DNA variation of the mammalian major histocompatibility complex reflects genomic diversity and population history

(class I genes/restriction fragment length polymorphism/felids)

NAOYA YUHKI AND STEPHEN J. O'BRIEN
Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick, MD 21701-1013

Communicated by Bruce Wallace, October 30, 1989

ABSTRACT The major histocompatibility complex (MHC) is a multigene complex of tightly linked homologous genes that encode cell surface antigens that play a key role in immune regulation and response to foreign antigens. In most species, MHC gene products display extreme antigenic polymorphism, and their variability has been interpreted to reflect an adaptive strategy for accommodating rapidly evolving infectious agents that periodically afflict natural populations. Determination of the extent of MHC variation has been limited to populations in which skin grafting is feasible or for which serological reagents have been developed. We present here a quantitative analysis of restriction fragment length polymorphism of MHC class I genes in several mammalian species (cats, rodents, humans) known to have very different levels of genetic diversity based on functional MHC assays and on allozyme surveys. When homologous class I probes were employed, a notable concordance was observed between the extent of MHC restriction fragment variation and functional MHC variation detected by skin grafts or genome-wide diversity estimated by allozyme screens. These results confirm the genetically depauperate character of the African cheetah, Acinonyx jubatus, and the Asiatic lion, Panthera leo persica; further, they support the use of class I MHC molecular reagents in estimating the extent and character of genetic diversity in natural populations.

The amount of genetic variability detected in nearly 1000 natural populations has been studied with biochemical methods since the introduction of allozyme electrophoresis over two decades ago (1, 2). Although most populations have appreciable variability, there are certain exceptional populations that appear relatively monomorphic for a variety of reasons. A common explanation for genetic uniformity is a history of inbreeding, which can eliminate normal genetic variation at a rapid rate. In several observed cases, inbreeding also causes deleterious physiological effects termed inbreeding depression, which include increased juvenile mortality (3–5), morphological asymmetry (6, 7), reproductive impairments (3, 8), and increased population susceptibility to pathogens (9–11).

The major histocompatibility complex (MHC) is an important genetic system that is considered to be critical to the acquisition of immune defenses (12–17). The MHC gene products are encoded by two distinct classes (I and II) of extremely polymorphic loci. Over 70 different alleles have been described for the human HLA class I loci and over 100 class I alleles have been described for the murine H-2 locus (12). These antigens, in different combinations, function in T-cell recognition of cells infected with invading viruses (11–17). Since many etiologic agents have the capacity to evolve phenotypes that abrogate the MHC-T-cell interaction process (10, 18, 19), it has been hypothesized that the extreme genetic diversity of the MHC is adaptive in providing a variety of heterogeneous host targets for these agents in natural populations (11–19). Because of the critical role of the MHC loci in immune recognition and defense, it would be useful to examine polymorphism of these loci in free-ranging mammals. We present here an analysis of DNA variation in class I MHC loci of three feline species (domestic cat, African cheetah, and lion) previously characterized by allozyme screens and in two cases by transplantation analysis. We compare these observations to similar estimates of MHC variation in mouse, human, and other populations of mammals.

MATERIALS AND METHODS
DNA Samples and Southern Blot Analysis. Samples were obtained from six inbred mouse strains, from 54 human patients, and from 16 domestic cats. Two populations of African cheetah, Acinonyx jubatus jubatus (South African) and A. jubatus raineyi (East African), and three free-ranging populations of lions (Panthera leo), Serengeti lions, Ngorongoro Crater lions, and Asiatic lions, were also examined. DNA samples were extracted from peripheral blood leukocytes, from frozen tissues (liver, kidney, spleen), and from fibroblast cultures established from skin biopsies by the method described elsewhere (20). For the blood samples, erythrocytes were lysed with 2 volumes of ACK lysing buffer (Quality Biologicals, Gaithersburg, MD): 8.023 g of NH₄Cl, 1 g of KHCO₃, and 0.037 g of EDTA per liter were added and incubated at 4°C for 10 min. Leukocytes were collected by centrifugation. For the tissue samples, samples were frozen in liquid nitrogen and pulverized with a pestle. Each sample was suspended in 10 mM Tris, pH 7.4/10 mM EDTA/150 mM NaCl. After SDS was added in 1% final concentration, protease K (200 μg/ml), Boehringer Mannheim) and RNase A (200 μg/ml; Sigma) were added and incubated for 2 hr at 55°C. The samples were extracted repeatedly with Tris (pH 7.4)-saturated phenol/chloroform, 1:1, until no interface material was visible and then were extracted once with chloroform. After addition of 2 volumes of 100% ethanol, DNAs were collected by centrifugation. Samples (10 μg) of these DNAs were digested with four or five appropriate restriction enzymes, electrophoresed through 0.8–1.0% agarose gels, and transferred to nitrocellulose filters (Schleicher & Schuell). After prehybridization of the filters in 1 M NaCl/50% formamide/50 mM Pipes, pH 7.0/0.02% salmon testis DNA (200 μg/ml)/0.1% Ficoll/0.1% polyvinylpyrrolidone/0.1% bovine serum albumin/0.2% sarcosyl/1 mM EDTA for 10–12 hr at 42°C, 32P-labeled MHC class I cDNA clones were added to the solution and incubated for another 10–12 hr at 42°C. The following MHC class I gene molecular cDNA clones were used: 5'DP (21) for mouse, HLA8T (22) for human, and pFLA24 (23) for domestic cat, African cheetahs.

Abbreviations: MHC, major histocompatibility complex; RFLP, restriction fragment length polymorphism; APD, average percent difference; MAPD, mean APD; P, fraction of polymorphic loci; H, average heterozygosity.
and three populations of lions. Heterologous probes can be used for such analyses (see ref. 49); however, resolution of DNA fragments is increased appreciably when homologous probes from more closely related species are employed (see Fig. 1). Membranes were washed first in 2 × SSC/0.1% SDS for 30 min at 37°C and then washed in either 0.1 × SSC/0.1% SDS (for mouse, human, and domestic cat) or 1 × SSC/0.1% SDS (for cheetahs and lions) for 30 min at 50°C. (SSC is 0.15 M NaCl/0.015 M sodium citrate, pH 7.) XRP (Kodak) films were used for autoradiography.

Quantitative Estimation of Genetic Variation of MHC Class I Genes. Estimates of three parameters, mean average percent of difference (MAPD), frequency of polymorphic loci (P), and average heterozygosity (H) were derived based on the results of Southern blot analyses using homologous class I cDNA probes in various species (24). These parameters were calculated in the following way.

Mean average percent difference (MAPD).

Percent difference (PD) = \( \frac{V_{ab}}{F_a + F_b} \times 100 \),

where \( V_{ab} \) is the total number of variable fragments between two individuals for a single restriction enzyme, \( F_a \) is the number of fragments resolved in individual A with that enzyme, and \( F_b \) is the number of fragments resolved in individual B with that enzyme.

Average percent difference (APD) = \( \frac{1}{C} \sum_{i=1}^{C} PD_i \),

where \( C \) is the number of pairwise comparisons in a population.

MAPD = \( \frac{1}{R} \sum_{i=1}^{R} \frac{C}{\sum_{i=1}^{C} PD_i} \),

where \( R \) is the number of restriction enzymes used.

The following estimates of population genetic parameters (P and H) are based upon the simplifying assumption that each restriction fragment represents a distinct class I gene or an allele of a class I gene.

Fraction of loci that are polymorphic, P.

\[ P = \frac{L_p}{L_m + L_p} \]

where \( L_m \) is the number of monomorphic loci (estimated as the number of restriction fragments that are invariant in the test population) and \( L_p \) is the number of polymorphic loci; we estimate

\[ L_p = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} V_i \]

where \( V_i \) is the number of variable bands for the \( i \)th individual and \( N \) is the number of individuals in the population.

Average heterozygosity (H) for the populations. Under the assumption of Hardy-Weinberg equilibrium and independence of allelic (fragment) association in the population, \( H \) is derived from the knowledge of \( s \), the frequency of individuals that express a polymorphic fragment in a population. Therefore,

\[ s = p^2 + 2pq = 1 - q^2 \]

where \( p \) is the allele frequency of a polymorphic fragment and \( q \) is the allele frequency or sum of allele frequencies of alternative alleles at a particular class I locus; therefore,

\[ q = (1 - s)^{1/2} \]

and the heterozygosity at one locus,

\[ h = 2pq = 2(1 - s)^{1/2}[1 - (1 - s)^{1/2}] \]

Thus,

\[ H = \frac{1}{L_T} \sum_{i=1}^{L_T} h \]

where \( F_T \) is the total number of fragments resolved in a population screen and \( L_T \) is the number of loci screened (\( = L_p + L_m \)).

RESULTS AND DISCUSSION

We first examined three groups of species that had been reported as having different levels of MHC polymorphism. The first species group included mouse and human, species with abundant functional MHC variability (12). The second group comprised 16 random-source domestic cats (Felis catus), a species that displays antigenic diversity at the MHC, termed FLA, but less than that observed in mouse or man (25). The third group included two populations of African cheetah subspecies, A. jubatus jubatus (South African) and A. jubatus raineyi (East African). A previous study of this species revealed a relative genetic uniformity based on allozymes, two-dimensional gel electrophoresis, and increased morphological asymmetry (3, 26, 27). Furthermore, reciprocal skin grafts between 12 unrelated and 2 sibling cheetahs all failed to be rejected acutely, suggesting functional allelic identity at the cheetah’s MHC (3). Apparent physiological consequences of the genetic uniformity include a high degree of spermatocozoal abnormality, reproductive impairment, and increased juvenile mortality compared to other captive-bred species. In addition, several severe epizootics of a normally low-mortality domestic cat virus (feline infectious peritonitis virus) revealed that the species had an extreme and nearly homogeneous vulnerability to this agent (3, 28). O’Brien et al. (27) have suggested an extreme demographic contraction, or population bottleneck, followed by inbreeding had occurred in the natural history of this species, perhaps toward the end of the late Pleistocene (ca. 10,000 years ago).

The results of a Southern analysis of each tested population are illustrated in Fig. 1 and tabulated in Table 1. The extent of restriction fragment length polymorphism (RFLP) was estimated by three quantitative parameters in each population: (i) MAPD, defined as the average percentage of DNA fragments that differ between individuals for multiple restriction enzymes; (ii) \( P \), an estimate of the percent of detected class I homologous loci that display restriction fragment variation in the population; and (iii) \( H \), defined as the fraction of all restriction site loci that are heterozygous over all individuals in the population. Algebraic derivation of these parameters is described in Materials and Methods (see also ref. 24).

The results of the species comparisons plus comparable RFLP estimates from other species whose MHCs have been studied immunologically are presented in Table 1. Two general conclusions are evident: First, the two species that do not reject allogeneic skin grafts, cheetah (3) and Syrian hamster (39), have the lowest amount of detected DNA variation in each family and genera. Conversely, species that have abundant MHC antigenic variability (mouse, pig, rat, and mole rat) show appreciable DNA variation. The domestic cat, which has an intermediate level of functional MHC polymorphism (25), ranks with humans as an intermediate in the extent of MHC RFLP. Second, there is also agreement between the results of allozyme surveys and the MHC diversity measured as RFLP. The two cheetah subspecies have been shown to have 10- to 100-fold less overall genetic.
diversity than 10 species of felids or other mammals similarly typed with 50 allozyme markers (26, 27). The relative genetic uniformity of the cheetah is also apparent in the low degree of RFLP detected at MHC class I loci (Table 1). Other species listed in Table 1 all show appreciable allozyme and MHC variation. The exception to this trend is the Syrian hamster, which is monomorphic at the MHC but rather polymorphic at allozyme loci (39, 40). The explanation for this discordance is not obvious but may imply specific selective or adaptive events relating to the MHC locus in the natural history of the species.

In an attempt to measure MHC-associated DNA variation within a single species, we examined three free-ranging populations of lions (P. leo), each with very different levels of allozyme-based genetic diversity and correlative reproductive–endocrine consequences (41, 42). The Serengeti lions constitute a large group of about 3000 lions living in the Serengeti plains in Tanzania. The Ngorongoro Crater population is a small group of about 100 lions living adjacent to, but geographically isolated from, the larger Serengeti population. Occasionally, resident lions emigrate, but no immigration into the crater has been recorded since 1975. The Ngorongoro

### Table 1. DNA variation in MHC class I genes in different species with varying extent of functional MHC variability

<table>
<thead>
<tr>
<th>Species</th>
<th>APD</th>
<th>RFLP</th>
<th>Allozyme variation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Species</strong></td>
<td><strong>MAPD</strong></td>
<td><strong>P, %</strong></td>
<td><strong>H, %</strong></td>
</tr>
<tr>
<td><strong>Mus musculus</strong></td>
<td>6</td>
<td>32.4</td>
<td>40.5</td>
</tr>
<tr>
<td>(mouse)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rattus norvegicus</strong></td>
<td>8</td>
<td>—</td>
<td>42.5</td>
</tr>
<tr>
<td>(rat)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sus scrofa</strong></td>
<td>13</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>(pig)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Spalax ehrenbergii</strong></td>
<td>22</td>
<td>16.26</td>
<td>30.7</td>
</tr>
<tr>
<td>(mole rat)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Felis catus</strong></td>
<td>16</td>
<td>8.3</td>
<td>10.7</td>
</tr>
<tr>
<td>(cat)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Homo sapiens</strong></td>
<td>54</td>
<td>8.1</td>
<td>13.2</td>
</tr>
<tr>
<td>(human)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mesocricetus auratus</strong></td>
<td>7</td>
<td>9.0</td>
<td>7.2</td>
</tr>
<tr>
<td>(Syrian hamster)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Acinonyx jubatus</strong></td>
<td>13</td>
<td>0.0</td>
<td>2.3</td>
</tr>
<tr>
<td>(raineyi)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(East African cheetah)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Acinonyx jubatus</strong></td>
<td>9</td>
<td>0.0</td>
<td>1.9</td>
</tr>
<tr>
<td>(South African cheetah)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Number of individuals analyzed.

*Species in which RFLP measurements were made between inbred strains, precluding an estimate of \( P \) and \( H \).
Crater lions are descended from a documented population bottleneck in 1962, when the number of lions dropped to 10–15 as the result of an epizootic of biting flies (Stomoxys calcitrans) (43). The third sample was from the relict population of Asiatic lions (P. leo persica), which reside in the Gir Forest Sanctuary in the Gujarat State in western India. The Asiatic lions today number about 250 animals, but they had suffered a demographic contraction to less than 20 animals at the turn of the century due to over-hunting (44). These three lion populations differ dramatically in their genetic diversity (Table 2) and in reproductive traits (41, 42). The Serengeti lions have abundant allozyme genetic diversity (comparable to domestic cat, mouse, and man) and a low amount of abnormal sperm (average 25%). The Asiatic lions exhibit no allozyme variants (out of 50 loci typed), have high levels of abnormal sperm (average 66%), and have severely diminished circulating testosterone levels. The moderately bottlenecked Ngorongoro lion population fell in between these two extremes in each measurement (41).

Estimates of MHC class I DNA variation for these three populations of lions are illustrated in Fig. 2 and summarized in Table 2. The extent of MHC variation shows a striking correlation (r = 0.98, MAPD vs. allozyme H) to the level of detectable allozyme variation and the associated reproductive consequences. The Asiatic lions were unique in our study in revealing no RFLP whatsoever. In the Asiatic lion samples, we monitored a total of 73 MHC fragments by using five restriction enzymes with 6-nucleotide recognition sites, for a minimum of 438 invariant nucleotides scored in each of 15 lions.

The present results affirm the utility of MHC RFLP screens as a measure of overall genetic diversity of populations as well as providing a quantitative correlate to functional variation at the MHC locus itself. Although this calibration is rough because of the complexity of the class I gene family, our results suggest MHC RFLP screening is applicable at least to comparisons among species of the same family or genera. This latter aspect would be particularly useful for monitoring species where skin grafting is impractical or for which serological reagents are unavailable. Since this is the case for the majority of endangered species, the DNA method may be a very practical means to evaluate genomic and MHC diversity in heretofore unstudied species and populations. By way of example, O’Brien and Evermann (10) have reviewed a series of 10 mammalian species that have experienced demographic contractions in their recent history. Several of these have been affected by post-bottleneck epizootics whose course may have been influenced by the genetic structure of the population. Knowledge of the relative level of MHC diversity would be a useful datum in the development of management strategies for these populations.

The maintenance of functional antigenic MHC diversity may be important for species survival not only to provide genetic heterogeneity against parasite adaptation in an immediate sense but also in the long-term prognosis. Hughes and Nei (45) have presented compelling evidence that allelic variation at class I MHC alleles can only be explained by postulating over-dominance, or heterozygote advantage, and not by increased rates of mutation or genomic strategies to increase endemic genetic diversity. Further, two recent studies of comparative MHC allelism in distantly related species showed that human and murine allelic variation is largely ancient, having been inherited from ancestors of these groups that lived as long as 10 million years ago (46–48). These results eliminate the necessity for proposals or models assuming high mutation rate, gene conversion, inter se recombination, etc., to explain the abundant variation observed at the MHC of most species. In the absence of such mechanisms,

Table 2. RFLP detected in lions with the feline MHC class I probe

<table>
<thead>
<tr>
<th>Population</th>
<th>APD</th>
<th>RFLP</th>
<th>Allozyme variation*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAPD</td>
<td>P, %</td>
<td>H, %</td>
</tr>
<tr>
<td>Serengeti</td>
<td>18</td>
<td>7.8</td>
<td>9.7</td>
</tr>
<tr>
<td>Ngorongoro crater</td>
<td>15</td>
<td>7.3</td>
<td>9.1</td>
</tr>
<tr>
<td>Gir Forest</td>
<td>15</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

*Number of individuals analyzed.

†Allozyme results from ref. 42.

Fig. 2. Southern blot analysis of MHC class I genes in three lion populations. DNA (10 μg per lane) was digested with Pst I (20 units). Arrows indicate polymorphic restriction fragments. kb, Kilobases.
for rapid generation of MHC diversity, the consequence of genetic uniformity at the MHC would be long-lasting, possibly for millions of years. Such an extended time of species vulnerability to disease outbreaks would certainly increase the chance of extinction in such a population's future.

We are grateful to Mary Eichelberger and Janice Martenson for excellent technical assistance and to Drs. David Derse, Gisela Heidecker, Peter Dratch, Cheryl Winkler, Dean Mann, and Bruce Weir for critical reading of early versions of the manuscript. We thank Drs. David Wildt, Michell Bush, Craig Packer, and Tim Caro for assistance in obtaining samples from free-ranging lion and cheetah populations. Blood and tissue culture specimens of nondomestic animals were collected in full compliance with specific permits (CITES; Endangered and Threatened Species; Captive Breeding) issued to the National Cancer Institute, National Institutes of Health, principal officer S.J.O., by the U.S. Fish and Wildlife Service of the Department of the Interior.