon gene (Nishida *et al.*, 1981) was digested

Table III. Divergence of amino acid sequences between mouse immunoglobulin epsilon chains and other H chains.

	mu	delta	gamma 1	gamma 2b	gamma 2a	alpha
C _{H1}	84	80	57	56	57	73
C _{H2}	83		(80)	(83)	(80)	(84)
C _{H3}	79	82	71	73	72	81
C _{H4}	64	71	59	60	59	66
M1	60	70	49	47	47	—
M2	—	—	71	68	64	—

Amino acid sequences of corresponding exons of immunoglobulin H chains were compared with those of the epsilon chain. Hinge regions of the gamma and alpha chains were paired with the C_{H2} domain of the epsilon chain, and shown in parenthesis. Numbers are percentages of divergence as compared with the epsilon chain.

with *Bam*HI and ligated with pBR322 (Boliver *et al.*, 1977) which had been cleaved with *Bam*HI and treated with bacterial alkaline phosphatase. The ligated DNA was then used to transform an *Escherichia coli* strain χ 1776. Transformants that contained the epsilon gene were identified by restriction cleavage site mapping. All cloning experiments were carried out in a P3 facility according to the Japanese guidelines of recombinant DNA.

mRNA hybridization

Cytoplasmic RNA from the hybridoma cell line, IgE-53-569 (Böttcher *et al.*, 1980), was twice enriched with an oligo-(dT) cellulose column. Glyoxal-treated RNA was fractionated by electrophoresis on a 1.2% agarose gel. RNA was transferred to nitrocellulose filters (Schleicher & Schuell B85) according to Thomas (1980). Hybridization was carried out as described by Southern (1975) except that RNA blots were washed as described by Thomas (1980).

DNA sequencing and other procedures

DNA sequencing was performed according to the method of Maxam and Gilbert (1980). Plasmid DNA was prepared as described (Clewett and Helinski, 1969). Southern blot hybridization was carried out as described previously (Southern, 1975). Nick-translation of probe DNA was carried out as described by Maniatis *et al.* (1975). The computer program used for calculation of nucleotide divergence rates (K_a , K_s , and K_p) was described previously (Miyata and Yasunaga, 1980). k_{aa}^C represents the amino acid difference per residue after correction for multiple substitutions, namely, $k_{aa}^C = -\ln(1 - k_{aa})$. The rate of evolution defined as divergence per residue per year (v_{aa}) was calculated according to the following equation; $v_{aa} = k_{aa}^C/2T$, where we assumed divergence time (T) between man and mouse as 75×10^6 years.

Acknowledgements

We are grateful to Y.Nakayama and S.Nishida for their excellent technical assistance and to F.Oguni for preparing this manuscript. This investigation

was supported in part by grants from Mitsubishi Science Foundation, from Torey Science Foundation, from Naito Foundation and from the Ministry of Education, Science and Culture of Japan.

References

- Bennich, H., and von Bahr-Lindstrom, H. (1974) *Progress in Immunology II*, 1, 49-58.
- Boliver, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneken, H.L., and Boyer, H.W. (1977) *Gene*, 2, 95-113.
- Bothwell, A.-L.M., Paskind, M., Schwartz, R.C., Sonenschein, G.E., Gefter, M.L., and Baltimore, D. (1981) *Nature*, 290, 65-67.
- Breathnach, R., Benoist, C., O'Hare, K., Gannon, F., and Chambon, P. (1978) *Proc. Natl. Acad. Sci. USA*, 75, 4853-4857.
- Böttcher, I., Ulrich, M., Hirayama, N., and Ovary, Z. (1980) *Int. Arch. Allergy Appl. Immunol.*, 61, 248-250.
- Cheng, H.-L., Blattner, F.R., Fitzmaurice, L., Mushinski, J.F., and Tucker, P.W. (1982) *Nature*, 296, 410-415.
- Clewell, D. and Helinski, D. (1969) *Proc. Natl. Acad. Sci. USA*, 62, 1159-1166.
- Dayhoff, M.O. (1978) *Atlas of Protein Sequence and Structure*, published by National Biomedical Research Foundation, Washington, DC.
- Dickerson, R.E. (1971) *J. Mol. Evol.*, 1, 26-45.
- Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R., and Hood, L. (1980) *Cell*, 20, 313-319.
- Ellison, J., Buxbaum, J., and Hood, L. (1981) *DNA*, 1, 11-18.
- Flanagan, J.G., and Rabbitts, T.H. (1982) *EMBO J.*, 1, 655-660.
- Hieter, P.A., Hollis, G.F., Korsmer, S.J., Waldmann, T.A., and Leder, P. (1981) *Nature*, 294, 536-540.
- Hieter, P.A., Max, E.E., Seidman, J.G., Maizel, J.V., and Leder, P. (1980) *Cell*, 22, 197-207.
- Honjo, T., Obata, M., Yamawaki-Kataoka, Y., Kataoka, T., Kawakami, T., Takahashi, N., and Mano, Y. (1979) *Cell*, 18, 559-568.
- Ishizaka, K., and Ishizaka, T. (1975) *Prog. Allergy*, 19, 60-121.
- Ishizaka, K., Tomioka, H., and Ishizaka, T. (1970) *J. Immunol.*, 105, 1459-1467.
- Kawakami, T., Takahashi, N., and Honjo, T. (1980) *Nucleic Acids Res.*, 8, 3933-3945.
- Kimura, M., and Ohta, T. (1974) *Proc. Natl. Acad. Sci. USA*, 71, 2848-2852.
- Konkel, D.A., Tilgman, S.M., and Leder, P. (1978) *Cell*, 15, 1125-1132.
- Maniatis, T., Jeffreys, A., and Kleid, D.G. (1975) *Proc. Natl. Acad. Sci. USA*, 72, 1184-1188.
- Max, E., Battey, J., Ney, R., Kirsch, I.R., and Leder, P. (1982) *Cell*, 29, 691-699.
- Maxam, A.M., and Gilbert, W. (1980) *Methods Enzymol.*, 65, 499-560.
- Miyata, T., and Yasunaga, T. (1980) *J. Mol. Evol.*, 16, 23-36.
- Miyata, T., Yasunaga, T., and Nishida, T. (1980) *Proc. Natl. Acad. Sci. USA*, 77, 7328-7332.
- Nakai, S., Oi, V.T., Herzenberg, L.A., Yamagishi, H., and Honjo, T. (1981) *Biomed. Res.*, 3, 37-45.
- Nikaido, T., Yamawaki-Kataoka, Y., and Honjo, T. (1982) *J. Biol. Chem.*, 257, 7322-7329.
- Nishida, Y., Kataoka, T., Ishida, N., Nakai, S., Kishimoto, T., Böttcher, I., and Honjo, T. (1981) *Proc. Natl. Acad. Sci. USA*, 78, 1581-1585.
- Nishida, Y., Miki, T., Hisajima, H., and Honjo, T. (1982) *Proc. Natl. Acad. Sci. USA*, 79, 3833-3837.
- Ollo, R., and Rougeon, F. (1982) *Nature*, 296, 761-763.
- Rabbitts, T.H., Foster, A., and Milstein, C.P. (1981) *Nucleic Acids Res.*, 9, 4509-4524.
- Rogers, J., Choi, E., Souza, L., Carter, C., Word, C., Kuel, M., Eisenberg, D., and Wall, R. (1981) *Cell*, 26, 19-27.
- Rogers, J., Early, P., Carter, C., Calame, K., Bond, M., Hood, L., and Wall, R. (1980) *Cell*, 20, 303-312.
- Schreier, P.H., Bothwell, A.-L.M., Muller-Hill, B., and Baltimore, D. (1981) *Proc. Natl. Acad. Sci. USA*, 78, 4495-4499.
- Sharp, P.A. (1981) *Cell*, 23, 643-646.
- Southern, E.M. (1975) *J. Mol. Biol.*, 98, 503-517.
- Thomas, P.S. (1980) *Proc. Natl. Acad. Sci. USA*, 77, 5201-5205.
- Tsuzukida, Y., Wang, C., and Putnam, F.W. (1979) *Proc. Natl. Acad. Sci. USA*, 76, 1104-1108.
- Tucker, P.W., Liu, C.-P., Mushinski, J.F., and Blattner, F.R. (1980) *Science (Wash.)*, 209, 1353-1360.
- Tucker, P.W., Slighton, J.L., and Blattner, F.R. (1981) *Proc. Natl. Acad. Sci. USA*, 78, 7684-7688.
- Tyler, B.M., Cowman, A.F., Gerondakis, S.D., Adams, J.M., and Bernard, O. (1982) *Proc. Natl. Acad. Sci. USA*, 79, 2008-2012.
- Wilson, A.C., Carlson, S.S., and White, T.J. (1977) *Annu. Rev. Biochem.*, 46, 573-639.
- Yamawaki-Kataoka, Y., Kataoka, T., Takahashi, N., Obata, M., and Honjo, T. (1980) *Nature*, 283, 786-789.
- Yamawaki-Kataoka, Y., Miyata, T., and Honjo, T. (1981) *Nucleic Acids Res.*, 9, 1365-1381.
- Yamawaki-Kataoka, Y., Nakai, S., Takahashi, N., Miyata, T., and Honjo, T. (1982) *Proc. Natl. Acad. Sci. USA*, 79, 2623-2627.