Major histocompatibility complex (MHC) variation in the endangered Mexican wolf and related canids[1]

PHILIP W. HEDRICK*, RHONDA N. LEE & KAREN M. PARKER†

Department of Biology, Arizona State University, Tempe, AZ 95287, U.S.A.

We have examined in Mexican wolves and related canids the amount of genetic variation for a class II gene in the major histocompatibility complex (MHC), thought to be part of the most important genetic basis for pathogen resistance in vertebrates. In Mexican wolves, descended from only seven founders over three lineages, there were five different alleles. These were in three phylogenetic groups, only one of which was shared between lineages. Using single stand conformation polymorphism (SSCP), we found that in samples of animals from the two polymorphic lineages, the observed heterozygosity was 0.74 and the genotypes were not different statistically from Hardy–Weinberg proportions. The Ghost Ranch lineage of Mexican wolves was monomorphic for the locus, consistent with the lower level of variation found previously for microsatellite loci and predicted from pedigree analysis. Samples of grey wolves, red wolves, and coyotes had 16 additional alleles. One Mexican wolf allele was also found in grey wolves and another allele was shared between grey and red wolves. Most of the nucleotide variation resulted in amino acid variation and there were five different amino acids found at two different positions. Only two of the 21 variable amino acid positions had solely synonymous nucleotide variation. The average heterozygosity for eight individual amino acid positions in the Mexican wolves was greater than 0.4. The estimated rate of nonsynonymous substitution was 2.5 times higher than that for synonymous substitution for the putative antigen binding site positions, indicative of positive selection acting on these positions. Examination of the known dog sequences for this locus showed that one of the Mexican wolf alleles was found in dogs and that the allele found in both grey and red wolves is also found in dogs.

Keywords: adaptive variation, dogs, heterozygosity, phylogenetic tree, substitution rate.

Introduction

The grey wolf, *Canis lupus*, once had a distribution throughout much of the northern hemisphere. However, hunting and eradication programmes in the late nineteenth and throughout the twentieth centuries resulted in its extirpation from most of its original range. In the U.S.A., the remaining wolves have been listed as endangered and the Mexican grey wolf subspecies, *C. l. baileyi*, was listed as endangered in 1976. Mexican wolves are thought to have been extirpated from the

U.S.A., and greatly reduced in Mexico, by 1970. Because there have been no confirmed sightings in over 15 years throughout their ancestral range, they are thought to be extinct in the wild.

The extant Mexican wolves are in captivity, except for approximately 20 wolves that have been reintroduced into east central Arizona over the past few years. A survey of 20 microsatellite loci (Garcia-Moreno *et al.*, 1996; Hedrick *et al.*, 1997) was used as a basis to combine the three captive Mexican wolf lineages — McBride (formerly known as Certified), Ghost Ranch, and Aragon. The combined captive population is now descended from seven founders — three from the McBride, two from Ghost Ranch, and two from Aragon lineages. All of the released wolves to date descend entirely from three founders of the McBride lineage.

The ancestral range of the Mexican wolf subspecies in northern Mexico and south-western U.S.A. is isolated

geographically from the other extant subspecies of the grey wolf. Mexican wolves are one of the smaller grey wolf subspecies in North America, but there is great range in size in both Mexican wolves and other subspecies. However, investigations using microsatellite variation clearly differentiated between Mexican and northern grey wolves (Garcia-Moreno *et al.*, 1996; Hedrick *et al.*, 1997).

The major histocompatibility complex (MHC) is one of the most important genetic systems for infectious disease resistance in vertebrates (Hill, 1998; Hedrick & Kim, 2000). The association of disease resistance and MHC variation has often been difficult to document for various reasons (Hedrick & Kim, 2000), but a number of human studies have documented its importance for resistance to malaria (Hill *et al.*, 1991), hepatitis (Thurz *et al.*, 1997), and AIDS (Carrington *et al.*, 1999). In addition, O'Brien & Evermann (1988) suggested that low MHC variation might be related to higher susceptibility to infectious disease. For small populations and endangered species, it is recognized that there may be lowered levels of genetic variation; it also appears that mortality from pathogens and parasites may be a significant extinction threat (Lyles & Dobson, 1993; Laurenson *et al.*, 1998; Murray *et al.*, 1999).

Therefore, in the present contribution we characterize the genetic variation in a highly variable class II MHC gene in the three lineages of the Mexican wolf, and, for comparison, examine samples from Canadian grey wolves, red wolves, and coyotes. We also compare our data to the known dog alleles for this gene.

Materials and methods

Sample sources and DNA techniques

DNA samples were obtained from R. Wayne and J. Leonard, University of California-Los Angeles. The 36 Mexican wolf samples were from the McBride lineage (18 individuals), Ghost Ranch lineage (10), and Aragon lineage (8). In addition, 15 samples of Canadian grey wolves, *C. lupus*, 12 samples of red wolves, *C. rufus*, and 10 samples of coyotes, *C. latrans*, were obtained from the same source.

Primers for amplifying a 228 base pair (bp) fragment of MHC class II *DRB*1 exon 2 used for single strand conformation polymorphism (SSCP) analysis and subsequent cloning were designated DM-1 (5′-aag tcc gag tgc tat ttc acc-3′) and DM-2 (5′-tcg ccg ctg cac cgt gaa gct-3′) and were designed based on the dog sequences for the homologous gene summarized by Kennedy *et al.* (1998). Amplification conditions consisted of 35 cycles of PCR in a 10 μL reaction containing 10 ng of genomic DNA, 0.5 μM of

each primer, 3 mM $MgCl_2$, 1× polymerase buffer and 1 U Taq polymerase (Promega, Madison, WI) with an annealing temperature of 65°C. PCR conditions for SSCP analysis were the same but included 1 μCi of ^{32}P in a 5 μL reaction volume. Five μL of stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol) was added to each sample. Samples were denatured by heating at 95–100°C for 3 min, cooling in ice water for 3 min and then loading on a 6% polyacrylamide gel with 2.6% cross-linking. The SSCP gel was electrophoresed at 50 watts for 3 h 30 min at 4°C, transferred to 3 MM Whatman paper, dried and exposed to X-ray film overnight.

PCR amplification products made with the primers DM-1 and DM-2 were purified on an Ultrafree-MC centrifugal filter unit as recommended by the manufacturer (Millipore, Bedford, MA), and ligated into pUC18-*Sma*-T overhang vector (Bayou Biolabs, Harahan, LA). Positive colonies were identified by insertional inactivation and picked and boiled in 50 μL HPLC grade water. One μL of this solution was used as template for subsequent PCR. Subclones with the correct size insert were screened by SSCP using DM-1 and DM-2 and clones with band profiles identical to that individual's genomic SSCP pattern were chosen for sequencing. Heterozygotes were identified with SSCP as the sum of the profiles for two allelic subclones. Sequencing templates were prepared by amplifying with the plasmid primers in a 100 μL reaction and then the products were purified using an Ultrafree-MC centrifugal filter unit. Subclones were sequenced on both strands on an Applied Biosystems 377 automated sequencer (division of Perkin-Elmer, Foster City, CA). Sequences were identical in all replicates and from both strands.

Data analysis

In order to align the sequences, the sequence editor ESEE 3.25 (Eric Cabot, 1998) was used. The MEGA program (Kumar *et al.*, 1993) was used to construct the neighbour-joining tree using the genetic distance of Jukes & Cantor (1969) and to obtain bootstrap confidence intervals (500 replicates). MEGA was also used to calculate the relative rate of nonsynonymous and synonymous substitutions according to Nei & Gojobori (1986) and applying the correction of Jukes & Cantor (1969) for multiple hits. Expected heterozygosity was calculated after Nei (1987) with the small sample size correction. Amino-acid heterozygosity for individual sites was calculated by weighting the amino acids in given sequences by their population frequency as in Hedrick *et al.* (1991).

Results

For the Mexican wolves, all of the samples successfully amplified except for one sample from the McBride lineage. SSCP gels representing the class II *DRB1* exon 2 resulted in band patterns consistent with two alleles in the McBride lineage, one allele in the Ghost Ranch lineage, and three alleles in the Aragon lineage (Table 1). The only overlap in alleles between lineages was the Ghost ranch allele *Calu*-3, which was also the most frequent allele in the Aragon lineage. The observed heterozygosities for the McBride and Aragon lineages were higher than that expected under Hardy–Weinberg proportions (Table 2). However, the observed heterozygosities in these relatively small samples were not significantly different statistically from that expected (or were they for the grey wolf, red wolf, and coyote samples discussed below).

For the grey wolves, 13 out of 15 samples amplified, for the red wolves only 3 out of 12 samples amplified, and for the coyotes, all 10 samples amplified. The amount and quality of the red wolf DNA in these samples was not adequate for good amplification. Ten of the amplified samples of the grey wolf were from the Northwest Territories, Canada, and three from Alberta,

Table 2 The observed and expected (using Hardy–Weinberg proportions) heterozygosities at the *DRB1* locus for the different lineages and species

Lineage or species	Heterozygosity	
	Observed	Expected
Mexican wolf		
McBride	0.59	0.47
Ghost Ranch	0.00	0.00
Aragon	0.88	0.66
Grey wolf	0.62	0.81
Red wolf	1.00	0.73
Coyote	0.50	0.81

Canada. The 10 coyote samples were all from Los Angeles County. The SSCP patterns were consistent with seven alleles in the grey wolves (one of which, allele *Calu*-5, was in the Mexican wolf Aragon lineage and one of which, allele *Calu*-11, was in red wolves where it is designated as *Caru*-1), three alleles in the red wolves (allele *Caru*-1 also found in grey wolves), and eight alleles in coyotes, none of which were found in the other taxa. The observed and expected heterozygosities for the

Table 1 The observed frequencies of the 21 alleles at the *DRB1* locus over the four taxa (three lineages of Mexican wolves) where *N* is the sample size. The abbreviations for the alleles, *Calu*, *Caru*, and *Cala*, indicate the taxa *C. lupens*, *C. rufus*, and *C. latrans*, respectively

	Mexican wolf					
Allele	McBride (N = 17)	Ghost (N = 10)	Aragon (N = 8)	Grey wolf (N = 13)	Red wolf (N = 3)	Coyote (N = 10)
Calu-1	0.35	—	—	—	—	—
Calu-2	0.65	—	—	—	—	—
Calu-3	—	1.0	0.50	—	—	—
Calu-4	—	—	0.19	—	—	—
Calu-5	—	—	0.31	0.08	—	—
Calu-6	—	—	—	0.31	—	—
Calu-7	—	—	—	0.04	—	—
Calu-8	—	—	—	0.04	—	—
Calu-9	—	—	—	0.23	—	—
Calu-10	—	—	—	0.04	—	—
Calu-11 (*Caru*-1)	—	—	—	0.27	0.17	—
Caru-2	—	—	—	—	0.50	—
Caru-3	—	—	—	—	0.33	—
Cala-1	—	—	—	—	—	0.20
Cala-2	—	—	—	—	—	0.10
Cala-3	—	—	—	—	—	0.40
Cala-4	—	—	—	—	—	0.05
Cala-5	—	—	—	—	—	0.10
Cala-6	—	—	—	—	—	0.05
Cala-7	—	—	—	—	—	0.05
Cala-8	—	—	—	—	—	0.05

grey wolves, red wolves, and coyotes are given in Table 2. Interestingly, for the three groups of captive animals which were polymorphic, the McBride and Aragon lineages of Mexican wolves and the red wolves, the observed heterozygosity is higher than expected while for the wild caught animals, grey wolves and coyotes, the observed heterozygosity is lower than expected. Although none of these values is statistically significant, a Wahlund effect caused by combining several populations may reduce observed heterozygosity in the wild caught samples and the avoidance of inbreeding may increase observed heterozygosity in the captive samples.

Figure 1 presents a neighbour-joining tree with the 21 *DRB*1 sequences (there are actually 24 sequences in the figure because three are in more than one lineage or taxa). The sequences from any one taxa are generally widely dispersed in the tree. Furthermore, in almost all cases, a sequence from another taxa is most closely related to a given sequence. Table 3 presents the nucleotide and amino-acid sequence for the variable positions for the 21 alleles (the complete sequences have been deposited in GenBank, accessions AY009508–

AY009530). The order of the sequences in Table 2 from top to bottom is the same as in Fig. 1 for ease of comparison.

The five alleles from the Mexican wolf were in three different groups, allele *Calu*-1, alleles *Calu*-4 and *Calu*-5, and alleles *Calu*-2 and *Calu*-3, spread throughout the tree. Alleles *Calu*-4 and *Calu*-5 differed by only one amino acid (one nucleotide) at position 56. Both of these alleles were found in the Aragon lineage, but allele *Calu*-5 was also found in the grey wolf sample. Alleles *Calu*-2 and *Calu*-3 differed by three amino acids (six nucleotides) spread throughout the exon. *Calu*-2 was found only in the McBride lineage while *Calu*-3 was found in both the Ghost Ranch and Aragon lineages. The average number of amino acid differences between these groups was between 10.7 and 11.5 and the average number of nucleotide differences was between 16.8 and 17.3. For four positions, 26, 28, 30, and 37, there were three different amino acids present in the five Mexican wolf alleles.

The other most similar alleles, besides those discussed above, were *Cala*-1 and *Cala*-5, differing only by one amino acid (one nucleotide), both found in coyotes, and alleles *Calu*-1 and *Calu*-7, differing by one amino acid (one nucleotide) and found in Mexican and grey wolves, respectively. The most divergent alleles were *Calu*-1 and *Calu*-6, which differed by 13 amino acids (23 nucleotides), but there were many other pairs of alleles that diverged nearly as much.

Of the 69 amino acids, 19 (27.5%) were variable and of the 207 nucleotides sequenced, 34 (16.4%) were variable. At two amino acid positions, 28 and 37, there were five different amino acids present over the 21 sequences while there were four different amino acids observed at position 71. There were only two codon positions, positions 73 and 77 that had only silent variation. Interestingly, the silent difference from the consensus sequence at position 73 was found in only two sequences, *Calu*-11 (*Caru*-1) and *Cala*-6, which cluster together in the phylogenetic tree, and the silent difference at position 77 was found in five sequences, four of which, *Calu*-2, *Calu*-3, *Calu*-6, and *Caru*-3, cluster together.

The amino acid positions, documented using X-ray crystallography in humans to be important in the antigen-binding site (ABS) (Brown *et al.*, 1993), are indicated here by an asterisk in Table 2. Of the 13 ABS positions, 11 (84.6%) are variable over the 21 alleles, while for the remaining 56 positions that are not thought to interact with the antigen-binding site, only 10 (17.8%) are polymorphic and two of these are variable only for synonymous variation.

The three lineages of Mexican wolves are being merged upon recommendation of the Genetics
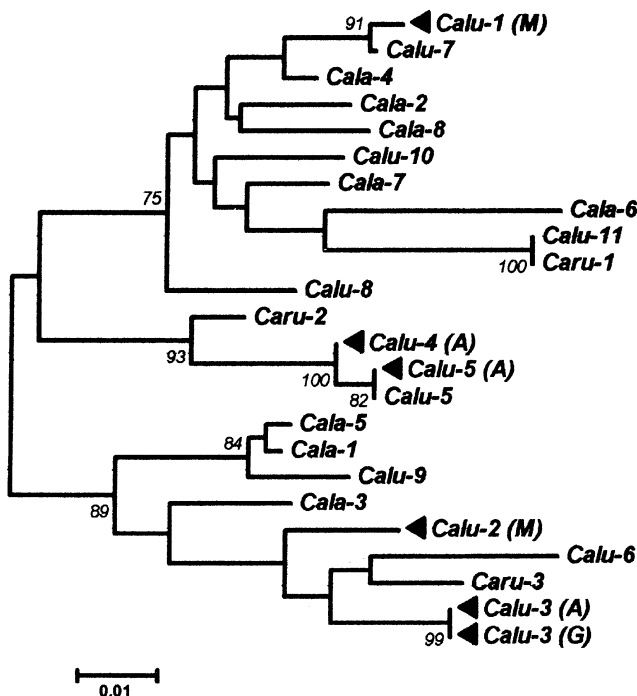

Fig. 1 A neighbour-joining tree giving the relationships for the 21 sequences found in this study. The triangles indicate the six different Mexican wolf sequences, the vertical lines indicate identical sequences found in different lineages of the Mexican wolf (M, McBride; G, Ghost Ranch; A, Aragon) or taxa, and the numbers indicate bootstrap significance values of 75 or greater.

Table 3 The nucleotide and amino acid sequence for the 21 variable positions for the 21 sequences found in Mexican, grey, and red wolves and coyotes

Position	26	27	28*	30*	32	37*	47	56	57*	60*	63	67*	70*	71*	73	74*	75	77	78*	84	86*
Consensus	ttc	ctg	gaa	tac	tat	aac	tac	ccc	gac	tcc	cgg	ctc	cgg	agg	gcc	gcg	gtg	acc	tac	ggg	ggc
	F	L	E	Y	Y	N	Y	P	D	S	R	L	R	R	A	A	V	T	Y	G	G
Calu-1	—	g—	—	—	—	t—	—	—	—	-a-	-c-	—	—	gc-	—	—	—	—	—	—	—
	–	V	–	–	–	Y	–	–	–	Y	P	–	–	A	–	–	–	–	–	–	–
Calu-7	—	g—	—	—	—	t—	—	—	—	—	-c-	—	—	cg-	—	—	—	—	—	—	—
	–	V	–	–	–	Y	–	–	–	–	P	–	–	A	–	–	–	–	–	–	–
Cala-4	—	g—	—	—	c—	—	—	—	—	—	—	—	—	gc-	—	—	—	—	—	—	—
	–	V	–	–	H	–	–	–	–	–	–	–	–	A	–	–	–	–	–	–	–
Cala-2	—	g—	—	—	c—	—	—	—	—	-a-	g—	—	—	-a-	—	-a-	—	—	—	—	—
	–	V	–	–	H	–	–	–	–	Y	G	–	–	K	–	E	–	–	–	–	–
Cala-8	—	g—	—	—	c—	—	—	—	—	—	—	—	—	-a-	—	—	—	—	—	—	att
	–	V	–	–	H	–	–	–	–	–	–	–	–	K	–	–	–	–	–	–	I
Calu-10	—	g—	—	—	c—	tt-	—	—	—	—	—	t—	-a-	—	—	-a-	—	—	—	—	—
	–	V	–	–	H	F	–	–	–	–	–	F	Q	–	–	E	–	–	–	–	–
Cala-7	—	g—	—	—	c—	tt-	—	—	—	—	g—	—	-a-	ga-	—	—	—	—	—	c—	—
	–	V	–	–	H	F	–	–	–	–	G	–	Q	E	–	–	–	–	–	R	–
Cala-6	—	g—	—	—	—	t—	—	—	—	—	—	a—	-a-	ga-	–a	a—	—	—g	gtg	—	—
	–	V	–	–	–	Y	–	–	–	–	–	I	Q	E	–	T	–	–	V	–	–
Calu-11	—	g—	—	—	c—	tt-	—	—	-t-	—	g—	a—	-a-	ga-	–a	a—	—	—	—	—	att
(Caru-1)	–	V	–	–	H	F	–	–	V	–	G	I	Q	E	–	T	–	–	–	–	I
Calu-2	–a	g—	—	—	c—	—	-t-	—	—	—	—	—	-a-	—	—	—	—	—	—	c—	—
	L	V	–	–	H	–	F	–	–	–	–	–	Q	–	–	–	–	–	–	R	–
Caru-2	–t	—	-cg	ag-	—	tt-	—	—	—	-a-	—	—	-a-	-a-	—	—	—	—	—	c—	—
	–	–	A	S	–	F	–	–	–	Y	–	–	Q	K	–	–	–	–	–	R	–
Calu-4	c-t	—	-cg	ag-	—	tt-	-t-	—	—	—	—	t—	-a-	—	—	—	—	—	—	c—	—
	L	–	A	S	–	F	F	–	–	–	–	F	Q	–	–	–	–	–	–	R	–
Calu-5	c-t	—	-cg	ag-	—	tt-	-t-	-g-	—	—	—	t—	-a-	—	—	—	—	—	—	c—	—
	L	–	A	S	–	F	F	G	–	–	–	F	Q	–	–	–	–	–	–	R	–
Cala-5	-at	—	-tg	g—	—	c—	-t-	—	—	-a-	g—	—	-a-	-a-	—	—	—	—	—	—	att
	Y	–	V	D	–	H	F	–	–	Y	G	–	Q	K	–	–	–	–	–	–	I
Cala-1	-at	—	-tg	g—	—	—	-t-	—	—	-a-	g—	—	-a-	-a-	—	—	—	—	—	—	att
	Y	–	V	D	–	–	F	–	–	Y	G	–	Q	K	–	–	–	–	–	–	I
Calu-9	-at	—	atg	g—	—	—	-t-	-g-	—	—	g—	—	-a-	-a-	—	—	—	—	—	—	att
	Y	–	M	D	–	–	F	R	–	–	G	–	Q	K	–	–	–	–	–	–	I
Cala-3	–t	—	ctg	g—	—	c—	—	—	—	-a-	—	—	—	—	—	-a-	—	—	—	—	att
	–	–	L	D	–	H	–	–	–	Y	–	–	–	–	–	E	–	–	–	–	I
Calu-2	-at	—	-tg	g—	—	c—	—	—	—	-a-	—	—	—	—	—	-a-	—	—g	gtg	c—	–
	Y	–	V	D	–	H	–	–	–	Y	–	–	–	–	–	E	–	–	V	R	–
Calu-6	-at	—	-tg	g—	—	ttg	-t-	—	—	-a-	g—	a—	—	—	—	-a-	c—	—g	gtg	—	att
	Y	–	V	D	–	L	F	–	–	Y	G	I	–	–	–	E	L	–	V	–	I
Caru-3	c-t	—	-tg	g—	—	c—	—	—	—	-a-	g—	—	—	—	—	-a-	—	—g	gtg	c—	att
	L	–	V	D	–	H	–	–	–	Y	G	–	–	–	–	E	–	–	V	R	I
Calu-3	-at	—	atg	g—	—	c—	-t-	—	—	-a-	g—	—	—	-a-	—	-a-	—	—g	gtg	—	att
	Y	–	M	D	–	H	F	–	–	Y	G	–	–	K	–	E	–	–	V	–	I

*indicates sites that are putative antigen binding site positions.

Committee of the Mexican Wolf recovery team (Hedrick et al., 1997). A tentative recommended goal was to combine the lineages in proportions 0.5 McBride, 0.25 Ghost Ranch, and 0.25 Aragon. Using these proportions and the frequencies of the sequences in the different lineages, we calculated the overall frequency of amino-acid heterozygosity for the examined 69 codons (Fig. 2). Here two amino-acid positions, 26 and 63, have heterozygosities over 0.6, and 10 positions in the ABS are variable, six of them with heterozygosities greater than 0.4.

In addition, the estimated rate of nonsynonymous (d_N) and synonymous (d_S) substitutions, for ABS and non-ABS amino-acid positions are given in Table 4. For
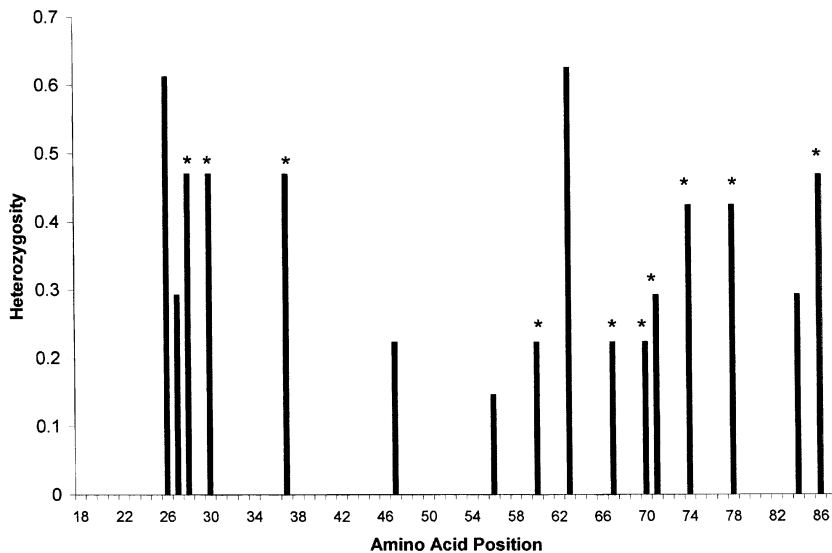
Fig. 2 The heterozygosity at individual amino-acid positions in the Mexican wolf where the McBride, Ghost Ranch, and Aragon lineages are combined in the proportions 0.5, 0.25, and 0.25, respectively. *indicates sites that are putative antigen binding site positions.

Table 4 The estimated rates of nonsynonymous and synonymous substitutions for antigen and nonantigen binding amino-acid positions and their ratio. N is the number of codons in each category and P is the probability that d_N and d_S are different

Positions	N	d_N	d_S	d_N/d_S	P
Antigen binding	13	0.374 ± 0.072	0.152 ± 0.086	2.46	< 0.025
Nonantigen binding	56	0.030 ± 0.009	0.026 ± 0.016	1.15	> 0.35
All	69	0.082 ± 0.012	0.046 ± 0.019	1.78	< 0.05

the antigen binding site positions. d_N (0.374) is significantly greater than d_S (0.152) and the ratio d_N/d_S is 2.46. For the nonantigen binding site positions, the difference is not significant and the ratio is only slightly greater than unity.

Discussion

The samples of Mexican wolves, founded from a total of only seven animals and with a history of inbreeding, have five different alleles for the *DRB1* locus. The Ghost Ranch lineage, consistent with the low level of genetic variation found for microsatellite loci (Garcia-Moreno *et al.*, 1996; Hedrick *et al.*, 1997) and high inbreeding coefficient from pedigree analysis (Hedrick *et al.*, 1997) was monomorphic for one allele, *Calu-3*. *Calu-3* was the only allele shared between Mexican wolf lineages. These findings support the recommendation by Hedrick *et al.* (1997) that combining the three Mexican wolf lineages (adding the Ghost Ranch and Aragon lineages to the McBride lineage) would increase genetic variation in the captive population.

The samples from grey and red wolves, and coyotes had 16 additional alleles. Even the small sample of three red wolves that we were able to examine had three quite different alleles. It has been suggested that red wolves

are a recent hybrid from grey wolves and coyotes (Wayne & Jenks, 1991; Roy *et al.*, 1994, 1996) because all the mtDNA haplotypes and microsatellite alleles found in red wolves were also found in either grey wolves or coyotes. In our dataset of MHC sequences, one of the red wolf alleles was found in the grey wolves and none were found in the coyote sample. In other words, two of the alleles were found only in red wolves and not in the two putative ancestral taxa. However, the sample sizes of these three taxa were so small that alleles may have been missed by chance. In addition, the source of the grey wolves (Canada) and coyotes (Los Angeles) may be sufficiently removed geographically from the location of the putative hybridization, south-eastern United States or Texas, that the exact alleles may differ.

Examination of the known dog, *C. famililaris*, MHC *DRB1* sequences in Genbank, showed that two alleles found in this study are identical over the 207 base pairs to dog alleles (Fig. 3). For Mexican wolves, *Calu-1*, found in the McBride lineage, is identical to dog *DRB1**00601 (top of Fig. 3). This is one of the more common alleles in dogs and was found in 10 different breeds at an overall frequency of 0.22 by Kennedy *et al.* (1998). Allele *Calu-11*, found in both grey and red wolves (*Caru-1*), is identical to *DRB1**00101 (middle of Fig. 3), another common dog allele that Kennedy *et al.*
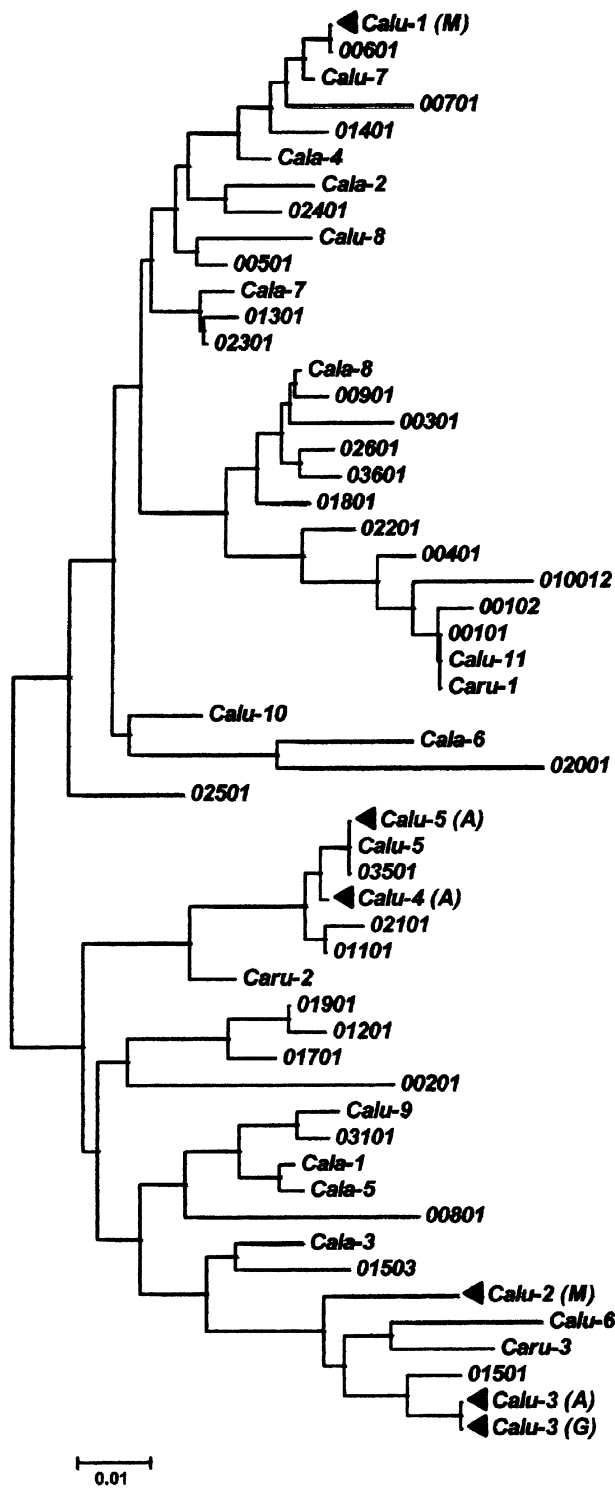
Fig. 3 A neighbour-joining tree giving the relationships for the sequences from this study and the available dog sequences indicated only by numbers. The triangles indicate the six different Mexican wolf sequences, the circles the three different red wolf sequences, and the vertical lines indicate identical sequences found in different lineages of the Mexican wolf (M, McBride; G, Ghost Ranch; A, Aragon) or taxa.

(1998) found in seven breeds with an overall frequency of 0.24. Four other alleles differed by only one amino acid (one nucleotide) from dog alleles. These are allele *Calu*-5, found in both the Aragon lineage and the grey wolf sample, and *DRB1**02101, allele *Calu*-7 in grey wolves and *DRB1**00601, allele *Cala*-8 in coyotes and *DRB1**00901, and allele *Cala*-7 in coyotes and *DRB1**02301. In addition, J. Angles (pers. comm.) allowed us to compare our sequences with three unpublished sequences that have been found in Alaskan grey wolves but have been given canid names under the dog major histocompatibility nomenclature (Kennedy *et al.*, 1999). These are *DRB1**03101, which differs by two base pairs from our coyote sequence *Calu*-9, *DBR1**03501, identical to our sequence for allele *Calu*-5, found in both Mexican and grey wolves, and *DRB1**03601, which differs by two base pairs from *Cala*-8, found in our coyote sample (see Fig. 3).

The intermingling of MHC alleles from different species within a phylogenetic tree from related taxa has been found in a number of organisms (see Edwards & Hedrick, 1998 for discussion). Often the closest sequence for MHC genes is from another species and not from another allele within the same species. Klein (1987) termed this phenomenon 'transpecies polymorphism', and suggested that the variability from an ancestral species is retained even when the species splits into two or more descendant species. The mechanism thought to be responsible for the retention of ancestral variation, is a form of balancing selection that results in alleles in one species having closest ancestry to alleles in another. The higher estimate of nonsynonymous than synonymous substitution, observed in the amino-acid positions involved in antigen binding, is consistent with this type of selection.

From the fossil record, the estimated time of common ancestry of the two most divergent taxa examined here, coyotes and grey wolves, is approximately one million years ago (Wayne *et al.*, 1991). Examining mtDNA sequence from the control region, Vila *et al.* (1997) found that all the coyote sequences clustered together and are clearly divergent from wolves and dogs. This contrasts with the pattern we found for MHC, in which all the coyote alleles were intermixed with wolf and dog alleles in the phylogenetic tree. In other words, one million years between wolves and coyotes does not appear to be sufficient time to have resulted in divergence into species-specific branches of the MHC tree.

The time of origin of dogs from wolves, based on the archeological record, is approximately 14 000 years before present (Olsen, 1985), but based on an estimate from mtDNA is more than 100 000 years before present (Vila *et al.*, 1997). As mentioned above, although the red wolf is considered a separate species, it appears to be

of more recent origin than the divergence of coyotes and wolves. In the phylogenetic tree in Fig. 1, all three of the major branches contain alleles for the grey, Mexican, and red wolves. The known sequences from dogs in Fig. 3 are spread throughout the phylogenetic tree and are also included in the three major branches. This is consistent with a broad origin from wolves, as suggested from the mtDNA data (Vila *et al.*, 1997), retention of the ancestral wolf polymorphism by selection, and, furthermore, may reflect continued gene flow from wolves into dog populations until fairly recent times.

Acknowledgements

References

BROWN, J. H., JARDETZKY, T. S., GORGA, J. C., STERN, L. J., URBAN, R. G., STROMINGER, J. L. AND WILEY, D. C. 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature*, **364**, 33–39.

CARRINGTON, M., NELSON, G. W., MARTIN, M. P., KISSNER, T., VLAHOV, D., GOEDERT, J. J. ET AL. 1999. *HLA* and HIV-1: Heterozygote advantage and *B*35-Cw*04* disadvantage. *Science*, **238**, 1748–1752.

EDWARDS, S. AND HEDRICK, P. W. 1998. Evolution and ecology of MHC molecules: from genomics to sexual selection. *Trends Ecol. Evol.*, **13**, 305–311.

GARCIA-MORENO, J., ROY, M. S., GEFFEN, E. AND WAYNE, R. K. 1996. Relationships and genetic purity of the endangered Mexican wolf based on analysis of microsatellite loci. *Conserv. Biol.*, **10**, 376–387.

HEDRICK, P. W. AND KIM, T. J. 2000. Genetics of complex polymorphisms: Parasites and maintenance of the major histocompatibility complex variation. In: Singh, R. S. and Krimbas, C. B. (eds) *Evolutionary Genetics: from Molecules to Morphology*, pp. 204–234. Cambridge University Press, Cambridge.

HEDRICK, P., WHITTAM, T. S. AND PARHAM, P. 1991. Heterozygosity at individual amino acid sites: extremely high levels for HLA-A and B genes. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 5897–5901.

HEDRICK, P. W., MILLER, P. S., GEFFEN, E. AND WAYNE, R. 1997. Genetic evaluation of the three captive Mexican wolf lineages. *Zool. Biol.*, **16**, 47–69.

HILL, A. V. S. 1998. The immunogenetics of human infectious diseases. *Ann. Rev. Immunol.*, **16**, 593–617.

HILL, A. V. S., ALLSOP, C. E. M., KWIATDOWSKI, D., ANTSEY, N. M., TWUMASI, P., ROWE, P. A. ET AL. 1991. Common West African HLA antigens are associated with protection from severe malaria. *Nature*, **352**, 595–600.

JUKES, T. H. AND CANTOR, C. R. 1969. Evolution of protein molecules. In: Munro, H. N. (ed.) *Mammalian Protein Metabolism*, pp. 21–132. Academic Press, New York.

KENNEDY, L. J., CARTER, S. D., BARNES, A., BELL, S., BENNETT, D., OLLIER, W. E. R. ET AL. 1998. Nine new dog *DLA-DRB1* alleles identified by sequence-based typing. *Immunogenetics*, **48**, 296–301.

KENNEDY, L. J., ALTET, L., ANGLES, J. M., BARNES, A., CARTER, S. D., FRANCINO, O. ET AL. 1999. Nomenclature for factors of the dog major histocompatibility system (DLA), 1998. First report of the ISAG DLA nomenclature committee. *Tissue Antigens*, **54**, 312–321.

KLEIN, J. 1987. Origin of major histocompatibility complex polymorphism: the trans-species hypothesis. *Hum. Immunol.*, **19**, 155–162.

KUMAR, S., TAMURA, K. AND NEI, M. 1993. MEGA: Molecular Evolutionary Genetics Analysis, Version 1.01. Pennsylvania State University, University Park, PA.

LAURENSON, K., SILLERO-ZUBIRI, C., THOMPSON, H., SHIFERAW, F., THIRGOOD, S. AND MALCOM, L. J. 1998. Disease as a threat to endangered species: Ethiopian wolves, domestic dogs and canine pathogens. *Anim. Cons.*, **1**, 273–280.

LYLES, A. M. AND DOBSON, A. P. 1993. Infectious disease and intensive management: population dynamics, threatened hosts, and their parasites. *J. Wild. Med.*, **24**, 315–326.

MURRAY, D. L., KAPKE, C. A., EVERMANN, J. F. AND FULLER, T. K. 1999. Infectious disease and the conservation of free-ranging large carnivores. *Anim. Cons.*, **2**, 241–254.

NEI, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.

NEI, M. AND GOJOBORI, T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.*, **3**, 418–426.

O'BRIEN, S. J. AND EVERMANN, J. F. 1988. Interactive influence of infectious disease and genetic diversity in natural populations. *Trends Ecol. Evol.*, **3**, 254–259.

OLSEN, S. J. 1985. *Origins of the Domestic Dog*. University of Arizona, Tucson, AZ.

ROY, M. S., GEFFEN, E., SMITH, D., OSTRANDER, E. A. AND WAYNE, R. K. 1994. Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci. *Mol. Biol. Evol.*, **11**, 553–570.

ROY, M. S., GEFFEN, E., SMITH, D. AND WAYNE, R. K. 1996. Molecular genetics of pre-1940 red wolves. *Conserv. Biol.*, **10**, 1413–1424.

THURZ, M. R., THOMAS, H. C., GRENNWOOD, B. M. AND HILL, A. V. S. 1997. Heterozygote advantage for HLA class-II type in hepatitis B virus infection. *Nat. Genet.*, **17**, 11–12.

VILA, C., SAVOLAINEN, P., MALDONADO, J. E., AMORIM, I. R., RICE, J. E., HONEYCUTT, R. L. ET AL. 1997. Multiple and ancient origins of the domestic dog. *Science*, **276**, 1687–1689.

WAYNE, R. K. AND JENKS, S. M. 1991. Mitochondrial DNA analysis implying extensive hybridization of the endangered red wolf *Canis rufus*. *Nature*, **351**, 565–568.

WAYNE, R. K., VAN VALDENBURGH, B. AND O'BRIEN, S. J. 1991. Molecular distance and divergence time in carnivores and primates. *Mol. Biol. Evol.*, **8**, 297–319.