Published in final edited form as:

Am J Primatol. 2010 June ; 72(7): 587–599. doi:10.1002/ajp.20811.

Genetic Characterization of Specific Pathogen-Free (SPF) Rhesus Macaque (*Macaca mulatta*) Populations at the California National Primate Research Center (CNPRC)

SREE KANTHASWAMY^{1,2,*}, ALEX KOU², JESSICA SATKOSKI², MARIA CECILIA T. PENEDO³, THEA WARD³, JILLIAN NG², LEANNE GILL¹, NICHOLAS W. LERCHE¹, BETHANY J-A ERICKSON², and DAVID GLENN SMITH^{1,2}

¹California National Primate Research Center, University of California, Davis, California

²Department of Anthropology, University of California, Davis, California

³Veterinary Genetics Laboratory, University of California, Davis, California

Abstract

A study based on 14 STRs was conducted to understand intergenerational genetic changes that have occurred within the CNPRC's regular SPF (Specific pathogen-free) and super-SPF captive rhesus macaque populations relative to their conventional founders. Intergenerational genetic drift has caused age cohorts of each study population, especially within the conventional population, to become increasingly differentiated from each other and from their founders. While there is still only minimal stratification between the conventional population and either of the two SPF populations, the separate derivation of the regular SPF and super-SPF animals from their conventional founders has caused the two SPF populations to remain marginally different from each other. The regular SPF and, especially, the super-SPF populations have been influenced by the effects of differential ancestry, sampling and lost rare alleles, causing a substantial degree of genetic divergence between these subpopulations. The country of origin of founders is the principal determinant of the MHC haplotype composition of the SPF stocks at the CNPRC. Selection of SPF colony breeders bearing desired genotypes of Mamu-A*01 or –B*01 has not affected the overall genetic heterogeneity of the conventional and the SPF research stocks.

Because misclassifying the ancestry of research stocks can undermine experimental outcomes by excluding animals with regional-specific genotypes or phenotypes of importance, understanding founder/descendent genetic relationships is crucial for investigating candidate genes with distinct geographic origins. Together with demographic management, population genetic assessments of SPF colonies can curtail excessive phenotypic variation among the study stocks and facilitate successful production goals.

Keywords

microsatellites; major histocompatibility complex (MHC); genetic structure; genetic management; colony management

*Correspondence to: S. Kanthaswamy, Department of Anthropology, University of California, Davis, California 95616. skanthaswamy@ucdavis.edu.

INTRODUCTION

The NIH sponsored National Primate Research Centers (NPRCs) define and derive their specific pathogen-free (SPF) animals in different ways. These differences are mainly due to varied housing, social enrichment, husbandry and management approaches that stem from each NPRC's unique historical and environmental characteristics [Schapiro et al. 1994; Schapiro and Bloomsfield 1995]. In general, SPF rhesus macaques are defined as animals that are antibody negative to herpes virus-1 (B virus), simian immunodeficiency virus (SIV), Type D simian retrovirus (SRV) and simian T-lymphotropic virus (STLV-1). The SPF status of the animals not only improves animal health and occupational safety but also removes potential confounding variables in research. For instance, retroviral infected animals are not desirable for SIV/AIDS-related research because studies on SIV/AIDS pathogenesis and vaccine development that utilize these animals may be compromised. Although the use of herpes B virus seropositive animals does not undermine SIV/AIDS-related investigations, these animals do present significant health risks to primate colony personnel.

The SPF program of the California National Primate Research Center (CNPRC) was first established in 1993 and now includes two types of SPF populations: the regular SPF and super-SPF, ranging in age from mature adults to newborns. The regular SPF animals, now numbering 746 individuals, are derived directly from the conventional, non-SPF breeding stock through nursery-rearing of neonates (produced by natural breeding or embryo transfer) and are seronegative for B virus, simian type D retrovirus (1 through 5), SIV, STLV and *Mycobacterium tuberculosis* [Lerche and Simmons 2008]. Unlike other NPRC SPF derivation strategies, the CNPRC's super-SPF (also known as 'super clean') populations, composed of 397 animals, are derived directly from the conventional (non-SPF) stock rather than being double-derived from the regular SPF stock. The CNPRC super-SPF population, which was first established in 2001, is further subdivided into two subpopulations; 95 animals derived from non-SPF founders from China and 302 animals that descend from Indian-derived conventional animals. In addition to pathogens absent in the regular SPF animals, both the Chinese and Indian super-SPF animals are also free of rhesus cytomegalovirus (RhCMV), rhesus rhadinovirus (RRV), and simian foamy virus (SFV), but are not readily available to most investigators. Both the SPF and super-SPF animals are repeatedly tested for target pathogens and are either housed individually or in contact with animals with the same virological status.

Behavioral, biochemical, immunological and molecular genetic characteristics of both the regular SPF and super-SPF populations at the CNPRC have been well-characterized [Lerche et al. 1991, 1994, Barry et al. 1996, Capitanio and Lerche 1998, Capitanio et al. 1998a b, Viray et al. 2001, Huff et al. 2003, Lerche and Osburn 2003, Capitanio et al. 2006, Khan et al. 2008, Barry and Strelow 2008]. The demographic composition, population growth, birth and death rates, genetic structure and experimental needs have been evaluated to improve animal well-being, safety and to best meet programmatic production goals. Three published studies have investigated the genetic diversity of the regular SPF population at the CNPRC. Smith [1994] used ten serum and erythrocyte enzyme polymorphisms to estimate allele frequencies, gene diversity and the genetic structure of regular SPF groups in five different breeding facilities, including the CNPRC, and Penedo et al. [2005] evaluated the informativeness of a panel of eight STRs in the rhesus (*Mamu*) major histocompatibility complex (MHC) class I and II regions on rhesus chromosome 4. Based on a sample set consisting of SPF animals from the CNPRC and six other NIH-funded primate centers, Kanthaswamy et al. [2006] evaluated a panel of 16 STRs proposed for standardized testing across the NPRCs. However, no DNA-based study has compared the genetic structure of the SPF and super-SPF populations at the CNPRC with that of their conventional, non-SPF forbearers.

Social structure, colony management and husbandry strategies can all effect changes in the genetic composition of the SPF rhesus macaque colonies [Kanthaswamy and Smith 2002, Ely et al. 1994]. It is difficult to predict future outcomes of closed colony management of a captive population without an understanding of the history of its population structure that addresses at least three important questions. First, have the CNPRC regular SPF and super-SPF populations already differentiated genetically from the conventional stock and from each other and, if so, how rapidly and to what degree has this differentiation occurred? Such differentiation can inflate the phenotypic variance of traits under study by increasing the genetic contribution to that variance, confounding interpretations of experimental treatment effects. Closed colonies are subject to rapid changes in allele frequencies due to genetic drift and an overall loss of genetic diversity, the extent of which is inversely proportional to the colony effective population size. The high variance in male reproductive success and low sex ratios characteristic of captive colonies of rhesus macaques diminish the effective population size, accelerating this loss of genetic diversity [Kanthaswamy and Smith 2002]. Second, undesirable results of artificial selection have been recorded in laboratory animal science for decades [Roberts 1979], but it is currently unknown if previous selective breeding practices at the CNPRC have had an effect on present population structure. For instance, it remains unclear if breeding strategies designed to increase numbers of rhesus macaques positive for *Mamu-A*01* and *-B*01* alleles among the SPF stocks to meet specific research needs have impacted the genetic composition of the colony at large. Such “enrichment” of alleles in the SPF portions of the colony provides SPF animals with genetic characteristics desirable for specific research objectives. However, it correspondingly depletes these alleles, and those linked to them, in the conventional colony, and might reduce genetic heterogeneity in the MHC region in the SPF colonies, because these animals are likely to share higher kinship coefficients than animals randomly selected from the colony, limiting the usefulness of these colonies for identifying genes that influence risk of specific diseases. Third, animals of Chinese origin at the CNPRC have been randomly assimilated into the conventional and regular SPF colonies since 1980 to counter anticipated loss of genetic variation among its Indian-derived rhesus macaques after exports from India were discontinued [Roberts et al. 2000]. The result of the introgression of Chinese genes into the predominantly Indian gene pool has been the production of hybrid Indian-Chinese, or *Chindian*, rhesus macaques, which now number more than a thousand animals and represent a significant part of the CNPRC colony [Kanthaswamy et al. 2008a]. Such intentional hybridization is unique among all NPRCs in the US. Unrecognized hybrid animals and their atypical genetic characteristics (e.g., enhanced level of genetic diversity and linkage disequilibrium), while uniquely useful for some purposes, such as admixture mapping of genes contributing to specific diseases and characterizing the haplotype block structure of the rhesus genome, may lead to false positive or false negative results when individuals of different ancestry proportions are distributed unequally between experimental and control groups under study.

This study addresses the three questions cited above by comparing the genetic structure and population characteristics of the regular SPF and super-SPF populations with those of the conventional population, from which both groups were independently derived at different times. The study group comprised 7536 animals representing multi-generational pedigrees that were established since 1970, and which include great-great-grandparents, great-grandparents, grandparents, parents and offspring. The complex demography of the study populations over a period of 40 years complicates drawing inferences about large-scale patterns of multigenerational genetic stratification based solely on pedigree information. Therefore, this study directly assessed the population structure based on the STR genotypes and the *Mamu-A*01* and *-B*01* phenotypes of animals. It has been demonstrated that population-level genome-wide comparisons based on autosomal markers such as STRs,

including MHC-linked loci, provide reliable estimates of population admixture and stratification [Falush et al. 2003; Kayser et al. 2008].

We predict that limited population size, long-term reproductive barriers, artificial selection, differential distribution of Chinese-Indian hybrid animals and genetic isolation produced by the independent derivation of the SPF colony and the two super-SPF populations have made these groups less genetically similar to each other as well as less genetically diverse than, and more differentiated from, the founding population from which they were both derived. To test for the presence of drift-driven differentiation, genetic diversity and subdivision were calculated for single year age cohorts in each of these populations and compared to estimates obtained from animals in the conventional population. We also predict that the enrichment of the two SPF populations for A*01 and B*01 positive animals (i.e., artificial selection), who are likely to share higher kinship than any two randomly chosen animals, has reduced genetic diversity at these three MHC-linked loci relative to diversity at other loci in the SPF populations and at these loci in the conventional population.

METHODS

STR genotyping, parentage determination, *Mamu-A*01* and *Mamu-B*01* typing were performed over a period of 10 years at the Veterinary Genetics Laboratory (VGL), University of California, Davis, CA. DNA was extracted by an alkaline lysis procedure [SanCristobal-Gaudy et al. 2000] and as detailed in Andrade et al. (2004). PCR amplification of STRs was done in two 25 μ L multiplexed reactions (groups A and B in Appendix A) each containing 1 μ L of DNA template, 2.5 mM MgCl₂, 200 μ M dNTPs, 1X PCR buffer II (Applied Biosystems), 0.5 U Amplitaq (Applied Biosystems) and fluorescence-labeled primers in final concentrations given in Appendix A. PCRs were carried out in PTC100 (MJ Research) and ABI 2720 (Applied Biosystems) thermal cyclers with cycling conditions that consisted of four cycles of 1 minute at 94°C, 30 seconds at 58°C, 30 seconds at 72°C; 25 cycles of 45 seconds at 94°C, 30 seconds at 58°C, 30 seconds at 72°C, and a final extension at 72°C for 30 minutes. PCR products from each multiplex group were separated on Applied Biosystems' ABI 373, ABI 377 PRIZM or ABI 3730 DNA Analyzer instruments according to the manufacturer's instructions. Fragment size analysis and genotyping was done with STRand software [Hughes 2000]. Automated genotype calls by STRand were checked and corrected as needed by means of two independent reads of electropherograms. Samples that yielded incomplete or low quality profiles (ca. 1-2%) were re-extracted and retested. Inclusion of a reference sample in all PCRs allowed for standardization of allele sizes within and across different electrophoresis instruments in order to maintain continuity of genotype data. Three of the 14 informative STRs used in this study (see Appendix A), D6S291, D6S276 and D6S1691 flank the core MHC region on rhesus macaque chromosome 4 [Penedo et al. 2005]. D7S513 (np 114,658,176 – 114,658,387) and D7S794 (np 181,448,827 – 181,448, 977) are 67 Mb apart on rhesus chromosome 3, and D8S1106 (np 12,737,294 – 12,737,516) and MFGT21 (np 111,063,327-111,063,442) are 98 Mb apart on chromosome 8. A chromosome location of MFGT22 could not be determined, presumably because it falls on a gap of the rhesus whole genome sequence. The remaining markers have been mapped to different chromosomes and, therefore, are physically unlinked.

*Mamu-A*01* and *-B*01* typing was performed by PCR using the sequence-specific primers described by Knapp et al. (1997) and Loffredo et al. (2005), respectively. PCR amplification and agarose gel electrophoresis with ethidium bromide staining were carried out as described in Loffredo et al. (2005), with the modification that rhesus beta-actin primers *ACTB-F*: ACCCCAGCCATGTACGTGGCCATCC and *ACTB-R*:

GCCTCAGGGCAGCGGAACCGCTCA were used to amplify a 395 bp fragment as the internal PCR control.

Only animals with genotype data for at least 12 of the 14 STRs were included in this study. Table I presents the total numbers of animals from the conventional, regular SPF and super-SPF populations included in the genetic diversity and genetic differentiation estimates. PopGene [v 1.32, Yeh and Boyle 1997] was used to calculate allele frequencies, observed heterozygosity (OH) and gene diversity (expected heterozygosity, EH) within each age cohort of each group of rhesus macaques at the CNPRC colony. The same program was used to examine the intergenerational genetic structure of each (total) population using F-statistics (Fst, Fis and Fit, Wright 1978) while the program GENEPOP [Rousset and Raymond 1995] was used to estimate the level of pairwise genetic differentiation between age cohorts of the conventional, regular SPF and super-SPF populations based on Weir and Cockerham's pairwise Fst [1984]. Summary F-statistics and annual pairwise Fst estimates were also computed to assess the degree of divergence between pairs of age cohorts within and between the Chinese and Indian-derived super-SPF subpopulations at the CNPRC.

To assess the ability of the allelic frequencies at 14 STRs to determine the degree of regional representation of the animals that comprise each of the study populations, an assignment analysis using the program STRUCTURE [Pritchard et al. 2000, Falush et al. 2003] was conducted at sweeps of 2×10^4 iterations after a burn-in period of 1×10^4 on a population subset of 737 animals, including individuals from the three study groups whose degree of Chinese or Indian ancestry were confirmed with their or their ascendants' import documentation that date back to 1970. To confirm the expected full blood Indian, full blood Chinese and hybrid structure at the CNPRC, STRUCTURE runs were performed assuming that between two and nine ($2 < K < 10$) genetically distinct classes of individuals exist among the study populations as proposed by Kanthaswamy et al. (2008). Each assumed K value was replicated 10 times with each set of assumptions so that the regional population or ancestry assignments had the highest possible probability. Analyses under the assumption of admixture between genetically distinct classes and their correlated allele frequencies were conducted without *a priori* geographic information to assure that the ancestry assignments based on the 14 STRs were not biased. Assignment of individuals into inferred ancestral populations was conducted using the probability Q, the proportion of an individual's genome that originated from the K^{th} subpopulation (Pritchard et al. 2000; Falush et al. 2003) and only individuals supported by at least 90% probability were assigned to any full blood class. Similarly, to establish a quantitative measure of population structure of the conventional (N = 6216), regular SPF (N = 993) and super-SPF (N = 327) animals of the entire study population, runs of clustering analyses of the same numbers of iterations and burn-in periods were performed without prior information on population membership to ascertain the probability of affiliation of each animal with the conventional, regular SPF and super-SPF populations (i.e., the ability of STRUCTURE to correctly assign each animal, based on its genotypes for the 14 STRs, to the population of which it is a member).

In order to determine if the Chinese or the Indian animals were the principal cause of genetic differentiation among the conventional, regular SPF and super-SPF populations at the CNPRC, pairwise Fst values between 20 randomly selected, STRUCTURE-assigned (see above for details on the STRUCTURE runs) full blood Indian and Chinese rhesus macaques pooled across years from each of the three study populations were calculated. Because each of the study groups contained significantly varying numbers of animals, the randomized selection of samples alleviated the effects of sample size on estimates of genetic differentiation.

To assess the possibility that the selection process to derive regular SPF and super-SPF animals with a high frequency of A*01 and B*01 positive phenotypes has produced non-random patterns of genetic variation within the rhesus MHC region or physically close to it, the range of divergence of the MHC types and the generational changes since 2001 in allelic diversity of three STRs that flank the rhesus MHC, i.e., D6S276, D6S291 and D6S1691, were compared with the same values for all 14 STRs (given in Table 1) among the conventional (N = 5891), regular SPF (N= 993) and the two super-SPF populations (Indian N= 243; Chinese N = 82). If artificial selection has reduced diversity in the SPF colonies, the average level of heterozygosity at the three MHC-linked STRs in those colonies should be lower than both that for all 14 loci and that for the conventional colony.

RESULTS

The average number of STR alleles (n) and the average observed (OH) and expected (EH) proportion of heterozygous genotypes in each age cohort in each of the three populations (including age cohorts of the Chinese and Indian super-SPF subpopulations) are presented in Table I. The table illustrates the fluctuations in the distribution of colony-wide STR heterozygosity in each group. Overall estimates of allele numbers show that an average of 5 rare alleles present in the conventional founder population were lost in both the regular SPF and super-SPF populations. While the loss of these rare alleles can preclude the identification of their potential association with rare diseases, it has not significantly reduced estimates of heterozygosity in either of these populations due to their low frequencies in all the colonies. Concordant with several previous studies of conventional Indian and Chinese rhesus macaque colonies [e.g., Kanthaswamy et al. 2008a], the Indian super-SPF animals were less genetically diverse than the Chinese super-SPF animals as measured by allelic numbers (n) and observed (OH) and expected (EH) heterozygosity estimates.

Table II shows varying pairwise differentiation among the age cohorts of each of the three populations (including age cohorts of the Chinese and Indian super-SPF animals) from 1993 to 2008. Lower pairwise F_{st} estimates are indicative of overlaps in allele frequencies between cohorts of populations being compared. All three populations exhibit increasing divergence from their respective founder generations due to genetic drift (i.e., from 0.0017 in 1997 to 0.0112 in 2008, from 0.0061 in 2001 to 0.0156 in 2008 and from 0.0101 in 2005 to 0.0138 in 2008 in the conventional, regular SPF and super-SPF populations, respectively). However, of the 66 possible pairwise comparisons between age cohorts in the three populations, the estimate of F_{st} exceeded 0.02, still a relatively low level of differentiation, only once (when the 2001 regular SPF and 2006 super-SPF populations were compared).

Periodic estimates of pairwise genetic differentiation among the age cohorts in the total population of the conventional, regular SPF and super-SPF age cohorts in 2001, 2005 and 2008 suggests that both regular and super-SPF populations remain only moderately differentiated from the conventional population ($F_{st} < 0.02$). By the end of the study period the pairwise differentiation between the age cohorts of the full blood Chinese super-SPF subpopulation and the conventional and regular SPF populations were approximately 10 times greater than those between the Indian super-SPF subpopulation and the conventional and regular SPF populations (Table II). Analysis based on twenty randomly selected full blood Chinese and Indian samples from all three populations shows that differences are greatest between the full blood animals in the super-SPF Chinese and Indian subpopulations while no differentiation was detected between the Indian origin SPF and super-SPF populations (Table III). Tables II and III also indicate that the Chinese rather than the Indian animals are responsible for the significant genetic differentiation between the super-SPF population and the conventional and regular SPF populations at the CNPRC. Differentiation among the cohorts of the full blood Chinese rhesus macaques in all three study groups is

significant while the full blood Indian-derived animals show either no ($F_{st} = -0.006$) or only scant ($F_{st} = 0.001 - 0.003$) differentiation (Table III).

The genetic structure of the conventional population from 1993 to 2008, as measured by overall estimates of F_{is} , F_{it} and F_{st} among cohorts were -0.0451 , 0.0557 and 0.0965 . The summary F -statistics for the regular SPF group from 1993 to 2008 were -0.0308 , 0.0273 and 0.0564 , respectively, while estimates for the super-SPF from 2001 to 2008 were -0.0386 , 0.0136 and 0.0502 , respectively. These values suggest low levels of excess homozygosity due to inbreeding (F_{is}) and only moderate levels of genetic drift-driven differentiation (F_{st}) in each of the three groups. Despite showing some early effects of sampling size on the estimates, the increasing degree of divergence ($F_{st} \sim 0.08$ in the 2005-07 period to 0.1 by the end of 2008) between the Chinese and Indian fractions of the super-SPF population reveals the deepening subdivision as they developed (Table II). The summary values of F_{is} , F_{it} and F_{st} for the Chinese super-SPF subpopulation were -0.05 , 0.02 and 0.06 , respectively, while those for the Indian subpopulation were -0.03 , 0.00 and 0.03 .

At $K = 3$, the STRUCTURE analysis of the 737 animals with known Chinese, Indian or hybrid ancestry, as supported by their or their ancestors' import documentation, indicated that there are animals of full blood Indian, full blood Chinese and hybrid Chinese-Indian (i.e., Chindian) animals in varying proportions within the CNPRC population (Figure 1). As K values increased to 10, the full blood Indian and Chinese clusters among the colony-wide dataset remained tightly homogenous, while increasing subdivision was observed among the Chinese-Indian hybrid animals. Of the entire dataset of 7536 individuals, STRUCTURE was able to assign only about 20%, 47% and 53% of the conventional, regular SPF and super-SPF animals, respectively, to their appropriate populations based on their genotypes, revealing a low level of genetic stratification among these populations (data not shown) in agreement with the low-to-moderate estimates of pairwise F_{st} between the conventional, regular SPF and super-SPF age cohorts shown in Table II.

A large fraction of the animals in the entire population ($N = 7536$) was assigned varying proportions of Chinese-Indian admixture, in close agreement with demographic records of animals with known $1/8$, $1/4$, $3/8$, $1/2$, $5/8$, $3/4$, and $7/8$ Chinese ancestry (see Kanthaswamy et al. 2008 and Figure 1). The regular SPF population contains more Indian-derived and hybrid animals than the conventional population while the super-SPF population contains more full blood Chinese animals than either the conventional or regular SPF populations; the super-SPF Chinese-Indian hybrid animals were exclusively restricted to the Indian super-SPF subpopulation. A regression analysis of the known level of admixture and the proportions of Indian and Chinese origin assigned by STRUCTURE produced an R^2 value of 0.87 (data not shown) indicating that the clustering analysis using the STRs reported in this study is able to estimate levels of admixture of both full blood and Chinese-Indian hybrid rhesus macaques with a relatively high level of accuracy.

By the end of the study period the estimates of diversity among the rhesus MHC-linked STR alleles in each of the populations were at their highest values (see Appendix B). Sampling effects in the early years of regular SPF and super-SPF development were clearly evident, but years after each of these populations was established and their animal numbers increased, their allelic diversity and heterozygosity rose dramatically and matched levels of the conventional population by 2008. The loss of rare alleles at the three MHC-linked loci was equivalent to that at non-MHC-linked loci in all three study populations and by 2008, all three study populations carried equal amounts of heterozygosity at the MHC-linked loci. In 2007 and 2008, the super-SPF population carried the greatest degree of genetic variation as reflected by its estimates of observed and expected heterozygosity. During the same period, the allele diversity at each of these three loci was roughly equivalent between the

Indian and Chinese super-SPF subpopulations, but the observed and expected heterozygosities of the Chinese subpopulation were much higher than those of the Indian subpopulation.

Across all three study populations, animals of Indian ancestry are the primary source of A*01 alleles (with a frequency of 27.8%, compared to 3.0% for those of Chinese ancestry), while B*01 alleles are present in animals of Chinese and Indian ancestry in more similar proportions (16.2% and 27.0%, respectively). Figure 2 illustrates the annual counts of A*01 animals within the super-SPF population according to their country of origin. About 1% of the full blood Chinese super-SPF animals were A*01 positive compared to 23% of the Indian animals whereas 14.5% of the Chinese and 23% of the Indian super-SPF animals were B*01 positive.

DISCUSSION

The present study represents a first direct assessment of the population structure of the SPF populations at the CNPRC based on DNA-based genotype data from successive generations relative to the conventional population from which they were established. Both the regular SPF and super-SPF populations at the CNPRC are under intense development and their numbers have significantly increased since their inception in 1993 and 2001, respectively. The CNPRC SPF populations have been influenced by artificial selection, demographic composition, including hybridization, and management differences while still maintaining high gene diversity. For example, both the SPF and super-SPF populations exhibit loss of some rare alleles, including those at MHC-linked loci, as they expanded from a few founding animals initially to 746 and 397 living animals, precluding investigation of their potential association with diseases or other phenotypes. As observed by Satkoski et al. [2007] and Kanthaswamy et al. [2008a], sample size, as represented in founder effects, intergenerational genetic drift and genetic bottlenecks, impact allelic diversity more profoundly than heterozygosity due to the disproportionate loss of rare, compared with common, alleles. Similar stochastic effects can also be expected in specific regions of interest in the rhesus genome as reflected by the loss of alleles at MHC-linked STR loci.

The Bayesian statistical application featured in STRUCTURE has been successfully used for identifying hybrid individuals and their approximate proportions of ancestry [Satkoski et al. 2007, Kanthaswamy and Smith 2008a, b]. Given the accurate (approximately 87% concordant) estimates of these animals' ancestry, assessments of the colony-wide sample-set, whose ancestry is not completely known from demographic records, revealed that a large number of previously undetected animals of partial Chinese descent make up both the regular SPF population and the Indian super-SPF populations. The occurrence of Chinese and Chinese-Indian hybrid animals in the SPF population as well as the conventional colony reflects the large numbers of Chinese rhesus macaques imported into the US in recent years [NCRR/OAR SPF colonies committee meeting, 2003] and the decision of the CNPRC not to completely isolate the Indian and Chinese portions of the breeding colonies from each other. Although the CNPRC super-SPF population was originally designed to contain two genetically distinct subpopulations, one each consisting of full blood Chinese and full blood Indian animals, the occurrence of Chinese-Indian hybrid animals in the Indian super-SPF subpopulation suggests that many previously undetected hybrid animals from the conventional colony have been inadvertently introduced into this subpopulation. Although comparable data from other regional primate centers are unavailable, it is possible that similar circumstances have led to unacknowledged hybridization between Indian and Chinese rhesus macaques at other breeding facilities in the US. Because the Chinese portion of the super-SPF population was founded by animals imported directly from China, no occurrences of Chinese-Indian hybrid animals were detected in that subpopulation.

Low to moderate amounts of genetic drift have led to some differentiation among the study populations, but the impact of genetic drift is not solely dependent on annual population size of birth cohorts and stochastic effects of breeding. A deficit of observed relative to expected heterozygosity within most age cohorts (Table I) in the conventional population reflects genetic substructure induced by non-random mating. Effects of genetic drift between adjacent age cohorts (intergenerational drift) of the regular SPF population and the super-SPF population emerged as these populations became more established at the CNPRC. These trends in all three populations led to sustained reductions in genetic diversity within the individual age cohorts relative to the population as a whole. However, the yearly pairwise F_{st} estimates between the gene pools of founders of each study population and their descendents suggest that differentiation among successive age cohorts within each study population is still nearly negligible. The conventional population exhibited some intergenerational genetic drift, probably because of the varied housing conditions of the non-SPF animals, including outdoor corrals, corncribs and indoor housing units, and the size and composition of their multiple breeding groups. Management policies including selectively harvesting females to minimize over-representation of one or only a few matriline and cross-fostering week-old infants between breeding corrals have prevented deeper subdivisions [Smith 1986]. Since the conventional population is the foundation of both the regular SPF and super-SPF populations, effects of the selective harvesting and cross-fostering programs in the non-SPF breeding groups are also evident in the SPF descendants as shown by the lack of excessive cross-generational homozygosity (i.e., negative F_{is} values). The twice greater intergenerational genetic drift experienced by the Chinese super-SPF subpopulation than the Indian super-SPF subpopulation ($F_{st} = 0.06$ vs. 0.03) which has accelerated the deepening genetic schism between the Chinese super-SPF animals and the other populations studied here exemplifies the strength of this evolutionary force at the CNPRC. For instance, by 2008, the degree of differentiation between the Indian component of the super-SPF population and its pure Chinese component was twice Kanthaswamy et al.'s [2008a] estimate of divergence between the colony-wide full blood Indian and pure Chinese animals, exceeding that differentiating well-recognized breeds of dogs [Kanthaswamy et al. 2009]. This is an indication that the two super-SPF subpopulations that are demographically different by design are becoming even more genetically distinct from each other. Thus, the increasingly common use of SPF animals in biomedical research carries a correspondingly greater risk of inflating phenotypic variance in treatment effects due solely to the escalation in genetic differentiation among research subjects.

The strategy to derive regular SPF and super-SPF animals directly from the conventional population forms a barrier to direct gene flow between the SPF and super-SPF populations and limits gene flow between both SPF populations and the conventional population. Had the super-SPF population been derived from the regular SPF population (i.e., double-derived), the potential for fixed differences between the two SPF populations would have been reduced, albeit even more drastic declines in allele numbers and heterozygosity would be expected in the super-SPF population compared to those in the conventional population. The impact of separate derivation of the regular SPF and super-SPF stocks is even more critical for programs that aim to convert entire rhesus colonies to seronegative status. The lack of animal movement between the two SPF stocks prevents recovery of unique alleles lost in each group leaving recurrent mutations, or subsequent gene flow from the conventional colony, as the only source of such recovery. The significant fragmentation of the genetic diversity as a consequence of bottleneck effects can lead to an increase in inbreeding coefficients (and possibly inbreeding depression) accelerating the loss of valuable alleles (and/or the long-term viability) of the SPF populations. The extent of divergence between the Indian and Chinese super-SPF age cohorts as well as the differentiation between the super-SPF Chinese subpopulation and the conventional and SPF

populations underscore the importance of sampling considerations when animals from these populations are assigned to specific research projects.

Understanding the consequences of past artificial directional selection for A*01 and B*01 animals that might be disproportionately culled from colonies to serve research needs can forecast differentiation among the regular SPF, super-SPF and conventional stocks particularly if these animals are prevented from interbreeding, i.e., are reproductively isolated. The “enrichment” of small founder populations with such alleles, many of which likely descend from a much smaller number of common ancestors, can increase inbreeding and founder effects in SPF colonies. However, by 2008 all three study populations carried approximately equal levels of heterozygosity at the MHC-linked loci that approximated those for non-MHC-linked loci, and the loss of rare alleles at the three MHC-linked loci studied was equivalent to that at non-MHC-linked loci. While observed levels of heterozygosity at the three MHC-linked loci were consistently, albeit marginally, lower than that at all 14 loci combined in the SPF colonies, other parameters were equivalent. Thus, there is no strong evidence that selecting SPF colony breeders bearing the desired genotypes of A*01 or B*01 have influenced the overall genetic heterogeneity and structure of either the conventional or SPF research stocks other than to enrich and deplete the numbers of A*01 and B*01 animals in the SPF and conventional colonies, respectively. Since the primary source for A*01 animals are the Indian-derived fraction of the CNPRC colony [Muhl et al. 2002, Vogel et al. 1995], the geographic origins of founder animals have a pronounced impact on the MHC haplotype composition of the SPF stocks at the CNPRC and other NPRCs.

Combining insights from population genetic assessment with those from demography to understand the genetic history of the SPF populations at the CNPRC permits colony managers to predict outcomes of the long-term maintenance of self-propagating SPF populations. Comparing genetic data from founder animals with that of their descendents sheds light on the population changes that occur within each generation of genetic management and facilitates future production goals. It is also crucial to determine the regional representation of imported animals and the strength and consequences of founder effects on the SPF populations that include them. Just as the presence of full blood Chinese and Chinese-Indian hybrid animals have inflated the genetic diversity of the rhesus macaque population at the CNPRC, the misclassification of full blood and hybrid animals can hamper research outcomes by excluding animals with regional-specific genotypes or phenotypes of interest. Having demonstrated that there is evidence for a high degree of intentional interregional admixing of rhesus macaques in Chinese breeding facilities, Satkoski et al. [2007] have argued that unnatural admixing of historically isolated taxa could irreversibly confound the genetic structure of domestic stocks. Assessing the degree of genetic variability within and between SPF populations will help curb excessive phenotypic variation among the study subjects that could complicate studies of disease pathogenesis and vaccine development [Kanthaswamy et al. 2008a, 2008b].

Appendix A

The 14 STRs used in this study.

Locus	Rhesus chromosome location**	Repeat motif	Multiplex Group	Dye Label-Forward Primer (5'-3')****	Reverse Primer (5'-3')****	μ M in PCR*****
D3S1768	2	Tetranucleotide	A	FAM-GGTTGCTGCCAAGATTAGA	CACTGTGATTTGCTGTGGGA	0.07
D6S276	4	Dinucleotide	B	NED-TTCCAGTGATACATCAATCAAAATCA	GGGTGCAACTTGTCTCTCCT	0.12
D6S291	4	Dinucleotide	B	VIC-CTCAGAGGATGCCATGTCTAAAAATA	GGGGATGACGAAATTAATCACTAACT	0.10
D6S1691	4	Dinucleotide	B	FAM-AGGACAGAAATTTGCTC	GCTGCTCTGTATAAGTAATAAAC	0.10
D7S513	3	Dinucleotide	A	VIC-AGTGTGTTTGAAGGTGTAGGTTAAT	ATATCTTTCAGGGGAGCAGG	0.07
D7S794	3	Tetranucleotide	A	FAM-ACCATACTCTCAGCCTCCA	GTGTTGGGGTTCTTCCAAAGA	0.07
D8S1106	8	Tetranucleotide	B	PET-GCGGCGCATGTTTTCCTACTTT	TTCTCAGAAATTGCTCATAGTGC	0.06
D10S1412	9	Trinucleotide	B	VIC-TGCTTAGCTCTGCATACCTGA	GGGACAGTTCTTCTCCCTCCA	0.10
D11S925	14	Dinucleotide	A	VIC-GCTCTCTCCAGTAATTCTGTC	TTAGACCAATTA TGGGGGCAA	0.09
D13S765	17	Tetranucleotide	A	NED-TGTAACTTACTTCAAAATGGCTCA	TTGAAACTTACAGACAGCTTGC	0.10
D16S403	20	Dinucleotide	A	VIC-GTTTCTCTCCCTGGGACATTT	TATTCATTTGTGTGGGCATG	0.17
D18S72	2	Dinucleotide	B	NED-GCTAGATGACCCAGTTCCC	CTGCAGAAAGGTTACATATCCA	0.05
MFGT21*	8	Dinucleotide	B	FAM-AACTTCAGTAAGATAAGGACC	CCTGAGGCTCTGGACTTTAT	0.10
MFGT22*	Unknown	Dinucleotide	B	VIC-CAACATAGAGAGATTCCATCTC	CGTTAAGTATGATGTTAGCTAG	0.15

* STR and primers developed from *Macaca fuscata* sequence [Domingo-Roura et al. 1997].

** Chromosome locations are based on the NCBI Map Viewer (<http://www.ncbi.nlm.nih.gov/mapview/>) and the Southwest Foundation Rhesus Genome Map (www.snprc.org/linkage/index.html).

*** Based on cloned and whole genome sequences

**** Primer sequences shown in bold font were redesigned from Jan. 2006 rheMac2 draft assembly v.1.0 available on the UCSC Genome Browser (<http://www.genome.ucsc.edu>). Primers for STRs with name prefix starting with "D" are based on human sequences that are fully, or nearly so, conserved in rhesus.

***** Primer concentration used for detection with ABI 3730 capillary electrophoresis instrument.

Appendix B

Allelic diversity and observed (Ho) and expected (He) heterozygosity estimates of three MHC-linked STRs among age cohorts in all three study populations. Estimates for the Indian and Chinese super-SPF age cohorts are italicized and underlined, respectively.

Year of Birth	Loci	Conventional 2001-2008			SPF 2001-2008			super-SPF 2001-2008					
		Ho	He	n	Ho	He	n	Ho		He		n	
2001	D6S291	0.64	0.71	10	0.64	0.69	10	1.00		0.83		3	
	D6S276	0.80	0.84	16	0.85	0.82	14	1.00		0.83		3	
	D6S1691	0.77	0.83	18	0.80	0.85	16	0.50		0.83		3	
2002	D6S291	0.72	0.73	12	0.73	0.74	8	0.88	0.86	0.77	0.79	6	6
	D6S276	0.82	0.85	15	0.85	0.84	11	0.88	0.86	0.84	0.90	8	8
	D6S1691	0.87	0.84	17	0.69	0.73	11	0.63	0.57	0.78	0.77	5	6
2003	D6S291	0.69	0.78	11	0.69	0.75	8	0.54	0.67	0.74	0.73	0.80	7
	D6S276	0.82	0.84	14	0.90	0.86	11	0.85	0.78	0.84	0.75	0.73	10
	D6S1691	0.81	0.84	21	0.72	0.75	10	0.69	0.56	0.77	0.56	0.93	9
2004	D6S291	0.67	0.77	13	0.72	0.76	9	0.63	0.50	0.81	0.73	0.84	9
	D6S276	0.79	0.85	14	0.81	0.82	10	0.79	0.60	0.90	0.73	0.90	12
	D6S1691	0.78	0.83	19	0.78	0.78	15	0.96	0.90	0.93	0.86	0.94	15
2005	D6S291	0.70	0.77	12	0.78	0.81	11	0.75	0.63	0.93	0.79	0.66	0.81
	D6S276	0.83	0.85	14	0.75	0.78	11	0.73	0.63	0.85	0.91	0.84	0.91
	D6S1691	0.82	0.83	19	0.80	0.82	17	0.78	0.67	0.93	0.91	0.85	0.90
2006	D6S291	0.73	0.78	13	0.71	0.77	10	0.74	0.73	0.74	0.79	0.74	0.81
	D6S276	0.81	0.85	14	0.76	0.81	14	0.82	0.75	0.96	0.87	0.77	0.86
	D6S1691	0.84	0.85	17	0.68	0.73	17	0.79	0.70	0.91	0.83	0.72	0.89
2007	D6S291	0.69	0.78	12	0.73	0.80	12	0.74	0.73	0.78	0.76	0.72	0.86
	D6S276	0.81	0.84	14	0.80	0.83	14	0.82	0.76	0.88	0.88	0.83	0.88
	D6S1691	0.82	0.84	20	0.78	0.79	17	0.83	0.79	0.89	0.85	0.77	0.89
2008	D6S291	0.70	0.78	12	0.69	0.79	10	0.72	0.69	0.91	0.74	0.70	0.82
	D6S276	0.82	0.83	14	0.84	0.82	12	0.88	0.84	0.89	0.88	0.83	0.89
	D6S1691	0.81	0.83	20	0.78	0.79	17	0.81	0.75	0.82	0.86	0.75	0.87
Overall		0.77	0.81	21	0.76	0.79	17	0.78	0.71	0.79	0.83	0.76	0.86

Acknowledgments

This study was supported by the California National Primate Research Center base grant (No. RR000169-48), and by grants number RR16023-01, RR005090, RR025781 and RR18144-01 of the National Center for Research Resources (NCRR) of the National Institutes of Health (NIH). This research adhered to the American Society of Primatologists principles for the ethical treatment of nonhuman primates. These animals were managed in compliance with Institutional Animal Care and Use Committee (IACUC) regulations or in accordance with the National Institutes of Health guidelines or the US Department of Agriculture regulations prescribing the humane care and use of laboratory animals. The authors wish to thank two anonymous reviewers for their insightful recommendations for improving this manuscript.

References

- Anderson EC, Thompson EA. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*. 2002; 160:1217–1229. [PubMed: 11901135]
- Andrade MCR, Penedo MCT, Ward T, Silva VF, Bertolini LR, Roberts JA, Leite JPG, Cabello PH. Determination of genetic status in a closed colony of rhesus monkeys (*Macaca mulatta*). *Primates*. 2004; 45:183–186. [PubMed: 15103562]
- Barry PA, Alcendor DJ, Power MD, Kerr H, Luciw PA. Nucleotide sequence and molecular analysis of the rhesus cytomegalovirus immediate-early gene and the UL121–117 open reading frames. *Virology*. 1996; 215:61–72. [PubMed: 8553587]
- Capitanio JP, Lerche NW. Social separation, housing relocation, and survival in simian AIDS: a retrospective analysis. *Psychosom Med*. 1998; 60:235–244. [PubMed: 9625208]
- Capitanio JP, Mendoza SP, Lerche NW. Individual differences in peripheral blood immunological and hormonal measures in adult male rhesus macaques (*Macaca mulatta*): evidence for temporal and situational consistency. *Am J Primatol*. 1998a; 44:29–41. [PubMed: 9444321]
- Capitanio JP, Mendoza SP, Lerche NW, Mason WA. Social stress results in altered glucocorticoid regulation and shorter survival in simian acquired immune deficiency syndrome. *P Natl Acad Sci USA*. 1998b; 95:4714–4719.
- Capitanio, JP.; Mason, WA.; Mendoza, SP.; Del Rosso, LA.; Roberts, JA. Nursery rearing and biobehavioral organization. In: Sackett, GP.; Ruppenthal, G., editors. *Nursery rearing of nonhuman primates in the 21st century*. Kluwer; Plenum: 2006.
- Council, NR. International perspectives: the future of nonhuman primate resources. National Academy Press; Washington DC: 2003.
- Domingo-Roura X, Lopez-Giraldez T, Shinohara M, Takenaka O. Hypervariable microsatellite loci in the japanese macaque (*Macaca fuscata*) conserved in related species. *Am J Primatol*. 1997; 43:357–360. [PubMed: 9403100]
- Ely J, Manis GS, Keeling ME, Stone WH. Maintenance of genetic variability in a specific pathogen-free breeding colony. *Lab Anim Sci*. 1994; 44:207–212.
- Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*. 2003; 164:1567–1587. [PubMed: 12930761]
- Huff JL, Eberle R, Capitanio J, Zhou S, Barry PA. Differential detection of B virus and rhesus cytomegalovirus in rhesus macaques. *J Gen Virol*. 2003; 84:83–92. [PubMed: 12533703]
- Hughes, SS. STRand nucleic acids analysis. University of California; Davis, CA: 2000. Available at (<http://www.vgl.ucdavis.edu/informatics/strand.php>)
- Kanthaswamy S, Smith DG. Use of microsatellite polymorphisms for paternity exclusion in rhesus macaque (*Macaca mulatta*). *Primates*. 1998; 39:135–145.
- Kanthaswamy S, Smith DG. Assessment of genetic management at three captive specific-pathogen-free rhesus macaque (*Macaca mulatta*) colonies. *Journal of Comparative Medicine*. 2002; 52:414–423.
- Kanthaswamy S, Von Dollen A, Kurushima JD, Alminas O, Rogers J, Ferguson B, Lerche NW, Allen PC, Smith DG. Microsatellite markers for standardized genetic management of captive colonies of rhesus macaques (*Macaca mulatta*). *Am J Primatol*. 2006; 68:73–95. [PubMed: 16419121]

- Kanthaswamy, Gill L, Satkoski J, Goyal V, Malladi V, Kou A, Basuta K, Sarkisyan L, George D, Smith DG. The development of a Chinese-Indian hybrid (Chindian) rhesus macaque colony at the California National Primate Research Center (CNPRC) by introgression. *J Med Primatol*. 2008a; 38:86–96. [PubMed: 18715266]
- Kanthaswamy, Satkoski J, George D, Kou A, Erickson BJ-A, Smith DG. Interspecies hybridization and the stratification of nuclear genetic variation of rhesus (*Macaca mulatta*) and long-tailed macaques (*Macaca fascicularis*). *Int J Primatol*. 2008b; 29:1295–1311. [PubMed: 19122840]
- Kanthaswamy S, Tom BK, Mattila AM, Johnston E, Dayton M, Halverson J, et al. A Canine STR Kit and Population Database for Use in Forensic Casework. *Journal of Forensic Science*. *J Forensic Sci*. 2009; 54:829–840.
- Kayser M, Lao O, Saar K, Brauer S, Wang X, Nürnberg P, Trent R, Stoneking M. Genome-wide analysis indicates more Asian than Melanesian ancestry of Polynesians. *Am J Hum Genet*. 2008; 82:194–198. [PubMed: 18179899]
- Khan IH, Ravindran R, Yee J, Ziman M, Lewinsohn DM, Gennaro ML, Flynn JL, Goulding CW, DeRiemer K, Lerche NW, Luciw PA. Profiling antibodies to *Mycobacterium tuberculosis* by multiplex microbead suspension arrays for serodiagnosis of tuberculosis. *Clin Vaccine Immunol*. 2008; 15:433–438. [PubMed: 18077619]
- Lerche LW, Osborn KG. Simian retrovirus infections: potential confounding variables in primate toxicology studies. *Toxicology Pathology*. 2003; 31:103–110.
- Lerche NW, Max PA, Gardner MB. Elimination of type D retrovirus infection from group housed rhesus monkeys using serial testing and removal. *Lab Anim Sci*. 1991; 41:123–127. [PubMed: 1658439]
- Lerche NW, Yee JL, Jennings MB. Establishing specific retrovirus-free breeding colonies of macaques: an approach to primary screening and surveillance. *Lab Anita Sci*. 1994; 44:217–221.
- Lerche NW, Simmons JH. Beyond specific pathogen-free: biology and effect of common viruses in macaques. *Comparative Med*. 2008; 58:8–10.
- Muhl T, Krawczak M, Haaft P, Hunsmann G, Sauermann U. MHC class I alleles influence set-point viral load and survival time in simian immunodeficiency virus-infected rhesus monkeys. *J Immunol*. 2002; 169:3438–3446. [PubMed: 12218167]
- Penedo MCT, Bontrop RE, Heijmans CMC, Offing N, Noort R, Rouweler AJM, deGroot N, deGroot NG, Ward T, Doxiadis GGM. Microsatellite typing of the rhesus macaque MHC region. *Immunogenetics*. 2005; 57:198–209. [PubMed: 15900491]
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000; 155:945–959. [PubMed: 10835412]
- Raymond M, Rousset F. Genepop (version 1.2), population genetics software for exact tests and ecumenicism. *Heredity*. 1995; 86:248–249.
- Roberts RC. Side effects for selection for growth in lab animals. *Livest Prod Sci*. 1979; 6:93–104.
- Roberts JA, Smith DG, Hendrickx A. Managing the rhesus supply. *Science*. 2000; 287:1591. [PubMed: 10733423]
- Sancristobal-Gaudy M, Renand G, et al. Individual traceability of bovine meat using molecular markers. *INRA Prod Anim*. 2000; 13:269–276.
- Satkoski J, George D, Smith DG, Kanthaswamy S. Genetic characterization of wild and captive rhesus macaques in China. *J Med Primatol*. 2007; 37:67–80. [PubMed: 18333917]
- Smith DG. Genetic heterogeneity in five captive specific pathogen-free groups of rhesus macaques. *Lab Anim Sci*. 1994; 44:200–210. [PubMed: 7933964]
- Smith S. Infant cross-fostering in captive rhesus monkeys (*Macaca mulatta*). *Am J Primatol*. 1986; 11:229–237.
- Viray J, Rolfs B, Smith DG. Comparison of the frequencies of major histocompatibility (MHC) class-II DQA1 and DQB1 alleles in Indian and Chinese rhesus macaques (*Macaca mulatta*). *Comp Med*. 2001; 51:555–561. [PubMed: 11924820]
- Vogel T, Norley S, Beer B, Kurth R. Rapid screening for Mamu-A1-positive rhesus macaques using a SIVmac Gag peptide-specific cytotoxic T-lymphocyte assay. *Immunology*. 1995; 84:482–487. [PubMed: 7751033]

- Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. *Evolution*. 1984; 38:1358–1370.
- Wright, S. *Evolution and the genetics of populations vol 4, variability within and among natural populations*. University of Chicago Press; Chicago, IL: 1978.
- Yeh FC, Boyle TJB. Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belg J Bot*. 1997; 129:157.

Fig. 1. Proportions of Indian and Chinese ancestry assigned by STRUCTURE to rhesus macaques in each of the study populations ($N = 737$). Note the predominance of full blood Chinese, Chinese-Indian hybrid and full blood Indian animals in the conventional, regular SPF and Indian super-SPF populations, respectively. All Chinese super-SPF animals were assigned as full-blood Chinese rhesus macaques.

Fig. 2.

Graph displays the frequency of *Mamu-A*01* and *-B*01* positive animals by birth cohort in the Indian and Chinese super-SPF subpopulations from 2003 to 2008.

TABLE I

Comparison of sample size, mean number of alleles averaged across loci (n), mean observed (OH) and expected (EH) heterozygosities averaged across 14 STR loci for each of the study populations' age cohorts from 1993 to 2008 and estimates of genetic differentiation between the pure Chinese and Indian-derived super-SPF subpopulations at the CNPRC from 2002 to 2008 (Data from 7211 of the 7536 study animals are presented below. No data were available for the Chinese portion of the super-SPF population in 2002).

Conventional (N = 5891)				SPF (N = 993)				super-SPF (N = 327)				Indian super-SPF (N = 243)				Chinese super SPF (N = 82)				
Year of Birth	N	n	OH	EH	N	n	OH	EH	N	n	OH	EH	N	n	OH	EH	N	n	OH	EH
1993	74	15.5	0.79	0.84	4	3.71	0.57	0.69												
1994	97	16	0.78	0.84	10	7.29	0.79	0.78												
1995	86	16.21	0.79	0.85	3	3.43	0.71	0.72												
1996	93	15.57	0.78	0.82	2	2.86	0.64	0.77												
1997	121	14.79	0.76	0.81	6	5.43	0.77	0.76												
1998	181	14.93	0.77	0.8	15	8	0.82	0.79												
1999	197	15.79	0.79	0.81	36	11.36	0.75	0.78												
2000	260	14.93	0.77	0.8	60	11.5	0.76	0.78												
2001	367	14.07	0.76	0.79	86	12.86	0.8	0.79	2	3.14	0.93	0.88								
2002	316	15.71	0.79	0.8	51	10.79	0.79	0.78	8	6.29	0.82	0.79	8	5.79	0.81	0.78	-	-	-	-
2003	466	14.36	0.78	0.8	29	10.29	0.78	0.78	13	8.36	0.77	0.8	10	6.42	0.79	0.76	3	3.71	0.73	0.77
2004	622	16.79	0.77	0.8	53	11.43	0.78	0.78	24	12.07	0.83	0.85	10	6.57	0.78	0.78	14	10.07	0.86	0.86
2005	748	17.14	0.77	0.8	75	12.43	0.79	0.8	40	12.93	0.8	0.83	27	7.93	0.75	0.77	13	9.43	0.86	0.84
2006	644	17.29	0.77	0.81	137	12.86	0.78	0.78	77	14.21	0.8	0.83	54	10.00	0.76	0.77	23	11.14	0.85	0.85
2007	795	17.36	0.77	0.81	172	13.64	0.78	0.79	110	14.93	0.78	0.82	92	11.14	0.76	0.78	18	10.57	0.80	0.84
2008	824	16.79	0.77	0.8	254	13.64	0.78	0.78	53	13.57	0.79	0.82	42	9.07	0.70	0.76	11	8.64	0.85	0.84
Overall:	5891	21.71	0.77	0.81	993	16.71	0.78	0.79	327	17.14	0.79	0.83	243	11.92	0.76	0.78	82	15.29	0.84	0.85

TABLE II

Pairwise Fst estimates between cohorts in the three study populations from 1993 to 2008. Underlined estimates* reflect pairwise conventional-SPF, conventional-super-SPF and SPF-super-SPF differentiation over time (* only estimates of samples sizes greater than N = 40 are shown in this table). Bold and italicized fonts are estimates of divergence between the Indian super-SPF animals and the conventional and SPF animals while bold and underlined fonts are estimates of divergence between the Chinese super-SPF animals and the conventional and SPF animals. Italicized estimates in shaded boxes are pairwise differentiation between the pure Chinese and Indian-derived super-SPF subpopulations at the CNPRC from 2003 to 2008 (Data from 7211 of the 7536 study animals are presented below. No data were available for the Chinese portion of the super-SPF population in 2002).

Year of Birth	1993	1997	2001	2005	2008	2001 SPF	2005 SPF	2008 SPF	2005 super-SPF	2006 super-SPF	2007 super-SPF	2008 super-SPF
1997	0.0017											
2001	0.0049	0.0018										
2005	0.0138	0.0118	0.0098						<i><u>0.0009</u></i> <i><u>0.0627</u></i>			
2008	0.0112	0.0094	0.0082	0.0007								<i><u>0.0072</u></i> <i><u>0.0673</u></i>
2001 SPF	0.0061	0.0028	<u>0.001</u>	0.0121	0.0098							
2005 SPF	0.009	0.0059	0.0061	<u>0.0042</u>	0.0038	0.0056			<i><u>0.0106</u></i> <i><u>0.0646</u></i>			
2008 SPF	0.0156	0.0114	0.0087	0.0052	<u>0.0055</u>	0.0074	0.0006					<i><u>0.0059</u></i> <i><u>0.0858</u></i>
2005 super- SPF	0.0101	0.0135	0.0159	<u>0.0041</u>	0.0056	0.0185	<u>0.0099</u>	0.0135	<i><u>0.0814</u></i>			
2006 super- SPF	0.013	0.0164	0.019	0.0069	0.0058	0.0211	0.0094	0.0149	0.0014	<i><u>0.0793</u></i>		
2007 super- SPF	0.0116	0.0125	0.0135	0.0025	0.0028	0.0152	0.0061	0.0092	0.002	0.0035	<i><u>0.0792</u></i>	
2008 super- SPF	0.0138	0.0136	0.0143	0.0037	<u>0.004</u>	0.0158	0.0044	<u>0.007</u>	0.0016	0.0048	0.003	<i><u>0.0959</u></i>

Pairwise Fst estimates between 20 randomly selected, STRUCTURE-assigned full blood Indian and Chinese rhesus macaques pooled across years from the three study populations

TABLE III

	Conventional full blood Chinese	Full blood Chinese SPF	Full blood Chinese super-SPF	Conventional full blood Indian	Full blood Indian SPF
Full blood Chinese SPF	0.0018				
Full blood Chinese super-SPF	0.0118	0.0185			
Conventional full blood Indian	0.0465	0.0493	0.0741		
Full blood Indian SPF	0.0627	0.0645	0.0798	0.0011	
Full blood Indian super-SPF	0.0648	0.0674	0.0864	0.0027	-0.0064