Comparative Medicine
Copyright 2005
by the American Association for Laboratory Animal Science

Vol 55, No 3
June 2005
Pages 227-230

Genetic Characterization of Indian–Origin and Chinese–Origin Rhesus Macaques (*Macaca mulatta*)

David Glenn Smith, PhD

Genetic differences between Indian-origin and Chinese-origin rhesus macaques are as great as those between some primate species and can influence the results of experiments in which both are used as animal models for the study of the same human diseases. Unfortunately, many breeding facilities do not know with certainty the origin of the founders of their rhesus breeding colonies. Here I summarize the most definitive of the genetic traits among the microsatellite (STR) loci and mitochondrial DNA sequences that my laboratory previously reported to characterize Indian-origin and Chinese-origin rhesus macaques and then estimate the frequencies of these traits and their reliability as indicators of country of origin. The expression of diagnostic traits at two or more of four different unlinked loci provides a nearly 100% reliability in distinguishing rhesus macaques of Indian and Chinese origin.

Rhesus macaques comprise the most frequently used models for the study of human diseases (22). Whereas many such models were developed in Indian-origin rhesus macaques, Chinese-origin rhesus macaques have been used with increasing frequency in studies of the same diseases after the foreign supply of Indian-origin rhesus macaques was discontinued (3, 5), because the demand for rhesus macaques for use in biomedical research exceeded the supply provided by domestic breeding programs. Presently, rhesus macaques from these two countries of origin comprise virtually all of the rhesus macaques bred domestically for biomedical research in the United States.

Recent studies of protein coding (17), microsatellite markers (STRs; 7, 13, 18, 19), class I (2) and class II (2, 23) major histocompatibility (MHC) loci, and mitochondrial DNA (mtDNA; 8, 12, 20) have shown that Indian and Chinese rhesus macaques exhibit noteworthy genetic differences. As a result, some of the differences between Indian and Chinese rhesus macaques in phenotypic responses to experimental protocols might solely be due to genetic differences between subjects from these different countries. For example, the distributions of alleles at most MHC loci, which influence immune responses, are restricted to either Indian or Chinese rhesus macaques, and Chinese rhesus macaques exhibit many more alleles per locus than do Indian rhesus macaques, reflecting a much higher level of genetic heterogeneity in Chinese than in Indian rhesus macaques (2, 23). Chinese and Indian rhesus macaques also display differences in their responses to experimental infection with simian immunodeficiency virus (6, 10) and other infectious agents; again, genetic differences at MHC loci might be responsible for these difference.

Because breeding facilities sometimes do not know the origin of the rhesus macaques comprising their founder breeding stock, the domestic supply of rhesus macaques increasingly will experience admixture between Indian and Chinese rhesus macaques. If the population used as experimental subjects is found to influence the results of studies of some human diseases, it will be crucial to genetically characterize individual rhesus macaques used in these studies to identify their country of origin. Although previous studies, such as the cited MHC studies, have identified frequencies of alleles segregating at functional polymorphic loci that differ markedly between Indian and Chinese rhesus macaques, such as those for the serum protein albumin (16) and the erythrocyte enzyme isocitrate dehydrogenase (17), loci that evolve more rapidly than coding regions, such as microsatellite (also called short tandem repeat [STR]) loci (18) and the first hypervariable segment (HVSI) of the mtDNA genome (20), are much more effective in differentiating among closely related populations of the same species. Moreover, because these loci are predominantly selectively neutral, they provide more valid evidence of the evolutionary history of and phylogenetic relationships among taxa than do functional loci. I provide here a genetic definition of the two regional varieties of rhesus macaques most frequently used in biomedical research in light of studies of a geographically representative sample of Indian and Chinese rhesus macaques.

Materials and Methods

DNA was isolated from the stored buffy coats, serum, or DNA samples of a total of 132 rhesus macaques alleged to be solely of Indian ancestry and 168 rhesus macaques alleged to be solely of Chinese ancestry. These samples were chosen to provide a study sample that was as geographically representative as possible across the entire geographic range of rhesus macaques. Figure 1 identifies the locations in India where the founders of the populations from which the Indian samples used in this study originated and the locations of the breeding centers in China from which the Chinese rhesus used were acquired. The rhesus macaques providing samples for this study all are managed in accordance with the United States Public Health Service *Policy on the Humane Care and Use of Laboratory Animals* (14). DNA was extracted from EDTA blood samples using a Qiagen QIAmp Blood Mini Kit (Valencia, Calif.). An approximately 835-bp fragment flanked by nts 15167 and 16050 in the homologous Bar-

Received: 11/24/04. Revision requested: 12/30/04. Accepted: 1/14/05.
Molecular Anthropology Laboratory, Department of Anthropology and California National Primate Research Center, University of California, Davis, California 95616.

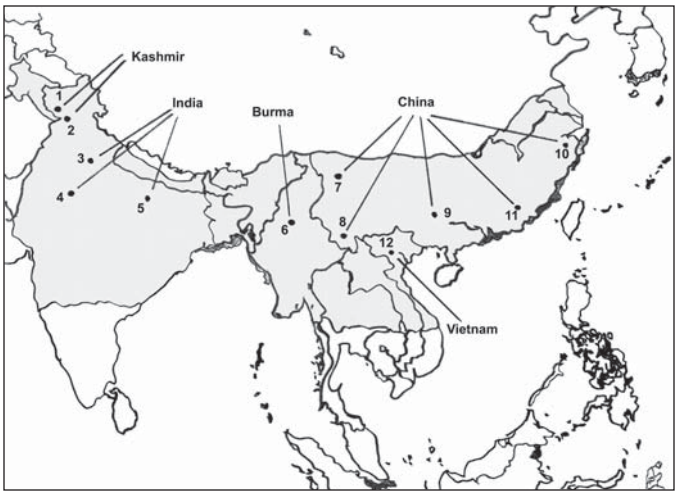

Figure 1. The geographic range of *M. mulatta* (shaded area) is depicted, and the origin locations of founders of Indian rhesus populations sampled for this study and location of breeding centers in China from which Chinese rhesus samples used in this study were acquired are identified. The westernmost (locations 1 and 2) and easternmost (location 5) sites in India are near the cities of Jammu and Lucknow, respectively. The breeding centers in China depicted as sites 7 through 11 are in Sichuan, Yunnan, Guangxi, Guangdong, and Jiangsu provinces, respectively.

bary ape (*Macaca sylvanus*) mtDNA sequence (1) was amplified using the primers and methods previously described (20). The sequences of Indian and Chinese rhesus macaques were aligned with Sequencher (Gene Codes Corporation, Ann Arbor, Mich.) and compared to identify mutations that consistently distinguished between the two populations. Approximately two-thirds (190) of these 300 mtDNA sequences were reported in a previous study (20; accession numbers AY646930 through AY647139 in GenBank). The additional samples were evaluated to provide a larger sample of sequences for the present population that better represents the entire geographic range of Indian and Chinese rhesus (Fig. 1).

In addition, my laboratory has studied several hundred tetranucleotide microsatellite loci to identify those most useful for genetic management of captive colonies of rhesus macaques. One of the criteria for selecting these loci was their ability to differentiate between Indian-origin and Chinese-origin rhesus macaques. We selected 24 loci for further consideration and used them to screen a sample of 186 Indian rhesus macaques and 206 Chinese rhesus macaques (18). This sample included most of the 300 animals whose mtDNAs were previously studied. For the present study, I compared allele frequencies in Indian and Chinese rhesus macaques for each of the 24 STR loci, which were previously reported (18), were compared to identify particular alleles that were very common in rhesus macaques from one of the two countries but rare or absent in those from the other country. The effectiveness of each locus individually, including the mtDNA sequence, and of multiple combinations of these loci in distinguishing between Indian and Chinese rhesus macaques was evaluated.

Results

The 300 mtDNA sequences revealed multiple polymorphic sites that uniquely characterize Indian and Chinese rhesus macaques. These sequences represent two (Ind1 and Ind2) clusters of haplotypes, or "haplogroups" (20), that are almost exclusively restricted to Indian-origin rhesus macaques and four (ChiE, ChiW1, ChiW2, and ChiW3) haplogroups restricted to Chinese-origin rhesus macaques. All mutations previously reported to be diagnostic of country of origin (20) were confirmed using the additional 110 sequences included in the present study. When the *M. sylvanus* sequence (1) is used as a reference sequence, Indian haplogroup Ind1 is uniquely defined by A>G transitions at nts 15215, 15246, 15350, 15377, 15722, and 15723; a C>T transition at nt 15535; a T>C transition at nt 15639; and an A>T transversion at nt 15765. The remaining Indian haplotypes, all members of haplogroup Ind2, are defined by a C>T transition at nt 15845, which sometimes also is found in members of Indian haplogroup Ind1 but which never occurs in Chinese rhesus macaques. Therefore, Indian rhesus macaques are defined by the presence of the described suite of nine mutations or a C>T mutation at nt 15845 or both. Members of the Chinese haplogroup ChiE, representing approximately half of the Chinese rhesus macaques studied, are uniquely defined by an A>G transition at nt 15400. Members of Chinese haplogroup ChiW1 have a unique C>T transition at nt 15959; those of ChiW2 show an A>G transition at nt 15329 (in addition to a C>T transition at nt 15296); and members of ChiW3 have an A>G transition at nt 15201. Thus, Chinese rhesus macaques are genetically defined by their possession of one of the aforementioned four mutations. In a study of more than 1000 rhesus macaques, approximately equal numbers of which were Indian-origin and Chinese-origin, all Indian rhesus macaques exhibited a restriction site associated with a mutation diagnostic of haplogroups Ind1 or Ind2, whereas all but three Chinese rhesus macaques exhibited a restriction site associated with a mutation diagnostic of haplogroups ChiE, ChiW1, ChiW2, or ChiW3 (18). Two of the three atypical samples exhibited the mutations characteristic of haplogroup Ind1 and were later found to descend from the same maternal granddam, and the other displayed the mutations diagnostic of haplogroup Ind2. Although the three atypical samples cited above might represent errors in reporting or recording the origin of these animals or their ancestors, in which case the mtDNA mutations cited above and in Table 1 would be able to distinguish the origin of fullblood Indian or Chinese rhesus macaques with 100% accuracy, we have chosen to be conservative in the present study and assume that only 99% of rhesus macaques from India and China can be genetically defined by the mtDNA mutations described.

Three STR loci—D1s548, D18s547, and DXs2506—showed the most marked frequency differences between Indian and Chinese rhesus macaques. The frequencies of population-specific genetic markers that define Indian–origin and Chinese–origin rhesus macaques are summarized in Table 1. The 185-bp allele is the most common allele at D1s548 in Indian rhesus macaques, reaching a frequency of 0.51, and it did not occur in any Chinese rhesus macaque. Therefore, under equilibrium conditions, more than three-quarters of all Indian rhesus macaques, but no Chinese rhesus macaques, are expected to exhibit at least one 185-bp allele at this locus. Two alleles, the 155- and 166-bp alleles, account for 79% of all alleles found at the D18s537 locus in Indian rhesus macaques but only 13% of the alleles at this locus in Chinese rhesus macaques. Correspondingly, the most common allele at this locus in Chinese rhesus macaques, the 170-bp allele, accounted for 74% of all alleles in those animals but occurs with a frequency of only 21% in Indian rhesus macaques. Therefore, 63% of Indian rhesus macaques but only 1% of Chinese animals

Table 1. Genetic definition of Indian-origin and Chinese-origin rhesus macaques (*Macaca mulatta*)

	Locus	Trait	Frequency in population	
			Indian	Chinese
Indian origin				
	D1s548	at least one 185-bp allele	0.76	0.00
	D18s537	only 155- or 166-bp alleles	0.63	0.01
	DXs2506	at least one 264-bp allele	0.96	< 0.01
	mtDNA	15350G or 15845T	> 0.99	< 0.01
	at least one of the above four traits		> 0.99995	< 0.03
	at least two of the above four traits		> 0.995	< 0.0003
	at least three of the above four traits		> 0.890	< 0.000001
Chinese origin				
	D1s548	only 193- to 197-bp alleles	0.02	0.95
	D18s537	homozygous for 170-bp allele	0.04	0.55
	DXs2506	only ≥ 268-bp alleles	0.01	0.99
	mtDNA	15201G, 15296T, 15400G, or 15959T	< 0.01	> 0.99
	at least one of the above four traits		< 0.08	> 0.999995
	at least two of the above four traits		< 0.002	> 0.9995
	at least three of the above four traits		< 0.00003	> 0.970

exhibit only the 155- or 166-bp alleles, whereas 55% of Chinese rhesus macaques but only 4% of Indian monkeys have only the 170-bp alleles at the D18s537 locus. Finally, the 264-bp allele at the DXs2506 locus demonstrated frequencies of 0.92 in Indian rhesus macaques and 0.01 in Chinese animals. Therefore, about 96% of Indian rhesus macaques but fewer than 1% of Chinese animals are expected to exhibit at least one 264-bp allele at this locus under equilibrium conditions.

The probabilities of Indian and Chinese rhesus macaques exhibiting one, two, or three of the four defining traits at the mtDNA, D1s548, D18s537, and Dxs2506 loci of each population are summarized in Table 1. These estimates are predicated on the assumption of Hardy–Weinberg equilibrium among the four loci rather than on estimates of the frequencies of multilocus genotypes. We chose this method because accurate estimates of these frequencies would have required a far larger sample size than could be used in this study. The hypothesis of multilocus equilibrium is justifiable because all four loci segregate on different chromosomes, or in different genomes, and are not believed to influence fitness; therefore, the genotypes at the four loci are statistically independent of each other and should exhibit independent assortment. Moreover, insofar as the estimates of these probabilities are subject to sampling error, they are offered as approximations, rather than precise predictions, of the level of success to be expected in assigning rhesus to their correct regions of origin. The possession of at least two of the traits diagnostic of country of origin provides the optimal level of reliability in identifying country of origin of rhesus macaques. The probability that both Indian and Chinese rhesus macaques exhibit at least two of the defining traits of their country of origin exceeds 99.5%, whereas the probability that each exhibits at least two of the defining traits of rhesus from the other country of origin is less than 0.2%. mtDNA is more effective for identifying country of origin than is any of the STR loci. Dxs2506 is the most effective of the STR loci, followed closely by D1s548.

Discussion

The traits that genetically define Indian-origin and Chinese-origin rhesus macaques ensure that use of the genetic markers cited in Table 1 readily distinguishes between the unmixed members of these two regional populations. The few exceptions to the discrete differences in the mtDNA of fullblood Indian and fullblood Chinese rhesus macaques that were detected might not result from cases of mistaken identity, because all three samples alleged to be of Chinese origin that belonged to the Indian haplogroups Ind1 or Ind2 also exhibited typical Chinese alleles at all three STR loci described in Table 1. Because mtDNA is solely maternally inherited, it cannot detect admixture between Indian and Chinese rhesus macaques that will probably occur with increasing frequency in some breeding facilities. Recently developed computer programs, such as STRUCTURE (15), that assign geographic origin to samples in light of their genotypes at many nuclear loci can be used to estimate the relative probability or proportion of Indian and Chinese ancestry of any rhesus macaque. My laboratory previously showed that STRUCTURE detects first-generation Indian–Chinese hybrids with nearly 100% probability because they exhibit, on average (but with a range of about 25 to 70%), equal number of alleles that are especially characteristic of Indian rhesus macaques as that number of alleles that are particularly indicative of Chinese animals (18). However, reliable detection of admixture of second-generation hybrid Indian–Chinese rhesus macaques or products of backcrosses of first-generation hybrids with fullblood Indian or fullblood Chinese rhesus macaques will require the discovery of many more loci with alleles uniquely characteristic of one or the other country of origin. Admixture between second generation hybrids will lead to, on average, a much larger range of probabilities for assignment of samples to either country of origin than the 25-70% range for first generation hybrids, whereas backcrosses will display either more or less than half of the alleles indicative of each of the two alternative origins. The need for additional diagnostic loci arises because the resulting range of probabilities of Indian and Chinese ancestry for these hybrids overlaps that for fullblood rhesus from one or the other of the two countries.

Because of the marked genetic differences between Indian and Chinese rhesus macaques (or between two groups of admixed rhesus with very different levels of Indian, or Chinese, ancestry), the various populations might not provide strictly comparable results when used as animal models for the study of the same human diseases. These genetic differences exceed those among regional populations, subspecies, or even species of many other mammalian taxa. For example, Indian and Chinese rhesus macaques are more divergent from each other than are Bornean

and Sumatran orangutans, which are regarded as separate species for the purposes of conservation (9). In addition, the species *M. mulatta* is paraphyletic with respect to other species of the fascicularis group of macaque species, such that Chinese rhesus macaques are related to *M. cyclopis* (Taiwanese macaques) and *M. fuscata* (Japanese macaques) more closely than to Indian rhesus macaques, with whom they are regarded as being conspecific (4, 11, 12, 21). Because both Japan and Taiwan were connected to the Chinese mainland as recently as 15,000 years ago, *M. cyclopis* and *M. fuscata* evolved from Chinese rhesus macaques long after Indian and Chinese rhesus macaques became reproductively isolated from each other (one million years ago or more [4]). Both species probably are best regarded as regional variants of Chinese rhesus macaques, but, like Chinese rhesus macaques, are quite genetically different from Indian rhesus macaques.

As we have previously demonstrated (18, 20), without genetic testing, the country of origin of rhesus macaques cannot be known with certainty, even if the country from which the monkeys were acquired is known without question. All rhesus macaques acquired from abroad and all domestic-bred rhesus whose ancestry has not been confirmed by genetic methods should be genetically characterized to confirm their country of origin. Studies that use rhesus macaques as experimental subjects can minimize the variance in phenotypic traits under study that are due solely to genetic differences among research subjects by selecting one or the other of these two populations as the animal model of choice. That variance is least when Indian-origin rhesus macaques are used as animal models (18, 20). Comparisons among studies that use rhesus macaques whose ancestry derives from different halves (eastern versus western) of the geographical distribution of rhesus macaques (e.g., Indian versus Chinese rhesus macaques) should be tempered by recognition of the key genetic differences between the respective subjects of research.

Acknowledgments

This study was supported by grant RR05090 from the National Center for Research Resources, National Institutes of Health.

References

1. **Arnason, U., A. Gullberg, A. S. Burquete, and A. Janke.** 2000. Molecular estimates of primate divergences and new hypotheses for primate dispersal and the origin of modern humans. Hereditas **133:**217-228.
2. **Doxiadis, G. G. M., N. Otting, N. G. deGroot, N. deGroot, A. J. M. Rouweler, R. Noort, E. J. Verschoor, I. Bontjer, and R. E. Bontrop.** 2003. Evolutionary stability of MHC class II haplotypes in diverse rhesus macaque populations. Immunogenetics **55:**540-551.
3. **Goodwin, W. J. and J. Augistin.** 1975. The primate research centers program at the National Institutes of Health. Fed. Proc. **34:**1641-1642.
4. **Hayasaka, K., K. Fujii, and S. Horai.** 1996. Molecular phylogeny of macaques: implications of nucleotide sequences from an 896-base-pair region of mitochondrial DNA. Mol. Biol. Evol. **13:**1044-1053.
5. **Held, J. R.** 1980. Breeding and use of non-human primates in the USA. Int. J. Stud. Anim. Prob. **2:**27-37.
6. **Joag, S. V., E. B. Stephens, R. J. Adams, L. Foresman, and O. Narayan.** 1994. Pathogenesis of SIVMAC infection in Chinese and Indian rhesus macaques: effects of spenectomy on virus burden. J. Virol. **200:**436-446.
7. **Kanthaswamy, S. and D. G. Smith.** 1998. Use of microsatellite polymorphisms for paternity exclusion in rhesus macaques (*Macaca mulatta*). Primates **39:**135-145.
8. **Kanthaswamy, S. and D. G. Smith.** 2004. Effects of geographic origins on captive *Macaca mulatta* mtDNA variation. Comp. Med. **54:**193-201.
9. **Kanthaswamy, S., D. G. Smith, and J. D. Kurushima.** Submitted for publication.
10. **Ling, B., R. S. Veazey, A. Luckay, C. Penedo, K. Xu, J. D. Lifson, and P. Marx.** 2002. SIVmac pathogenesis in rhesus macaques of Chinese and Indian origin compared with primary HIV infections in humans. AIDS **16:**1489-1496.
11. **Marmi, J., J. Bertranpetit, J. Terradas, O. Takenaka, and X. Domingo-Roura**. 2004. Radiation and phylogeography in the Japanese macaque, *Macaca fuscata*. Mol. Phylogenet. Evol. **30:**676-685.
12. **Melnick, D. J., G. A. Hoelzer, R. Absher, and M. V. Ashley.** 1993. mtDNA diversity in rhesus monkeys reveals overestimates of divergence time and paraphyly with neighboring species. Mol. Biol. Evol. **10:**282-295.
13. **Morin, P. A., S. Kanthaswamy, and D. G. Smith.** 1997. Simple sequence repeat (SSR) polymorphisms for colony management and population genetics in rhesus macaques (*Macaca mulatta*). Am. J. Primatol. **42:**199-213.
14. **National Research Council.** 1996. Guide for the care and use of laboratory animals. National Academic Press, Washington, D.C.
15. **Pritchard, J. K. and W. Wen.** 2003. Software for genetic analysis. [Online] Available at http://pritch.bsd.uchicago.edu). Accessed 3/02/05.
16. **Smith, D. G.** 1980. Paternity exclusion in six captive groups of rhesus monkeys (*Macaca mulatta*). Am. J. Phys. Anthropol. **53:**243-249.
17. **Smith, D. G.** 1994. Genetic heterogeneity in five captive specific pathogen-free groups of rhesus macaques. Lab. Anim. Sci. **44:**200-210.
18. **Smith, D. G., D. A. George, S. Kanthaswamy, and J. W. McDonough.** Genetic tests to verify the country of origin of and level of admixture between Indian and Chinese rhesus macaques (*Macaca mulatta*). Int. J. Primatol., in press.
19. **Smith, D. G., S. Kanthaswamy, J. Viray, and L. Cody.** 2000. Additional highly polymorphic microsatellite (STR) loci for estimating kinship in rhesus macaques (*Macaca mulatta*). Am. J. Primatol. **50:**1-7.
20. **Smith, D. G. and J. W. McDonough.** 2005. Mitochondrial DNA variation in Chinese and Indian rhesus macaques (*Macaca mulatta*). Am. J. Primatol. **65:**1-25.
21. **Smith, D. G., J. W. McDonough and D. A. George.** 2005. Unpublished data.
22. **VandeBerg, J. and S. Williams-Blangero.** 1997. Advantages and limitations of nonhuman primates as animal models in genetic research on complex diseases. J. Med. Primatol. **26:**113-119.
23. **Viray, J., B. Rolfs, and D. G. Smith**. 2001. Comparison of the frequencies of major histocompatibility (MHC) class-II DQA1 and DQB1 alleles in Indian and Chinese rhesus macaques (*Macaca mulatta*). Comp. Med. **51:**555-561.