8-1-1990

DNA Recombination and Natural Selection Pressure Sustain Genetic Sequence Diversity of the Feline MHC Class I Genes

Naoya Yuhki
National Cancer Institute at Frederick

Stephen J. O'Brien
National Cancer Institute at Frederick, sobrien1@nova.edu

Published August 1, 1990

DNA Recombination and Natural Selection Pressure Sustain Genetic Sequence Diversity of the Feline MHC Class I Genes

Naoya Yuhki and Stephen J. O'Brien

From the Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick, Maryland 21701

Summary

Sequence comparisons of seven distinct MHC class I cDNA clones revealed that feline class I molecules have a remarkable similarity to human HLA genes in their organization of functional domains as well as in the nonrandom partitioning of genetic variability according to the functional constraints ascribed to different regions of the MHC molecule. The distribution of the pattern of sequence polymorphism in the cat as compared with genetic diversity of human and mouse class I genes provides evidence for four coordinate factors that contribute to the origin and sustenance of abundant allele diversity that characterizes the MHC in the species. These include: (*a*) a gradual accumulation of spontaneous mutational substitution over evolutionary time; (*b*) selection against mutational divergence in regions of the class I molecule involved in T cell receptor interaction and also in certain regions that interact with common features of antigens; (*c*) positive selection pressure in favor of persistence of polymorphism and heterozygosity at 57 nucleotide residues that comprise the antigen recognition site; and (*d*) periodic intragenic (interallelic) and intergenic recombination within the class I genes. We describe a highly conserved 23-bp nucleotide sequence within the coding region of the first α-helix that separates two relatively polymorphic segments located in the α1 domain that may act as a template or "hot spot" for homologous recombination between class I alleles.

The MHC genes encode two classes of cell surface molecules that can present immunological peptides to TCRs (class I and class II molecules) (1, 2). These molecules have extremely polymorphic features in most species (1). According to the x-ray crystallographic model of the human HLA-A2 class I molecule (3), the majority of polymorphic residues of this molecule in human and mouse are located on the molecular surface that comprises the putative antigen binding site (4). This finding, together with the evidence that HLA class I molecules have differential binding capacities for immunological peptides (5), suggests that the polymorphism of MHC class I molecules reflects the capacity of these molecules to bind a wide spectrum of immunological determinants.

Several competing hypotheses have been proposed to explain the evolutionary origin of the high degree of polymorphism; these include (*a*) a high rate of mutation (6), (*b*) interlocus recombination or gene conversion (7, 8), (*c*) overdominant selection (9–13), (*d*) frequency-dependent selection (11), or (*e*) some combinations of the above factors. Pursuant to this question, recent comparative analyses of sequence diversity within human and chimpanzee coding regions have revealed that a large fraction of polymorphic residues were found in both human and chimpanzee MHC alleles (14–18). Of 198 different polymorphic sites found in the chimpanzee, 167 (84%) were also found in one or more HLA molecules, suggesting that most of the modern sequence diversity predates the divergence of these two species over 5 million years ago. A similar analysis revealed that murine class II alleles present today in rodents have shared primitive sequence stretches that must have occurred before the divergence of mouse and rat over 10 million years ago (19). The accumulation and persistence of extensive polymorphism for such extended periods of evolutionary time would preclude the necessity for increased mutation rate as an explanation for abundant diversity.

A different approach by Hughes and Nei (12, 13) tested the increased mutation vs. selection hypotheses for MHC polymorphism by analyzing the theoretical predictions of the two competing hypotheses in the context of the functional domains of the MHC molecule. These investigators examined the ratio of synonymous vs. nonsynonymous nucleotide substitutions in the well-defined antigen recognition site (ARS)[1] of human and murine class I allelic transcripts. They reasoned that, if the ARS was the object of overdominant selection in light of its function to recognize, bind, and present

[1] *Abbreviations used in this paper:* ARS, antigen recognition site; FeLV, feline leukemia virus; FIV, feline immunodeficiency virus.

Published August 1, 1990

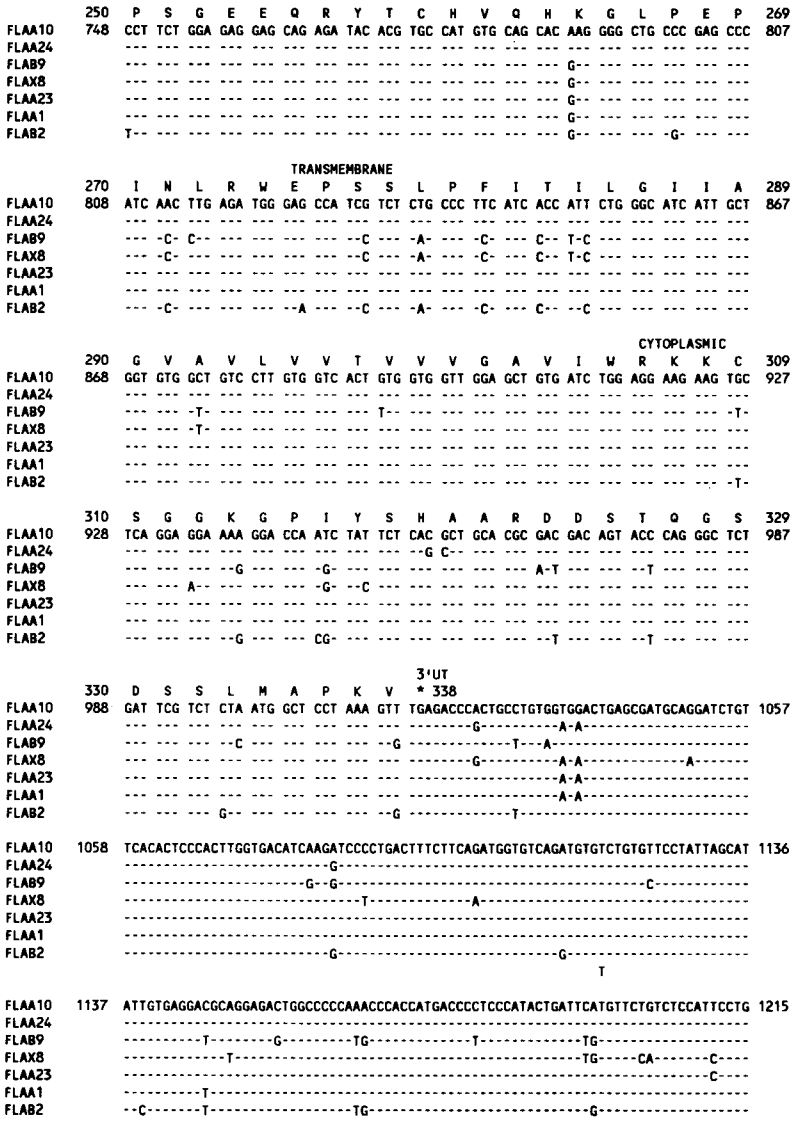

Figure 1. Nucleotide sequences of seven distinct feline class I cDNA clones. The complete sequences are shown for the FLAA10 clone. The dashed line and single letter indicate the identical and distinct nucleotide from the FLAA10 clone. The deduced amino acid sequence of the FLAA10 cDNA clone is shown above the nucleotide sequence of FLAA10. Asterisk, plus, and dot symbols above the sequence of FLAA10 indicate residues toward the antigen binding site of antigen recognition, on an α-helix pointing up, and on an α-helix pointing away from the antigen binding site and TCR recognition site, respectively, according to the homologous structure of the human HLA-A2 molecule (3, 4).

Published August 1, 1990

various peptides to TCRs, then the nonsynonymous (amino acid altering) substitutions would be more abundant than synonymous nucleotide substitutions in the same region. Pairwise comparisons of 12 human and 9 murine class I cDNA allelic transcripts revealed a marked excess in nonsynonymous substitutions compared with synonymous changes in the ARS. Since a model of increased spontaneous mutation under a presumption of selective neutrality would not favor either type of nucleotide substitutions, they interpreted their results as strong evidence against the increased mutation explanation and in favor of overdominant selection.

Parham et al. (15), in an extensive comparative analysis of 40 human HLA class I transcripts and 9 chimpanzee ChLA class I transcripts, presented further evidence for the ancient origin of allelic diversity. Their results also revealed a "patchwork" pattern of diversity between allelic transcripts where common stretches of sequence identity between transcripts were apparent. These results suggested that interallelic recombination likely played a major role in generation of novel class I alleles in human-ape evolution. Intergenic recombination, sometimes referred to as "gene conversion" in the context of MHC diversity observed in rodent class I genes, was thought to be of minor importance since transcripts of each human locus, i.e., HLA-A, -B, and -C, each had numerous shared derived (or synapomorphic) nucleotide residues that were characteristic of the three loci (14, 15). The logic argued that if interlocus exchanges were frequent then diversity between loci would be "homogenized", but that was clearly not the case in man. In contrast, the rodent class I locus were highly homogenized apparently by the inter-locus sequence exchanges (7, 20).

We present in this paper a reexamination of these hypotheses using a new set of class I MHC transcripts derived from the domestic cat. This species has several advantages for studies of MHC diversity. Cats have a moderately polymorphic MHC, called FLA, that has been studied from both serological (21) and molecular perspectives (22–24). In addition, the cat is afflicted by epidemics of two pathological RNA viruses: feline leukemia virus, FeLV (25); and feline immunodeficiency virus, FIV (26, 27). Finally, the cat and its free-ranging relatives have been the object of ongoing studies of population genetics and molecular evolution (28–30). The present results affirm and extend, in a third mammalian order, the critical roles of positive selection and DNA recombination in the origins and maintenance of genetic diversity of the MHC. In addition, although locus specificity of FLA class I transcripts are evident in the conserved transmembrane and cytoplasmic domains, occasional mosaic sequences indicate that inter-locus sequence exchange has played a notable role in the evolution of feline MHC class I genes.

Materials and Methods

Total RNA was isolated from three unrelated domestic cat spleen cell preparations by the hot phenol-guanidium isothiocyanate method (31). Poly(A)$^+$ RNAs were isolated by oligo(dT)–cellulose columns (32). cDNAs were synthesized by a modification of the method of Gubler and Hoffman (33). These cDNAs were ligated with a λ ZAP II vector (34) and in vitro packaged using a Gigapack plus reagent (Stratagene). 2×10^5 unamplified plaques from each library were screened with a ^{32}P-labeled feline class I cDNA insert (pFLA24) (23) at low stringency. Positive single plaques were isolated by the second screening and cDNA clones were autoexcised into phagemid forms by the infection of R408 helper phages (34). The phagemid clone DNAs were electrophoresed on 1% agarose gel, stained with ethidium bromide and the sizes were measured by comparing them with the full-length cDNA clone, pFLA24. The full-length or nearly full-length clones were then transfected into JM101 *Escherichia coli* cells, superinfected with M13K07 helper phages or the cDNAs were subcloned into M13 vectors and single-strand DNAs for sequence analysis were purified (35). The sequences of each clone were determined by a dideoxynucleotide sequencing method using [^{35}S]dATP (New England Nuclear, Boston, MA) and T7 DNA polymerase (36) (U.S. Biochemical Corp., Cleveland, OH), and seven sequencing primers for the conserved sequences of two previously isolated feline class I cDNA clones (23). These sequences were compared using the Wilbur and Lipman algorithm (37).

The names of two previously described transcripts, FLA2, FLA24 (23), were changed to FLAB2 and FLAA24, respectively, to reflect locus-specificity in this paper. The seven cDNA clones described here were isolated from libraries made from four different cats as follows: cat no. 1, FLAB2 and FLAA24; cat no. 2, FLAA10, FLAA23 and FLAB9; cat no. 3, FLAA1; and cat no. 4, FLAX8.

Results

Five distinct full-length or nearly full-length feline MHC class I cDNA clones were identified from three unrelated cat spleen cDNA libraries by a screen of over 60 cDNA clones. Together with two previously described, distinct feline MHC class I cDNA clones (23), we determined the nucleotide sequence of seven distinct cDNA transcripts in order to examine the diversity and the evolution of class I molecules. MHC class I genes had been reported to consist of eight coding exons which encode the following protein domains: (*1*) leader sequence; (*2*) and (*3*) α1 and α2 polymorphic extracellular domains which are involved in antigen and TCR recognitions; (*4*) α3 extracellular domain; (*5*) transmembrane domain; and (*6–8*) three cytoplasmic domains. According to the x-ray crystallographic analysis of human HLA-A2 class I molecules, Bjorkman et al. (4) have described 57 amino acid residues located in the α1 and α2 domains that are involved in antigen recognition and are termed the antigen recognition site (ARS). Further, the ARS occurs in two α-helix stretches that are separated by one platform of β-pleated sheets in the class I molecule. In the present analysis we compare the complete sequences for the seven feline cDNA clones (Figs. 1 and 2). The regions of protein topography (α helices and β-pleated sheets) are indicated in the sequence of Fig. 2 according to the structure of the HLA-A2 molecule.

We first examined the first two extracellular domains. Analysis of the pattern of sequence polymorphism in FLA transcripts revealed that, as in the human and mouse, there are clearly recognizable residues that are invariant and others that are highly polymorphic. Further, the variant and invariant amino acids are generally at the same positions in the three species. Table 1 lists 32 residues that are invariant among 40

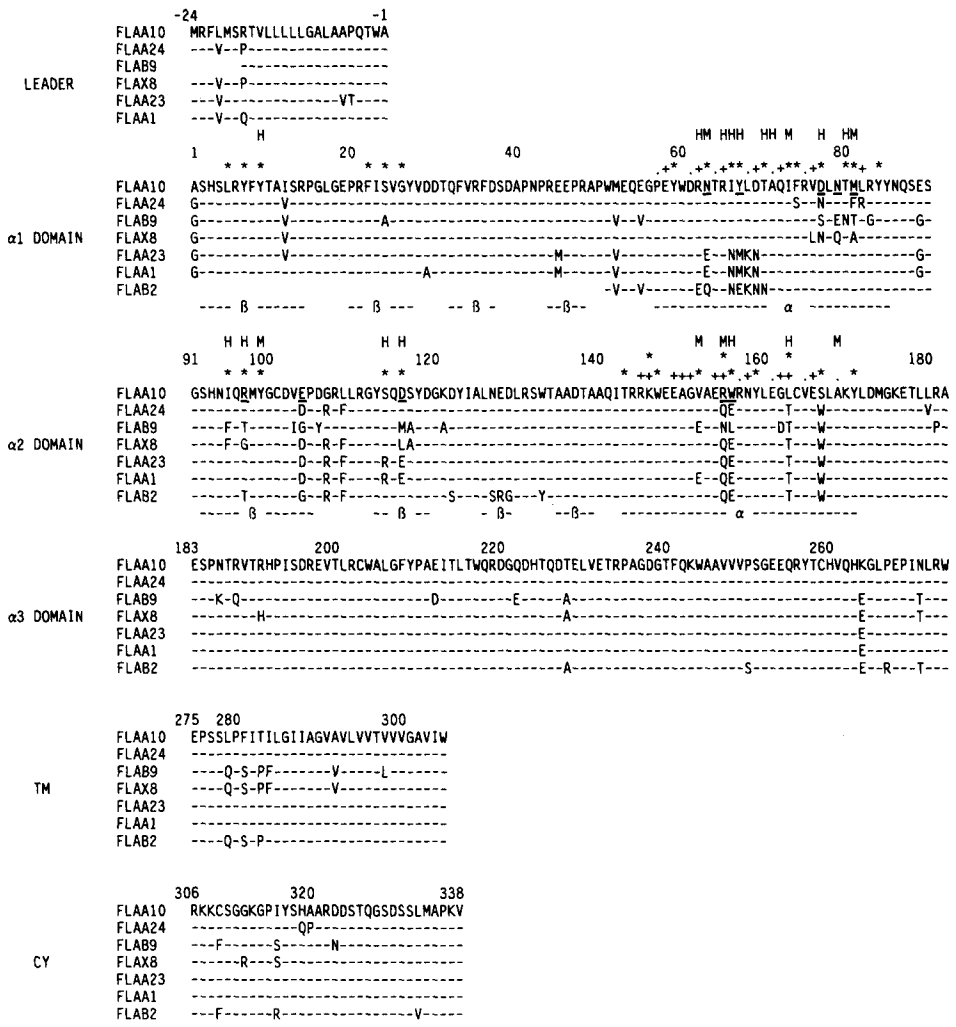

Figure 2. Amino acid sequences of the seven distinct feline class I molecules. The complete sequence of FLAA10 is shown. Single letters and dashed lines indicate the different residues and the identical residues with those of FLAA10, respectively. Two α-helix coding regions and eight β-sheet coding regions are indicated below the sequences. The same definitions as for Fig. 1 are used for the asterisk, plus, and dot symbols above the amino acid sequence. H and M above the number indicate polymorphic sites defined in human and mouse class I molecules respectively. Highly polymorphic sites of feline class I molecules are defined as the site that showed over three different residues and are indicated as underlines of the FLAA10 sequence.

human transcripts. Of these, 31 are also invariant in cats and 27 of the 32 conserved residues have identical amino acids in human and feline transcripts. Four cat residues were different from the human residues but were still conserved within the feline molecules (L5, I73, N158, A169). The single highly polymorphic feline residue that was invariant in humans (Q155) is also polymorphic in mouse H-2 transcripts (23). The majority of highly polymorphic residues of feline class I molecules (9 of 10 highly polymorphic sites which show over three different residues) are located on the site facing the antigen binding groove according to the model of the HLA-A2 molecule.

The pattern and distribution of sequence substitution between allelic transcripts can be used to test for the role of mutation versus the selection of MHC diversity based upon the procedures of Hughes and Nei (12, 13). If the MHC molecules are subject to positive selection, and if the 57 residues of the ARS that participate directly in antigen recognition are the objects of selective pressure, the rate of nonsynonymous mutation will be higher than that of synonymous substitution in ARS. The logic is predicated upon the theoretical notion that positive selection would enhance the rate of "sensible" codon substitution compared with that for synonymous substitutions because synonymous mutations are more or less selectively neutral. By contrast, if high polymorphism is a consequence of increased mutation, the rates of synonymous vs. nonsynonymous mutation would be equivalent in all regions. In a comparison of FLA transcripts, there occurred nearly twice as many nonsynonymous substitutions as synonymous substitutions in the ARS region (Table 2). Of the 21 pairwise comparisons between the 7 FLA transcripts, in every case there was a greater number of nonsynonymous substitutions in this functional region. By contrast, the same analysis of the 125 residues that are in the α1 and α2 domains, but are not in the ARS, resulted in the opposite observation, namely that the frequency of synonymous substitutions were greater by nearly twofold. These findings are in precise agreement with the same analysis of human and murine nucleotide substitutions (Table 2). Taken together the data provide strong evidence for the role of overdominant selection or heterozygote advantage in the ARS region of the MHC molecules in the three species.

A close examination of the highly polymorphic region in the α1 domain that encodes the first α-helix (amino acid

Published August 1, 1990

Table 1. *Highly Conserved and Polymorphic Residues in the α1 and α2 Domains of Human and Cat Class I Molecules*

	Criteria	Human	Cat
Highly conserved residues	Recognition of common features of Ag	M5, Y7, F22, G26, Y59, Y84, T143, K146, Y159, Y171	L5, Y7, F22, G26, Y59, Y84, T143, K146, Y159, Y171
	Formation of conserved faces of α-helixes	P57, E58, D61, T64, Q72, R75, L78, Y84, K146, A150, E154, Q155, G162, V165	P57, E58, D61, T64, Q72, R75, L78, Y84, K146, A150, E154, (G162), V165
		T73, A158, R169	I73, N158, A169
	N-Glycosylation site	N86, Q87, S88	N86, Q87, S88
	Disulfide bonding	C101, C164	C101, C164
Highly polymorphic residues	β-sheets in α1 domain	9	None
	First α-helix	62, 65, 66, 67, 70, 71, 77, 80	63, 67, 77, 79*, 81
	β-sheets in α2 domain	95, 97, 114, 116	97, 104, 116
	Second α-helix	156, 163	155, 156

Probable functional constraints are based on homology to human residues (3, 4). Uniquely conserved residues in feline class I molecules are underlined. One residue (G162) is conserved in six of seven feline molecules. 9 of 10 highly polymorphic residues (over 3 different residues at the site) are located on the site facing the antigen binding groove according to the human HLA-A2 model (with the exception of one residue, 79*).

residues 58–86) demonstrated a patchwork or mosaic pattern between FLA transcripts (Fig. 2). FLAA24, FLAB9, and FLAX8 were identical to FLAA10 in the NH_2-terminal half of the α-helix while the same molecules diverged appreciably in the COOH-terminal half of this region. The remaining three transcripts (FLAA23, FLAA1, and FLAB2) were divergent from FLAA10 in the NH_2-terminal half, while they were identical to FLAA10 in the COOH-terminal

Table 2. *Mean Numbers of Nucleotide Substitutions* per 100 Synonymous Sites (dS) and per 100 Nonsynonymous Sites (dN) of MHC Class I Molecules*

Species	No. of sequences	No. of comparisons	Antigen recognition sites (ARS) (n = 57) dS	Antigen recognition sites (ARS) (n = 57) dN	Remaining codons in exon 2 & 3 (n = 125) dS	Remaining codons in exon 2 & 3 (n = 125) dN
Domestic cat	7	21	7.5 ± 2.4	13.60 ± 1.9‡	7.34 ± 2.4	4.9 ± 0.8
		dS>dN:dN>dS		0:21		16:5
Human	12	66	6.8 ± 2.3	20.8 ± 2.3§	11.6 ± 2.1	5.2 ± 0.8‡
		dS>dN:dN>dS		0:66		63:3
Mouse	9	36	13.8 ± 3.2	21.4 ± 2.4	8.3 ± 1.6	6.5 ± 0.8
		dS>dN:dN>dS		6:30		29:7

* By the method of Nei and Gojobori (38). For each species the ratio of pairwise sequence comparisons when synonymous substitutions exceeded nonsynonymous substitutions (dS>dN) vs. the reverse situation is presented as dS>dN:dN>dS. Human and mouse data are derived from (12).
‡ Significant at 0.1% level.
§ Significant at 0.01% level.

Published August 1, 1990

region (Fig. 2). When the nucleotide sequences of these same transcripts were examined (Figs. 1 and 3), they revealed highly polymorphic regions in the 5′ and 3′ ends of the first α-helix region separated by a strongly conserved 23-nucleotide stretch. We tested the sequence divergence of the two divergent segments of the α-helix for evidence of recombination vs. mutation origins using the algorithm of Stephens (39). This method effectively discriminates random mutational substitution vs. block changes that occur by recombination between divergent sequence. In comparison of the 5′ and 3′ segments of FLAA10 transcript to homologous regions of each of the other transcripts, the results affirmed divergent clustering ($p \leq 0.05$) in 11 of 12 comparisons lending strong credence to the recombinational origin of the FLA transcripts. Additional vestiges of recombination were apparent in comparisons of the α2 domains where the cDNAs did not conform to partitioning of transcripts evident in the α1 domain (Fig. 2 and Table 3).

The 3′ domains of the mammalian MHC class I molecule, (α3 extracellular, transmembrane, and cytoplasmic) are generally conserved in their sequence and have been used to assign locus specificity of class I transcripts (15). The seven FLA transcripts were compared in this region and fell into three discrete classes (Table 3 and Fig. 4). The first group (FLAA10, A24, A23, and A1) were nearly identical throughout the three 3′ domains (only three nucleotide substitutions in 441 nucleotides), but quite divergent from the remaining transcripts. The second group contained FLAB2 and B9 which were similar, though not identical, in the TM and CY domains but more divergent in the α3 extracellular domain. Based upon the precedence in humans whereby alleles of the same class I B and C loci vary by a maximum of eight nucleotide substitutions (15), we hypothesize that these two classes of transcripts are products of different loci, designated A and B, respectively. Support for this assignment is obtained because FLAB9, A10, and A23 were derived from a single individual cat. In that animal, we suggest the A locus expressed A10 plus A23 and the B locus expressed B9. The last transcript, FLAX8, has a transmembrane domain sequence that was nearly identical to the B locus FLAB9 (one substitution), a cytoplasmic domain that was more similar to the A locus FLAA1 and A23 sequences (3 substitutions), and an α3 domain that was equidistant from all other transcripts. The mosaic sequence structure of the FLAX8 transcript makes it difficult to assign it unequivocally to either locus. An illustration of the shared domains of the seven transcripts and their locus designation is presented in Fig. 4.

A reexamination of the mosaic sequence structure in the α-helix region (Fig. 3) armed with the knowledge of locus specificity of the transcripts is illuminating. The FLAA10 specific 5′ and 3′ sequences are found in transcripts from both A and B loci. If our locus designations are correct, this indicates that inter-allelic exchange occurs not only within, but also between, class I loci. Alternatively, the FLAA10–derived sequences in this region are primitive and the variants are derived by recombination with a different ancestral donor gene sequence. In either case, intra- or intergenic recombination has occurred and the conserved 23 bp intermediate sequence served as a junction for the exchange. Further, the mosaic structure within the α2, transmembrane, and cytoplasmic domains (Fig. 4) would also affirm a role for recombination in deriving these arrangements.

Discussion

The present sequence comparison of seven feline MHC class I transcripts to mouse and human genes revealed a remarkable conservation in the patterns of domain organization and in the sites of intraspecies polymorphisms. With only one exception, the FLA class I domains were precisely the length of the human HLA-A and -C domains (15, 22), and there was a striking conservation of sequence homology for 32 invariant human residues in α1 and α2 domains known to play a critical role in MHC class I function (Table 1). The cat molecules not only were genetically invariant at each of the sites, but they also retained identical residues at 27 of the 32 sites. In addition the feline molecules displayed a clustering of polymorphic residues within the two α-helix and two β-strand regions in a manner similar to the pattern of human and mouse sequences (4). Finally, strong evidence against increased mutation rate and for the force of selection of polymorphic sites in the 57 residues of the antigen recognition site was obtained by the demonstrable excess of codon-altering (as compared with synonymous) substitution rate in this region. An earlier analysis of human and mouse transcripts yielded a very similar result (12, 13). We conclude that the class I genes of the three mammalian species retain ancestral constraints both in their organization of functional domains as well as in their tolerance (or predilection) for polymorphism. These shared primitive (ancestral) constraints likely reflect a functional organization inherited from a common ancestor of the three mammalian orders.

Based on the cumulative results of our studies of the human, mouse, and cat sequences, we can recognize four major factors

FIRST α- HELIX CODING REGION

	22bp	23bp	21bp	LOCUS
FLAA10	CGGAACACCCGGATTTACTTGG	ACACCGCACAGATTTTCCGAGTG	GACCTGAACACGATGCTCCGC	A
FLAA24	--------G----------	----------C-------	A----------T-T-G----	A
FLAB9	--------G----------	----------	AG----G-G-AC-C--G-G-A	B
FLAX8	----------	----------T--	A-----C-G--CGCC------	AB
FLAA23	---G-G--G----ACATGAA-A	----------	----------	A
FLAA1	---G-G--G----ACATGAA-A	----------	----------	A
FLAB2	GA-C-G--G----ACG-GAA-A	---A----------	----------	B

Figure 3. Nucleotide sequences that encode the first long α-helix region. Dashed line indicates identical nucleotide to FLAA10 sequence.

Published August 1, 1990

Table 3. *The Number of Nucleotide Substitutions*

		A10	A24	A23	A1	B9	B2*	X8
Leader	FLAA10		2	3	3	0	—	2
	FLAA24			3	2	0	—	0
	FLAA23				4	2	—	3
	FLAA1					1	—	2
	FLAB9						—	0
	FLAB2							—
α1 Domain	FLAA10		9	20	21	19	(16)	11
	FLAA24			22	22	17	(20)	10
	FLAA23				1	27	(10)	25
	FLAA1					27	(10)	27
	FLAB9						(26)	16
	FLAB2							(24)
α2 Domain	FLAA10		16	16	18	28	25	20
	FLAA24			6	8	33	17	13
	FLAA23				2	29	15	9
	FLAA1					27	14	11
	FLAB9						38	23
	FLAB2							20
α3 Domain	FLAA10		0	1	1	10	7	7
	FLAA24			1	1	10	7	7
	FLAA23				0	8	6	7
	FLAA1					9	6	6
	FLAB9						8	9
	FLAB2							8
TM Domain	FLAA10		0	0	0	8	6	7
	FLAA24			0	0	8	6	7
	FLAA23				0	8	6	7
	FLAA1					8	6	7
	FLAB9						4	1
	FLAB1							3
CY Domain	FLAA10		2	0	0	8	6	7
	FLAA24			2	2	10	10	5
	FLAA23				0	8	8	3
	FLAA1					8	8	3
	FLAB9						4	9
	FLAB2							9
3′UT	FLAA10		4	3	4	20	31	15
	FLAA24			3	3	22	33	13
	FLAA23				3	24	34	12
	FLAA1					21	34	16
	FLAB9						26	28
	FLAB2							42

* B2cDNA clone begins within the α1 domain.

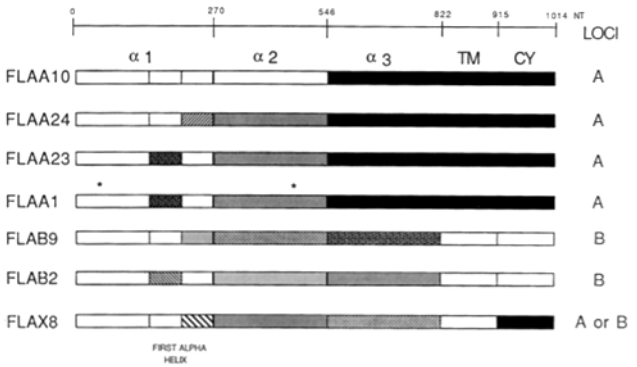

Figure 4. Schematic features of seven feline class I molecules. FLAA23 and A1 are subtypes that show only two amino acid substitutions (indicated by asterisks).

in the generation and maintenance of diversity. The first is a stochastic accumulation of spontaneous mutations throughout the class I molecule. There is little persuasive evidence for an increased frequency of mutations in these genes (12). The second key factor is a discernable negative selection pressure against increased polymorphism in two important regions; namely, the areas including the conserved outer face of the α-helix that interacts directly with the complementation determining regions (CDR1 and CDR2) of the TCR V gene product (2) and the conserved regions of the antigen binding site that recognize a common feature of immunological peptides (4). Residues from both of these regions are invariant within the three species and in most cases retain identical amino acids in each of these sites (see Table 1).

The third important factor is a positive selective pressure in favor of increased heterozygosity and polymorphism at loci in the 57 nucleotide residues in the ARS. Evidence for positive selection was obtained by comparing the ratio of codon-altering (nonsynonymous) substitution rate to synonymous substitution rate in the same region. As predicted by a model of overdominant selection by Hughes and Nei (12), the extent of nonsynonymous substitution rate in the ARS was significantly higher than the synonymous substitution rate (Table 2). The reverse was true in the neighboring residues not involved in antigen recognition, apparently because those regions are not subject to the intense selection for maintenance of genetic diversity required by the sites of antigen recognition. The findings with FLA class I molecules were in exact agreement with the results obtained by Hughes and Nei (12) who compared human and mouse class I transcripts. The results may also reveal a level of frequency dependent selection; however, it is difficult to implicate this precisely with the present data.

The fourth critical component for generation of MHC diversity is intragenic (interallelic) and intergenic recombination. Evidence for both types of recombination as an origin for allelic diversity was evident in the first α-helix where a patchwork or mosaic pattern of shared amino acid diversity was observed (Figs. 1 and 2). Examination of the seven cat nucleotide sequences revealed a 23-bp segment between recom-

binant stretches that is likely the junction of previous recombinant events that generated the previous transcripts (Fig. 3). In addition, shared stretches of mutational differences in several FLA transcripts in the $\alpha 2$ domain are likely derived by recombination between the $\alpha 1$ and $\alpha 2$ domain as well (Fig. 1). Finally, the mosaic structure in transmembrane and cytoplasmic domains of one FLA transcript are likely derived by intergenic recombination between these two domains (Fig. 1 and 4). Because the mosaic patterns do not partition between putative class I loci, these combined observations would suggest a pivotal role for both interallelic and intergenic recombination in the formulation of FLA class I allele diversity.

The recent extensive studies of human transcripts are resounding in their finding of within-locus association in contrast to between-locus similarity (15, 16). Locus specificity actually extends beyond species boundaries since HLA-A alleles more closely resemble chimpanzee ChLA-A alleles than either group resembles their respective B-locus alleles (15). Locus-specific partitioning of allele sequences would not be consistent with a large role for intergenic recombination because that process would tend to homogenize alleles rather than separate them. Murine class I transcripts tend to be more homogenized and lack clear locus-specific sequence variation (7, 20), suggesting a historic role for intergenic exchange in rodent gene evaluation. The feline class I sequences presented here support the occurrence of between locus but within coding region recombinants in the generation of class I sequence variations.

A role for recombination in the generation of diversity would probably be adaptive since the types of interallelic recombination or intergenic recombination between functional loci would be able to create numerous polymorphic, but functional, molecules since the functional constraints were already overcome in prototype molecules (15). Such a mechanism might actually be more efficient than a random mutational process for generating new diversity. It is tempting to speculate that certain sequences of notable conservation located between regions of high variation may act as templates for homologous intra- and intergenic recombination particularly for those regions conserved in the three species discussed here. The 23-bp region of the first α-helix would be a possible example of this phenomenon. This sequence may have been an important recombination junction and may still retain the potential to create further polymorphic molecules in the future by homologous recombination.

Received for publication 29 January 1990 and in revised form 2 May 1990.

References

1. Klein, J. 1986. Natural History of the Major Histocompatibility Complex. John Wiley and Sons Inc., New York, NY.
2. Davis, M.M., and P.J. Bjorkman. 1988. T-cell antigen receptor genes and T-cell recognition. *Nature (Lond.).* 334:395.
3. Bjorkman, P.J., M.A. Saper, B. Samraoui, W.S. Bennett, J.L. Strominger, and D.C. Wiley. 1987. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature (Lond.).* 329:506.
4. Bjorkman, P.J., M.A. Saper, B. Samraoui, W.S. Bennett, J.L. Strominger, and D.C. Wiley. 1987. The foreign antigen binding site and T-cell recognition regions of class I histocompatibility antigens. *Nature (Lond.).* 329:512.
5. Chen, B.P., and P. Parham. 1989. Direct binding of influenza peptides to class I HLA molecules. *Nature (Lond.).* 337:743.
6. Klein, J. 1978. H-2 mutations: their genetics and effect on immune functions. *Adv. Immunol.* 26:55.
7. Nathenson, S.G., J. Geliebter, G.M. Pfaffenbach, and R.A. Zeff. 1986. Murine major histocompatibility complex class-I mutants: molecular analysis and structure-function implications. *Annu. Rev. Immunol.* 4:471.
8. Mellor, A.L., E.H. Weiss, K. Ramachandran, and R.A. Flavell. 1983. A potential donor gene for the bm1 gene conversion event in the C57BL mouse. *Nature (Lond.).* 306:792.
9. Doherty, P.C., and R.M. Zinkernagel. 1975. Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature (Lond.).* 256:50.
10. Klein, J., and F. Figueroa. 1986. Evolution of the major histocompatibility complex. *CRC Crit. Rev. Immunol.* 6:295.
11. Bodmer, W.F. 1972. Evolutionary significance of the HL-A system. *Nature (Lond.).* 237:139.
12. Hughes, A.L., and M. Nei. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature (Lond.).* 335:167.
13. Hughes, A.L., and M. Nei. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proc. Natl. Acad. Sci. USA.* 86:958.
14. Parham, P., C.E. Lomen, D.A. Lawlor, J.P. Ways, N. Holmes, H.L. Coppin, R.D. Salter, A.M. Wan, and P.D. Ennis. 1988. Nature of polymorphism in HLA-A, -B, and -C molecules. *Proc. Natl. Acad. Sci. USA.* 85:4005.
15. Parham, P., D.A. Lawlor, C.E. Lomen, and P.D. Ennis. 1989. Diversity and diversification of HLA-A, -B, -C alleles. *J. Immunol.* 142:3937.
16. Mayer, W.E., M. Jonker, D. Klein, P. Ivanyi, G. van Seventer, and J. Klein. 1988. Nucleotide sequences of chimpanzee MHC class I alleles: Evidence for trans-species mode of evolution. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:2765.
17. Lawlor, D.A., F.E. Ward, P.D. Ennis, A.P. Jackson, and P. Parham. 1988. HLA-A and -B polymorphisms predate the divergence of humans and chimpanzees. *Nature (Lond.).* 335:268.
18. Klein, J., J. Gutknecht, and N. Fischer. 1990. The major histocompatibility complex and human evolution. *Trends Genet.* 6:7.
19. Figueroa, F., E. Gunther, and J. Klein. 1988. MHC polymorphism pre-dating speciation. *Nature (Lond.).* 335:265.
20. Rada, C., R. Lorenzi, S-J. Powis, J. van den Bogaerde, P.

Parham, and J.C. Howard. 1990. Concerted evolution of class I genes in the major histocompatibility complex of murine rodents. *Proc. Natl. Acad. Sci. USA.* 87:2167.
21. Winkler, C., A. Shultz, S. Cevario, and S.J. O'Brien. 1989. Genetic characterization of FLA, the cat major histocompatibility complex. *Proc. Natl. Acad. Sci. USA.* 86:943.
22. Yuhki, N., and S.J. O'Brien. 1988. Molecular characterization and genetic mapping of class I and class II MHC genes of the domestic cat. *Immunogenetics.* 16:339.
23. Yuhki, N., G.F. Heidecker, and S.J. O'Brien. 1989. Characterization of MHC cDNA clones in the domestic cat. Diversity and evolution of class I genes. *J. Immunol.* 142:3676.
24. Yuhki, N., and S.J. O'Brien. 1990. DNA variation of the mammalian major histocompatibility complex reflects genomic diversity and population history. *Proc. Natl. Acad. Sci. USA.* 87:836.
25. Cotter, S.M., W.D. Hardy, Jr., and M. Essex. 1975. Association of feline leukemia virus with lymphosarcoma and other disorders in the cat. *J. Am. Vet. Med. Assoc.* 166:449.
26. Pedersen, N.C., E.W. Ho, M.L. Brown, and J.K. Yamamoto. 1987. Isolation of a T-lymphotropic virus from domestic cats with an immunodeficiency-like syndrome. *Science (Wash. DC).* 235:790.
27. Olmsted, R.A., A.K. Barnes, J.K. Yamamoto, V.M. Hirsch, R.H. Purcell, and P.R. Johnson. 1989. Molecular cloning of feline immunodeficiency virus. *Proc. Natl. Acad. Sci. USA.* 86:2448.
28. O'Brien, S.J., D.E. Wildt, D. Goldman, C.R. Merril, and M. Bush. 1983. The cheetah is depauperate in genetic variation. *Science (Wash. DC).* 221:459.
29. O'Brien, S.J., M.E. Roelke, L. Marker, A. Newman, C.A. Winkler, D. Meltzer, L. Colly, J.F. Evermann, M. Bush, and D.E. Wildt. 1985. Genetic basis for species vulnerability in the cheetah. *Science (Wash. DC).* 227:1428.
30. O'Brien, S.J. 1986. Molecular genetics in the domestic cat and its relatives. *Trends Genet.* 2:137.
31. Maniatis, T., E.-F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. 545 pp.
32. Davis, L.G., M.D. Dibner, and J.F. Battey. 1986. Basic Methods in Molecular Biology. Elsevier, New York. 139 pp.
33. Gubler, U., and B.J. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. *Gene (Amst.).* 25:263.
34. Short, J.M., J.M. Fernandez, J.A. Sorge, and W.D. Huse. 1988. λ ZAP: a bacteriophage λ expression vector with in vivo excision properties. *Nucleic Acids Res.* 16:7583.
35. Vieira, J., and J. Messing. 1987. Production of single-stranded plasmid DNA. *Methods Enzymol.* 153:3.
36. Tabor, S., and C.C. Richardson. 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA.* 84:4767.
37. Wilbur, W.J., and D.J. Lipman. 1983. Rapid similarity searches of nucleic acid and protein data banks. *Proc. Natl. Acad. Sci. USA.* 80:726.
38. Nei, M., and T. Gojobori. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* 3:418.
39. Stephens, J.C. 1985. Statistical methods of DNA sequence analysis: detection of intragenic recombination or gene conversion. *Mol. Biol. Evol.* 2:539.