Published in final edited form as:

Immunogenetics. 2013 September ; 65(9): 675–689. doi:10.1007/s00251-013-0711-z.

Polymorphisms and Tissue Expression of the Feline Leukocyte Antigen Class I Loci FLAI-E, -H and -K

Jennifer C. Holmes¹, Savannah G. Holmer¹, Peter Ross¹, Adam S. Buntzman², Jeffrey A. Frelinger², and Paul R. Hess^{1,*}

¹Immunology Program, and Department of Clinical Sciences, North Carolina State University College of Veterinary Medicine, Raleigh, North Carolina, United States of America

²Department of Immunobiology, University of Arizona, Tucson, Arizona, United States of America

Abstract

Cytotoxic CD8⁺ T-cell immunosurveillance for intracellular pathogens, such as viruses, is controlled by classical major histocompatibility complex (MHC) class Ia molecules, and ideally, these antiviral T-cell populations are defined by the specific peptide and restricting MHC allele. Surprisingly, despite the utility of the cat in modeling human viral immunity, little is known about the Feline Leukocyte Antigen class I complex (FLAI). Only a few coding sequences with uncertain locus origin and expression patterns have been reported. Of 19 class I genes, 3 loci - FLAI-E, -H and -K – are predicted to encode classical molecules, and our objective was to evaluate their status by analyzing polymorphisms and tissue expression. Using locus-specific, PCR-based genotyping, we amplified 33 FLAI-E, -H, and -K alleles from 12 cats of various breeds, identifying, for the first time, alleles across 3 distinct loci in a feline species. Alleles shared the expected polymorphic and invariant sites in the 1/ 2 domains, and full-length cDNA clones possessed all characteristic class Ia exons. Alleles could be assigned to a specific locus with reasonable confidence, although there was evidence of potentially confounding interlocus recombination between FLAI-E and -K. Only FLAI-E, -H and -K-origin alleles were amplified from cDNAs of multiple tissue types. We also defined hypervariable regions across these genes, which permitted the assignment of names to both novel and established alleles. As predicted, FLAI-E, -H, and -K fulfill the major criteria of class Ia genes. These data represent a necessary prerequisite for studying epitope-specific antiviral CD8⁺ T-cell responses in cats.

Keywords

Cat; Major histocompatibility complex; Polymorphism; Classical class I genes

Introduction

Effective immunosurveillance of infected cells by CD8⁺ T cells depends on T-cell receptor (TCR) recognition of specific epitopes that are derived from proteasome-degraded cytosolic proteins and presented by major histocompatibility complex (MHC) class I molecules. Remarkably, only a small fraction of the peptides that potentially can be generated from a given antigen – on the order of 1-3% – are capable of being processed and binding with sufficient affinity to stabilize the surface peptide-MHC complex (Assarsson et

*Corresponding author: Paul R. Hess, Department of Clinical Sciences, North Carolina State University College of Veterinary Medicine, Box 8401, 1060 William Moore Drive, Raleigh, North Carolina 27607, United States of America; Telephone: 919-513-6183; Fax: 919-513-6336; paul_hess@ncsu.edu.

Conflict of Interest Statement: The authors declare that they have no conflicts of interest.

al. 2007). Despite this limitation, the potential vulnerability of CD8⁺ T-cell defenses to immunoevasion through the mutation of epitope-encoding sequences is greatly reduced by the highly polymorphic nature of classical class I (Ia) molecules. In humans, for example, >4500 Human Leukocyte Antigen (HLA) class Ia alleles are known (Robinson et al. 2011). Because different alleles preferentially bind peptides with specific amino acid types at anchor positions (the binding motif), the collective pool of variants provides much greater coverage of the pathogen proteome than the restricted repertoires of individual alleles would suggest. Class Ia diversity is therefore a barrier within the population to the spread of intracellular organisms that are selected under immune pressure to evade detection. Even within an individual, protective polymorphisms are in force, as class Ia molecules are usually polygenic, with high levels of heterozygosity and co-dominant expression of alleles. For example, nucleated human cells typically display six different class Ia molecules – two alleles each at the HLA-A, -B and -C loci – that collectively can present an array of pathogen-derived peptides.

Surprisingly, despite the long-standing value of the domestic cat as a model for human retroviral immunity (reviewed in Elder et al. 2010), the understanding of the MHC class I complex in this species remains comparatively rudimentary. Class I gene products can be detected on the surface of feline cells (Pollack et al. 1988) and transcripts can be found in a variety of tissues (Yuhki et al. 1989). An early estimate of gene number by Southern blotting suggested 20-30 class I sequences per haploid genome (Yuhki and O'Brien 1988). Polymorphisms in class I genes have been found by isoelectric focusing of immunoprecipitated lymphocyte proteins (Neefjes et al. 1986), restriction fragment length polymorphism assay (Yuhki and O'Brien 1988; Yuhki and O'Brien 1990b), and serologic analysis using lymphocytotoxic alloantibodies (Winkler et al. 1989). Initially, such polymorphisms were thought to originate from a single (Yuhki et al. 1989), or possibly two or more (Neefjes et al. 1986), unnamed loci. In a later study, comparison of seven distinct feline class I cDNA sequences obtained from a T-cell lymphoma line and three unrelated cats suggested two loci, designated A and B (Yuhki and O'Brien 1990a). The homology of these clones to HLA-A2 permitted identification of the eight domains characteristic of class Ia molecules, as well as polymorphic residues within the antigen recognition site (ARS). Since then, additional class I alleles have been reported in the cat with tentative assignment to the A and B loci (Sato and Yamamoto 2009) or without regard to specific origin (so-called "Z" locus) (Smith and Hoffman 2001). More recently, the genomic structure of the Feline Leukocyte Antigen (FLA) complex has been detailed (Beck et al. 2005; Yuhki et al. 2007), and large sequences have been annotated (Yuhki et al. 2008). Nineteen class I genes, designated FLAI-A through FLAI-S, have been identified. Of these, seven were gene fragments, nine were assigned as non-classical genes, and three loci – FLAI-E, -H and -K – were considered to be class Ia genes. The assignment of classical status was based on the presence of full-length exons, intact splice sites, specific invariant residues in the 1/ 2 domains, and a complete upstream promoter motif (enh.A-ISRE-W/S-X1-X2/site -Y/enh.B) known to regulate MHC class Ia expression.

There a number of viruses afflicting humans, including human immunodeficiency virus (HIV) and human T-cell leukemia virus (HTLV), that have naturally occurring homologs in cats (reviewed in O'Brien et al. 2012). Moreover, some of these feline viruses threaten wild cat species, such as the lion, panther and lynx (Brown et al. 2008; Meli et al. 2009; Roelke et al. 2009). Given the potential usefulness of the domestic cat to understanding adaptive immune responses to such important pathogens, we sought to better characterize the three FLAI loci that presumably restrict antiviral CD8⁺ T-cell effectors by evaluating polymorphisms and transcription of these genes. Data obtained from 12 cats of various breeds showed high diversity at each locus, with 12, 11 and 10 alleles at FLAI-E, -H and -K, respectively, and cDNAs from all three genes were found in diverse tissue samples. These

results constitute strong support for the designation of FLAI-E, -H and -K as class Ia genes, and provide the groundwork for the study of epitope-specific, cytotoxic T-cell responses in the cat.

Materials and Methods

Isolation and preparation of nucleic acids from cats in the study

Anticoagulated venous blood samples (200 μ L) were obtained with owner consent from clinical specimens collected from 11 cats during evaluation at the North Carolina State University Veterinary Teaching Hospital. Blood and tissue samples were harvested from a feral adult domestic shorthair (DSH) cat (DSH7) euthanized at a community animal shelter. Genomic DNA (gDNA) was extracted using a QIAamp DNA Blood Mini Kit (Qiagen, Valencia, CA), and RNA was purified with an RNeasy Plus Mini Kit (Qiagen) from peripheral blood mononuclear cells (PBMCs) that had been isolated by density gradient centrifugation over Histopaque-1077 solution (Sigma-Aldrich, St. Louis, MO). Tissue samples from DSH7 were placed in RNA*later* solution (Ambion/Life Technologies, Carlsbad, CA) and stored at -80°C until homogenization and RNA isolation were performed.

PCR amplification and cloning of FLAI sequences

Genotyping primers (Supplemental Table S1) were designed using NCBI PrimerBlast software (Rozen and Skaletsky 2000) to generate complete reads of exons 2 and 3 from FLAI-E, -H and -K, based on the genomic sequence EU153401 (Yuhki et al. 2008), as well as earlier sequences and exon assignments (Yuhki and O'Brien 1990a). Amplifications were performed in 25 μ L reaction volumes using 1 μ M primers (Invitrogen/Life Technologies) and a HotStar HiFidelity DNA Polymerase Kit with Q solution (Qiagen). Cycling conditions are specified in Supplemental Table S2. To minimize *in vitro* recombination artifacts, we monitored template quality, and the amount per reaction (40 ng) was kept as low as possible, extension times were not longer than recommended, and excessive cycles were avoided (Judo et al. 1998; Kanagawa 2003; Lenz and Becker 2008; Lukas and Vigilant 2005). Gel-purified amplimers were TA cloned using the pGEM-T Easy Vector System (Promega, Madison, WI). Plasmid DNA samples isolated from transformed *Escherichia coli* clones (GC10 Chemical Competent Cells, Genesee Scientific, San Diego, CA) were screened via *Eco*RI restriction digest. Insert-positive plasmids were sequenced in both directions by Eurofins MWG Operon (Huntsville, AL) using either standard SP6 and T7 primers, or internal sequencing primers flanking FLAI exons 2 and 3.

For obtaining FLAI full-length (fl) sequences, primer pairs fl-E and fl-H/K (Supplemental Table S1) were designed using PrimerBlast, with the forward primer at the exon 1 initiating ATG and reverse primers in the 3' UTR. Amplimers were generated in two separate PCRs, using each primer pair and 1 μ L cDNA prepared from 250 ng PBMC RNA with Superscript III RT (Life Technologies, Grand Island, NY), and then TA cloned.

Evaluation of tissue-specific gene transcription

For expression analysis, universal ("U") primers were designed to anneal to conserved sequences in exons 1 and 4, and amplify most, if not all, FLAI genes predicted to have full-length transcripts. Test sequences from gDNA demonstrated that products from at least seven different loci (data not shown; GenBank accession KC763068-KC763076), including FLAI-E, -H and -K, were recovered. cDNA was prepared using Superscript III RT from 2.5 mg total RNA isolated from each of 8 different tissues from DSH7. Amplimers of tissue-specific cDNAs and gDNA from this same individual were cloned as described above. A minimum of 36 insert-positive cDNA clones from each tissue were screened by both *StyI*

and *PstI* restriction digests, and grouped by digestion band patterns. For each tissue, at least 3 clones from each pattern group, as well as any non-conforming clones, were sequenced, and a minimum of 11 insert sequences were obtained for analysis.

Data analyses and allele nomenclature

Sequences were analyzed and edited with A Plasmid Editor (ApE v2.0.41 (Davis 2012)) and Geneious v5.6.3 (Drummond et al. 2012). To validate a novel sequence, we always sought to obtain 3 representative clones per cat, and ideally, to recover that allele from more than one individual, or minimally, from more than one PCR. In a few cases, despite repeated attempts, these goals were not met (see caption, Supplemental Fig. S1). Sequences ultimately represented by a single clone were discarded. Obtaining six clones of the identical allele from an individual was considered sufficient to declare homozygosity at that locus with acceptable degree ($>96\%$; $[1 - (0.5^{n-1})] \times 100$) of confidence, with the assumption that amplification and allele recovery were unbiased. In our analyses, sequences initially were assumed to originate from the locus that was targeted by the specific primer pair; to confirm this, we compared identity bit scores to EU153401 using a BLASTN search (Zhang et al. 2000). If highest score was to the target locus, then the assignment was considered valid. If the sequence had a very high identity to a non-FLAI-E, -H or -K locus, then that non-target identity was assigned. FLAI-E, -H or -K are more similar to each other than to all other FLAI loci (see Fig. 1), so if assignment to one of the three was ambiguous, then the originating locus was deduced by elimination, considering all other sequences discovered in that individual, and assuming that the sequence in question belonged to one of the putative class Ia genes, with no more than 2 alleles at any locus.

Alleles were named according to the 5 digit nomenclature system developed for the Dog Leukocyte Antigen (DLA) (Kennedy et al. 1999; Kennedy et al. 2001). Specifically, residues in the hypervariable regions (HVRs) define a type (digits 1-3), expected to recapitulate the serotype, and coding differences outside the HVRs define the subtype (digits 4-5). A 6th digit differentiates alleles that have non-coding differences, and a lower case “n” is used to denote non-functional sequences. At each locus, we used 00101 to designate the allele matching the genomic sequence EU153401, e.g. FLA-E*00101 (note: the “T” in FLAI signifying class I is omitted in allele names). Initially, alleles were provisionally named according to sequence of discovery; after feline HVRs were established by variability analysis of our larger data set, the correct names were assigned for use throughout the paper.

Sequences were deposited in GenBank, and alleles that meet acceptance criteria will be submitted to the Immune Polymorphism Database (IPD)-MHC database (<http://www.ebi.ac.uk/ipd/mhc/>) (Robinson et al. 2005). Observed and expected heterozygosities were calculated using Arlequin (Excoffier et al. 2005). To test for positive selection of amino acid sites, the exon 2 and 3 nucleotide sequences of all known FLAI alleles were aligned by codons and subjected to Bayesian analysis, using MrBayes 3.1.2 (Huelsenbeck and Dyer 2004). Two parallel runs of 10^6 cycles were performed, with the burnin percentage set to 10%. Codons were presumed to be under positive selection when posterior probability values, analyzed by Tracer 1.5 (Rambaut and Drummond 2007), exceeded 0.95.

Results

The domestic cat genome contains 19 MHC class I gene candidates, of which 12 are predicted, by comparison to clone FLAA24, to maintain full-length exons, and 3 of those loci are anticipated to encode classical molecules: FLAI-E, -H, and -K (Yuhki et al. 2008). The goal of this study was to assess this claim by analyzing polymorphisms and transcription of these genes in tissues. Secondly, our intent was to examine the degree of

diversity at these loci in a number of cats comprising several breeds, and to characterize the relationship of newly discovered FLAI alleles to those found in earlier studies.

Development and evaluation of FLAI-E, -H, and -K locus-specific PCR primer pairs

Polymorphisms in classical MHC genes, which may be driven by one or more evolutionary mechanisms, such as selection of rare advantageous alleles or overdominant positive selection (reviewed in Pirotney and Oliver, 2006), are concentrated accordingly in the antigen recognition site (ARS). The sequences of the α 1/2 domains, which contain the ARS, are therefore sufficient for allelic differentiation, and typically are the only reported portion of the allele (Bjorkman et al. 1987; Castro-Prieto et al. 2011b; Kennedy et al. 1999; Parham et al. 1995). As in other species, FLAI exons 2 and 3 encode these domains and contain the majority of polymorphic sites (Yuhki and O'Brien 1990a). To efficiently genotype our cats, we sought to design locus-discriminatory primers by identifying unique sites that flanked these exons and avoided areas of nucleotide variability. The obvious limitations to this effort were that only a single representative allele for each gene was available from genomic data, precluding any estimates of intronic variability; the originating locus for each of the handful of reported full-length cDNA clones (Sato and Yamamoto 2009; Yuhki and O'Brien 1990a) was uncertain; and nucleotide differences between the three loci at almost all candidate primer sites were minimal, often only one or two base pairs. Ultimately, five primer pairs (designated E1, E2, E3, H, K; Supplemental Table S1), were designed that appeared likely to have good discriminatory power, although some degree of cross-locus annealing or resistance to amplification was anticipated. We assessed the capability of our first set of primer pairs (E1, H, K) with gDNA from cat DSH1. A phylogram of all sequences obtained from this individual demonstrate tight clustering of the alleles retrieved with each primer pair with the reference sequences for their expected locus, with a few exceptions (Fig. 1). After analysis of additional samples, the E1 primer pair was found to have a fairly high rate of cross-locus amplification, and was discarded in favor of the more specific E2 and E3 pairs (Supplemental Table S1, right column). Overall, the E2/E3, H and K primer pairs, particularly when used as a set, preliminarily appeared capable of discriminating between these extremely similar genes.

Characterization of allelic variation at FLAI-E, -H, and -K

We genotyped an additional 11 unrelated cats and, in total, found 33 distinct FLAI sequences (Table 1). Eight of these alleles have been reported previously (Table 2) (Sato and Yamamoto 2009; Smith and Hoffman 2001; Yuhki et al. 2008; Yuhki and O'Brien 1990a), while 25 are novel (Supplemental Fig. S1). In no cat did we obtain more than two alleles at any locus. In our samples, FLAI-E had the highest number of alleles, 12, encoding 12 unique amino acid sequences. At the H locus, we detected 11 alleles, encoding 9 unique amino acid sequences. Ten alleles were found at the K locus. Two of these alleles, K*00601n and K*00602n (which differ by 1 bp), have a single nucleotide substitution in exon 2 that introduces a premature stop at position 63 in the α 1 domain. These variants otherwise conform to the expected class Ia pattern of polymorphic and conserved sites, and hence, appear to be mutations (null ["n"] alleles), not pseudogenes. Excluding these two, eight FLAI-K amino acid sequences were found. Our data show that FLAI-E, -H and -K are polymorphic genes, consistent with their putative classical status.

Two common allele types were observed: FLAI-E*003 and -K*003. The FLAI-E*003 representatives, E*00302 and E*00303, were found in 6 individuals; E*00303 was carried by 5 of the 12 cats. The FLAI-K*003-type alleles, K*00302 and K*00303, were also detected in 6 cats. Their prevalence is further supported by the fact that two previously reported clones belong to these allele types: FLAA10 (designated E*00301 in our analysis) and FLAB2 (designated as K*00301) (Yuhki and O'Brien 1990a). Sixteen alleles were not

shared among the cats in our study population (Table 1, boldface type), although one (E*00101) has been previously identified (Yuhki et al. 2008). In fact, all three reference alleles found in the RPCI86 cat BAC library were recovered here.

For the four cats that were homozygous at two of the three loci, eight haplotypes could be deduced; of these, two (E*00303, H*00101, K*00801 and E*01001, H*00101, K*00801, both carried by DSH2) were potentially shared with another individual, DSH5. The Persian and Himalayan cats also shared a haplotype, as these individuals only varied by a single FLAI-E allele. The two breeds are very closely related (Himalayans are a “variant breed” originally selected from pairings, and occasionally outcrossed to Persians (Menotti-Raymond et al. 2008)), so common alleles are not surprising. Moreover, it is possible that class I allelic diversity is restricted within feline breeds (purebred cats have lower overall genetic variation than do outbred cats (Lipinski et al. 2008)), although perhaps not to the same degree as observed in other domestic animal breeds, such as the dog (Ross et al. 2012) or miniature swine (Ando et al. 2003), where breeding barriers are stronger. As a measure of genetic diversity at each loci, we computed the observed and expected heterozygosities for FLAI-E, -H and -K (Table 3). These values are similar to an average heterozygosity of 0.85 calculated for 10 polymorphic simple tandem repeat loci in outbred cats (Menotti-Raymond et al. 2012). For two loci, FLA-E and -H, there was a significant departure from Hardy-Weinberg equilibrium (a deficit of heterozygotes), which could reflect, in part, our small sample size (Pruett and Winker 2008) or amplification resistance of some alleles, but presumably also is due to inbreeding in cats. Consistent with this supposition, genetic variance in outbred cats is lower than that in humans (Lipinski et al. 2008).

Evaluation of locus assignment

To assess whether our sequences were accurately assigned to their locus of origin, we constructed a phylogenetic tree of the 33 FLAI alleles, using exon 2-intron 2-exon 3 nucleotide sequences; if attributed correctly, then alleles from the same locus should partition into monophyletic clusters (Gu and Nei 1999). Figure 2 shows that all H alleles segregate into a highly supported clade, while partitions between the tentative E and K alleles are less well resolved. Notably, one allele, K*00801, which was found in 3 individuals, can be seen to group with other presumed E locus-origin sequences (Fig. 2, arrow), and therefore would seem to be misplaced, or possibly, to represent a PCR amplification artifact – an unlikely explanation, as this allele has been reported previously as a full-length transcript (FLAB9) (Yuhki and O'Brien 1990a). In two of the cats (DSH2 and DSH5), K*00801 was carried with E*01001, which is contained within a well-supported clade of other E alleles and is identical to FLAA23, a clone that clearly originates from a locus (A, by the previous designation) distinct from that of FLAB9 (B locus). Moreover, both cats possessed a second allele (E*00303) with strong homology to the E locus. Collectively, these observations argue that K*00801 is the product of a non-E locus. In further support, the other B locus-origin clone identified in the study of Yuhki et al. (1990a), FLAB2, best matches the FLAI-K alleles found in this study. Therefore, we maintained the assignment of K*00801 to the K locus, recognizing that its phylogenetic placement suggests that this is a hybrid allele generated by interlocus (intergenic) recombination. In fact, this process may explain the imperfect segregation of the FLAI-E and -K alleles; the expectation for monophyly of same-locus alleles assumes that interlocus recombination or gene conversion is a very small contributor to diversity. MHC class I intronic data can also be phylogenetically informative (Cereb et al. 1995). In intron 2, we observed 35 variable sites among the FLAI alleles, including a 5 bp insertion/deletion (indel) in the 3' GGG trinucleotide splicing enhancer region (Supplemental Figure S1, nucleotides 458-462). However, when a phylogram is constructed only with our intronic sequence data (i.e., disregarding those regions expected to be under positive selection pressure), then most

partitions were poorly supported, with only 7 H alleles forming a distinct clade (98% consensus support; data not shown). This intron homogenization supports the idea that interlocus recombination has occurred. Further, an analysis of a limited number (seven) of class I transcripts found clear evidence for interlocus (and interallelic) recombination in FLA, specifically identifying in the first α -helix region (in the α 1 domain) a mosaic pattern of two shared polymorphic sequences separated by a conserved 23-bp stretch (Supplemental Fig. S1, bp 206-228) (Yuhki and O'Brien 1990a). This presumptive recombination hotspot is also found in non-domestic felids – ocelot, cheetah, lion, tiger and extinct saber-toothed cat – and therefore is likely ancient in origin (Pokorny et al. 2010; Sachdev et al. 2005; Yuhki and O'Brien 1994). These same motifs are observed in our sequences. There are two: a motif exemplified by FLAA10/E*00301 and common to the E-origin alleles in one partition (Fig. 2, marked *), and a second pattern represented by FLAA23/E*01001 that is found in a separate partition shared by E- and K-origin alleles (marked **). Hence, the incomplete partitioning of our alleles originating from the FLA-E and -K loci may be the result of interlocus recombination that was previously observed in their A and B analogs (and in two clones of seeming hybrid lineage, given an “AB” locus designation (Yuhki and O'Brien 1990a; Yuhki and O'Brien 1994). Consequently, there is an acknowledged uncertainty for our locus assignments that lies somewhere between humans, where virtually all alleles cluster by loci, and mice, where locus specificity is not phylogenetically discernible (Gu and Nei 1999). Such difficulty in definitively identifying the originating locus is commonplace in MHC class I studies in animals, due to interlocus recombination and haplotype-variable gene expression (Babiuk et al. 2007; Ellis et al. 1999; Holmes et al. 2003; Miltiadou et al. 2005; Tallmadge et al. 2010); in cattle and sheep, for example, alleles are designated with the prefix N to reflect that uncertainty.

Assessment of structural features common to classical MHC class Ia molecules

The 29 unique predicted amino acid sequences found in this study are shown in Fig. 3. All alleles conform to the expected pattern of polymorphic and conserved sites in the α 1/2 domains, as detailed in Table 4. All important structural residues are present, including those involved in TCR contact; the C101 and C164 residues that form the disulfide linkage between the α -sheet and the α 2 domain α -helix; and amino acids in the ARS that are conserved across murine and human class Ia heavy chains, with the exception of an L160A substitution in two alleles, E*00501 and E*01001 (Bjorkman and Parham 1990; Kaufman et al. 1994; Yuhki and O'Brien 1990a). Similarly, 13 of the 19 residues in the α 1 and α 2 domains of HLA-A2 that contact β 2M are found in all our feline alleles (Saper et al. 1991). A few exceptions are noted: at F9, all but one allele has a Y, and at Y116, most residues use D or E. These same position 9 and 116 variants are also observed in the canine classical class Ia molecule, DLA-88 (Ross et al. 2012). Conversely, humans and dogs use a T at position 94, while our feline alleles possess N or S at this site. In addition to evaluating α 1/2 sequences, we sought to verify other characteristic MHC class Ia structural elements (leader peptide; α 3 extracellular, transmembrane and cytoplasmic domains) in one or more representative FLAI alleles. To make this determination, we amplified full-length transcripts from PBMCs. Alignment of the recovered sequences with FLAA23 shows that all expected domains are present (Supplemental Fig. S2). We also generated two full-length FLAI-E clones from two additional DSH cats (not shown); the same structural components were identifiable.

Evaluation of tissue expression of FLAI-E, -H and -K genes

MHC class I cDNA sequences with >99% identity to FLAI-H and -K have been obtained from a cell line derived from the Abyssinian cat used in the feline genome project (Yuhki et al. 2008), demonstrating that these two loci are transcribed. By definition, classical MHC molecules are widely expressed, and accordingly, we sought to determine whether

transcripts from any or all of the three putative class Ia loci could be isolated from multiple feline tissues. Using the U primer pair, FLAI-E and -K alleles were amplified from the cDNAs of all 8 tissue samples obtained from DSH7, and an FLAI-H allele was retrieved from 7 tissues (Table 5). An additional sequence that was identified in two clones isolated from the ovary matched E*00501 from the beginning of exon 1 to midway in exon 2, while the remainder matched K*00701. Most likely, this sequence was a PCR chimera, but given its discovery in ovarian tissue, this hybrid could also have resulted from a meiotic crossover event. Thus, all three loci are expressed across a range of tissue types, supporting the prediction that FLAI-E, -H, and -K are class Ia loci in the domestic cat.

Characterization of other FLAI loci

Based on genomic analysis, there are five additional full-length FLAI genes, FLAI-F, -J, -L, -M and -O, which have most, but not all, of the 31 residues in the 1 and 2 domains that are conserved across human and feline class Ia molecules (Yuhki and O'Brien 1990a). Hence, Yuhki et al. (2008) designated these loci as non-classical (class Ib) – functional MHC molecules with limited polymorphisms and tissue-restricted expression (reviewed in Rodgers and Cook, 2005); however, the true status of each of these feline genes is unknown. While some have complete class I promoter motifs, such as FLAI-O, no cDNA clones representing these loci have been reported. Further, we found no evidence that these genes were transcriptionally active in tissue samples when using primers (U) capable of amplifying FLAI-F, -J, -O and -Q-origin sequences from gDNA. Smith et al. (2001) reported an FLAI allele “1” from an undetermined (Z) locus that was carried by 3 of 7 domestic cats, as well as by 4 other species of felids, including the caracal and Geoffroy's cat. This sharing of a single allele of a normally highly polymorphic MHC gene across the three major cat lineages was recognized as “unusual”, and was therefore considered to denote high selective pressure for locus conservation. With the benefit of subsequent sequencing and annotation of the domestic cat MHC, it now appears that allele 1 is FLAI-O. With the U or K primer pairs, we retrieved sequences from the gDNA of 6 of the 12 cats (each represented by 2 colonies) that were virtually identical to FLAI-O, although most had 1 to 3 base pair deletions or additions in exon 3. An alignment of our FLAI-O sequences, as well as the homologous Z locus-origin sequences reported by Smith et al. (2001), is shown in Supplemental Fig. S3. Collectively, the low variation between individual cats and across species, the coding errors, and lack of evidence of expression suggests that FLAI-O may be an inactive, class I relic (also referred to as a “dead” gene) (Nei et al. 1997; Zemmour et al. 1990). Similarly, additional sequences represented by two or more colonies were recovered from DSH7 gDNA that were homologous to FLAI-F and -J – both full-length and variants with premature stops – which may suggest that these loci, too, are no longer active. Of course, the possibility that FLAI-F, -J and -O are (non-classical) class Ib genes expressed exclusively in tissues that were not sampled, or at levels below our detection capabilities, cannot be excluded.

Assessment of diversity across new and established classical FLAI sequences

Finally, we wished to investigate patterns of diversity within and between all known class Ia alleles of the domestic cat. Previous to this study, a number of putative classical FLAI sequences had been deposited in GenBank (Table 2); 12 were not observed in our cats. When aligned to the full-length class I genes, each of the 12 matched the reference FLAI-E, -H and -K sequences most strongly, although none had been specifically designated as products of these loci. Combined with the alleles found here, there are thus a total of 41 unique amino acid sequences discovered in the cat to date that presumably represent classical molecules. When genetic diversity was evaluated in HLA class Ia genes by comparing polymorphisms in exons 2, 3 and 4 of 39 HLA-A, -B and -C alleles, three clusters of high amino acid variability (now known as HVRs) could be identified, one in the -helix

of the 1 domain, and two in the α -helix and α -strand of the 2 domain. Not surprisingly, the polymorphic residues of these HVRs pointed into, or up from, the ARS, likely interacting with the bound peptide or TCR (Parham et al. 1988). The same HVRs can be observed in DLA-88 (Graumann et al. 1998), and the boundaries of the canine HVRs (positions 62-77, 91-116, 152-158) (Kennedy et al. 1999) contain the highly polymorphic positions specifically identified in seven feline class I clones (Yuhki and O'Brien 1990a), suggesting that these regions also constitute HVRs in cats. To confirm this supposition, we used the data set of 41 pooled alleles to calculate the variability index (vi) (Wu and Kabat 1970) for the 1 and 2 regions. Overall, as seen in Fig. 4, there were 15 positions where variability was defined as high ($vi \geq 4$), with a pattern of polymorphisms that was similar to that observed in the class I alleles of tigers (Pokorny et al. 2010). Three HVRs, which are defined as amino acid sequences containing and bounded at either end by highly variable residues, are apparent: HVR1, positions 62-81; HVR2, 94-116; and HVR3, 152-156. These HVRs occupy virtually the same residue stretches as those of HLA and DLA, and similarly, many (27) of the feline polymorphic sites are part of the ARS (Yuhki and O'Brien 1994). If a variability plot is generated from the alleles at each locus separately, then the E locus has 21 positions where $vi \geq 4$; the K locus has 8 positions, and the H locus, 6 (data not shown). The H locus-origin alleles are mainly variable in HVR1 only, and the diversity in HVR3 is due almost entirely to substitutions in E locus-origin alleles. The dominant contribution of one locus to variability at particular positions is also observed in HLA-A and -B molecules, and may indicate that interlocus recombination is limited (Parham et al. 1988).

The preferential accumulation of non-synonymous nucleotide substitutions in the ARS is considered strong evidence of overdominant selection as a substantial factor in maintaining MHC class I diversity. This observation has been reported previously in cats: by using a pairwise comparison between seven FLAI transcripts, a biased distribution of non-synonymous mutations was detected in the ARS, but not in regions of the 1/ 2 domains that lay outside of the ARS (Yuhki and O'Brien 1990a). We wished to confirm this finding in our larger set of 43 nucleotide sequences (including K*00601n and K*00602n). Using Bayesian analysis, 5 codons were identified in the HVRs that had a mean probability $>95\%$ of positive selection: positions 62, 63, 81, 97 and 156 (Fig. 4 [starred residues]). As expected, all positions occurred in the ARS. Interestingly, studies in the tiger and leopard have found evidence of positive selection at three of the same sites (residues 63, 97 and 156), as well as at other positions in the HVRs (residues 66, 67, 70, 74, 75, 77, 116, 152 and 155) (Castro-Prieto et al. 2011a; Pokorny et al. 2010).

We then investigated the phylogenetic relationship between all 41 FLAI alleles. For this analysis, the 12 previously reported class Ia sequences not detected here were re-named according to locus homology and interallelic relationships defined by synonymous and non-synonymous substitutions within and outside the HVRs, as described in the Materials and Methods. When the amino acid sequences of the 1/ 2 domains were grouped together in a consensus tree, most alleles segregated with $>50\%$ support into locus-specific clades (Fig. 5). As in the previous analysis that had been limited to our new nucleotide sequences, two E locus lineages are visible, and K*00801 is again placed among the E locus alleles (arrow). Seven sequences fall outside of locus groupings. These alleles include 4 sequences that are available only as GenBank submissions, so the details of their discovery are unpublished. One of these alleles, H*00901 (B*no5), carries a T substitution at K146, which is a highly conserved position ((Yuhki and O'Brien 1990a) and our data), suggesting either potential errors in sequencing or the possibility that this variant encodes a class Ib molecule. H*00901 and the other ungrouped sequences could also represent chimeric alleles, resulting from either *in vitro* or genuine interlocus or interallelic recombination.

Discussion

The intent of this study was to evaluate the prediction that FLAI-E, -H, and -K are classical MHC class Ia genes in the domestic cat, and to describe polymorphisms at these loci. Our design aligned with the goals advocated as necessary to characterize MHC diversity: define the number of expressed loci; assign alleles to loci and assess heterozygosity; and determine the extent of sequence variation in structurally important regions, i.e., the ARS (Piertney and Oliver 2006). Using sets of locus-specific primer pairs to amplify FLAI gene segments from 12 cats, we identified 25 novel alleles, and verified the sequences of 8 others residing in the GenBank database. One unexpected finding is that we failed to recover transcripts from non-classical genes in our expression analysis, even though there may be as many as nine full-length class Ib loci (Yuhki et al. 2008), and our primer pair was capable of amplifying at least four of those genes. Of course, this negative result may simply reflect a technical limitation; to find such evidence may require the design of primers specific for these genes or isolation of RNA from other tissues. Analysis of genomic FLAI-F, -J and -O sequences suggested that these three genes do not encode non-classical class Ib molecules, but rather, have become inactive relics.

This work continues the compilation and clarification of FLAI alleles in the domestic cat. Most sequences from earlier studies appeared to be restricted to only two loci of origin, potentially due to probe or primer mismatch, so presumably, alleles were previously underreported. The PCR conditions and primers developed here should prove to be a valuable tool for class Ia genotyping, although it should be emphasized that we also may have failed to detect some alleles, leading to an underestimation of heterozygosity, as our primer pairs were designed using a limited data set. Additionally, analogous studies in other species have established that some alleles are preferentially amplified (Miltiadou et al. 2005), and that rare transcripts may be missed by conventional Sanger sequencing methods (Budde et al. 2010; Kita et al. 2012). On the other hand, the use of multiple primer pairs with different locus specificities can potentially improve the recovery of sequences from genes transcribed at lower levels (Bettinotti et al. 2003).

Further work will be necessary to confirm the provisional locus assignments for FLAI alleles that were made here. Identifying the originating locus solely from the gene regions encoding the highly polymorphic 1/2 domains can be difficult. For example, when neighbor-joining trees of 10 HLA loci were constructed based on sequences from either the 5' flanking region, introns, exons 2 and 3, or exons 4 through 8 (which encode the 3, transmembrane and cytoplasmic domains), topologies were similar, but the confidence in partitioning based on exon 2 and 3 sequences was lower than the others (Sawai et al. 2004). Locus-specific motifs have been shown to reside in other regions of class I genes, including the 5' untranslated region (UTR) (Cereb and Yang 1994), introns (Cereb et al. 1997; Cereb et al. 1995), exons 4 through 8 (Birch et al. 2006; Ellis et al. 2005; Yuhki and O'Brien 1990a) and the 3' UTR (Koller et al. 1984; Tallmadge et al. 2010), and therefore, definitive FLAI locus interpretation will ultimately require full-length data. Ideally, next-generation sequencing methods can be used to expediently obtain complete heavy chain sequences; these techniques will also facilitate scanning larger samples of the domestic cat population to identify prevalent haplotypes, and promote the discovery of rare class Ia transcripts (Budde et al. 2010), which can clarify assignments. Finally, it should be noted that, despite the current uncertainty with locus assignment, this attribution is of secondary importance: as long as allele recovery is robust, then the downstream immunologic applications of such data – epitope discovery; disease association studies – can proceed unimpeded.

The final step in defining FLAI-E, -H and -K as classical loci will be to demonstrate their restriction of CD8⁺ T-cell activity. If highly prevalent alleles are used in making such

determinations, then immunodominant cytotoxic T-cell responses shared among cats carrying that allele can be defined, a critical prerequisite to understanding adaptive immunity to viruses in this species. The cat remains an important model of oncoviral and lentiviral diseases (O'Brien et al. 2008; O'Brien et al. 2012). Cytotoxic T lymphocyte (CTL) responses to experimental infection and vaccination have been investigated in feline sarcoma virus (McCarty and Grant 1983), feline leukemia virus (Flynn et al. 2000), and most notably, feline immunodeficiency virus (Bonci et al. 2009; Burkhard et al. 2001; Flynn et al. 1996; Gupta et al. 2007; Koksoy et al. 2001; Leutenegger et al. 2000; Pu et al. 1999; Song et al. 1992), and the role of these CTL in controlling infection has been demonstrated by challenge protection (Flynn et al. 2000; Flynn et al. 1996) and adoptive transfer of immunity (Flynn et al. 2002; Pu et al. 1999). Despite this body of work, the MHC class I molecules that control such CTL are unknown. Importantly, in this study, we not only identified prevalent alleles, but also common allele types (E*003 and K*003). While not genotypically identical, the alleles that comprise these types have the same HVRs, and potentially have overlapping peptide binding preferences and present the same epitopes. Such so-called supertypes (del Guercio et al. 1995; Sidney et al. 1995) effectively increase the prevalence of functionally equivalent alleles in populations under study, and hence, their recognition is valuable in developing peptide-based vaccines or studying peptide-specific Tcell responses with the broadest possible coverage.

Defining and surveying classical class Ia molecules in the domestic cat may also benefit wild cats. Recent viral epidemics with high mortality rates have adversely impacted a number of big cat populations, and there is concern for cross-species disease transmission (Brown et al. 2008; Heeney et al. 1990; Meli et al. 2009; O'Brien and Johnson 2005; Pearks Wilkerson et al. 2004; Roelke-Parker et al. 1996). Consequently, effective immunization programs are considered a critical part of the conservation efforts for these species. With a more complete understanding of FLA, the domestic cat could play a role in such vaccine development. Much of the genetic variation in the Felidae MHC appears to have been set before speciation (Castro-Prieto et al. 2011a; Wei et al. 2010; Yuhki and O'Brien 1994; Yuhki and O'Brien 1997), and some polymorphisms are ancient. Therefore, it seems likely that there are similarities in MHC class Ia alleles - and the CD8+ T-cell responses they restrict – across species and lineages of cats.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the Wake County Animal Center (Raleigh, NC), Edward Collins (University of North Carolina, Chapel Hill), and Greg Gojanovich, Jeff Thorne and Jeff Yoder (all at North Carolina State University) for their help. This work was supported by a National Institutes of Health grant (K08 DK082264) to PRH, and is dedicated to the memory of Inky.

References

- Ando A, Kawata H, Shigenari A, Anzai T, Ota M, Katsuyama Y, Sada M, Goto R, Takeshima SN, Aida Y, Iwanaga T, Fujimura N, Suzuki Y, Gojobori T, Inoko H. Genetic polymorphism of the swine major histocompatibility complex (SLA) class I genes, SLA-1, -2 and -3. *Immunogenetics*. 2003; 55(9):583–593. [PubMed: 14605775]
- Assarsson E, Sidney J, Oseroff C, Pasquetto V, Bui HH, Frahm N, Brander C, Peters B, Grey H, Sette A. A quantitative analysis of the variables affecting the repertoire of T cell specificities recognized after vaccinia virus infection. *J Immunol*. 2007; 178(12):7890–7901. [PubMed: 17548627]

- Babiuk S, Horseman B, Zhang C, Bickis M, Kusalik A, Schook LB, Abrahamsen MS, Pontarollo R. BoLA class I allele diversity and polymorphism in a herd of cattle. *Immunogenetics*. 2007; 59(2): 167–176. [PubMed: 17206412]
- Beck TW, Menninger J, Murphy WJ, Nash WG, O'Brien SJ, Yuhki N. The feline major histocompatibility complex is rearranged by an inversion with a breakpoint in the distal class I region. *Immunogenetics*. 2005; 56(10):702–709. [PubMed: 15592824]
- Bettinotti MP, Hadzikadic L, Ruppe E, Dhillon G, Stroncek DS, Marincola FM. New HLA-A, -B, and -C locus-specific primers for PCR amplification from cDNA: application in clinical immunology. *J Immunol Methods*. 2003; 279(1-2):143–148. [PubMed: 12969555]
- Birch J, Murphy L, MacHugh ND, Ellis SA. Generation and maintenance of diversity in the cattle MHC class I region. *Immunogenetics*. 2006; 58(8):670–679. [PubMed: 16807744]
- Bjorkman PJ, Parham P. Structure, function, and diversity of class I major histocompatibility complex molecules. *Annu Rev Biochem*. 1990; 59:253–288. [PubMed: 2115762]
- Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature*. 1987; 329(6139):512–518. [PubMed: 2443855]
- Bonci F, Zabogli E, Conti F, Merico A, Freer G, Bendinelli M, Pistello M. A novel method for producing target cells and assessing cytotoxic T lymphocyte activity in outbred hosts. *BMC Biotechnol*. 2009; 9:18–28. [PubMed: 19284578]
- Brown MA, Cunningham MW, Roca AL, Troyer JL, Johnson WE, O'Brien SJ. Genetic characterization of feline leukemia virus from Florida panthers. *Emerg Infect Dis*. 2008; 14(2): 252–259. [PubMed: 18258118]
- Budde ML, Wiseman RW, Karl JA, Hanczaruk B, Simen BB, O'Connor DH. Characterization of Mauritian cynomolgus macaque major histocompatibility complex class I haplotypes by high-resolution pyrosequencing. *Immunogenetics*. 2010; 62(11-12):773–780. [PubMed: 20882385]
- Burkhard MJ, Mathiason CK, Bowdre T, Hoover EA. Feline immunodeficiency virus Gag- and Env-specific immune responses after vaginal versus intravenous infection. *AIDS Res Hum Retroviruses*. 2001; 17(18):1767–1778. [PubMed: 11788028]
- Castro-Prieto A, Wachter B, Melzheimer J, Thalwitzer S, Sommer S. Diversity and evolutionary patterns of immune genes in free-ranging Namibian leopards (*Panthera pardus pardus*). *J Hered*. 2011a; 102(6):653–665. [PubMed: 21914667]
- Castro-Prieto A, Wachter B, Sommer S. Cheetah paradigm revisited: MHC diversity in the world's largest free-ranging population. *Mol Biol Evol*. 2011b; 28(4):1455–1468. [PubMed: 21183613]
- Cereb N, Hughes AL, Yang SY. Locus-specific conservation of the HLA class I introns by intra-locus homogenization. *Immunogenetics*. 1997; 47(1):30–36. [PubMed: 9382918]
- Cereb N, Maye P, Lee S, Kong Y, Yang SY. Locus-specific amplification of HLA class I genes from genomic DNA: locus-specific sequences in the first and third introns of HLA-A, -B, and -C alleles. *Tissue Antigens*. 1995; 45(1):1–11. [PubMed: 7725305]
- Cereb N, Yang SY. The regulatory complex of HLA class I promoters exhibits locus-specific conservation with limited allelic variation. *J Immunol*. 1994; 152(8):3873–3883. [PubMed: 8144957]
- Davis, MW. ApE - A Plasmid Editor. 2.0.41. 2012.
- del Guercio MF, Sidney J, Hermanson G, Perez C, Grey HM, Kubo RT, Sette A. Binding of a peptide antigen to multiple HLA alleles allows definition of an A2-like supertype. *J Immunol*. 1995; 154(2):685–693. [PubMed: 7529283]
- Drummond, AJ.; Ashton, B.; Buxton, S.; Cheung, M.; Cooper, A.; Duran, C.; Field, M.; Heled, J.; Kearse, M.; Markowitz, S.; Moir, R.; Stones-Havas, S.; Sturrock, S.; Thierer, T.; Wilson, A. Geneious. 5.6. 2012.
- Elder JH, Lin YC, Fink E, Grant CK. Feline immunodeficiency virus (FIV) as a model for study of lentivirus infections: parallels with HIV. *Curr HIV Res*. 2010; 8(1):73–80. [PubMed: 20210782]
- Ellis SA, Holmes EC, Staines KA, Smith KB, Stear MJ, McKee DJ, MacHugh ND, Morrison WI. Variation in the number of expressed MHC genes in different cattle class I haplotypes. *Immunogenetics*. 1999; 50(5-6):319–328. [PubMed: 10630296]

- Ellis SA, Morrison WI, MacHugh ND, Birch J, Burrells A, Stear MJ. Serological and molecular diversity in the cattle MHC class I region. *Immunogenetics*. 2005; 57(8):601–606. [PubMed: 16086171]
- Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online*. 2005; 1:47–50. [PubMed: 19325852]
- Flynn JN, Dunham SP, Watson V, Jarrett O. Longitudinal analysis of feline leukemia virus-specific cytotoxic T lymphocytes: correlation with recovery from infection. *J Virol*. 2002; 76(5):2306–2315. [PubMed: 11836409]
- Flynn JN, Hanlon L, Jarrett O. Feline leukaemia virus: protective immunity is mediated by virus-specific cytotoxic T lymphocytes. *Immunology*. 2000; 101(1):120–125. [PubMed: 11012762]
- Flynn JN, Keating P, Hosie MJ, Mackett M, Stephens EB, Beatty JA, Neil JC, Jarrett O. Env-specific CTL predominate in cats protected from feline immunodeficiency virus infection by vaccination. *J Immunol*. 1996; 157(8):3658–3665. [PubMed: 8871667]
- Graumann MB, DeRose SA, Ostrander EA, Storb R. Polymorphism analysis of four canine MHC class I genes. *Tissue Antigens*. 1998; 51(4 Pt 1):374–381. [PubMed: 9583809]
- Gu X, Nei M. Locus specificity of polymorphic alleles and evolution by a birth-and-death process in mammalian MHC genes. *Mol Biol Evol*. 1999; 16(2):147–156. [PubMed: 10028282]
- Gupta S, Leutenegger CM, Dean GA, Steckbeck JD, Cole KS, Sparger EE. Vaccination of cats with attenuated feline immunodeficiency virus proviral DNA vaccine expressing gamma interferon. *J Virol*. 2007; 81(2):465–473. [PubMed: 17079309]
- Heeney JL, Evermann JF, McKeirnan AJ, Marker-Kraus L, Roelke ME, Bush M, Wildt DE, Meltzer DG, Colly L, Lukas J. Prevalence and implications of feline coronavirus infections of captive and free-ranging cheetahs (*Acinonyx jubatus*). *J Virol*. 1990; 64(5):1964–1972. [PubMed: 2157864]
- Holmes EC, Roberts AF, Staines KA, Ellis SA. Evolution of major histocompatibility complex class I genes in Cetartiodactyls. *Immunogenetics*. 2003; 55(4):193–202. [PubMed: 12756505]
- Huelsenbeck JP, Dyer KA. Bayesian estimation of positively selected sites. *J Mol Evol*. 2004; 58(6):661–672. [PubMed: 15461423]
- Judo MS, Wedel AB, Wilson C. Stimulation and suppression of PCR-mediated recombination. *Nucleic Acids Res*. 1998; 26(7):1819–1825. [PubMed: 9512558]
- Kanagawa T. Bias and artifacts in multitemplate polymerase chain reactions (PCR). *J Biosci Bioeng*. 2003; 96(4):317–323. [PubMed: 16233530]
- Kaufman J, Salomonsen J, Flajnik M. Evolutionary conservation of MHC class I and class II molecules—different yet the same. *Semin Immunol*. 1994; 6(6):411–424. [PubMed: 7654997]
- Kennedy LJ, Altet L, Angles JM, Barnes A, Carter SD, Francino O, Gerlach JA, Happ GM, Ollier WE, Polvi A, Thomson W, Wagner JL. Nomenclature for factors of the dog major histocompatibility system (DLA), 1998. First report of the ISAG DLA Nomenclature Committee. International Society for Animals Genetics. *Tissue Antigens*. 1999; 54(3):312–321. [PubMed: 10519375]
- Kennedy LJ, Angles JM, Barnes A, Carter SD, Francino O, Gerlach JA, Happ GM, Ollier WE, Thomson W, Wagner JL. Nomenclature for factors of the dog major histocompatibility system (DLA), 2000: second report of the ISAG DLA Nomenclature Committee. *Anim Genet*. 2001; 32(4):193–199. [PubMed: 11531697]
- Kita YF, Ando A, Tanaka K, Suzuki S, Ozaki Y, Uenishi H, Inoko H, Kulski JK, Shiina T. Application of high-resolution, massively parallel pyrosequencing for estimation of haplotypes and gene expression levels of swine leukocyte antigen (SLA) class I genes. *Immunogenetics*. 2012; 64(3):187–199. [PubMed: 21932051]
- Koksoy S, Phipps AJ, Hayes KA, Mathes LE. SV40 immortalization of feline fibroblasts as targets for MHC-restricted cytotoxic T-cell assays. *Vet Immunol Immunopathol*. 2001; 79(3-4):285–295. [PubMed: 11389962]
- Koller BH, Sidwell B, DeMars R, Orr HT. Isolation of HLA locus-specific DNA probes from the 3' untranslated region. *Proc Natl Acad Sci U S A*. 1984; 81(16):5175–5178. [PubMed: 6591185]
- Lenz TL, Becker S. Simple approach to reduce PCR artefact formation leads to reliable genotyping of MHC and other highly polymorphic loci—implications for evolutionary analysis. *Gene*. 2008; 427(1-2):117–123. [PubMed: 18848974]

- Leutenegger CM, Boretti FS, Mislin CN, Flynn JN, Schroff M, Habel A, Junghans C, Koenig-Merediz SA, Sigrist B, Aubert A, Pedersen NC, Wittig B, Lutz H. Immunization of cats against feline immunodeficiency virus (FIV) infection by using minimalistic immunogenic defined gene expression vector vaccines expressing FIV gp140 alone or with feline interleukin-12 (IL-12), IL-16, or a CpG motif. *J Virol*. 2000; 74(22):10447–10457. [PubMed: 11044089]
- Lipinski MJ, Froenicke L, Baysac KC, Billings NC, Leutenegger CM, Levy AM, Longeri M, Niini T, Ozpinar H, Slater MR, Pedersen NC, Lyons LA. The ascent of cat breeds: genetic evaluations of breeds and worldwide random-bred populations. *Genomics*. 2008; 91(1):12–21. [PubMed: 18060738]
- Lukas D, Vigilant L. Reply: facts, faeces and setting standards for the study of MHC genes using noninvasive samples. *Mol Ecol*. 2005; 14(5):1601–1602.
- McCarty JM, Grant CK. Feline cytotoxic immune mechanisms against virus-associated leukemia and fibrosarcoma. *Cell Immunol*. 1983; 81(1):157–168. [PubMed: 6311439]
- Meli ML, Cattori V, Martinez F, Lopez G, Vargas A, Simon MA, Zorrilla I, Munoz A, Palomares F, Lopez-Bao JV, Pastor J, Tandon R, Willi B, Hofmann-Lehmann R, Lutz H. Feline leukemia virus and other pathogens as important threats to the survival of the critically endangered Iberian lynx (*Lynx pardinus*). *PLoS One*. 2009; 4(3):e4744. [PubMed: 19270739]
- Menotti-Raymond M, David VA, Pflueger SM, Lindblad-Toh K, Wade CM, O'Brien SJ, Johnson WE. Patterns of molecular genetic variation among cat breeds. *Genomics*. 2008; 91(1):1–11. [PubMed: 17964115]
- Menotti-Raymond M, David VA, Weir BS, O'Brien SJ. A population genetic database of cat breeds developed in coordination with a domestic cat STR multiplex. *J Forensic Sci*. 2012; 57(3):596–601. [PubMed: 22268511]
- Miltiadou D, Ballingall KT, Ellis SA, Russell GC, McKeever DJ. Haplotype characterization of transcribed ovine major histocompatibility complex (MHC) class I genes. *Immunogenetics*. 2005; 57(7):499–509. [PubMed: 16028041]
- Neefjes JJ, Hensen EJ, de Kroon TI, Ploegh HL. A biochemical characterization of feline MHC products: unusually high expression of class II antigens on peripheral blood lymphocytes. *Immunogenetics*. 1986; 23(5):341–347. [PubMed: 3458670]
- Nei M, Gu X, Sitnikova T. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc Natl Acad Sci U S A*. 1997; 94(15):7799–7806. [PubMed: 9223266]
- O'Brien SJ, Johnson W, Driscoll C, Pontius J, Pecon-Slaterry J, Menotti-Raymond M. State of cat genomics. *Trends in Genetics: TIG*. 2008; 24(6):268–279. [PubMed: 18471926]
- O'Brien SJ, Johnson WE. Big cat genomics. *Annu Rev Genomics Hum Genet*. 2005; 6:407–429. [PubMed: 16124868]
- O'Brien SJ, Troyer JL, Brown MA, Johnson WE, Antunes A, Roelke ME, Pecon-Slaterry J. Emerging viruses in the Felidae: shifting paradigms. *Viruses*. 2012; 4(2):236–257. [PubMed: 22470834]
- Parham P, Adams EJ, Arnett KL. The origins of HLA-A,B,C polymorphism. *Immunol Rev*. 1995; 143:141–180. [PubMed: 7558075]
- Parham P, Lomen CE, Lawlor DA, Ways JP, Holmes N, Coppin HL, Salter RD, Wan AM, Ennis PD. Nature of polymorphism in HLA-A, -B, and -C molecules. *Proc Natl Acad Sci U S A*. 1988; 85(11):4005–4009. [PubMed: 3375250]
- Pearks Wilkerson AJ, Teeling EC, Troyer JL, Bar-Gal GK, Roelke M, Marker L, Pecon-Slaterry J, O'Brien SJ. Coronavirus outbreak in cheetahs: Lessons for SARS. *Curr Biol*. 2004; 14(6):R227–R228. [PubMed: 15043830]
- Piertney SB, Oliver MK. The evolutionary ecology of the major histocompatibility complex. *Heredity (Edinb)*. 2006; 96(1):7–21. [PubMed: 16094301]
- Pokorny I, Sharma R, Goyal SP, Mishra S, Tiedemann R. MHC class I and MHC class II DRB gene variability in wild and captive Bengal tigers (*Panthera tigris tigris*). *Immunogenetics*. 2010; 62(10):667–679. [PubMed: 20821315]
- Pollack MS, Hayes A, Mooney S, Pedersen NC, Cook RG. The detection of conventional class I and class II I-E homologue major histocompatibility complex molecules on feline cells. *Vet Immunol Immunopathol*. 1988; 19(1):79–91. [PubMed: 3176336]

- Pruett CL, Winker K. The effects of sample size on population genetic diversity estimates in song sparrow *Melospiza melodia*. *J Avian Biol.* 2008; 39:252–256.
- Pu R, Omori M, Okada S, Rine SL, Lewis BA, Lipton E, Yamamoto JK. MHC-restricted protection of cats against FIV infection by adoptive transfer of immune cells from FIV-vaccinated donors. *Cell Immunol.* 1999; 198(1):30–43. [PubMed: 10612649]
- Rambaut, A.; Drummond, AJ. Tracer v1 5. 2007. Available from: <http://beast.bio.ed.ac.uk/>
- Robinson J, Mistry K, McWilliam H, Lopez R, Parham P, Marsh SG. The IMGT/HLA database. *Nucleic Acids Res.* 2011; 39(Database issue):D1171–1176. [PubMed: 21071412]
- Robinson J, Waller MJ, Stoeckl P, Marsh SGE. IPD--the Immuno Polymorphism Database. *Nucleic Acids Res.* 2005; 33:D523–526. Database issue. [PubMed: 15608253]
- Rodgers JR, Cook RG. MHC class Ib molecules bridge innate and acquired immunity. *Nat Rev Immunol.* 2005; 5(6):459–471. [PubMed: 15928678]
- Roelke-Parker ME, Munson L, Packer C, Kock R, Cleaveland S, Carpenter M, O'Brien SJ, Pospischil A, Hofmann-Lehmann R, Lutz H, Mwamengele GL, Mgasia MN, Machange GA, Summers BA, Appel MJ. A canine distemper virus epidemic in Serengeti lions (*Panthera leo*). *Nature.* 1996; 379(6564):441–445. [PubMed: 8559247]
- Roelke ME, Brown MA, Troyer JL, Winterbach H, Winterbach C, Hemson G, Smith D, Johnson RC, Pecon-Slattery J, Roca AL, Alexander KA, Klein L, Martelli P, Krishnasamy K, O'Brien SJ. Pathological manifestations of feline immunodeficiency virus (FIV) infection in wild African lions. *Virology.* 2009; 390(1):1–12. [PubMed: 19464039]
- Ross P, Buntzman AS, Vincent BG, Grover EN, Gojanovich GS, Collins EJ, Frelinger JA, Hess PR. Allelic diversity at the DLA-88 locus in the Golden Retriever and Boxer breeds is limited. *Tissue Antigens.* 2012; 80(2):175–183. [PubMed: 22571293]
- Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol.* 2000; 132:365–386. [PubMed: 10547847]
- Sachdev M, Sankaranarayanan R, Reddanna P, Thangaraj K, Singh L. Major histocompatibility complex class I polymorphism in Asiatic lions. *Tissue Antigens.* 2005; 66(1):9–18. [PubMed: 15982252]
- Saper MA, Bjorkman PJ, Wiley DC. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution. *J Mol Biol.* 1991; 219(2):277–319. [PubMed: 2038058]
- Sato E, Yamamoto JK. Direct submission to GenBank. 2009
- Sawai H, Kawamoto Y, Takahata N, Satta Y. Evolutionary relationships of major histocompatibility complex class I genes in simian primates. *Genetics.* 2004; 166(4):1897–1907. [PubMed: 15126407]
- Sidney J, del Guercio MF, Southwood S, Engelhard VH, Appella E, Rammensee HG, Falk K, Rotzschke O, Takiguchi M, Kubo RT, et al. Several HLA alleles share overlapping peptide specificities. *J Immunol.* 1995; 154(1):247–259. [PubMed: 7527812]
- Smith MM, Hoffman SM. A class I MHC locus compared among felid species. *Mamm Genome.* 2001; 12(5):394–396. [PubMed: 11331949]
- Song W, Collisson EW, Billingsley PM, Brown WC. Induction of feline immunodeficiency virus-specific cytolytic T-cell responses from experimentally infected cats. *J Virol.* 1992; 66(9):5409–5417. [PubMed: 1323704]
- Tallmadge RL, Campbell JA, Miller DC, Antczak DF. Analysis of MHC class I genes across horse MHC haplotypes. *Immunogenetics.* 2010; 62(3):159–172. [PubMed: 20099063]
- Wei K, Zhang Z, Wang X, Zhang W, Xu X, Shen F, Yue B. Lineage pattern, trans-species polymorphism, and selection pressure among the major lineages of feline MHC-DRB peptide-binding region. *Immunogenetics.* 2010; 62(5):307–317. [PubMed: 20372886]
- Winkler C, Schultz A, Cevario S, O'Brien S. Genetic characterization of FLA, the cat major histocompatibility complex. *Proc Natl Acad Sci U S A.* 1989; 86(3):943–947. [PubMed: 2492667]
- Wu TT, Kabat EA. An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J Exp Med.* 1970; 132(2):211–250. [PubMed: 5508247]
- Yuhki N, Beck T, Stephens R, Neelam B, O'Brien SJ. Comparative genomic structure of human, dog, and cat MHC: HLA, DLA, and FLA. *J Hered.* 2007; 98(5):390–399. [PubMed: 17675392]

- Yuhki N, Heidecker GF, O'Brien SJ. Characterization of MHC cDNA clones in the domestic cat. Diversity and evolution of class I genes. *J Immunol.* 1989; 142(10):3676–3682. [PubMed: 2715636]
- Yuhki N, Mullikin JC, Beck T, Stephens R, O'Brien SJ. Sequences, annotation and single nucleotide polymorphism of the major histocompatibility complex in the domestic cat. *PLoS One.* 2008; 3(7):e2674. [PubMed: 18629345]
- Yuhki N, O'Brien SJ. Molecular characterization and genetic mapping of class I and class II MHC genes of the domestic cat. *Immunogenetics.* 1988; 27(6):414–425. [PubMed: 2897330]
- Yuhki N, O'Brien SJ. DNA recombination and natural selection pressure sustain genetic sequence diversity of the feline MHC class I genes. *J Exp Med.* 1990a; 172(2):621–630. [PubMed: 1695669]
- Yuhki N, O'Brien SJ. DNA variation of the mammalian major histocompatibility complex reflects genomic diversity and population history. *Proc Natl Acad Sci U S A.* 1990b; 87(2):836–840. [PubMed: 1967831]
- Yuhki N, O'Brien SJ. Exchanges of short polymorphic DNA segments predating speciation in feline major histocompatibility complex class I genes. *J Mol Evol.* 1994; 39(1):22–33. [PubMed: 8064870]
- Yuhki N, O'Brien SJ. Nature and origin of polymorphism in feline MHC class II DRA and DRB genes. *J Immunol.* 1997; 158(6):2822–2833. [PubMed: 9058818]
- Zemmour J, Koller BH, Ennis PD, Geraghty DE, Lawlor DA, Orr HT, Parham P. HLA-AR, an inactivated antigen-presenting locus related to HLA-A. Implications for the evolution of the MHC. *J Immunol.* 1990; 144(9):3619–3629. [PubMed: 2329283]
- Zhang Z, Schwartz S, Wagner L, Miller W. A greedy algorithm for aligning DNA sequences. *J Comput Biol.* 2000; 7(1-2):203–214. [PubMed: 10890397]

Abbreviations

ARS	Antigen recognition site
CTL	Cytotoxic T lymphocyte
DLA	Dog leukocyte antigen
DSH	Domestic shorthair
FLA	Feline leukocyte antigen
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
HTLV	Human T-cell leukemia virus
HVR	Hypervariable region
MHC	Major histocompatibility complex
PBMC	Peripheral blood mononuclear cellsTCR T-cell receptor
UTR	Untranslated region

Fig. 1.

Neighbor-joining consensus tree of FLAI sequences that were amplified from the gDNA of DSH1 with putative locus-specific primer pairs relative to the 12 class I genes with full-length exons in the cat genomic sequence, EU153401. For clarity, allele names are not shown for reference sequences, but would be 00101 at each locus, according to our convention. Definitive names for alleles isolated from DSH1 were assigned at the end of the study, once relationships to other alleles were established. Sixteen clones were generated by primer pair E1; 7 clones were FLA-E*00101, 8 were E*00701, and 1 was H*00701. Ten clones were produced by primer pair H; 6 clones were FLA-H*00701 and 4 clones were H*008011. Ten clones were generated by primer set K; 3 clones were FLA-K*00302, 5 clones were K*00303 and 2 clones aligned with FLAI-O, a putative non-classical gene (amplification of FLAI-O with this primer pair was later found in 5 additional cats – see Supplemental Fig. S3), and are designated FLAO3. For phylogenetic analysis, all sequences were trimmed to the region spanning exon 2 through exon 3. The tree was constructed in Geneious (Drummond et al. 2012) using the Tamura-Nei genetic distance model with bootstrap resampling, and outgrouped to HLA-A2. Numbers at nodes are consensus support (%). Excluding the minor exceptions noted above, the alleles obtained with each locus-specific primer pair partition the reference locus sequence with strong consensus support.

Fig. 2.

Neighbor-joining consensus tree of all FLAI sequences that were amplified from gDNA samples from the 12 study cats with locus-specific primer pairs. Sequences were trimmed and the phylogram constructed as described in Fig. 1. Numbers at nodes are consensus support (%). Most alleles partition into locus-specific clades with good consensus support. Alleles FLA-E*00701, E*01101, E*01201 and K*00801 had ambiguous identity (by BLASTN search (Zhang et al. 2000)) to the intended reference locus, but only K*00801 (arrow) clustered with alleles from a different putative locus.

Fig. 3.

The alignment of the predicted translations of all FLAI alleles amplified from gDNA samples from the 12 study cats with locus-specific primer pairs. Exons were deduced by comparison to cDNA sequences (from our data and (Yuhki and O'Brien 1990a)). Dots signify identities to the 00101 allele (top rows) representing each locus, while letters indicate amino acid differences. Domain names are shown in the grey bar above the sequences. The * symbol indicates sites expected to be conserved. The + symbol indicates positions identified as highly polymorphic in felines (Castro-Prieto et al. 2011a; Castro-Prieto et al. 2011b; Pokorny et al. 2010; Yuhki and O'Brien 1990a; Yuhki and O'Brien 1994); i.e., those positions having at least 3 different amino acid substitutions. Residues in bold are predicted to constitute the ARS (Bjorkman et al. 1987; Yuhki and O'Brien 1994).

Fig. 4.

The variability index (vi) plotted versus residue position in the 1 and 2 domains of classical FLA class Ia molecules. For all 41 unique amino acid sequences, vi was calculated according to the equation: $vi = (\text{\# of different residues at position } x) / (\text{frequency of most common residue at position } x \text{ per total alleles considered})$. An invariant position has a $vi=1$; a highly variable position was defined as $vi \geq 4$ (dotted line). Three HVRs are observed. Residues indicated by a star were determined by Bayesian analysis to be positively selected with >95% probability.

Fig. 5.

Neighbor-joining consensus tree of deduced FLAI amino acid sequences obtained in this study and from the GenBank database. A phylogram of the 1 and 2 domains was constructed with Geneious, using the Jukes-Cantor genetic distance model with bootstrap resampling, and was rooted by including HLA-A2 as an outgroup. Numbers at nodes are consensus support (%). In this analysis, FLA-H*00801 represents allele subtypes H*008011 and H*008012, and FLA-H*00301 represents allele subtypes H*003011 and H*003012 – these subtypes are distinguished by non-coding nucleotide substitutions. Most alleles partition into locus-specific clades with good consensus support, although seven sequences appear ungrouped; K*00801 is indicated with an arrow.

Table 1

MHC class Ia genotypes of cats in this study^a

Cat	FLA-I-E	FLA-I-H	FLA-I-K
DSH1	E*00101 ^b	H*00701	H*008011 K*00302 K*00303
DSH2	E*00303	H*00101	K*00801
DSH3	E*00902	H*003011	H*00601 K*00401 K*00602n
DSH4	E*01401	H*003011	H*00601 K*00101 K*00501
DSH5	E*00303	H*00101	H*00102 K*00501 K*00801
DSH7	E*00501	H*003012	K*00401 K*00701
DLH	E*00501	H*00101	K*00601n K*00801
Tonkinese	E*01301	H*00501	K*00201 K*00302
Himalayan	E*00302	H*00201	H*00401 K*00302 K*00501
Persian	E*00302	H*00201	H*00401 K*00302 K*00501
Savannah Cat	E*00303	E*00601	H*00601 K*00101
Siamese	E*00303	E*01201 H*008012	K*00303 K*00602n

^a GenBank accession KC763013-KC763045

^b Alleles in boldface are unique to one individual

Table 2

Previously discovered FLAI alleles

Allele	Previous name	GenBank accession	Source ^a	Number ^b	Reference
<i>Found in this study</i>					
E*00101	FLAI-E	EU153401	BAC clones	1/1	Yuhki et al., 2008
E*00501	FLA-Z*12	AF283076, AF283093	gDNA	2/3	Smith and Hoffman, 2001
E*01001	FLAA23	U07669	mRNA	3/1	Yuhki and O'Brien, 1990a
E*01201	FLA-I-A*n01	EU915355	mRNA	1/1?	Sato and Yamamoto, 2009
H*00101	FLAI-H	EU153401	BAC clones	3/1	Yuhki et al., 2008
K*00101	FLAI-K	EU153401	BAC clones	2/1	Yuhki et al., 2008
K*00701	FLA-I-B*n03	EU915357	mRNA	1/1?	Sato and Yamamoto, 2009
K*00801	FLAB9	U07672	mRNA	3/1	Yuhki and O'Brien, 1990a
<i>Not found in this study</i>					
E*00201	FLAA24	U7670	mRNA	0/1	Yuhki and O'Brien, 1990a
E*00301	FLAA10	U7668	mRNA	0/1	Yuhki and O'Brien, 1990a
E*00401	FLA-Z*11	AF283080, AF283097	gDNA	0/1	Smith and Hoffman, 2001
E*00801	FLAX8	U07674	mRNA	0/1	Yuhki and O'Brien, 1990a
E*00901	FLAA1	U07667	mRNA	0/1	Yuhki and O'Brien, 1990a
E*01501	FLA-I-A*n02	EU915356	mRNA	0/1?	Sato and Yamamoto, 2009
E*01601	FLA-I-B*n04	EU915358	mRNA	0/1?	Sato and Yamamoto, 2009
E*01701	FLAX10	U07673	mRNA	0/1	Yuhki and O'Brien, 1994
E*01801	FLA-I-B*n06	EU915360	mRNA	0/1?	Sato and Yamamoto, 2009
H*00202	FLA-Z*9	AF283084, AF283101	gDNA	0/1	Smith and Hoffman, 2001
H*00901	FLA-I-B*n05	EU915359	mRNA	0/1?	Sato and Yamamoto, 2009
K*00301	FLAB2	U07671	mRNA	0/1	Yuhki and O'Brien, 1990a

^aThe BAC clones used by Yuhki et al. (2008) were obtained from a DSH (<http://bacpac.chori.org/mfeline86.htm>); the breeds from which all other previously reported sequences were isolated is specified

^bThe total number of cats in which the allele has been found: this study/the referenced study. The 1? designation is given to the data from Sato and Yamamoto (2009), as the number of individuals from which sequences were derived is not stated

Table 3
Heterozygosity at classical FLA class I loci

Locus	No. of alleles	Observed heterozygosity	Expected heterozygosity
E	12	0.750 ^a	0.924
H	11	0.500 ^a	0.920
K	10	0.833	0.913
Mean +/- s.d.	11	0.694 +/- 0.173	0.919 +/- 0.006

^aSignificantly deviated from Hardy-Weinberg equilibrium (P<0.05)

Table 4
Conserved and polymorphic sites in the 1/ 2 domains of FLA class Ia alleles

Description	Residues/sites ^a	Exceptions in this study
Very highly conserved across multiple species (Bjorkman and Parham 1990; Hughes and Nei 1989; Kaufman et al. 1994); see also: (http://www.ebi.ac.uk/ipd/mhc/)	H3 Y7 V25 G26 V28 D29 W51 Y59 W60 T64 Q72 Y84 N86 Q87 H93 Q96 M98 G100 C101 G112 Q115 Y118 D119 G120 A125 T143 K146 W147 E154 Y159 L160 C164 V165 L168 Y171 L179 R181	S125 in K*00701 A160 in E*00501, E*01101
Highly conserved across multiple species (Bjorkman and Parham 1990; Kaufman et al. 1994); see also: (http://www.ebi.ac.uk/ipd/mhc/)	T10 P20 Y27 S38 R44 P47 E55 D61 G91 D122 E166 K170 G175	K44 in K*00701 Q47 in E*00501 V55 in 9 alleles
Highly polymorphic (>3 residues per site) in felines (Castro-Prieto et al. 2011a; Castro-Prieto et al. 2011b; Pokorny et al. 2010; Yuhki and O'Brien 1990; Yuhki and O'Brien 1994)	9 12 31 41 45 49 50 52 62 63 66 67 69 70 73 74 75 76 77 79 80 81 89 94 95 97 99 103 104 105 107 114 116 125 128 131 145 149 152 154 155 156 163 169 174	Only 2 residues at sites: 9 31 45 52 66 69 73 74 80 89 94 103 107 114 125 128 131 145 163 Invariant residues: P41 A49 P50 R75 V76 Y99 P105 E149 E154 M174

^aNumbering begins with the first amino acid of the 1 domain

Table 5
Recovery of FLAI transcripts from tissues^a

Tissue	E*00501 ^b	E*01401	H*003012	K*00401	K*00701	Other
Heart	4 ^c	3	3	0	3	none
Kidney	4	2	5	0	4	none
Lymph Node	1	6	1	1	1	none
Lung	5	3	1	0	1	none
PBMC	2	6	3	2	5	none
Ovary	1	1	5	0	3	Recombinant allele? ^d
Jejunum	4	3	0	0	3	none
Spleen	4	4	2	0	2	none

^aThe identical FLAI-E, -H and -K alleles, as well as sequences representing FLAI-F and -O (GenBank accession KC763059-KC763067) were amplified from this cat's gDNA

^bcDNA clones are listed under GenBank accession KC763047-KC763051

^cNumber of clones from which the sequence was obtained

^dGenBank accession KC763046