Immunogenetics
DOI 10.1007/s00251-009-0413-8

ORIGINAL PAPER

Major histocompatibility complex variation and evolution at a single, expressed DQA locus in two genera of elephants

Elizabeth A. Archie · Tammy Henry · Jesus E. Maldonado · Cynthia J. Moss · Joyce H. Poole · Virginia R. Pearson · Suzan Murray · Susan C. Alberts · Robert C. Fleischer

Received: 9 November 2009 / Accepted: 12 November 2009

Abstract Genes of the vertebrate major histocompatibility complex (MHC) are crucial to defense against infectious disease, provide an important measure of functional genetic diversity, and have been implicated in mate choice and kin recognition. As a result, MHC loci have been characterized for a number of vertebrate species, especially mammals; however, elephants are a notable exception. Our study is the first to characterize patterns of genetic diversity and natural selection in the elephant MHC. We did so using DNA sequences from a single, expressed DQA locus in elephants. We characterized six alleles in 30 African elephants (*Loxodonta africana*) and four alleles in three Asian elephants (*Elephas maximus*). In addition, for two of the African alleles and three of the Asian alleles, we characterized complete coding sequences (exons 1–5) and nearly complete non-coding sequences (introns 2–4) for the class II DQA loci. Compared to DQA in other wild mammals, we found moderate polymorphism and allelic diversity and similar patterns of selection; patterns of non-synonymous and synonymous substitutions were consistent with balancing selection acting on the peptides involved in antigen binding in the second exon. In addition, balancing selection has led to strong trans-species allelism that has maintained multiple allelic lineages across both genera of extant elephants for at least 6 million years. We discuss our results in the context of MHC diversity in other mammals and patterns of evolution in elephants.

Keywords African elephant · Asian elephant · DQA · Major histocompatibility complex · Coding sequence · Molecular evolution

E. A. Archie · T. Henry · J. E. Maldonado · R. C. Fleischer
Center for Conservation and Evolutionary Genetics,
National Zoological Park & National Museum of Natural History,
Smithsonian Institution,
Washington, DC, USA

E. A. Archie (✉)
Department of Biological Sciences, University of Notre Dame,
Notre Dame, IN 46556, USA
e-mail: earchie@nd.edu

T. Henry
Department of Environmental Science and Policy,
George Mason University,
Fairfax, VA, USA

C. J. Moss · J. H. Poole
Amboseli Trust for Elephants,
Nairobi, Kenya

V. R. Pearson
Philadelphia Zoo,
3400 Girard Avenue,
Philadelphia, PA, USA

S. Murray
Animal Health, National Zoological Park,
Washington, DC, USA

S. C. Alberts
Department of Biology, Duke University,
Durham, NC, USA

Introduction

Crucial to an animal's ability to resist disease, genes of the major histocompatibility complex (MHC) encode cell-surface glycoproteins that bind to and present foreign antigens for immune recognition (Klein 1986). MHC genes, especially those in class II, are among the most diverse in vertebrate genomes (Garrigan and Hedrick 2003; Gaudier et al. 2000). Each MHC allele is thought to respond to a class of potential antigens, and individuals and populations with

Published online: 08 January 2010

more diverse MHC loci are better able to cope with multiple infections (Meyer-Lucht and Sommer 2005; Paterson et al. 1998; Penn et al. 2002; Westerdahl et al. 2005). Genetic diversity at the MHC is likely to be maintained by positive balancing selection (reviewed in Garrigan and Hedrick 2003; Hughes 1999; Hughes and Yeager 1998; Piertney and Olivier 2006). In support, allelic lineages may be maintained over long evolutionary time scales (i.e., trans-species allelism; Klein 1980; Takahata and Nei 1990), and the codons responsible for antigen binding—located in the second exon in the case of class II loci—often have a higher rate of non-synonymous than synonymous nucleotide substitutions (e.g., Hughes and Nei 1989). However, patterns of diversity differ somewhat across exons in the same gene; for instance, the third exon of class II loci encodes an extracellular domain close to the transmembrane region, which may experience purifying selection (Hughes and Nei 1989). Hence, comparing different gene regions may inform us about evolutionary relationships and different patterns of selection (Hughes and Nei 1989).

Here, we characterize a DQA-like MHC locus in 30 African (*Loxodonta africana*) and three Asian elephants (*Elephas maximus*). The MHC is completely uncharacterized in elephants, yet these genes are important for understanding elephant evolution and conservation for three reasons. First, elephants are members of a distinct lineage of mammals, the superorder Afrotheria, which diverged from other placental mammals approximately 100 MYA (Kriegs et al. 2006; Murphy et al. 2001; Scally et al. 2001; Waddell and Shelley 2003). The MHC has yet to be characterized in any Afrotherian mammal, and understanding the evolutionary relationships between MHC loci in multiple species of elephants and across mammals may shed light on the evolution of mammalian MHC. Second, evidence suggests that elephants might recognize and avoid inbreeding with kin, including those that are apparently socially unfamiliar to them (Archie et al. 2007). MHC variation has been implicated in kin recognition in a few vertebrate species (Brown and Eklund 1994; Manning et al. 1992; Rajakaruna et al. 2006; Zelano and Edwards 2002; but see Cheetham et al. 2007), and characterizing MHC diversity is a first step toward understanding whether similar mechanisms operate in elephants. Third, the MHC is a particularly informative measure of functional genetic diversity; such genetic diversity is a conservation concern for both wild and captive elephant populations (Armbruster and Lande 1993; Caughley et al. 1990). Characterizing MHC diversity is important for predicting how captive and wild elephant populations will respond to disease threats. In the last decade, herpes viruses killed at least 25 captive elephants (Richman et al. 1999; Richman et al. 2000; Ryan and Thompson 2001), and encephalomyocarditis, salmonellosis, and anthrax have all threatened wild African elephants (Grobler et al. 1995; Lindique and Turnbull 1994; Mbise et al. 1998). In addition, some diseases affect Asian and African elephants differently, and these species-level differences might be explained by differences in the MHC.

Our objectives were to describe patterns of allelic variation and confirm transcription in a DQA locus in African and Asian elephants. We did this by characterizing variation in a region spanning DQA's second exon, second intron, and third exon and confirming transcription by amplifying this locus from RNA. We then used cDNA from RNA transcripts to generate coding sequence and then designed primers to obtain non-coding sequence from genomic DNA. Finally, we tested for evidence of selection by examining the ratios of synonymous and non-synonymous changes and patterns of trans-species allelism. We discuss our results in the context of DQA diversity in other mammals and evolution in elephants.

Methods

Study subjects

We characterized DQA variation in 30 African elephants and three Asian elephants. Of the 30 African elephants, 23 were wild elephants living in and around Amboseli National Park, Kenya, and seven were captive elephants living in North American zoos. The 23 wild African elephants included seven parent–offspring pairs (mothers or fathers) that were previously confirmed through microsatellite genotyping (Archie et al. 2006, 2008; Hollister-Smith et al. 2007). The seven captive African elephants included one individual who was born in captivity (the offspring of two other captive elephants in our sample) and housed at the Six Flags Wild Safari Park in Jackson, NJ. The six other captive African elephants were born in the wild but housed in captivity; one was born in Zimbabwe and housed at the Philadelphia Zoo in Philadelphia, PA, another was born in Uganda and housed at the Gladys Porter Zoo in Brownsville, TX, and four were born in Uganda and housed at the Great Adventure Safari Park in Jackson, NJ. Of the three Asian elephants in our sample, all were born in the wild—one each in Sri Lanka, India, and Thailand—and all were housed in captivity at the Smithsonian National Zoological Park.

Because it can be difficult to obtain fresh blood samples for RNA isolation from elephants, the study subjects we used to confirm transcription of the DQA locus differed from those used for the genomic DNA analysis. Animals used to confirm transcription were three captive African elephants (two from the Baltimore Zoo and one from

Disney World's Animal Park) and one of the three Asian elephants from the Smithsonian National Zoo included in the sample above. RNA isolated from that same Asian elephant was used to amplify the full-length coding sequence. Internal primers were then designed to generate full sequences from genomic DNA samples from the other two Asian elephants from the National Zoo and the African Elephant from the Philadelphia zoo (see details below).

Genetic methods to characterize DQA variation across individuals

In order to characterize allelic diversity across the 30 African elephants and three Asian elephants in our sample, we amplified an 818 base pair region of DQA from genomic DNA spanning the second exon, second intron, and third exon. DNA extraction was carried out in a room physically separate from the lab in which PCR amplification or post-PCR analyses were conducted. We extracted DNA from 35 samples (for two elephants we extracted DNA from two samples); 29 of the samples (from 27 individuals, three Asian elephants and 24 African elephants) were either whole blood, tissue collected from biopsy darts, or buccal cells collected from cheek swabs. The remaining six samples, all from African elephants, were dung. DNA was extracted from blood, tissue, and buccal samples using Qiagen's DNeasy Tissue Kits according to the manufacturer's instructions (Qiagen, Valencia, CA, USA). DNA was extracted from feces according to the methods described in Archie et al. (2003, 2006). Briefly, feces were collected within 10 min of defecation, preserved in ethanol, and DNA was extracted using a modified protocol (Archie et al. 2003) for the QIAmp DNA Stool Kit (Qiagen).

Each DNA extract was amplified via PCR using the primers MDQA1 and MDQA2 (Table 1; Slade et al. 1993; Fig. 1). Each 10 μl reaction contained 1 μl of DNA extract, 0.25 μl of each 10 μM primer, 1 μl of 2 mM dNTP mix (Invitrogen, Carlsbad, CA, USA), 1 μl of 100 mg ml^{-1} BSA, 1 μl 10× PCR buffer without $MgCl_2$, 1.0 μl of 1.5 mM $MgCl_2$, 0.15 μl of AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA, USA), and 4.35 μl of water. A negative control (with water replacing template DNA) was run with all PCR reactions. Reactions were amplified in a MJ Research PTC-200 Thermal cycler (MJ Research, Waltham, MA, USA). Amplification was preceded by a 10-min denaturation and polymerase activation step at 95°C, followed by 40 cycles of 1 min each at 56°C annealing, 72°C extension, and 95°C denaturation. These cycles were followed by a 5-min extension step at 72°C.

To assess the total number of alleles amplified by MDQA1 and MDQA2, ligation and cloning were conducted using the TOPO-TA Cloning® Kit for Sequencing (Invitrogen). Positive PCR products were ligated into a pCR®4-TOPO cloning vector and transformed into competent Mach1™-T1^R *Escherichia coli* cells by standard procedures. Colonies were cultured for 6 h in LB broth and 1 μl of each culture was used as template in a PCR reaction using T3 and T7 universal primers. Each 20 μl reaction contained 1 μl of cultured *E. coli*, 0.5 μl of each 10 μM primer, 2 μl of 2 mM dNTP mix (Invitrogen), 2 μl 10× PCR buffer without $MgCl_2$, 1.6 μl of 1.5 mM $MgCl_2$, 0.2 μl of AmpliTaq Gold DNA polymerase (Applied Biosystems), and 12.2 μl of water. Amplification was preceded by 10 min at 95°C, followed by 35 cycles of 1 min each at 56°C annealing, 72°C extension, and 95°C denaturation. These cycles were followed by 5 min at 72°C. Positive PCR reactions of the expected length (~800 base pairs) were sequenced in both the 3′ and 5′ directions using an ABI PRISM® 3100 DNA Analyzer using Dye Terminator Cycle Sequencing (Applied Biosystems).

To control for the mutation errors inherent in sequencing cloned products, the risk of heteroduplex formation in PCR (Borriello and Krauter 1990; L'Abbe et al. 1992), and the expected high heterozygosity at MHC loci, we cloned, on average, 2.0 independently generated PCR products (range= 1–3 PCR products) and sequenced 28.8 clones (range=16–56 clones) per individual. As reported by previous studies that used the primers DQA1 and DQA2, our cloned products each contained one of two very divergent sets of sequences that could not be aligned with each other (Decker et al. 2002;

Table 1 Primer sets, sequences, and annealing temperatures used in this study

Primer name	Primer sequence (5′ to 3′)	T_a (°C)	Source
E1F	CACAACCCTGGACAGCAAC	60–58[a]	This study
E1F77	TGAGCAACTGTGGAGGTGAA	60–58[a]	This study
E1R	AGCACAGCTATGTTCCTCAGTC	60–58[a]	This study
MDQA1	CCGGATCCCAGTACACCCATGAATTTGATGG	56	Slade et al. 1993
MDQA2	CCGGATCCCCAGTGCTCCACCTTGCAGTC	56	Slade et al. 1993
E45F	ACTTCCTCCCCAAGGATGAT	60–58[a]	This study
E5R	TGGGAAATTTATTGCTTCCA	60–58[a]	This study

[a] 60–58 indicates a touchdown protocol starting at 60°C and ending at 58°C (see text for details).

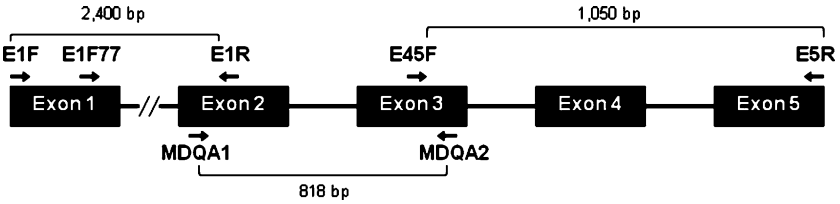


Fig. 1 Diagram of the MHC DQA locus characterized in this study. *Black boxes* indicate exons, *lines* indicate introns, and *arrows* indicate the relative locations of primers used in this study. From genomic DNA, primer pair E1F and E1R amplifies an approximately 2,400-bp product; primers E1F77 and E1R amplify an approximately 2,300-bp product (with cleaner sequences than E1F and E1R); primers MDQA1 and MDQA2 amplify an 818-bp product; primers E45F and E5R amplify a 1,050-bp product

Lehman et al. 2004). Comparison with sequences on NCBI GenBank revealed that the first set of sequences (519 clones, 55%) aligned closely to MHC sequences from the class II locus DOA, rather than DQA. The DOA locus is not directly involved in antigen binding (Decker et al. 2002; Naruse et al. 1999), and these sequences will be the focus of a separate analysis not discussed here. The second set of sequences (429 clones, 45%) aligned closely with MHC sequences from the class II locus DQA, and these sequences are the focus of our analysis in this paper. On average, we sequenced 12.6 DQA-like clones (range= 2–42 clones) per individual. A given DQA sequence was defined as an allele when copies with identical nucleotide sequences occurred in at least three clones (Table 2). In addition, all alleles occurred either in multiple individuals or in independently amplified PCR reactions from the same individual (Table 2).

Table 2 MHC DQA alleles from two species of elephants

Allele	No. of elephants with allele (no. of independent PCR reactions, no. of clones)	Frequency[a]	Genbank accession no.
Loxodonta africana			
LoafDQA*01	20 (20, 235)	0.611	GU369694
LoafDQA*02	4 (4, 4)	0.056	GU369695
LoafDQA*03	1 (2, 5)	0.028	GU369696
LoafDQA*04	10 (14, 25)	0.167	GU369697
LoafDQA*05	2 (4, 13)	0.028	GU369698
LoafDQA*06	7 (11, 51)	0.111	GU369699
Elephas maximus			
ElmaDQA*01	1 (2, 23)	–	GU369700
ElmaDQA*02	1 (2, 4)	–	GU369701
ElmaDQA*03	1 (3, 35)	–	GU369702
ElmaDQA*04	1 (4, 5)	–	GU369703

[a] For African elephants, allele frequencies are calculated from a subset of 18 individuals that were not genotyped from feces and for which we sequenced ≥6 clones. We do not present allele frequencies for Asian elephants because of the small sample size (N=3 individuals)

Genetic methods to confirm transcription

In order to confirm that the DQA locus we amplified was transcribed, we used the same primers as above (MDQA1 and MDQA2; Table 1) to amplify DQA from cDNA prepared from purified RNA. Specifically, blood from three African elephants and one Asian elephant was collected into EDTA vacutainer tubes. For the African elephants, 2 ml of blood in EDTA was added to 5 ml of RNALater (Qiagen) and stored at −80°C. The same procedure was followed for the Asian elephant, but the blood/EDTA mixture was preserved in 5 ml of Trizol (Invitrogen) instead of RNALater. For the African elephant samples, 800 μl of each sample was centrifuged, and 1 ml of Trizol with β-mercaptoethanol was added to the pellet. For the Asian elephant sample, 800 μl of the sample was added to 400 μl of Trizol with β-mercaptoethanol. RNA was extracted from all four samples using a standard phenol–chloroform separation followed by isopropanol precipitation and ethanol washes. The RNA pellet was resuspended in RNase-free water and stored at −20°C. To confirm transcription, the RNA extract was reverse transcribed into cDNA using an iScript cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA), and the primers MDQA1 and MDQA2 were used to amplify the second and third exons using the same amplification conditions as previously described. PCR products were run on an agarose gel and the presence of two bands constituted evidence for transcription: one 818-bp band, amplified from genomic DNA, and one 460-bp band, amplified from cDNA.

Genetic methods to generate full-length coding sequences

Using the RNA extracted from the Asian elephant sample above, we amplified the full-length DQA coding sequence using Invitrogen's GeneRacer Kit (Invitrogen). The resulting products were cloned using the TOPO-TA Cloning® Kit for Sequencing (Invitrogen) according to the methods described above and sequenced using the primers provided with the GeneRacer kit and MDQA1 and MDQA2.

From the resulting full-length coding sequence, we used Primer3 software (Rozen and Skaletsky 2000) to design

new primers to amplify complete coding and nearly complete non-coding sequences from genomic DNA (Table 1 and Fig. 1). We then used these primers to characterize exons 1–5 and introns 2–4 for five alleles from four animals from our main sample: the African elephant at the Philadelphia Zoo and the three Asian elephants housed at the Smithsonian National Zoo. Reaction conditions were identical for all primer sets; each 20 μl reaction contained 2 μl of template DNA, 0.5 μl of each 10 μM primer, 2 μl of 2 mM dNTP mix (Invitrogen), 2 μl 10× PCR buffer without $MgCl_2$, 2.0 μl of 1.5 mM $MgCl_2$, 2 μl of 100 mg ml^{-1} BSA, 0.3 μl of AmpliTaq Gold DNA polymerase (Applied Biosystems), and 8.7 μl of water. Amplification was preceded by 10 min at 95°C, followed by five cycles of 30 s at 95°C denaturation, 30 s at 60°C annealing, and 1 min at 72°C extension, followed by 30 cycles of 30 s at 95°C denaturation, 30 s at 58°C annealing, and 1 min at 72°C extension. These cycles were followed by 10 min at 72°C. The resulting PCR products were sequenced directly in both the 3′ and 5′ directions using an ABI PRISM® 3100 DNA Analyzer using Dye Terminator Cycle Sequencing (Applied Biosystems). DNA sequences were inspected and aligned using Sequencher software version 4.1 (Gene Codes Corporation). Polymorphism was identified by double peaks, and haplotypes were assigned using PHASE 2.1.1 software (Harrigan and Mazza 2008).

Data analyses

In order to characterize patterns of allelic diversity, we calculated nucleotide diversity (π) across all exons and introns as the average proportion of nucleotide differences between sequences using MEGA software version 2.1 (Kumar et al. 2001). For African elephants, allele frequencies and tests of Hardy–Weinberg equilibrium were carried out in Genepop (version 4.0; these analyses were not performed on Asian elephant genotypes because of small sample size). These analyses were performed on a subset of African elephant genotypes that met two conditions: their genotypes were all derived from tissue samples (as opposed to DNA extracted from feces in which dropout of alleles is more likely), and they were all from individuals for which we sequenced more than six clones, which is the minimum number of clones needed to detect two alleles at a locus with 95% certainty.

To investigate patterns of selection, we calculated the rates of non-synonymous (d_N) and synonymous (d_S) substitutions using the Nei and Gojobori (1986) method, with Jukes–Cantor correction using MEGA software version 2.1 (Kumar et al. 2001). We did this for all exons and for second exon codons in the putative antigen-binding region (ABR; defined according to Reche and Reinherz 2003). In addition, we used MEGA to perform *Z* tests of positive, neutral and purifying selection on the entire second exon, ABR codons, and the third exon, which encodes an extracellular domain close to the transmembrane region.

In order to investigate evolutionary relationships between DQA in elephants (Afrotheria) and other mammals (Boreoeutheria), we constructed phylogenetic trees using maximum parsimony (MP), maximum likelihood (ML), and neighbor joining (NJ) methods from PAUP v. 4.0b (Swofford 2002). For all analyses, we used a concatenated data set of the second and third exons. These exons were chosen because we were able to find representative sequences for the widest range of mammal species. Specifically, we included: (a) the African and Asian elephant DQA sequences from our study, (b) a representative selection of mammalian DQA sequences from GenBank (species and accession numbers=*Mirounga leonine*, MLUO3583; *Zalophus californianus*, AF502564; *Alopex lagopus*, Z26591; *Canis lupus*, NM_001011726; *Canis familiaris*, AJ311099; *Bos taurus*, D50454; *Ovis avies*, EE803522; *Sus scrofa*, AY285934; *Equus caballus*, L33909; *Mus musculus*, BC019721; *Aotus nancymaae*, AF201296; *Macacca mulatta*, EF362438; *Pan troglodytes*, AY663401; *Homo sapiens*, AY375903 and AK130811), and (c) the marsupial mammal, *Monodelphis domesticus* (XM_001376727), as an outgroup. The best model of evolution was chosen using the analyses generated with ModelTest version 3.7 (Posada and Crandall 1998). According to the results of ModelTest, ML trees were generated using the K80+G, Kimura two-parameter model with rate variation among sites. The ML analysis was conducted using a heuristic search under likelihood criteria, obtained via random stepwise addition with ten replicates starting from random trees, including tree bisection reconnection with multiple tree swapping. Nodal support was assessed with bootstrap resampling by using GARLI to create 1,000 ML replicates under the same search conditions as above, after which PAUP was used to calculate majority rule consensus.

Results

MHC DQA alleles in Asian and African elephants

From 429 cloned sequences of the second exon, second intron, and third exon of DQA (MDQA1 and MDQA2 amplicons), we identified ten unique alleles: six alleles in 30 African elephants and four alleles in three Asian elephants (Table 2). We chose five of these alleles, two derived from African elephants (LoafDQA*01, LoafDQA*02) and three from Asian elephants (ElmaDQA*01, ElmaDQA*02, and ElmaDQA*03) to characterize complete coding and nearly

complete non-coding sequences. Intron one sequences are not included here because intron one contained two microsatellites and several poly-T and poly-A regions, making it difficult to reliably sequence and align across individuals.

None of the nucleotide (Fig. 2) or amino acid sequences (Fig. 3) for any of the ten alleles have been previously described. Evidence supports the hypothesis that this locus is transcribed and expressed; cDNA amplification with MDQA1 and MDQA2 amplified the expected 460-bp fragment, none of the alleles contained stop codons in their expected coding regions, and the observed patterns of selection were inconsistent with neutral evolution (see below). These alleles are likely part of a single locus as no more than two alleles ever occurred in a single individual, and we observed Mendelian inheritance in six pairs of known parents and offspring whose genotypes were derived from tissue or blood samples (parentage confirmed by microsatellite genotyping; Archie et al. 2006, 2008; Hollister-Smith et al. 2007). We did not observe Mendelian inheritance in three additional pairs of parents and offspring that were genotyped from feces. This lack of Mendelian inheritance was probably due to allelic dropout, as all genotypes derived from feces were homozygous. Such allelic dropout is not surprising given that that DNA derived from fecal samples is often degraded into short fragments and even the short allele we amplified was still relatively long (818 base pairs). Primer sets that amplify smaller overlapping pieces may reduce the rate of allelic dropout in genotypes from fecal-derived DNA samples.

Allelic variation within species

African elephants In the six African elephant alleles, there were 70 variable nucleotide sites across all 1,893 coding and non-coding nucleotides (nucleotide diversity, π=0.027; Fig. 2). Across the entire coding sequence, 21 out of 255 amino acid residues were variable (8.24% variable amino acids; Fig. 3). Allelic diversity was highest in the second exon, with the highest nucleotide diversity of any region (π=0.039), and each second exon allele had a unique amino acid sequence, with 15.66% (13 of 83) variable amino acid residues. By comparison, nucleotide diversity was lower in all other exons and introns (exon 1 π=0; intron 2 π=0.026; exon 3 π=0.019; intron 3 π=0.008; exon 4 π=0.029; intron 4 π=0; exon 5 π=0.015). In addition, amino acid variability was lower in all other exons; no amino acids varied in the first exon, in the third exon 5.25% (5 of 95) of amino acids varied, while 5.88% (3 of 51) of amino acids varied in the fourth exon (the fifth exon is non-coding).

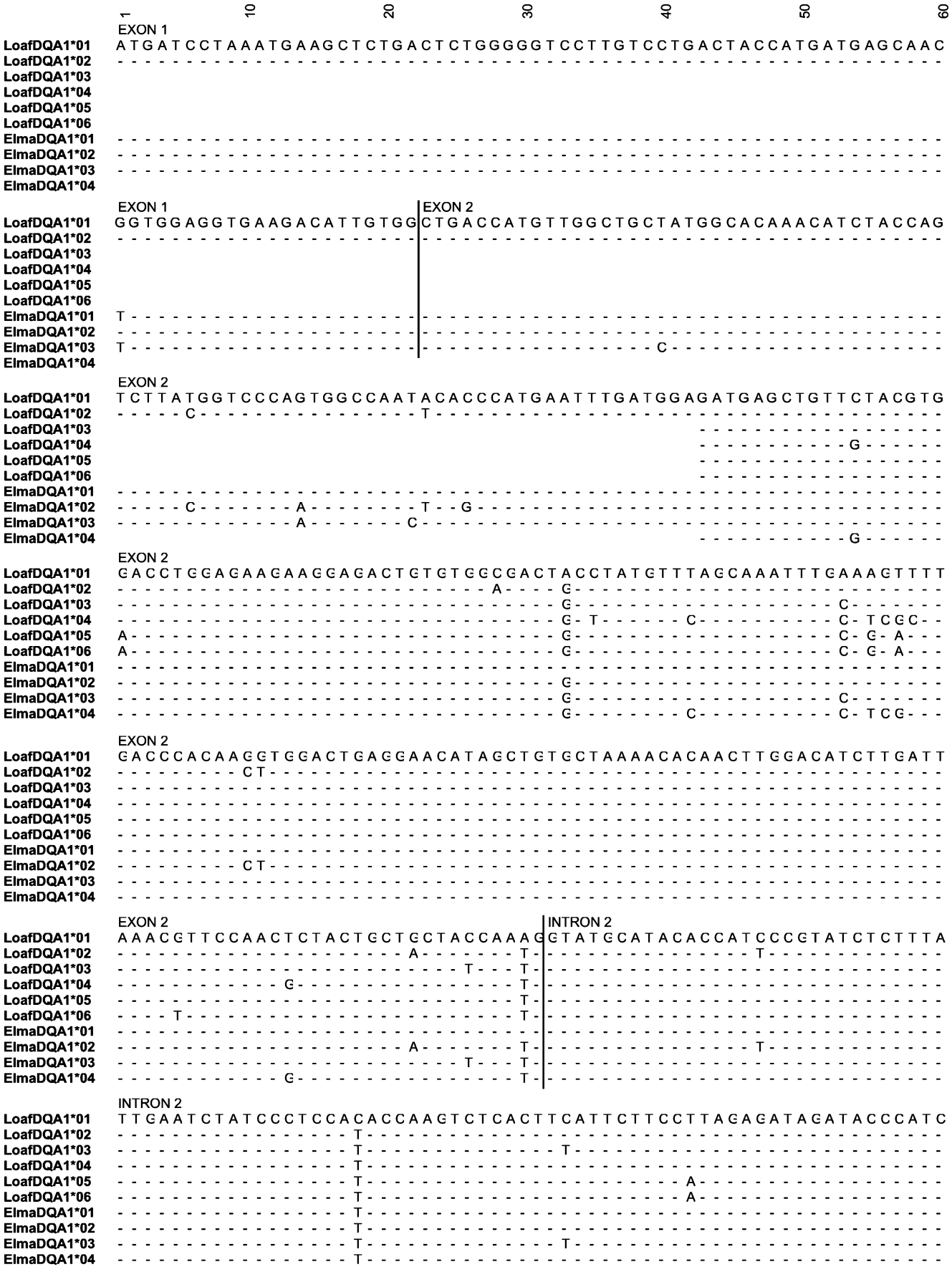
Asian elephants In the four alleles found in Asian elephants, there were 81 variable nucleotides across all 1,893 coding and non-coding bases (π=0.026; Fig. 2). Across the entire coding sequence, there were 18 variable amino acid residues in 255 amino acids (7.06% variable amino acids; Fig. 3). Allelic diversity was highest in the second exon; nucleotide diversity (π) was 0.039, and each second exon allele had a unique amino acid sequence, with 13.25% (11 of 83) of amino acid residues varying across the exon. By comparison, amino acid variability was lower in all other exons; one amino acid varied in the first exon (3.70%; one of 27), in the third exon 4.21% (four of 95) of amino acids varied, while 5.88% (three of 51) of amino acids varied in the fourth exon. Compared to the second exon, nucleotide diversity was also lower in all other exons and introns, except exon 4 (exon 1 π=0.008; intron 2 π=0.025; exon 3 π=0.017; intron 3 π=0.011; exon 4 π=0.040; intron 4 π=0.031; exon 5 π=0.023).

Allele frequencies and heterozygosity in African elephants

We calculated DQA allele frequencies from the 18 African elephants (15 wild and three captive) that were genotyped from six or more clones using only tissue-derived (no fecal-derived) DNA (Table 2). Among these individuals, mean heterozygosity was 0.611, and we observed one very common allele, LoafDQA1*01, which occurred in 16 of 18 samples (frequency of 0.611). Two other alleles, LoafDQA1*04 and LoafDQA1*06, were also relatively common (LoafDQA1*04 frequency=0.167 and LoafDQA1*06 frequency=0.111), while LoafDQA1*02, LoafDQA1*03, and LoafDQA1*05 were more rare (LoafDQA1*02 frequency=0.056 and LoafDQA1*05 frequency=0.028).

It is fairly unusual to observe a single common allele at an MHC locus, and LoafDQA1*01 might be common in multiple populations of African elephants; out of all 30 African elephants we genotyped, LoafDQA1*01 occurred in 20 individuals, including three of seven zoo elephants (one from Zimbabwe and two from Uganda), and 17 of 23 individuals from Amboseli National Park, Kenya. The high frequency of LoafDQA1*01 in Amboseli was not due to the fact that our sample contained close kin (parent–offspring pairs). Rather, with parent–offspring pairs removed, the frequency of LoafDQA1*01 in Amboseli was

Fig. 2 Alignment of MHC DQA sequences in Asian and African elephants relative to the most common allele in African elephants, LoafDQA1*01. Identities are shown by *dashes* and deletions are shown by *back slashes*. Alleles LoafDQA1*01, LoafDQA*02, ElmaDQA*01, ElmaDQA*02, and ElmaDQA*03 depict complete coding and non-coding sequences, except intron 1. Alleles LoafDQA1*03, LoafDQA*04, LoafDQA1*05, LoafDQA*06, and ElmaDQA*04 were amplified using the primers MDQA1 and MDQA2, which span the second exon, second intron, and third exon


1 10 20 30 40 50 60
EXON 1
LoafDQA1*01 ATGATCCTAAATGAAGCTCTGACTCTGGGGGTCCTTGTCCTGACTACCATGATGAGCAAC
LoafDQA1*02
LoafDQA1*03
LoafDQA1*04
LoafDQA1*05
LoafDQA1*06
ElmaDQA1*01
ElmaDQA1*02
ElmaDQA1*03
ElmaDQA1*04
EXON 1 EXON 2
LoafDQA1*01 GGTGGAGGTGAAGACATTGTGGCTGACCATGTTGGCTGCTATGGCACAAACATCTACCAG
LoafDQA1*02
LoafDQA1*03
LoafDQA1*04
LoafDQA1*05
LoafDQA1*06
ElmaDQA1*01 T
ElmaDQA1*02
ElmaDQA1*03 T C
ElmaDQA1*04
EXON 2
LoafDQA1*01 TCTTATGGTCCCAGTGGCCAATACACCCATGAATTTGATGGAGATGAGCTGTTCTACGTG
LoafDQA1*02 C T
LoafDQA1*03
LoafDQA1*04 G
LoafDQA1*05
LoafDQA1*06
ElmaDQA1*01
ElmaDQA1*02 C A T G
ElmaDQA1*03 A C
ElmaDQA1*04 G
EXON 2
LoafDQA1*01 GACCTGGAGAAGAAGGAGACTGTGTGGCGACTACCTATGTTTAGCAAATTTGAAAGTTTT
LoafDQA1*02 A G
LoafDQA1*03 G C
LoafDQA1*04 G T C C T C G C
LoafDQA1*05 A G C G A
LoafDQA1*06 A G C G A
ElmaDQA1*01
ElmaDQA1*02 G
ElmaDQA1*03 G C
ElmaDQA1*04 G C C T C G
EXON 2
LoafDQA1*01 GACCCACAAGGTGGACTGAGGAACATAGCTGTGCTAAAACACAACTTGGACATCTTGATT
LoafDQA1*02 C T
LoafDQA1*03
LoafDQA1*04
LoafDQA1*05
LoafDQA1*06
ElmaDQA1*01
ElmaDQA1*02 C T
ElmaDQA1*03
ElmaDQA1*04
EXON 2 INTRON 2
LoafDQA1*01 AAACGTTCCAACTCTACTGCTGCTACCAAAGGTATGCATACACCATCCCGTATCTCTTTA
LoafDQA1*02 A T T
LoafDQA1*03 T T
LoafDQA1*04 G T
LoafDQA1*05 T
LoafDQA1*06 T T
ElmaDQA1*01
ElmaDQA1*02 A T T
ElmaDQA1*03 T T
ElmaDQA1*04 G T
INTRON 2
LoafDQA1*01 TTGAATCTATCCCTCCACACCAAGTCTCACTTCATTCTTCCTTAGAGATAGATACCCATC
LoafDQA1*02 T
LoafDQA1*03 T T
LoafDQA1*04 T
LoafDQA1*05 T A
LoafDQA1*06 T A
ElmaDQA1*01 T
ElmaDQA1*02 T
ElmaDQA1*03 T T
ElmaDQA1*04 T

```
             1        10        20        30        40        50        60
             INTRON 2
LoafDQA1*01  TCACCATGCTATAGAACTCTCTCCTTTACAAGGAGTTCCCACTTCTTCTCATGGTAATTG
LoafDQA1*02  C-----CT----------------------------------A-----------------
LoafDQA1*03  ------CT--G------------------------------TA-----------------
LoafDQA1*04  C-----CT----------------------------------A-----------------
LoafDQA1*05  ------CT--G-------------------------------A-----------------
LoafDQA1*06  C-----CT----------------------------------A-G---------------
ElmaDQA1*01  ------------------------------------------A-----------------
ElmaDQA1*02  C-----CT----------------------------------A-----------------
ElmaDQA1*03  ------CT--G------------------------------TG-----------------
ElmaDQA1*04  C-----CT----------------------------------G-----------------

             INTRON 2
LoafDQA1*01  CTGAAACCTCATCATCACCCTTCTCACATTCCCATGTAATACAAGGGCCCTTACTGCCAT
LoafDQA1*02  --------------------G-------------G-------------------------
LoafDQA1*03  -----C------------------------------------------------------
LoafDQA1*04  -----C------------------------------------------------------
LoafDQA1*05  -----C------------------------------------------------------
LoafDQA1*06  -----C------------------------------------------------------
ElmaDQA1*01  ------------------------------------------------------------
ElmaDQA1*02  --------------------G-------------G-------------------------
ElmaDQA1*03  -----C------------------------------------------------------
ElmaDQA1*04  -----C------------------------------------------------------

             INTRON 2
LoafDQA1*01  CATGCATTCCTTGCCTCCTTCTAGGATGGGTCCCATAGACCTCCTAGCTTAAAGAAGGTA
LoafDQA1*02  ------------------------------------------------------------
LoafDQA1*03  ------------------------------------------------------------
LoafDQA1*04  ------------------------------------------------------------
LoafDQA1*05  ------------------------------------------------------------
LoafDQA1*06  ------------------------------------------------------------
ElmaDQA1*01  ------------------------------------------------------------
ElmaDQA1*02  ------------------------------------------------------------
ElmaDQA1*03  ------------------------------------------------------------
ElmaDQA1*04  ------------------------------------------------------------

             INTRON 2
LoafDQA1*01  CAGGGATAAAGTGCAGGCAGTACACAGTGATTCCCAAGCCAAAGGGAAACAAAAACCCTT
LoafDQA1*02  ----------------------------------------------------G-------
LoafDQA1*03  ----------------------------------------------------G-------
LoafDQA1*04  ----------------------------------------------------G------C
LoafDQA1*05  --A----------------------------------------------G--G-------
LoafDQA1*06  -------------------------------------------------G--G-------
ElmaDQA1*01  ------------------------------------------------------------
ElmaDQA1*02  ----------------------------------------------------G-------
ElmaDQA1*03  ----------------------------------------------------G-------
ElmaDQA1*04  ----------------------------------------------------G------C

             INTRON 2
LoafDQA1*01  CCACTTTCAGGTGGAGAAATATTGGTGGGAGGGCTCCCCCAAGATGCAATGCTATTAGCA
LoafDQA1*02  ------------------------------------------------------------
LoafDQA1*03  ---------------------------------------A-----------G--------
LoafDQA1*04  ----------------------------------------G-------------------
LoafDQA1*05  ------C-------------------------------------------AT--\-----
LoafDQA1*06  ------C--------------------------------------------T--\-----
ElmaDQA1*01  ----------T-------------------------------------------------
ElmaDQA1*02  ------------------------------------------------------------
ElmaDQA1*03  ---------------------------------------A-----------G--------
ElmaDQA1*04  ----------------------------------------G-------------------

             INTRON 2                      |EXON 3
LoafDQA1*01  ATTCATACCAGTGCCATGTCCTCACCACAG|AGGTTCCTGAAGTGACTGTGTTTCCCAAGT
LoafDQA1*02  --------------A-C-------------|------------------------------
LoafDQA1*03  ---T--------------------------|-------G----------------------
LoafDQA1*04  ------------------------------|------------------------------
LoafDQA1*05  -C----------------------------|------------------------------
LoafDQA1*06  -C----------------------------|------------------------------
ElmaDQA1*01  ------------------------------|------------------------------
ElmaDQA1*02  --------------A-C-------------|------------------------------
ElmaDQA1*03  ---T--------------------------|-------G----------------------
ElmaDQA1*04  ------------------------------|------------------------------

             EXON 3
LoafDQA1*01  CTCCCGTGACGCTGGGACAGCCCAACCTCCTCATCTGTTTTGTGGACAACATCTTTCCTC
LoafDQA1*02  ------------------------------------------------------------
LoafDQA1*03  ------------------------------------------------------------
LoafDQA1*04  --------------------------------------C---------------------
LoafDQA1*05  ----------C-------------------------------------------------
LoafDQA1*06  ----A-----C-------------------------------------------------
ElmaDQA1*01  ------------------------------------------------------------
ElmaDQA1*02  ------------------------------------------------------------
ElmaDQA1*03  ------------------------------------------------------------
ElmaDQA1*04  ----------A---------------------------C---------------------
```

Fig. 2 (continued)

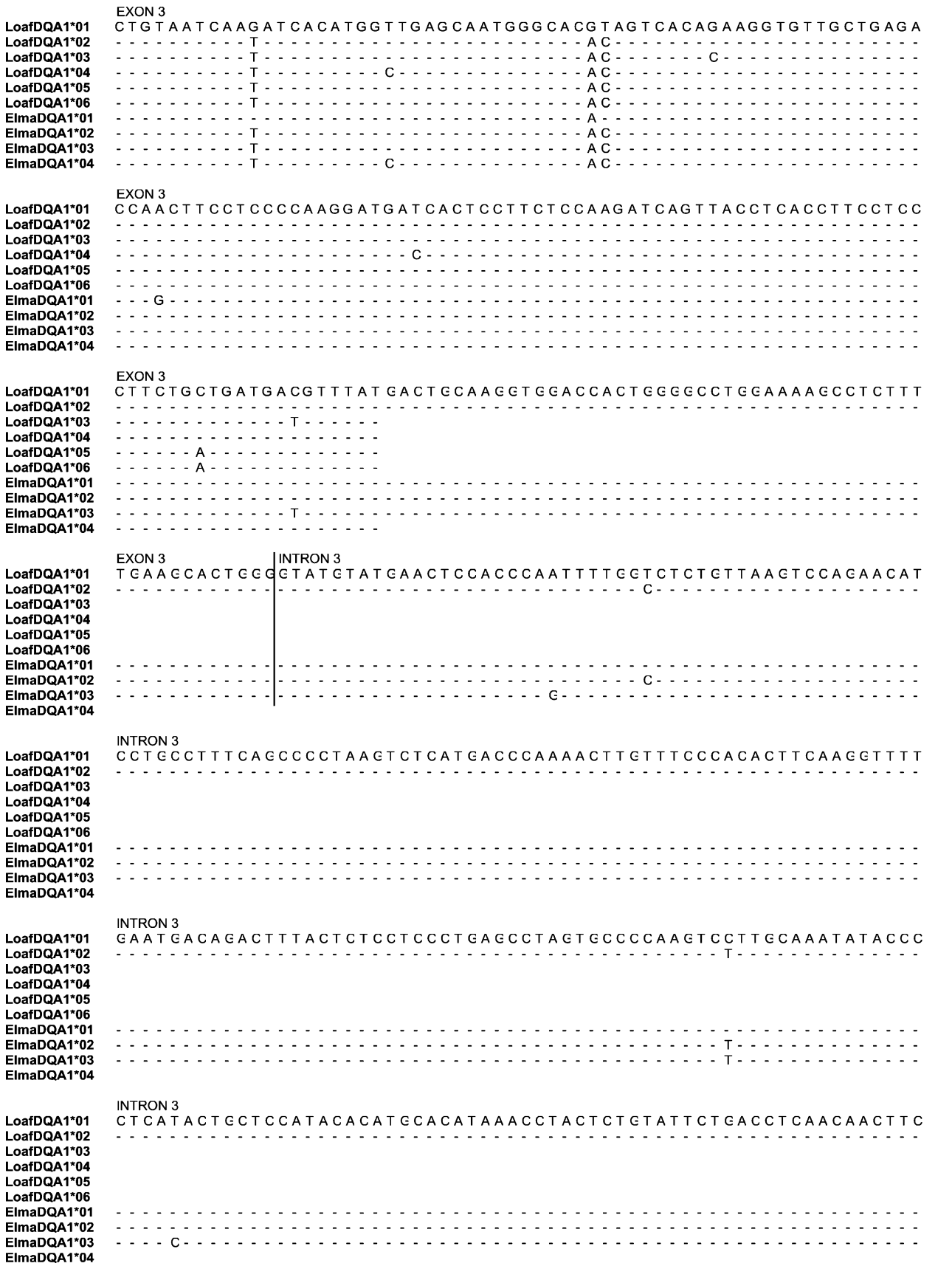


Fig. 2 (continued)

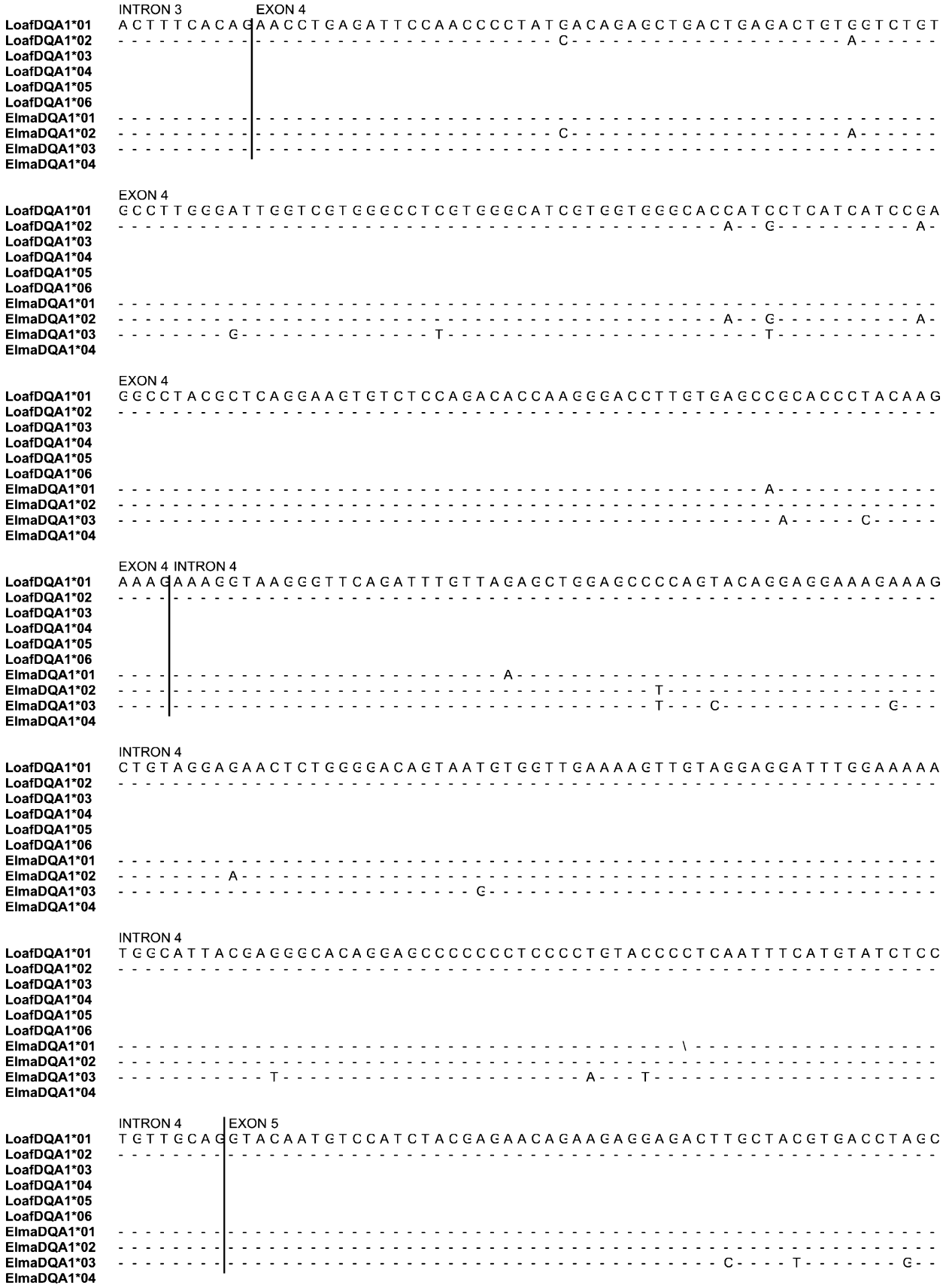


Fig. 2 (continued)

Fig. 2 (continued)

still 0.58. The high frequency of LoafDQA1*01 in Amboseli occurred in the context of Hardy–Weinberg equilibrium; heterozygosity in these 15 individuals was 0.6, which was not significantly different from expected heterozygosity (expected heterozygosity=0.58; exact test, P=0.71).

d_N/d_S supports balancing selection in the ABR of exon 2 and purifying selection in exon 3

The relative proportions of non-synonymous (d_N) and synonymous (d_S) substitutions reveal historical patterns of selection; a larger number of non-synonymous than synonymous changes constitutes evidence for balancing selection, while d_N/d_S ratios less than one are evidence for purifying selection (Hughes and Nei 1989). Consistent with the hypothesis that balancing selection operates on the ABR of the second exon, the d_N/d_S ratio in the putative ABR was 3.18 across all ten alleles (Table 3). Z tests of selection (Nei and Kumar 2000) indicated that this ratio was significantly different from neutrality (Z=2.17, P=0.032) and supported balancing selection (positive selection; Z=2.00, P=0.024).

In contrast to the ABR, the third exon is expected to be under purifying selection (Hughes and Yeager 1998). In support of this prediction, the rate of synonymous substitutions was similar across both the second and third exon (African elephants, second exon d_S=0.034, third exon d_S=0.042; Asian elephants, second exon d_S=0.034, third exon d_S=0.037; Table 3), but the rate of non-synonymous substitutions was much lower in the third exon as compared to the second exon (African elephants, second exon d_N=0.042, third exon d_N=0.013; Asian elephants, second exon d_N=0.039, third exon d_N=0.012; Table 3). As a result, the d_N/d_S ratio is markedly lower in the third exon (d_N/d_S=0.32). Z tests allowed us to exclude the possibility that the third exon experiences balancing selection (Z=−1.44, P=1.00); however, a test of purifying selection was not significant (Z=1.45, P=0.075), and Z

tests indicated that the d_N/d_S ratio does not differ from neutrality ($Z=-1.55$, $P=0.12$).

Phylogenetic patterns: trans-generic allelism supports balancing selection

An additional prediction of the hypothesis that balancing selection acts on this DQA locus is that allelic lineages should be maintained across species over long evolutionary time frames (Klein 1980; Takahata and Nei 1990). Topologies estimated from MP, ML, and NJ analyses of concatenated second and third exon data set were highly concordant and we found that ML produced the most representative relationships (Fig. 4). The phylogenetic analysis resolves the basal branching between the Afrotherian (e.g., elephants, aardvarks) and Boreoeutherian mammals (e.g. primates, carnivores, rodents, ruminants), and indicates that elephants do not share second exon alleles with any major groups of Boreoeutherian mammals. However, within the two extant species of elephants, we found strong evidence of trans-generic allelism (Fig. 4). In particular, the entire second exon, second intron, and third exon sequence of each Asian allele was closely matched to an allele found in African elephants. The nucleotide sequences at the second exon, second intron, and third exon for the African allele, LoafDQA1*03, and the Asian allele, ElmaDQA1*03, have identical nucleotide sequences across the second exon, second intron, with only a single

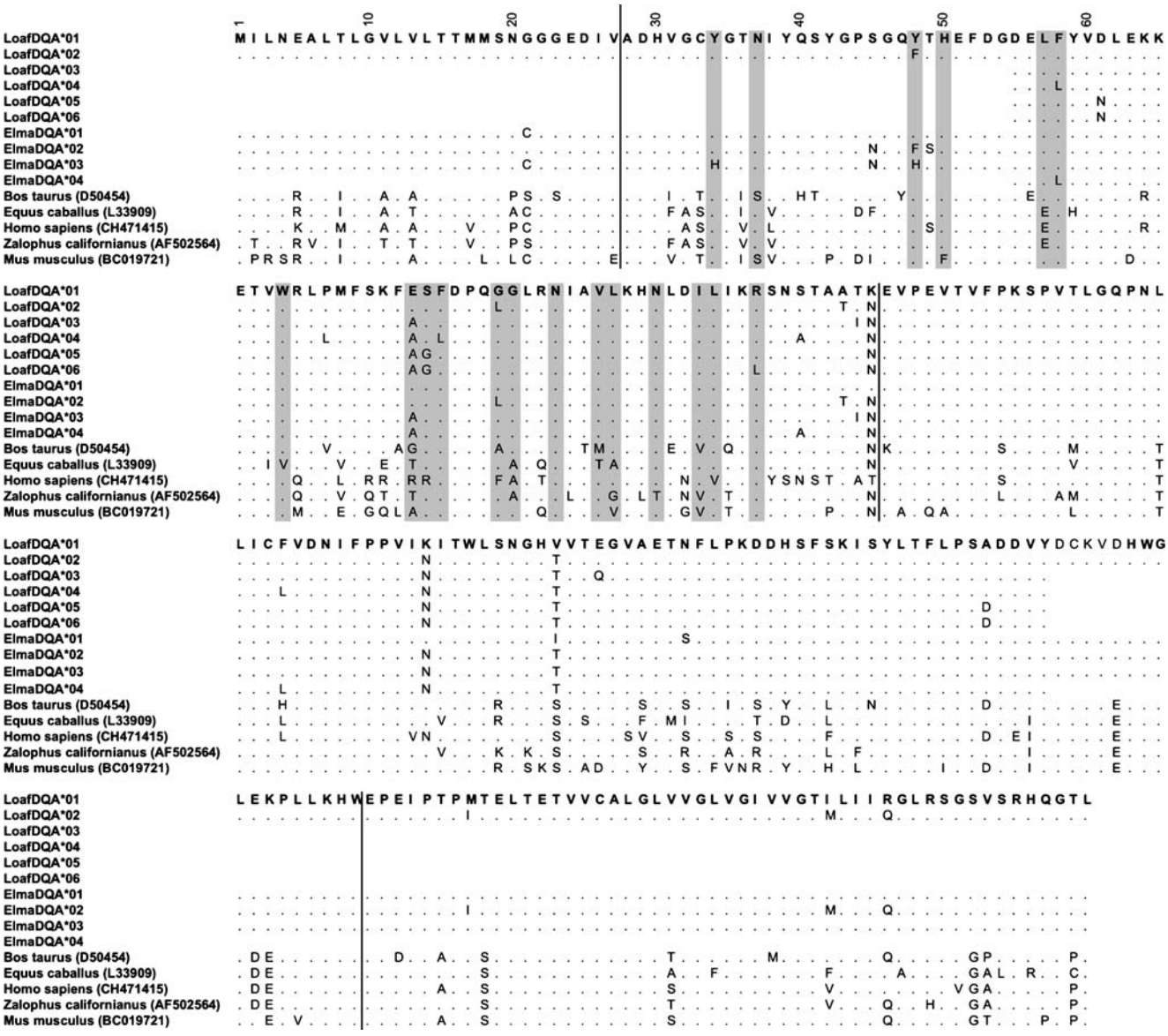


Fig. 3 Alignment of MHC DQA predicted amino acid sequences from Asian and African elephants and five other Boreoeutherian mammals. Amino acid positions included in the peptide binding region are *shaded gray* and were defined according to Reche and Reinherz (2003). Genbank accession numbers are in *parentheses*. *Bars* indicate divisions between exons; identities are shown by *dots*

Table 3 Estimates of d_N and d_S (±SE) in all DQA exons including the hypothesized antigen-binding region (ABR) of the second exon

Region	No. of codons	*Loxodonta africana*			*Elephas maximus*			Species combined
		d_N	d_S	d_N/d_S	d_N	d_S	d_N/d_S	d_N/d_S
Exon 1	27	0±0	0±0	0	0.017±0.016	0±0	NA	NA
Exon 2	83	0.042±0.015	0.034±0.018	1.24	0.043±0.016	0.034±0.021	1.26	1.15
Non-ABR	64	0.018±0.008	0.038±0.022	0.47	0.021±0.010	0.038±0.028	0.55	0.49
ABR	19	0.116±0.040	0.039±0.039	2.97	0.109±0.059	0.024±0.025	4.54	3.18
Exon 3	89	0.013±0.006	0.042±0.016	0.31	0.012±0.006	0.037±0.019	0.32	0.32
Exon 4	52	0.025±0.014	0.043±0.029	0.58	0.028±0.012	0.078±0.031	0.036	0.38

mutation leading to a single amino acid change in the third exon (Figs. 2 and 3). ElmaDQA1*02 and LoafDQA1*02 had identical nucleotide sequences across all exons, except for two amino acid changes in non-ABR regions of the second exon (Fig. 3). ElmaDQA1*01 and LoafDQA1*01 had identical amino acid sequences across all exons, except for two amino acid changes in the third exon (Fig. 3). ElmaDQA*04 and LoafDQA*04 had identical amino acid sequences except for a single amino acid change in the second exon (Fig. 3).

Discussion

Our characterization of elephant MHC loci represents the first description of an MHC locus in the superorder of Afrotherian mammals. Beyond the phylogenetic value of the analysis, it represents a first step towards characterizing genetic diversity at a set of loci that may be involved in the kin discrimination that we have previously documented in this species (Archie et al. 2007). In addition, MHC loci might provide an important measure of functional genetic

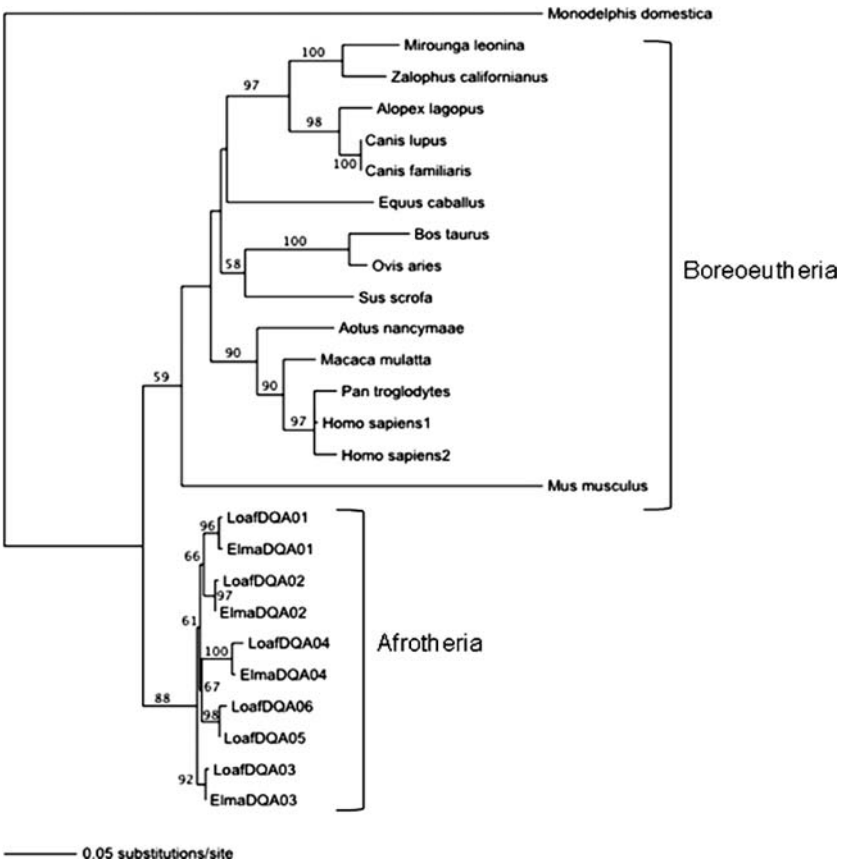


Fig. 4 Maximum likelihood phylogram depicting evolutionary relationships among MHC DQA haplotypes across Boreoeutherian mammals and elephants (Afrotheria) for concatenated exons 2 and 3. *Numbers* above internodes are bootstrap support values from ML analyses. See text for tree construction methods and sequence accession numbers

diversity in threatened elephant populations, especially with regard to response to infectious disease. The locus we characterized is most closely related to MHC DQA1 in other mammals.

Comparative genetic diversity

Patterns of MHC diversity across different species can reveal how differences in demography, life history, or disease threats may shape MHC evolution. Elephants have several life history traits that differ from those of typical mammals (e.g., extremely large body size and long life span), and these traits, as well as the complex sociality that characterizes elephants and a number of other mammal species, might influence their exposure to disease. For instance, because elephants are unusually long-lived, we might expect them to encounter relatively more infectious agents over the course of their lives and consequently require relatively high MHC diversity. However, the patterns of genetic diversity at the DQA locus in our study were—for the most part—qualitatively similar to DQA diversity in other wild mammals. For instance, compared to elephants, more alleles per study subject were observed in wild baboons and Weddell seals (Alberts 1999; Lehman et al. 2004), while fewer alleles were found in Ross seals, leopard seals, elephant seals, marmosets, and tuco tucos (Antunes et al. 1998; Cutrera and Lacey 2006; Lehman et al. 2004; Weber et al. 2004).

However, the African DQA locus in our study differed from other studies of wild mammals in that one allele was especially common. This allele, LoafDQA*01, occurred in over half of our samples and may be common in multiple populations. Such high frequencies of a single DQA allele have previously been reported only in species with low polymorphism; for instance, only species with two or fewer alleles have produced frequencies for a single DQA allele that are greater than 0.5 (Antunes et al. 1998; Lehman et al. 2004; Weber et al. 2004). One speculation is that LoafDQA*01 is common because it confers a selective advantage to individuals in resisting a disease that was or is highly prevalent. However, a larger sample size of individuals and populations is necessary to confirm that this allele is really as common as it appears in our sample. In addition, we might expect to observe evidence of a selective sweep in the region of the genome surrounding this allele.

Evidence for balancing and purifying selection

Genetic diversity at MHC loci is likely to be maintained by balancing selection (reviewed in Garrigan and Hedrick 2003; Hughes 1999; Hughes and Yeager 1998; Piertney and Olivier 2006). Such selection leaves a characteristic mutational signature in coding regions; in particular, codons that have experienced balancing selection should have more non-synonymous changes than synonymous changes, while codons that experienced purifying selection will have very few non-synonymous changes relative to the number of synonymous changes. We found evidence that is consistent with both balancing and purifying selection in elephants; the d_N/d_S ratios in the ABR of the second exon were significantly greater than one, indicating that positive evolution has acted on the ABR. In contrast, the third exon, which is not involved in antigen recognition, appears to have experienced purifying selection, as we found two or three times as many synonymous changes than non-synonymous changes in both African and Asian elephants. However, for the third exon, we were not able to reject the null hypothesis of neutral evolution. This might be because of our small sample size of alleles (and therefore high variance in our estimates of d_N and d_S); further study is needed to conclusively distinguish between the forces of selection and neutrality at these loci.

Balancing selection can also act to maintain lineages of alleles across species over long evolutionary time scales. While we did not find any evidence that Asian or African elephants share alleles with the analyzed Boreoeutherian mammals, we did find strong evidence of trans-species allelism acting across the two extant elephant genera, *Loxodonta* and *Elephas*. This pattern was clearly evident in the second exon, where the amino acid sequence of all four Asian elephant alleles matched very closely the amino acid sequences of four African elephant alleles. The trans-species allelism we observed in the second exon extended to the intron and third exon of the same pairs of allele, but was less extreme. These trans-species allelic lineages may have been maintained over relatively long evolutionary time frames, as Asian and African elephants are thought to have diverged around 6 million years ago (Krause et al. 2006). However, this phenomenon has been observed over even longer evolutionary time frames. For instance, in primates, trans-species allelic lineages have been conserved for over 30 million years (Geluk et al. 1993). As more information on the MHC of other mammal species becomes available, it will be interesting to know if trans-species allelism operates across Afrotheria.

Conclusions and implications

Our results provide the first view of sequence structure and diversity of MHC in elephants and Afrotheria. They provide full transcript sequences that should prove useful for further development of markers to assess genetic diversity at this locus, and indicate that diversity at the DQA locus in elephants is similar to that seen in other mammals. This diversity is likely to be a result of the

historical action of balancing selection; in particular, the proportions of synonymous to non-synonymous changes and the strong evidence for trans-species allelism suggests that balancing selection has acted to maintain multiple, diverse lineages of DQA alleles in Asian and African elephants over at least 6 million years. However, our study did not reveal the underlying mechanisms driving balancing selection at this locus. A more thorough characterization of diversity at this locus, and especially other loci, are needed to fully understand the evolutionary ecology of the MHC in elephants.

Acknowledgments We thank the Office of the President of Kenya for permission to work in Amboseli National Park under permit number MOES&T 13/001/30C 72/7. We thank the Kenya Wildlife Service for local sponsorship. We thank the Amboseli Elephant Research project for invaluable scientific and logistical support, particularly the team of N. Njiraini, K. Sayialel, and S. Sayialel who contributed greatly to the collection of genetic samples. We thank the Smithsonian National Zoo, the Philadelphia Zoo in Philadelphia, PA, the Gladys Porter Zoo in Brownsville, TX, and the Six Flags Wild Safari Park in Jackson, NJ for their support and cooperation in sample collection. This research was supported by the Smithsonian Institution Abbott Endowment Fund, the National Zoo's Institution Center for Conservation and Evolutionary Genetics, the Friends of the National Zoo, the National Science Foundation (IBN0091612 to SCA), the Amboseli Trust for Elephants, the Amboseli Elephant Research Project, and Duke University.

References

Alberts SC (1999) Thirteen Mhc-DQA1 alleles from two populations of baboons. Immunogenetics 49:825–827

Antunes SG, de Groot NG, Brok H, Doxiadis G, Menezes AAL, Otting N, Bontrop RE (1998) The common marmoset: a new world primate species with limited MHC class II variability. Proc Natl Acad Sci U S A 95:11745–11750

Archie EA, Moss CJ, Alberts SC (2003) Characterization of tetranucleotide microsatellite loci in the African savannah elephant (*Loxodonta africana africana*). Mol Ecol Notes 3:244–246

Archie EA, Moss CJ, Alberts SC (2006) The ties that bind: genetic relatedness predicts the fission and fusion of groups in wild African elephants (*Loxodonta africana*). Proc Roy Soc London 273:513–522

Archie EA, Hollister-Smith JA, Poole JH, Lee PC, Moss CJ, Maldonado JE, Fleischer RC, Alberts SC (2007) Behavioral inbreeding avoidance in wild African elephants. Mol Ecol 16:4138–4148

Archie EA, Maldonado JE, Hollister-Smith JA, Poole JH, Moss CJ, Fleischer RC, Alberts SC (2008) Fine-scale population genetic structure in a fission–fusion society. Mol Ecol 17:2666–2679

Armbruster P, Lande R (1993) A population viability analysis for African elephant (*Loxodonta africana*): how big should reserves be? Conserv Biol 7:602–610

Borriello F, Krauter KS (1990) Reactive site polymorphism in the murine protease inhibitor gene family is delineated using a modification of the PCR reaction. Nucleic Acids Res 18:5481–5487

Brown JL, Eklund A (1994) Kin recognition and the major histocompatibility complex: an integrative review. Am Nat 143:435–461

Caughley G, Dublin HT, Parker ISC (1990) Projected decline of the African elephant. Biol Conserv 54:157–164

Cheetham SA, Thom MD, Jury F, Ollier WER, Beynon RJ, Hurst JL (2007) The genetic basis of individual-recognition signals in the mouse. Curr Biol 17:1771–1777

Cutrera AP, Lacey EA (2006) Major histocompatibility complex variation in talas tuco-tucos: the influence of demography on selection. J Mammal 87:706–716

Decker DJ, Stewart BS, Lehman N (2002) Major histocompatibility complex class II DOA sequences from three Antarctic seal species verify stabilizing selection on the DO locus. Tissue Antigens 60:534–538

Garrigan D, Hedrick PW (2003) Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. Evolution 57:1707–1722

Gaudier S, Dawkins RL, Habara K, Kulski JK, Gojobori T (2000) SNP profile within the human major histocompatibility complex reveals and extreme and interrupted level of nucleotide diversity. Genome Res 10:1579–1586

Geluk A, Elferink DG, Slierendregt BL, Vanmeijgaarden KE, de Vries RRP, Ottenhoff THM, Bontrop RE (1993) Evolutionary conservation of major histocompatibility complex-DR/peptide/T cell interactions in primates. J Exp Med 177:979–987

Grobler DG, Raath JP, Braak LE, Keet DF, Gerdes GH, Barnard BJ, Kriek NP, Jardine J, Swanepoel R (1995) An outbreak of encephalomyocarditis-virus infection in free-ranging African elephants in the Kruger National Park. Onderstepoort J Vet Res 62:97–108

Harrigan RJ, Mazza ME (2008) Computation vs. cloning: evaluation of two methods for haplotype determination. Mol Ecol Resour 8:1239–1248

Hollister-Smith JA, Poole JH, Archie EA, Vance EA, Georgiadis NJ, Moss CJ, Alberts SC (2007) Age, musth and paternity in wild male African elephants, *Loxodonta africana*. Anim Behav 74:287–296

Hughes AL (1999) Adaptive evolution of genes and genomes. Oxford University Press, New York

Hughes AL, Nei M (1989) Nucleotide substitution at major histocompatibility complex class II loci evidence for overdominant selection. Proc Natl Acad Sci 86:958–962

Hughes AL, Yeager M (1998) Natural selection at major histocompatibility complex loci of vertebrates. Annu Rev Genet 32:415–435

Klein J (1980) Generation of diversity at MHC loci: implications for T-cell receptor repertoires. In: Fougereau M, Dausset J (eds) Immunology 80. Academic, London, pp 239–235

Klein J (1986) Natural history of the major histocompatibility complex, 1st edn. Wiley, New York

Krause J, Dear PH, Pollack JL, Slatkin M, Spriggs H, Barnes I, Lister AM, Ebersberger I, Paabo S, Hofreiter M (2006) Multiplex amplification of the mammoth mitochondrial genome and the evolution of elephantidae. Nature 439:724–727

Kriegs JO, Churakov G, Kiefmann M, Jordan U, Brosius J, Schmitz J (2006) Retroposed elements as archives for the evolutionary history of placental mammals. PLoS Biol 4(4):e91

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17:1244–1245

L'Abbe D, Belmaaza A, Decary F, Chartrand P (1992) Elimination of heteroduplex artifacts when sequencing HLA genes amplified by polymerase chain reaction. Immunogenetics 35:395–397

Lehman N, Decker DJ, Stewart BS (2004) Divergent patterns of variation in major histocompatibility complex class II alleles among Antarctic phocid pinnipeds. J Mammal 85:1215–1224

Lindique PM, Turnbull PCB (1994) Ecology and epidemiology of anthrax in the Etosha National Park, Namibia. Onderstepoort J Vet Res 61:71–83

Manning CJ, Wakeland EK, Potts WK (1992) Communal nesting patterns in mice implicate MHC genes in kin recognition. Nature 360:581–583

Mbise AN, Mlengeya TDK, Mollel JO (1998) Septicaemic salmonellosis of elephants in Tanzania. Bull Anim Health Prod Afr 46:95–100

Meyer-Lucht Y, Sommer S (2005) MHC diversity and the association to nematode parasitism in the yellow-necked mouse (*Apodemus flavicollis*). Mol Ecol 14:2233–2243

Murphy WJ, Eizrik E, O'Brien SJ, Madsen O, Scally M, Douady CJ, Teeling E, Ryder OA, Stanhope MJ, de Jong WW, Springer MS (2001) Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science 294:2348–2351

Naruse TK, Kawata H, Anzai T, Takashige N, Kagiya M, Nose Y, Nabeya N, Isshiki G, Tatsumi N, Inoko H (1999) Limited polymorphism in the HLA-DOA gene. Tissue Antigens 53:359–365

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418–426

Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press, New York

Paterson S, Wilson K, Pemberton JM (1998) Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population (*Ovis aries* L.). Proc Natl Acad Sci U S A 95:3714–3719

Penn DJ, Damjanovich K, Potts WK (2002) MHC heterozygosity confers a selective advantage against multiple-strain infections. Proc Natl Acad Sci U S A 99:11260–11264

Piertney SB, Olivier MK (2006) The evolutionary ecology of the major histocompatibility complex. Heredity 96:7–21

Posada D, Crandall KA (1998) MODEL TEST: testing the model of DNA substitution. Bioinformatics 14:817–818

Rajakaruna RS, Brown A, Kaukinen KH, Miller KM (2006) Major histocompatibility complex and kin discrimination in Atlantic salmon and brook trout. Mol Ecol 15:4569–4575

Reche PA, Reinherz EL (2003) Sequence variability analysis of human class I and class II molecules: functional and structural correlates of amino acid polymorphisms. J Mol Biol 331:623–641

Richman LK, Montali RJ, Garber RL, Kennedy MA, Lehnhardt J, Hildebrandt T, Schmitt D, Hardy D, Alcendor DJ, Hayward GS (1999) Novel endotheliotropic herpesviruses fatal for Asian and African elephants. Science 283:1171–1176

Richman LK, Montali RJ, Hayward GS (2000) Review of a newly recognized disease of elephants caused by endotheliotropic herpesviruses. Zoo Biol 19:383–392

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, NJ, pp 365–386 Source code available at http://fokker.wi.mit.edu/primer3/.

Ryan SJ, Thompson SD (2001) Disease risk and inter-institutional transfer of specimens in cooperative breeding programs: herpes and the elephant species survival plan. Zoo Biol 20:89–101

Scally M, Madsen O, Douady CJ, de Jong WW, Stanhope MJ (2001) Molecular evidence for the major clades of placental mammals. J Mamm Evol 8:239–277

Slade RW, Moritz C, Heideman A, Hale PT (1993) Rapid assessment of single-copy nuclear DNA variation in diverse species. Mol Ecol 2:359–373

Swofford DL (2002) Phylogenetic analysis using parsimony (*and other methods). Sinauer, Sunderland

Takahata N, Nei M (1990) Allelic genealogy under overdominant and frequency-dependent selection and polymorphism at the major histocompatibility complex loci. Genetics 124:967–978

Waddell PJ, Shelley S (2003) Evaluating placental inter-ordinal phylogenies with novel sequences including RAG1, [gamma]-fibrinogen, ND6, and mt-tRNA, plus MCMC-driven nucleotide, amino acid, and codon models. Mol Phylogenet Evol 28:197–224

Weber DS, Brent SS, Schienman J, Lehman N (2004) Major histocompatibility complex variation at three class II loci in the northern elephant seal. Mol Ecol 13:711–718

Westerdahl H, Walenstrom J, Hansson B, Hasselquist D, von Schantz T, and Bensch S (2005) Associations between malaria and MHC genes in a migratory songbird. Proceedings of the Royal Society of London 272:1511–1518

Zelano B, Edwards SV (2002) An MHC component to kin recognition and mate choice in birds: predictions, progress, and prospects. Am Nat 160:S225–S237