J Mol Evol (1994) 39:22–33

Journal of Molecular Evolution

Exchanges of Short Polymorphic DNA Segments Predating Speciation in Feline Major Histocompatibility Complex Class I Genes

Naoya Yuhki, Stephen J. O'Brien

Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick Cancer Research and Development Center, Frederick, MD 21702-1201, USA

Received: 12 May 1993 / Revised: 8 September 1993 / Accepted: 26 October 1993

Abstract. Sequence comparisons of 14 distinct MHC class I cDNA clones isolated from species representing the three major taxonomic lineages of Felidae (domestic cat lineage, ocelot lineage, and pantherine lineage) revealed that feline MHC class I alleles have highly mosaic structures with short polymorphic sequence motifs that are rearranged between alleles of individual MHC loci, between MHC class I genes within cat species, and between homologous MHC loci in different species. The pattern of sequence variation in felids supports the role of the following factors in production and maintenance of MHC variation: (1) gradual spontaneous mutation; (2) selective pressure to conserve certain residues but also to vary in hypervariable regions, notably residues that functionally participate in antigen recognition and presentation; and (3) recombination-mediated gene exchange between alleles and between related genes. The overall amount of genetic variation observed among MHC class I genes in the Felidae family is no greater than the amount of variation within any outbred cat species (i.e., domestic cat, ocelot). The occurrence of equivalent levels of polymorphism plus the simultaneous persistence of the same sequence motifs in divergent feline species suggest that most MHC class I nucleotide site polymorphism predated species divergences. Ancient polymorphisms have been transmitted through the speciation events and modern feline MHC class I alleles were derived by recombinational exchange of polymorphic sequence motifs. Moreover, some of these sequence motifs were found in other mammalian MHC class I genes, such as classical human HLA-B5, nonclassical human HLA-E class I genes, and bovine class I genes. These results raise the prospect of an ancient origin for some motifs, although the possibility of convergence in parallel mammalian radiations cannot be excluded.

Key words: Feline major histocompatibility complex class I genes — Molecular evolution

Correspondence to: N. Yuhki

Introduction

The major histocompatibility complex encodes two classes of antigen-presenting molecules for T-cell receptors (class I and class II molecules) (Klein 1986; Davis and Bjorkman 1988). Both molecules appear to have similar peptide binding grooves in their extracellular portions (Bjorkman et al. 1987a,b; Brown et al. 1988, 1993). Highly polymorphic features found in these molecules contribute to the formation of various shapes of these grooves (Bjorkman et al. 1987) and also to the capacity of each molecule for binding various spectra of immunological peptides (Van Bleek and Nathenson 1990; Rotzschke et al. 1990; Falk et al. 1991). To explore the origins of such polymorphic features of these molecules, we recently examined the pattern of MHC class I coding sequence variation among domestic cats (Yuhki and O'Brien 1990). Our results supported and extended parallel studies in human and rodent MHC variation (Lawlor et al. 1988, 1991; Mayer et al. 1988; Watkins et al. 1990, 1991; Nathenson et al. 1986) in implicating four coordinate factors that con-

tribute to the origin and sustenance of abundant MHC allelic diversity in these species. The factors include: (1) gradual accumulation of spontaneous nucleotide substitution; (2) negative selection in regions involved in T-cell receptor interaction and recognizing common features of peptides; (3) positive selection pressure to produce functional sequence variation in residues that bind to variable antigens; and (4) periodic intragenic (interallelic) and intergenic DNA recombinations within class I genes based on the nature of highly mosaic structure among domestic cat MHC class I sequences.

Recent comparative studies of primate MHC class I allelic sequences have revealed additional insight that is relevant to the evolutionary process of MHC class I gene development. First, a remarkable degree of shared polymorphisms between chimpanzee and human homologous class I alleles (Lawlor et al. 1988; Mayer et al. 1988) has been interpreted as evidence of a "trans-species" mode of evolution whereby the established class I alleles are transmitted as polymorphic entities from one ancestral species to descendent modern species. There is also evidence emerging for evolutionary recruitment or interaction of nonclassical class I genes with functional class I loci. For instance, New World primate MHC class I alleles of the cotton-topped tamarin show close homology to human non-classical HLA-G class I sequence (Watkins et al. 1990) while a portion of an Old World gorilla class I allele is quite similar to human nonclassical HLA-H sequence (Watkins et al. 1991; Lawlor et al. 1991).

In this report we examine directly the pattern of sequence divergence of feline MHC class I coding region in three species of felids: domestic cat (*Felis catus*), ocelot (*Leopardus pardalis*), and cheetah (*Acinonyx jubatus*). These three species are representative of distinct evolutionary lineages of the cat family, Felidae (O'Brien 1986; O'Brien et al. 1987). They share a common ancestor that lived between 10 and 15 million years before the present (Neff 1983; Savage and Russell 1983). The cat family has particular interest not only because of the widespread epidemiologic studies on felid viral pathogens (e.g., feline leukemia virus, feline infectious peritonitis virus, feline immunodeficiency virus) (Cotter et al. 1975; Heeney et al. 1990; Pederson et al. 1987; Olmsted et al. 1992), but also because the Felidae provides a useful evolutionary framework of some 37 distinct species as a context for approaching the origins and transmission of well-characterized feline MHC class I variation (Winkler et al. 1989; Yuhki and O'Brien 1988, 1990; Yuhki et al. 1989, 1991). Our results are consistent with an ancient origin of MHC alleles whereby polymorphisms are generated by both mutation and recombinational mechanisms. The occurrence of common sequence motifs in divergent felid species provides evidence for a "trans-species" mode of retention of ancient variation through speciation processes in Felidae.

Materials and Methods

Poly A^+ RNA was isolated from fibroblast culture cell lines of two South African cheetahs (*Acinonyx jubatus jubatus,* Aju 55 and 89) and two ocelots (*Leopardus pardalus,* Lpa 1 and 2) using a FASTTRACK reagent (Invitrogen).

Construction of cDNA libraries were described previously (Yuhki and O'Brien 1990). For PCR amplification of feline MHC class I gene transcripts, approximately 1 μg of poly A^+ RNA was primed with 0.5 μg of oligo $d(T)_{12-18}$ DNA (Stratagene) and first-strand complementary DNA was synthesized using a cloned MuLV Reverse transcriptase (BRL) for 1 h at 42°C. After two cycles of concentration to a 50-μl volume using centricon 100 (Amicon) to eliminate primer DNA, 5 μl of final cDNA solutions as subjected to the following PCR reaction. PCR amplification was designed to cover a mature peptide coding sequence for class I molecules (Yuhki and O'Brien 1990); 5′ primer MHCIPIAB (5′-TTAAGGATCCGCCGCGCCCCAGACCTGGGC-3′) and 3′ primer MHCIPHE (5′-AATTGAATTCACAGATCCTGCATCGCTCAG-3′) were synthesized using a DNA synthesizer (ABI, model 394). Primers MHCIPIAB and MHCIPHE contained *Bam*HI and *Eco*RI restriction sites shown as underline, respectively. PCR amplification was performed for one cycle of 94°C, 5 min, 30 cycles consisting of denaturation of 94°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 2 min, and at 72°C for 10 min after the final cycle. The amplification products (1,098 nts) were treated with proteinase K, polished with Klenow fragment, and digested with *Bam*HI and *Eco*RI restriction enzymes and ligated with *Bam*HI, *Eco*RI double-digested M13 mp18 RF DNA before transformation of DH5 αF′ *Escherichia coli* cells (BRL). Purified recombinant M13 phage DNAs were sequenced by a conventional method using T7 DNA polymerase (U.S.B.) with both dGTP and dITP or by a cycle-sequencing method using Taq DNA polymerase and an automated DNA sequencer (ABI, model 373A) according to manufacturer's instructions. Universal and internal sequencing primers were used to obtain complete sequences.

Sequence analyses and alignment were performed with the GCG package of the University of Wisconsin (Devereux et al. 1984) in the Vax cluster at NCI-FCRDC.

Results

Eight MHC class I cDNA clones were isolated by screening approximately 3×10^5 plaques of a cDNA library constructed from a cheetah (Aju 55) fibroblast cell line using a 1.5-kb domestic cat MHC class I sequence (FLAA24) as a probe (Yuhki et al. 1989); two distinct sequences (designated Ajui1 and Ajui3) were identified by sequencing all eight clones. Three recombinant M13 clones were isolated from another unrelated cheetah (Aju 89) fibroblast cell line by the method of PCR cloning and showed the identical sequences with the Ajui3 clone. Sixteen PCR-amplified class I cDNA clones were isolated from two individual ocelot (Lpa-1 and Lpa-2) fibroblast cell lines. We identified three distinct sequences (designated LPAI69HK, LPAI69I, and LPAI69J) from Lpa-1 and one unique sequence (LPAI70B) from Lpa-2. These two MHC class I sequences from cheetahs plus four sequences from ocelots were compared to seven previously reported domestic cat MHC class I sequences plus a new domestic cat sequence FLAX10, derived from a spleen cDNA library (Fig. 1).

Alignment of nucleic acid and deduced amino acid sequences of cheetah and ocelot MHC genes (Figs. 1 and 2) show near precise homology with domestic cat transcripts, permitting the identification of functional domains defined for human class I molecules—leader, α1, α2, α3, transmembrane, and cytoplasmic domains (Bjorkman et al. 1987a,b; Yuhki and O'Brien 1990). Each domain has been the same length for all transcripts with the exception of the transmembrane domain in ocelot class I molecules, which is two amino acids longer in two transcripts (LPAI69J and LPAI69HK).

We first examined the distribution of genetic variation in residues around the putative antigen recognition sites in α1 and α2 extracellular domains which participate in both T-cell receptor and antigen recognition (Davis and Bjorkman 1988). We previously reported that domestic cat class I molecules have highly conserved residues in this region involved in recognition of common features of antigens and for the conserved faces of α helixes recognized by T-cell receptor (Yuhki and O'Brien 1990). In addition, as for human and rodent class I transcripts, we could identify highly polymorphic residues in the domestic cat class I molecules located on the sites facing the putative antigen binding groove that interact with variable antigen features (Yuhki and O'Brien 1990). The exotic cat transcripts show a similar pattern with homologous conserved residues being virtually invariant among cat species; 29 of 31 conserved domestic cat residues were identical in all transcripts from the three species (Table 1, Fig. 3). In addition, a majority of highly polymorphic sites (more than three polymorphic residues per site) found in all feline molecules were located on the sites facing the antigen binding groove (21 of 25 highly polymorphic sites) (Table 1, Fig. 3). These results were consistent with the pattern for human class I alleles, suggesting that MHC class I molecules of these feline species have similar antigen binding sites to those of human molecules. The parallels in strictly conserved vs highly polymorphic residues reinforce the evolutionary inference of adaptive functional constraints on sequence diversity for this molecule (Yuhki and O'Brien 1990; Lawlor et al. 1988, 1991; Mayer et al. 1988).

The extent of nucleotide and amino acid sequence similarity between all pair-wise combinations of the 14 felid class I sequences is presented in Table 2. In general, the amount of similarity between species was equivalent to the similarity of transcripts from different individuals of the same species. For example, nucleotide sequence similarity within species was 93% for cheetah, 93–99% for ocelot, and 92–100% for domestic cat. These values are comparable to sequence similarities of transcripts from different species: 91–95% nucleotide sequence similarity between domestic cat and ocelot; 90–91% similarity between ocelot and cheetah; and 90–93% similarity between domestic cat and cheetah. Amino acid sequence comparisons are also equivalent within and between species (Table 2). Of a total of 450 polymorphic nucleotides (nts) among these 14 class I sequences, 308 nts (67%) were found in class I sequences of at least two species. Fifty-two nucleotide substitutions (13%) were species-specific and the remaining 90 nts (20%) were allele-specific. The near-equivalent distribution of shared variation between divergent feline species, both quantitatively and qualitatively, suggests that much of the variation is ancestral and had been transmitted en bloc to the derivative species.

We have previously reported that the pattern of amino acid variation in MHC transcripts from domestic cats is punctuated by a mosaic or patchwork pattern of residue motifs that are likely derived by recombinational gene exchange (Yuhki and O'Brien 1990). To examine this hypothesis for other felids, we illustrate only the polymorphic nucleotides from the 14 aligned transcripts in Fig. 4a (for the entire sequence) and a highly polymorphic first α-helix region in Fig. 4b. These comparisons revealed some additional mosaic patterns that reinforce both the ancient (prior to species divergence) and recombinational origin of transcript variation. For example, in exon 2 (Fig. 4b) there are two versions of an 22-nucleotides motif and each type is present in domestic cats, in cheetahs, and in ocelots. Further, shared motifs are found in transcripts from alternative class I A and B subloci of the domestic cat (i.e., FLAA10, A23, A24, and A1 are A locus transcripts while FLAB2 and B9 are B locus transcripts, see ref. Yuhki and O'Brien 1990). One of the ocelot transcripts, LPA70B, had an identical sequence in the first 22-bp region of the α helix with two domestic cat transcripts, FLAA10 and FLAX8 (Fig. 4b). The next 23-bp segment was similar between the species transcripts but the third 21-bp segment shows that LPA70B had a sequence similar to FLAB9 and AJUI3. Similar patterns were also observed for other sequence motifs highlighted in Fig. 4. Moreover, in addition to motifs shared between species there were also sequence motifs unique to certain species—e.g., AJUI3 and LPA69I (Fig. 4)—often on the same transcripts with an ancestral shared motif.

The cytoplasmic domain of domestic cat transcripts is a 99-nt stretch that exists in two recognizable motifs among eight domestic cat transcripts sequenced to date (Yuhki and O'Brien 1990). Based on the pattern in human HLA loci where the cytoplasmic domain is locus specific, it is likely that these two motifs derive from separate subloci FLA-A and FLA-B in the domestic cat (Yuhki and O'Brien 1990). The ocelot transcripts showed only the FLA-A-type cytoplasmic domains (green in Fig. 4a) while the cheetah transcripts display only the FLA-B-like motif (red in Fig. 4a) regardless of the types of mosaic structure in their first α helix coding regions (Fig. 4a).

A database search for homologous sequences to these

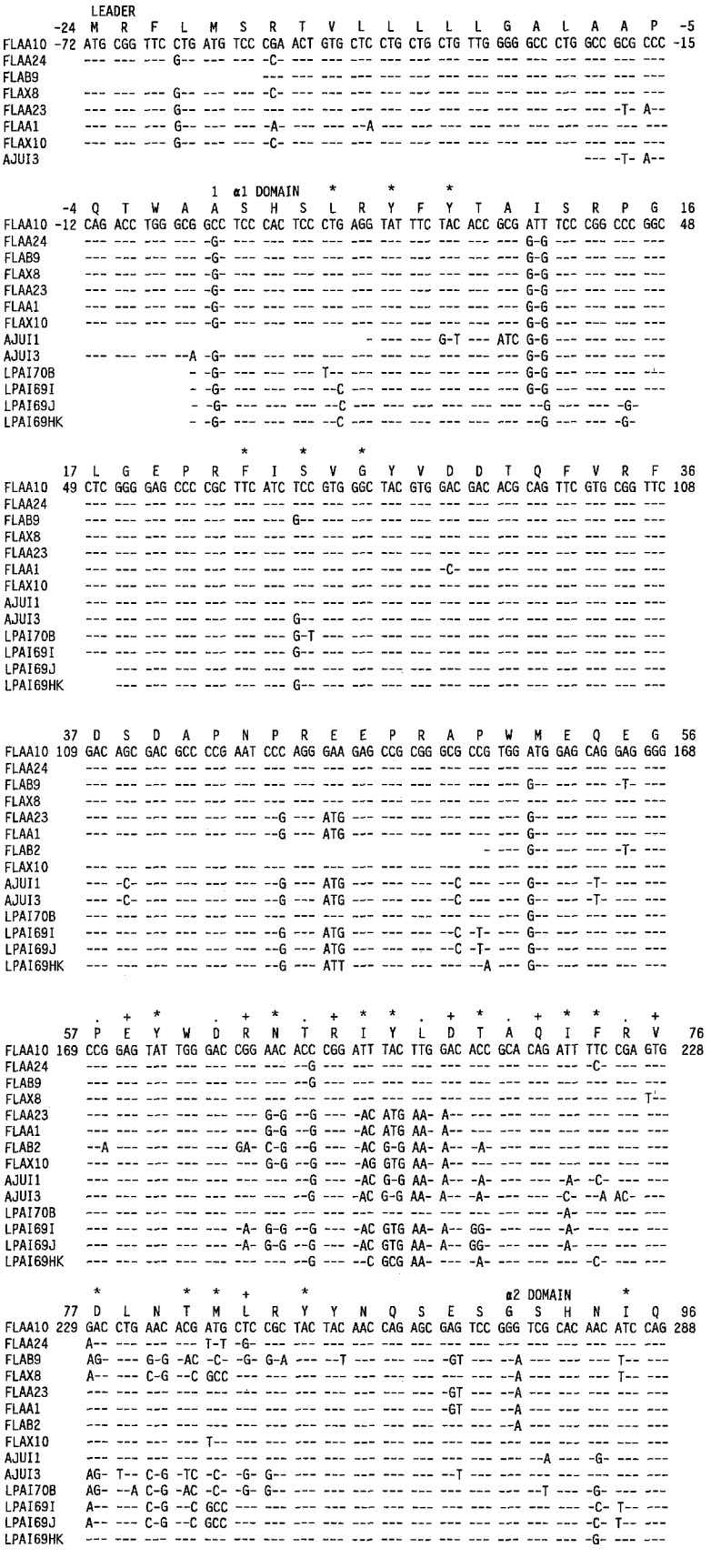

Fig. 1. Nucleotide sequences of 14 distinct feline class I cDNA clones: 8 domestic cat—FLA; 2 cheetah—AJU; and 4 ocelot—LPA. The complete sequences are shown for the FLAA10 clone. The *dashed lines* and *single letters* indicate the identical and distinct nucleotides from the FLAA10 clone. Gaps were introduced to maximize similarities. A deduced amino acid sequence is shown above the FLAA10 sequence. *Asterisk, plus, and dot symbols* above the sequence if FLAA10 indicate residues toward the antigen binding site, 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 (Bjorkman et al. 1987a,b).

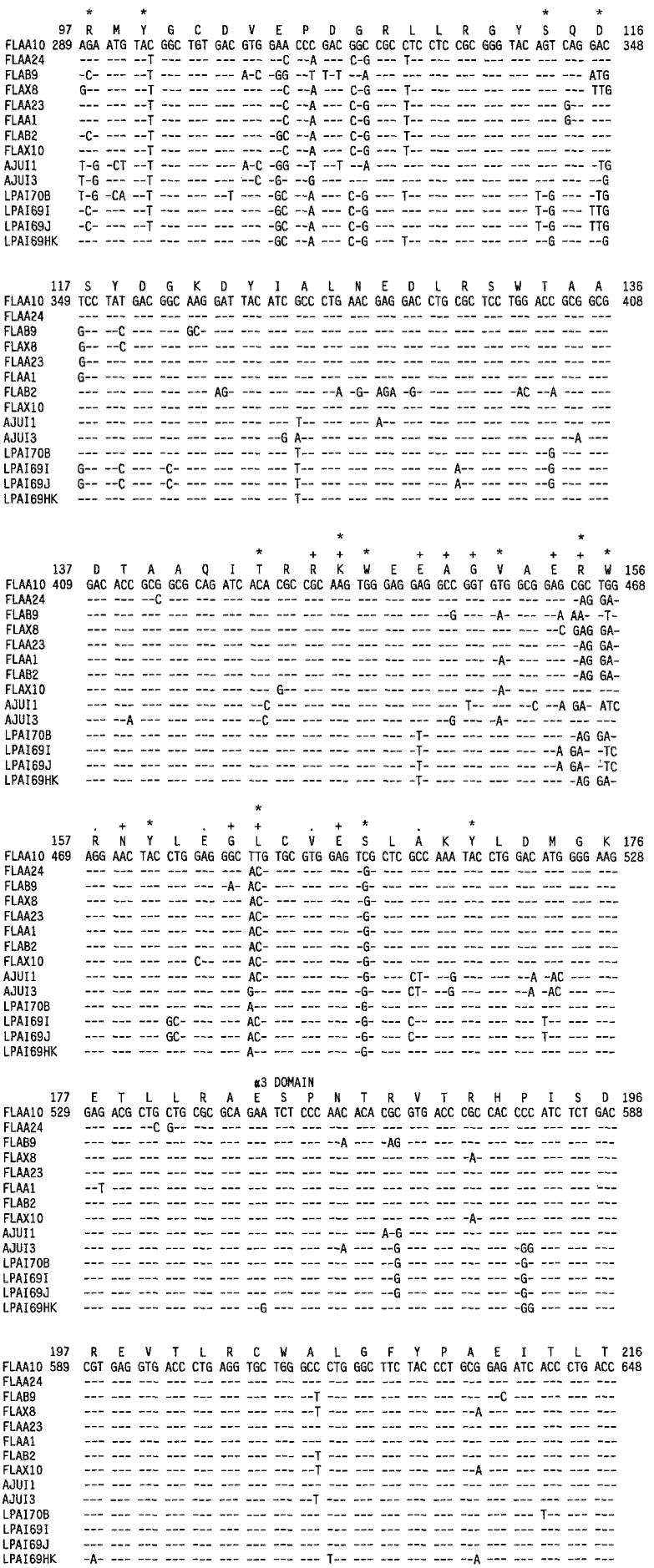

Fig. 1. Continued.

```
      217  W   Q   R   D   G   Q   D   H   T   Q   D   T   E   L   V   E   T   R   P   A  236
FLAA10 649 TGG CAG CGT GAT GGG CAG GAC CAC ACC CAG GAC ACA GAG CTT GTG GAG ACC AGG CCT GCG 708
FLAA24     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAB9      --- --- --- --- --- G-- --- --- --- --- --- G-- --- --- --- --- --- --- --- ---
FLAX8      --- --- --- --- --- --- --- --- --- --- --- G-- --- --- --- --- --- --- --- ---
FLAA23     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAA1      --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAB2      --- --- --- --- --- --- --- --- --- --- --- G-- --- --- --- --- --- --- --- ---
FLAX10     --- --- --- --- --- --- --- --- --- --- --- G-- --- --- --- --- --- --- --- ---
AJUI1      --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
AJUI3      --- --- --- --- --- G-- --- --- --- --- --- G-- --- --- --- --- --- --- --- ---
LPAI70B    --- --- --- A-- --- G-- --- --- --- --- --- --- --- T-- --- --- --- --- --- ---
LPAI69I    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
LPAI69J    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
LPAI69HK   --- --- -A- --- --- --- --- --- --- --- --- --- --- --- --- --C --- --- --A ---

      237  G   D   G   T   F   Q   K   W   A   A   V   V   V   P   S   G   E   E   Q   R  256
FLAA10 769 GGA GAT GGG ACC TTC CAG AAG TGG GCG GCT GTG GTG GTG CCT TCT GGA GAG GAG CAG AGA 768
FLAA24     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAB9      --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAX8      --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAA23     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAA1      --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAB2      --- --- --- --- --- --- --- --- --A --- --- --- --- T-- --- --- --- --- --- ---
FLAX10     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
AJUI1      --- --- --- -G- --- --- --- --- --- --- --- --- --- T-- --- --- --- --- --- ---
AJUI3      --- --- --A --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
LPAI70B    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
LPAI69I    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
LPAI69J    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
LPAI69HK   --- --- --A --- --- --- --- --- --A --- --- --- --- --- --- --- --- --- --- ---

                                                                                TRANSMEMBRANE
      257  Y   T   C   H   V   Q   H   K   G   L   P   E   P   I   N   L   R   W   E   P  276
FLAA10 709 TAC ACG TGC CAT GTG CAG CAC AAG GGG CTG CCC GAG CCC ATC AAC TTG AGA TGG GAG CCA 828
FLAA24     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAB9      --- --- --- --- --- --- --- G-- --- --- --- --- --- --- -C- C-- --- --- --- ---
FLAX8      --- --- --- --- --- --- --- G-- --- --- --- --- --- --- -C- --- --- --- --- ---
FLAA23     --- --- --- --- --- --- --- G-- --- --- --- --- --- --- --- --- --- --- --- ---
FLAA1      --- --- --- --- --- --- --- G-- --- --- --- --- --- --- --- --- --- --- --- ---
FLAB2      --- --- --- --- --- --- --- G-- --- --- -G- --- --- --- -C- --- --- --- --A ---
FLAX10     --- --- --- --- --- --- --- G-- --- --- --- --- --- --- -C- --- --- --- --- ---
AJUI1      --- --- --- --- --- --- --- G-- --- --C G-- --- --- --- -C- C-- --- --- --- ---
AJUI3      --- --- --- --- --- --- --- G-- --- --C G-- --- --- --- -C- C-- --- --- --- ---
LPAI70B    --- --- --- --- --- --- --- G-- --- --- --- --- --- --- -C- C-- --- --- --- ---
LPAI69I    --- --- --- --- --- --- --- G-- --- --- --- --- --- --- -C- C-- --- --- --- ---
LPAI69J    --- --- --- --- --- --- --- G-- --- --- --- --- --- --- -C- C-- --- --- --- ---
LPAI69HK   --- --- --- --- --- --- --- G-- --- --- --- --- --- --- -C- C-- --- --- --- --- GAG CCA

      277  S   S   L   P   F   I   T   I   L   G   I   I   A   G   V   A   V   L   V   V  296
FLAA10 829 TCG TCT CTG CCC TTC ATC ACC ATT CTG GGC ATC ATT GCT GGT GTG GCT GTC CTT GTG GTC 888
FLAA24     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAB9      --C --- -A- --- -C- --- C-- T-C --- --- --- --- --- --- --- -T- --- --- --- ---
FLAX8      --C --- -A- --- -C- --- C-- T-C --- --- --- --- --- --- --- -T- --- --- --- ---
FLAA23     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAA1      --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAB2      --C --- -A- --- -C- --- C-- --C --- --- --- --- --- --- --- --- --- --- --- ---
FLAX10     --C --- -A- --- -C- --- C-- T-C --- --- --- --- --- --- --- -T- --- --- --- ---
AJUI1      --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
AJUI3      --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
LPAI70B    --C --- -A- --- -C- --- C-- --C --- --- --- --- --- --- --- -T- --- --- --- ---
LPAI69I    --C --- -A- --- -C- --- C-- --C --- --- --- --- --- --- --- -T- --- --- --- ---
LPAI69J    --C --- -A- --- -C- --- C-- --C --- --- --- --- --- --- --- -T- --- --- --- ---
LPAI69HK   --C --- -A- --- -C- --- C-- T-C --- --- --- --- --- --- --- -T- --- --- --- ---

                                                CYTOPLASMIC
      297  T   V   V   V   G   A   V   I   W   R   K   K   C   S   G   G   K   G   P   I  316
FLAA10 889 ACT GTG GTG GTT GGA GCT GTG ATC TGG AGG AAG AAG TGC TCA GGA GGA AAA GGA CCA ATC 948
FLAA24     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAB9      --- T-- --- --- --- --- --- --- --- --- --- --- -T- --- --- --- --G --- --- -G-
FLAX8      --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- A-- --- --- --- -G-
FLAA23     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAA1      --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAB2      --- --- --- --- --- --- --- --- --- --- --- --- -T- --- --- --- --G --- --- CG-
FLAX10     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- A-- --- --- --- -G-
AJUI1      --- --- --- --- --- --- --- --- --- --- --- --- -TT --- --- --- --- --- --- -G-
AJUI3      --- --- --- --- --- --- --- --- --- --- --- --- -T- --- --- --- --G --- --- -G-
LPAI70B    --- T-- --- --- --- --- --- --- --- --- --- --- -TT --- --- --- --- --- --- -G-
LPAI69I    --- T-- --- --- --- --- --- --- --- --- --- --- -TT --- --- --- --G --- --- -G-
LPAI69J    --- T-- --- --- --- --- --- --- --- --- --- --- -TT --- --- --- --- --- --- -G-
LPAI69HK   --- T-- --- --- --- --- --- --- --- --- --- --- -TT --- --- --- --- --- --- -G-

      317  Y   S   H   A   A   R   D   D   S   T   Q   G   S   D   S   S   L   M   A   P  336
FLAA10 949 TAT TCT CAC GCT GCA CGC GAC GAC AGT ACC CAG GGC TCT GAT TCG TCT CTA ATG GCT CCT 1008
FLAA24     --- --- --G C-- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAB9      --- --- --- --- --- --- A-T --- --- --T --- --- --- --- --- --- --C --- --- ---
FLAX8      --C --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAA23     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAA1      --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FLAB2      --- --- --- --- --- --- --T --- --- --T --- --- --- --- --- G-- --- --- --- ---
FLAX10     --C --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
AJUI1      --C --- --T --- --- T-- --T --- --- --- --- --- --- --- --C --- --C --- --- -T-
AJUI3      --- --- --T --- --- T-- --T --- --- --- --- --T --- --- --- G-- --C --C --- ---
LPAI70B    --- --- --- --- --- --- --- --- --- --- --- --- --- --- -G- --- --- --- --- ---
LPAI69I    --- --- --- --- --- --- --- --- --- --- --- --- --- --- -G- --- --- --- --- ---
LPAI69J    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
LPAI69HK   --C --- --T --- --- -A- --T --- --- --- --- --- --- --- --- --C --- --- ---
```

Fig. 1. Continued.

```
        337  K   V  END 3'UT
FLAA10 1009 AAA GTT TGAGACCCACTGCCTGTGGTGGA
FLAA24      --- --- --------G----------A-A-
FLAB9       --- --G -------------T---A-----
FLAX8       --- --- --------G----------A-A-
FLAA23      --- --- --------------------A-A-
FLAA1       --- --- --------------------A-A-
FLAB2       --- --G -------------T---------
FLAX10      --- --- --------G----------A-A-
AJUI1       --- --G ----------------------C-
AJUI3       --- --- ----------------------
LPAI70B     C-- --G --------G----------A-A-
LPAI69I     C-- --G --------G----------A-A-
LPAI69J     --- --- ------T------T---------
LPAI69HK    --- --- ------T------T---------
```

Fig. 1. Continued.

polymorphic sequence motifs revealed unexpected sequence similarities in class I sequences of other distinct species. For example, highly conserved 23-bp-sequence motifs consisting of the first α-helix coding region in domestic cat sequences were also found in human nonclassical class I HLA-E gene and in bovine and orangutan class I sequences (green in Fig. 4b). A similar observation was obtained by comparing domestic cat sequences with human classical HLA-B5, A1, and chimpanzee B1 and A108 class I sequences (yellow in Fig. 4b). These results suggest, first, that these sequence motifs are ancient and established before speciation of these feline species and second, that extensive exchanges of these polymorphic motifs took place to generate mosaic structures not only within but also between two class I loci.

Finally, the pattern of sequence variation in the regions of the class I molecule involved in antigen recognition (57 amino acid residues termed the antigen recognition site, ARS) was examined to detect evidence for selective pressure in favor of persistent polymorphism (Bjorkman et al. 1987a,b; Yuhki and O'Brien 1990; Hughes and Nei 1988; Nathenson et al. 1986). We previously argued that positive selection pressure has contributed to extensive polymorphism and residue heterozygosity in the ARS of domestic cat molecules, since nearly all the nucleotide substitutions in this region are nonsynonymous while they are predominantly synonymous in other residues of the same molecule (Yuhki and O'Brien 1990; Hughes and Nei 1988). We examine here the pattern of 59 α-1 and α-2 domain nucleotide substitutions that were found as allele-specific and considered as newly generated substitutions in comparison of the 14 transcripts from the three felid species. A large excess of nonsynonymous nucleotide substitutions was observed in the ARS (20 nonsynonymous:2 synonymous nucleotide substitutions), while no tendency was found in nearby residues ($N = 125$) not involved in antigen recognition (16 nonsynonymous:21 synonymous nucleotide substitutions). These results extend our inference (Yuhki and O'Brien 1990) that positive selection pressure is operating to accumulate functional amino acid substitutions in the antigen recognition site of these molecules during Felidae evolution.

Discussion

The MHC class I genes of three feline species (domestic cat, ocelot, and cheetah), which were estimated to have diverged from a common ancestor 10–15 million years ago, display highly mosaic MHC class I molecules consisting of short polymorphic sequence motifs in their entire coding region. A majority of similar motifs were recognized in class I alleles of at least two feline species. However, none of these class I alleles had an identical combination of these motifs in their mosaic structures. These observations lead to several conclusions: (1) A majority of polymorphic sequence motifs were established before speciation of these feline species, and some of them may trace back to the origins of mammal orders, since identical or similar motifs were found in higher primate and bovine class I sequences; (2) since these motif sequences were recognized in two or three divergent feline species, ancestral species from which they descended had limited numbers of the motifs, but these were rearranged by recombination before/during species divergence; (3) although these motif sequences were well conserved, most new nucleotide substitutions found at 57 amino acid residues of the antigen recognition site were nonsynonymous, indicating that an accumulation of nucleotide substitutions and a subsequent positive selection pressure are other important factors in establishing new polymorphisms of feline MHC class I molecules.

So far two competing hypotheses have been proposed to account for the speed and mode of the evolution of MHC class I alleles based on analysis of murine and primate sequences. Extensive examinations of a series of murine H2-K^b mutant class I alleles revealed that these mutant alleles have various types of clustered nucleotide substitutions in polymorphic first and second extracellular-domain coding regions (Nathenson et al. 1986). A screening of complete sets of cosmid contigs for murine MHC class I genes indicated that H2-D and -Qa genes were likely origins for these clustered substitutions. These results suggested that new murine class I alleles can be created very rapidly by gene conversion or intergenic DNA recombination. In contrast, however, comparisons of 40 human HLA-A, -B, and -C and 10 chimpanzee class I allelic sequences reveal ancient intragenic (interallelic) DNA recombinations as a major factor for their mosaic structures because of intralocus similarities of polymorphic sequence motifs found in human HLA-A, -B, and -C alleles and high similarities between all chimpanzee and their homologous human class I allelic sequences (Lawlor et al. 1988; Mayer et al. 1988).

Our present data for feline class I alleles support both intergenic and intragenic (interallelic) DNA recombinations to generate mosaic structures of MHC class I allelic sequences. However, we suggest that, unlike murine H-$2K^b$ mutant alleles, a majority of inter-

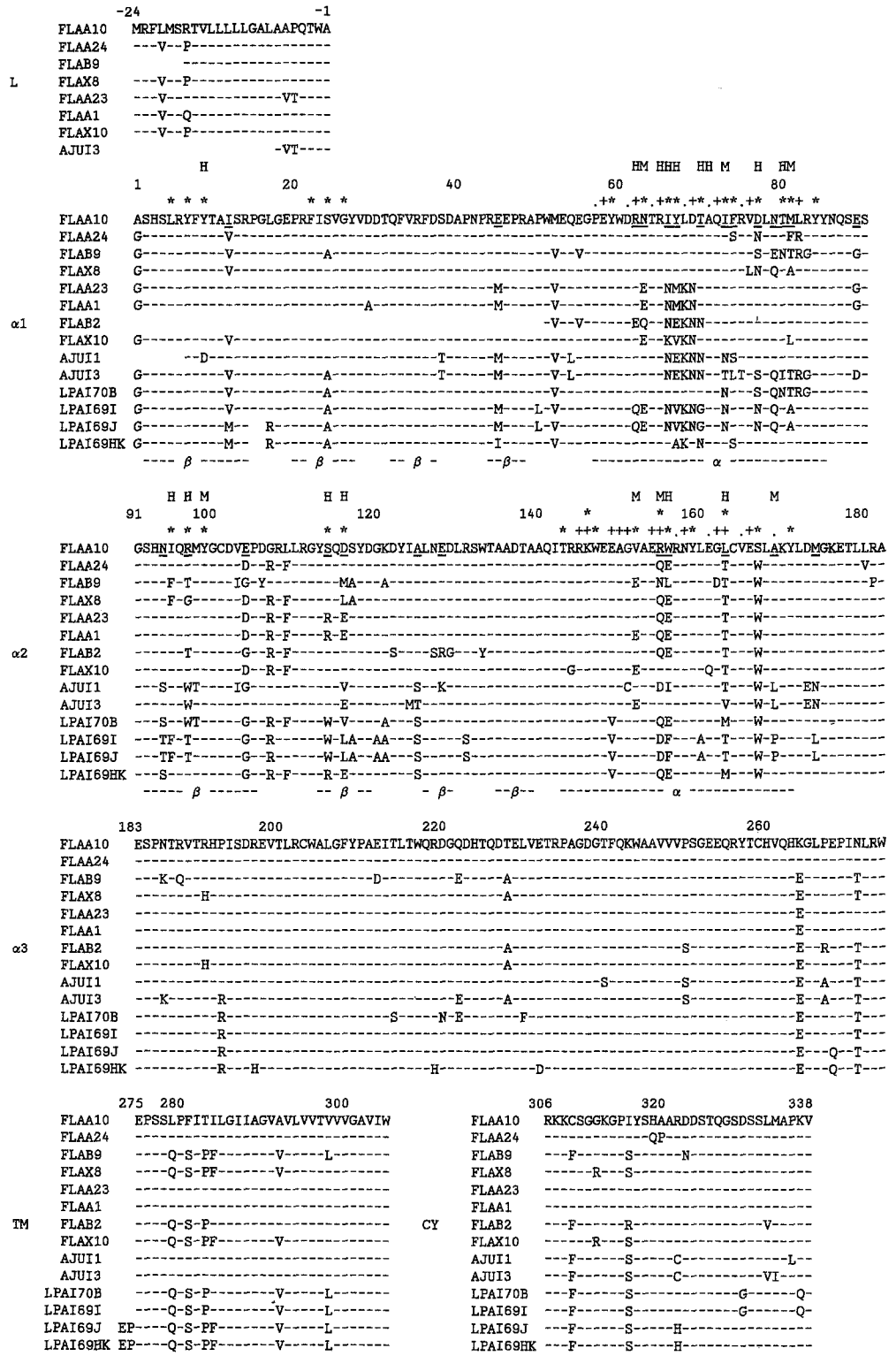

Fig. 2. Deduced amino acid sequences of the 14 distinct feline MHC class I sequences. The complete sequence for the FLAA10 molecule is shown. *Dashed lines and a single letter* indicate the amino acid residues identical to and different than those of FLAA10, respectively. Two α-helix coding regions and 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 sites that showed over three different residues and are indicated as *underlines* of the FLAA10 sequence.

Table 1. Highly conserved and highly polymorphic residues in α1 and α2 domains in MHC class I alleles of three feline species—domestic cat, cheetah, and ocelot*

Criteria	Positions
Highly conserved residues	
Recognition of common features of Ag	L5[a], Y7, F22, G26, Y59, Y84, T143, K146, Y159, Y171
Formation of conserved faces of α-helix	P57, E58, D61, T64, Q72, R75[b], L78, Y84, K146, A150, E154, G162[b], V165 (I73), H158[a], (A169)
N-glycosylation site	N86, Q87, S88, C101, C164
Disulfide bonding	C101, C164
Highly polymorphic sites[c]	
β sheets in α1 domain	12, 45
First α helix	62, 63, 66, 67, 70, 73, 74, 77, 79, 80, 81
β sheets in α2 domain	94, 97, 104, 114, 116, 128
Second α helix	155, 156, 163, 169
Other regions	89, 125, 128, 174

* All conserved residues except underlined residues were also conserved among human class I molecules. Underlined residues represent conserved residues in domestic cat class I molecules. Of these residues: [a]Only two residues, L5 and H158, were identical in other exotic cat class I molecules. [b]Only one allele showed a distinct residue at this position. [c]These sites of amino acid residues showed more than three distinct residues.

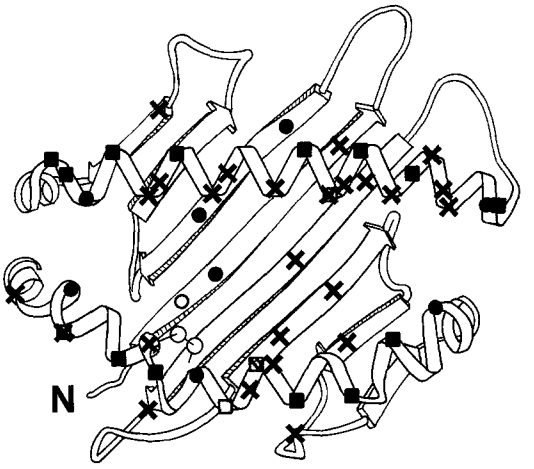

Fig. 3. Locations of highly conserved and highly polymorphic residues found in 14 feline MHC class I molecules on an X-ray crystallographic model of α1 and α2 domains of the human HLA-A2 molecule. *X* symbols indicate the position of highly polymorphic residues which showed more than three residues per site. *Solid symbols* indicate conserved and identical residues among human and feline class I molecules. *Open symbols* indicate conserved but distinct in human and feline molecules. *Circles* indicate conserved residues for common features of antigens. *Squares* indicate conserved residues for conserved faces of α helixes.

Table 2. Percentage nucleotide (top) and amino acid (bottom) sequence similarity between feline class I genes

	FLA								AJUI		LPAI			
	A10	A24	B9	X8	A23	A1	B2	X10	1	3	70B	69I	69J	69HK
FLAA10		97	93	95	96	96	93	96	92	92	91	91	93	93
FLAA24	95		93	96	97	97	93	96	92	91	92	91	94	94
FLAB9	88	87		94	92	92	91	93	90	92	93	92	91	94
FLAX8	91	93	90		95	95	93	93	90	91	94	93	93	95
FLAA23	94	95	87	91		100	95	97	93	93	94	93	95	93
FLAA1	93	94	87	91	99		95	96	93	92	93	93	94	93
FLAB2	89	88	85	88	91	91		95	91	90	92	91	93	91
FLAX10	92	92	89	94	93	93	90		91	92	93	93	94	93
AJUI1	88	87	83	85	90	89	88	86		93	91	91	91	91
AJUI3	87	86	87	85	88	88	85	86	90		91	90	91	91
LPAI70B	86	86	87	89	89	88	86	88	86	85		99	93	94
LPAI69I	84	84	86	88	87	86	86	87	85	84	97		94	93
LPAI69J	88	89	86	89	90	89	88	89	86	85	88	91		93
LPAI69HK	88	89	90	89	88	87	85	88	85	86	90	88	90	

genic DNA recombinations in feline class I alleles occurred only between (or among) functional or at least transcriptionally active class I genes, since the donor and recipient alleles for polymorphic motifs were easily recognized by the screening of feline cDNA class I sequences in this study.

We also present an interesting type of "trans-species" mode for evolution in feline MHC class I alleles which show a high degree of retention for only short polymorphic motifs but not for entire coding sequences in MHC class I alleles of three feline species (Fig. 4a,b). This pattern is in contrast with human and chimpanzee class I alleles in which a high degree of retention was observed in their entire coding sequences. There are two possible explanations for these observations—first, a majority of possible recombinations for mosaic sequences were completed before speciation of feline species. The repertoire of modern MHC class I alleles in each feline species would represent a subset of ancestral class I alleles which new species have inherited from a pool of polymorphic alleles in the ancestral feline species. The second hypothesis poses a more recent and active involvement of DNA recombination for the generation of modern class I alleles, as is observed in murine MHC class I mutant alleles. Although dynamic gene-segment exchange could conceivably generate mo-

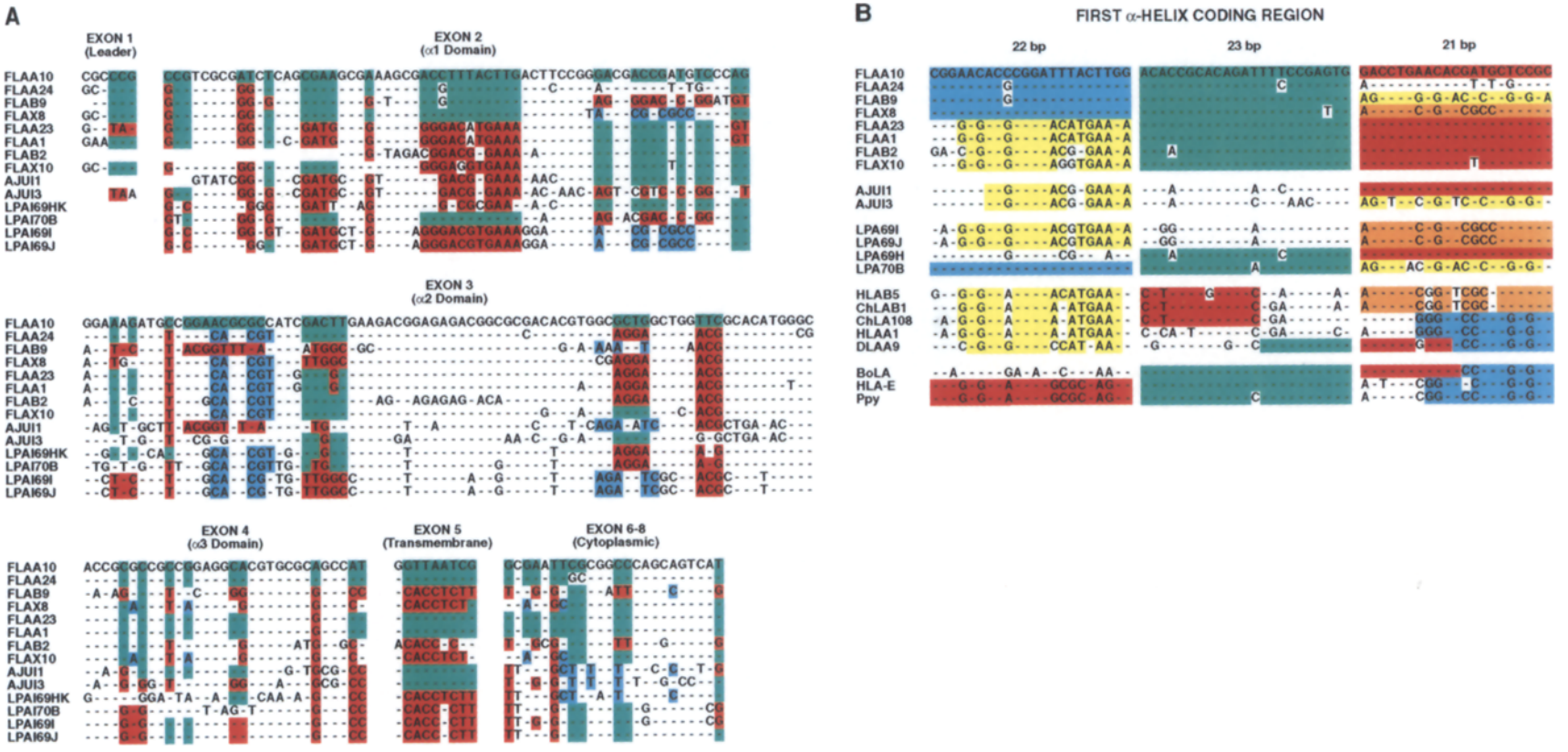

Fig. 4. **A** Polymorphic nucleotides in 14 feline MHC class I cDNA clones. All polymorphic nucleotides were extracted and aligned in each exon shown above the sequences. Nucleotides of the FLAA10 cDNA clone were used as reference sequences. *Dashed lines and single letters* indicate the identical and distinct nucleotides with respect to the FLAA10. Colors indicate motif identities distributed between class I transcripts. **B** Typical mosaic sequences in the first α-helix coding regions of MHC class I molecules. FLAA10 sequence was used as a reference sequence. *Dashed lines and single letters* indicate the identical and distinct nucleotides with reference to FLAA10 sequence, respectively. Nonfeline MHC class I sequences (ChLA: chimpanzee, DLA: dog, BoLA: cow, Ppy: orangutan) were obtained from GenBank

saic genes that are observed in feline MHC class I alleles, it is unlikely that ancestral retention does not play a primary role for at least two reasons. First, the size of overall variation between the three divergent Felidae lineages is no greater than intraspecies variation, suggesting rather slow rates of mutational divergence (Table 2). Second, we have recently obtained feline class I sequences from 14,000-year-old fossils of *Smilodon fatalis,* the extinct saber-toothed cat (Janczewski et al. 1992). The MHC class I gene segments of this species have motifs that are nearly identical to the sequences presented in Fig. 4b, and more importantly, combinations of these motifs are also similar to two of the modern domestic cat sequences, indicating the ancient origins of the motifs and the recombinational events (Janczewski et al. 1992). Recent accumulations of class I and class II allelic sequences from other species outside of the Felidae also reveal a similar mosaic structure in their MHC alleles and support our conclusions. For example, PCR sequencing of rodent Ab class II genes reveals rodent-wide retention of short polymorphic motifs in their mosaic β 1 domain (She et al. 1991). Moreover, 14 gorilla class I allelic sequences did not show any overall similarity to human or chimpanzee class I alleles, but the majority of polymorphic substitutions and sequence motifs of the gorilla class I alleles are shared with the human and chimpanzee systems (Lawlor et al. 1991).

Finally, highly conserved short polymorphic motifs consisting of the first α-helix coding sequence between domestic cat, cow, and human nonclassical HLA-E genes may provide another clue for recruitment or interaction of non-classical class I genes for functional class I alleles.

References

Bjorkman PJ, Saper MA, Samraoui B, Bennette WS, Strominger JL, Wiley DS (1987a) Structure of the human class I histocompatibility antigen, HLA-A2. Nature 329:506–512

Bjorkman PJ, Saper MA, Samraoui B, Bennette WS, Strominger JL, Wiley DC (1987b) The foreign antigen binding site and T-cell recognition regions of class I histocompatibility antigens. Nature 329:512–518

Brown JH, Jardetsky T, Saper MA, Samraoui B, Bjorkman PJ, Wiley DC (1988) A hypothetical model of the foreign antigen binding site of class II histocompatibility molecules. Nature 332:845–850

Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, Wiley DC (1993) Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature 364:33–39

Cotter SM, Hardy WD Jr, Essex M (1975) Association of feline leukemia virus with lymphosarcoma and other disorders in the cat. J Am Vet Med Assoc 166:449–454

Davis MM, Bjorkman PJ (1988) T-cell antigen receptor genes and T-cell recognition. Nature 334:395–402

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the Vax. Nucleic Acids Res 12: 387–395

Falk K, Rotzschke O, Stevanovic S, Jung G, Rammensee H-G (1991) Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. Nature 351:290–296

Heeney JL, Evermann JF, Mckeirnan AJ, Marker-Kraus L, Roelke ME, Bush M, Wildt DE, Meltzer DG, Cooly L, Lukas J, Manton VJ, Caro T, O'Brien SJ (1990) Prevalence and implications of feline coronavirus infections of captive and free-ranging cheetahs (*Acinonyx jubatus*). J Virol 64:1964–1972

Hughes AL, Nei M (1988) Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature 335:167–170

Janczewski DN, Yuhki N, Gilbert DA, Jefferson GT, O'Brien SJ (1992) Molecular phylogenetic inference from saber-toothed cat fossils of Rancho La Brea. Proc Natl Acad Sci USA 89:9769–9773

Klein J (1986) Natural history of the major histocompatibility complex. John Wiley and Sons, New York

Lawlor DA, Ward FE, Ennis PD, Jackson AP, Parham P (1988) HLA-A and -B polymorphisms predate the divergence of humans and chimpanzees. Nature 335:268–271

Lawlor DA, Warren E, Taylor P, Parham P (1991) Gorilla class I major histocompatibility complex alleles: comparison to human and chimpanzee class I. J Exp Med 174:1491–1509

Mayer WE, Jonker M, Klein D, Ivanyi P, van Seventer G, Klein J (1988) Nucleotide sequences of chimpanzee MHC class I alleles: evidence for trans-species mode of evolution. EMBO J 7:2765–2774

Nathenson SG, Geliebter J, Pfaffenbach GM, Zeff RA (1986) Murine major histocompatibility complex class I mutants: molecular analysis and structure-function implications. Annu Rev Immunol 4:471–502

Neff N (ed) (1983) The big cats: the paintings of Guy Coheleach. Harry N Abrams, New York, p 1007

O'Brien SJ (1986) Molecular genetics in the domestic cat and its relatives. Trends Genet 2:137–143

O'Brien SJ, Collier GE, Benveniste RE, Nash WG, Newman AK, Simonson JM, Eichelberger MA, Seal US, Bush M, Wildt DE (1987) Setting the molecular clock in Felidae: the great cats, *Panthera.* In: Tilson RL, Seal US (eds) Tigers of the world: the biology, bipolitics, management, and conservation of an endangered species. Noyes Publications, Park Ridge, NJ, p 10

Olmsted RA, Langley R, Roelke ME, Goekin RM, Adjar Johnson D, Goff J, Albert J, Packer C, Laurenson MK, Caro TM, Scheepers L, Wildt DE, Bush M, Martenson JS, O'Brien SJ (1992) Worldwide prevalence of lentivirus infection of wild feline species: epidemiologic and genetic aspects. J Virol 66:6008–6018

Pederson NC, Ho EW, Brown ML, Yamamoto JK (1987) Isolation of a T-lymphotropic virus from domestic cats with an immunodeficiency-like syndrome. Science 235:790–793

Rotzschke O, Falk K, Deres K, Schild H, Norda M, Metzger J, Jung G, Rammensee H-G (1990) Isolation and analysis of naturally processed viral peptides as recognized by cytotoxic T cells. Nature 348:252–254

Savage DE, Russell DE (eds) (1983) Mammalian paleofaunas of the world. Addison-Wesley, London, p 1779

She JX, Boehme SA, Wang TW, Bonhomme F, Wakeland EK (1991) Amplification of major histocompatibility complex class II gene diversity by intraexonomic recombination. Proc Natl Acad Sci USA 88:453–457

Van Bleek GM, Nathenson SG (1990) Isolation of an endogenously processed immunodominant viral peptide from the class I H-2K^b molecule. Nature 348:213–216

Watkins DI, Chen ZW, Hughes AL, Evans MG, Tedder TF, Letvin NL (1990) Evolution of the MHC class I genes of a New World primate from ancestral homologues of human non-classical genes. Nature 346:60–63

Watkins DI, Chen ZW, Garber TL, Hughes AL, Letvin NL (1991) Segmental exchange between MHC class I genes in a higher primate: recombination in the gorilla between the ancestor of a hu-

man non-functional gene and an A locus gene. Immunogenetics 34:185–191

Winkler CA, Schulz A, Cevario S, O'Brien SJ (1989) Genetic characterization of FLA, the cat major histocompatibility complex. Proc Natl Acad Sci USA 86:943–947

Yuhki N, O'Brien SJ (1988) Molecular characterization and genetic mapping of class I and class II MHC genes of the domestic cat. Immunogenetics 27:414–425

Yuhki N, Heidecker GF, O'Brien SJ (1989) Characterization of MHC cDNA clones in the domestic cat. Diversity and evolution of class I genes. J Immunol 142:3676–3682

Yuhki N, O'Brien SJ (1990a) DNA variation of the mammalian major histocompatibility complex reflects genomic diversity and population history. Proc Natl Acad Sci USA 87:836–840

Yuhki N, O'Brien SJ (1990b) DNA recombination and natural selection pressure sustains genetic sequence diversity of the feline MHC class I genes. J Exp Med 172:621–630

Yuhki N, Winkler C, O'Brien SJ (1991) The MHC genes of the domestic cat. In: Srivastaria R, Ram BP, Tyle P (eds) Immunogenetics of the major histocompatibility complex. VCH, Cambridge, p 348