

ORIGINAL ARTICLE

Population dynamics of rhesus macaques and associated foamy virus in Bangladesh

Mostafa M Feeroz¹, Khanh Soliven², Christopher T Small², Gregory A Engel^{3,4}, M Andreina Pacheco⁵, JoAnn L Yee⁶, Xiaoxing Wang², M Kamrul Hasan¹, Gunwha Oh³, Kathryn L Levine², SM Rabiul Alam¹, Karen L Craig², Dana L Jackson², Eun-Gyung Lee², Peter A Barry⁶, Nicholas W Lerche⁶, Ananias A Escalante⁵, Frederick A Matsen IV², Maxine L Linial² and Lisa Jones-Engel³

Foamy viruses are complex retroviruses that have been shown to be transmitted from nonhuman primates to humans. In Bangladesh, infection with simian foamy virus (SFV) is ubiquitous among rhesus macaques, which come into contact with humans in diverse locations and contexts throughout the country. We analyzed microsatellite DNA from 126 macaques at six sites in Bangladesh in order to characterize geographic patterns of macaque population structure. We also included in this study 38 macaques owned by nomadic people who train them to perform for audiences. PCR was used to analyze a portion of the proviral *gag* gene from all SFV-positive macaques, and multiple clones were sequenced. Phylogenetic analysis was used to infer long-term patterns of viral transmission. Analyses of SFV *gag* gene sequences indicated that macaque populations from different areas harbor genetically distinct strains of SFV, suggesting that geographic features such as forest cover play a role in determining the dispersal of macaques and SFV. We also found evidence suggesting that humans traveling the region with performing macaques likely play a role in the translocation of macaques and SFV. Our studies found that individual animals can harbor more than one strain of SFV and that presence of more than one SFV strain is more common among older animals. Some macaques are infected with SFV that appears to be recombinant. These findings paint a more detailed picture of how geographic and sociocultural factors influence the spectrum of simian-borne retroviruses.

Emerging Microbes and Infections (2013) 2, e29; doi:10.1038/emi.2013.23; published online 22 May 2013

Keywords: anthropogenic change; Bangladesh; emerging infectious diseases; macaques; simian foamy virus

INTRODUCTION

The discovery that human immunodeficiency virus evolved from enzootic nonhuman primate (NHP)-borne viruses has stimulated an interest in other NHP-borne retroviruses and their transmission to humans.¹ Human societies critically influence the ecological contexts in which zoonotic transmission occurs.^{2–5} Nowhere is this more evident than in South and Southeast Asia, where the world's densest human populations are situated close to some of the most diverse NHP fauna. Widespread destruction of NHP habitat, rapidly developing infrastructure, as well as social and economic forces, are changing the geographic landscape and bringing previously separate NHP populations into contact with each other, as well as into contact with humans. As a result, these new geographic relationships are likely to bring together various infectious agents harbored by these NHP populations, in new combinations, which may increase the potential for the emergence of zoonotic disease.^{3,6}

At the epicenter of these forces lies Bangladesh, home to millions of people and thousands of NHP. The major river systems and its tributaries of Bangladesh are its most prominent geographic feature and have for centuries provided the means for vital trade and the movement of humans and their belongings throughout the Indian subcontinent.

The human population density in Bangladesh is among the highest in the world (<https://www.cia.gov/library/publications/the-world-factbook/index.html>). Most of the 161 million people live in crowded urban or peri-urban areas. While predominately a Muslim country, ~10% of the population identify themselves as Hindu. Hindus consider NHP sacred, the 'children' of the monkey deity Hanuman. Because of these cultural/religious beliefs, NHP populations are often found in Hindu communities. Muslims takes a less sanguine view of monkeys, and since they are not hunted for bush meat, their presence is not always welcomed.⁷

While Bangladesh contains a rich NHP fauna including five species of macaque (Rhesus (*Macaca mulatta*), Pig-tail (*M. leonine*), Stumptail (*M. arctoides*), Long-tail (*M. fascicularis*) and Assamese (*M. assamensis*) macaques) and three species of langur (Hanuman (*Semnopithecus entellus*), Capped (*Trachypithecus pileatus*) and Phayre's langur (*Trachypithecus phayrei*)), only the rhesus macaque is still found in significant numbers though many of these free-ranging rhesus populations are now restricted to urban and peri-urban areas. With few exceptions the other NHP are found only in the isolated forest patches that remain along the borders with India and Myanmar.^{8,9} In addition to free-ranging rhesus populations, pet and performing monkeys can be found in Bangladesh. The latter are often associated with nomadic people

¹Jahangirnagar University, Savar, Dhaka 1342, Bangladesh; ²Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA; ³University of Washington, National Primate Research Center, Seattle, WA 98195, USA; ⁴Swedish Medical Center, Seattle, WA 98122, USA; ⁵Arizona State University, Tempe, AZ 85287, USA and ⁶University of California, Davis, CA 95616, USA

Correspondence: L Jones-Engel

E-mail: jonesengel@wanprc.org

Received 21 February 2013; revised 11 March 2013; accepted 14 March 2013

found throughout South Asia (known as the Bade, or in other countries as Qalandar) who train these rhesus macaques to perform for audiences.¹⁰ Relatively little is known about these people, their patterns of travel or their practices of animal husbandry.

Bangladeshi NHP come into contact with humans in several contexts: urban, shrine/temple, performing and wild, free-ranging populations in forested areas.^{11,12} Importantly, these NHP populations are not isolated from one another; they are dynamic and overlapping so that they can potentially intermix. Animals taken from the wild or urban areas become pets. Pets are released or escape into the wild. Abundant human/NHP contact makes Bangladesh a 'natural laboratory' for the study of cross-species transmission of infectious agents.¹³

Bangladeshi NHP harbor at least three retroviruses. While simian betaretrovirus, previously known as simian retrovirus D (SRV-D), and simian T-cell lymphotropic virus (Jones-Engel, unpublished data) are found in only a few, isolated populations of some species, simian foamy virus (SFV) is ubiquitous. SFV is a member of a subfamily of retroviruses, the *Spumaretrovirinae*, which are distinct from other retroviruses by their mode of replication. For example, the orthoretroviruses such as human immunodeficiency virus are RNA viruses that reverse transcribe their genomes into DNA when they infect new cells. In contrast, spumaretroviruses are DNA viruses that package RNA and reverse transcribe their genomes during assembly and/or release. In this manner, they resemble the hepadnaviruses such as hepatitis B virus.¹⁴ Foamy virus (FV) naturally infect cats, cows, some horses and all species of NHP tested. Phylogenetic analyses suggest that foamy-like viruses evolved over 400 million years ago, making them the oldest known animal viruses.¹⁵ Although FV have been shown to be highly cytopathic in some cell types in culture, such as fibroblast lines, there is no evidence that they cause disease, either in their natural hosts, or in other mammalian species.¹⁶ FV replicate only in the differentiating superficial epithelial cells of the oral mucosa *in vivo* and are efficiently secreted into saliva in natural hosts.^{17,18} However, most organs of FV-infected animals contain latent proviral DNA. Proviral DNA can readily be found in peripheral blood mononuclear cells.^{19,20} FV can be considered 'perfect' viruses in that they are transmitted very efficiently without inducing pathological changes that compromise the host.¹⁴ Since essentially all adult NHP are infected with FV, it can be considered part of the normal host flora.

SFV is highly prevalent and is efficiently transmitted through saliva among rhesus macaques (up to 100% of free ranging macaques are infected by age 3).²¹ Some human infections have been documented, but no human-to-human transmission has been reported.²² As such, it is an appropriate agent for studying how parenterally transmitted viruses move within and between NHP populations, and from NHP to human populations. Several research groups have examined zoonotic transmission of SFV from NHP to humans. Zoonotic transmission has been documented in monkey handlers and laboratory technicians and veterinarians, bush meat hunters in Africa and temple workers in South and Southeast Asia that live near monkeys.^{23–27}

In the present study, we initially screened for SRV-D and SFV; we then characterized SFV sequences from 164 rhesus macaques from six sites (four urban and two forested) in Bangladesh, as well as from performing monkeys that were sampled at several locations around the country. The differences in the SFV *gag* gene present in different populations allowed us to delineate strains. We used sequence analysis to investigate the following questions: (i) can rhesus macaques be infected with more than one SFV strain at a time? (ii) do SFV strains recombine in naturally infected rhesus macaques? and (iii) do both

natural geographic barriers and human influences contribute to SFV strain variation among Bangladesh rhesus macaques? Our data contribute to an understanding of the ecology of SFV in macaques at the human/primate interface. Further, when considered in conjunction with our accompanying paper for this project (Engel *et al.*, in review), which characterizes the zoonotic transmission of SFV in these populations, our results enhance our understanding of the viral genetic factors that can influence zoonotic transmission.

MATERIALS AND METHODS

Ethics statement

These research protocols have been reviewed and approved by the University of Washington Institutional Animal Care and Research Committee (4233-01). This research adhered to the legal requirements of the countries in which the research was conducted. Biological samples were shipped internationally under a Convention on International Trade in Endangered Species permit.

Sampling sites

In the present study, free-ranging populations of rhesus macaques were sampled in six geographically distinct areas of Bangladesh as shown in Figure 1. These six sites were selected following a country-wide survey designed to locate NHP populations that were synanthropic (species that thrive in human-altered environments).²⁸ Dokhola is a protected forest area in north central Bangladesh. Permanent human establishments are few in this 8500-hectare preserve, though the area is heavily used by villagers living on the periphery, who collect resources from the forest. It is estimated that at least several hundred rhesus macaques in two or three large groups range throughout the protected area. Bormi and Dhamrai are densely populated urban areas located 150 and 100 km respectively south of Dokhola. These two urban areas are approximately 50 km apart. No natural patches of forest remain in these urban areas, though isolated large trees and numerous fruiting trees and bushes are maintained by local people for cultural and economic benefits. Both Bormi and Dhamrai have substantial Hindu populations, which are concentrated in specific locations. Many of the macaque sleeping trees are found within these Hindu areas, though the monkeys range throughout the urban sites.

Narayanganj is a major port city located southeast of the capital, Dhaka. The macaque population in this area has dwindled significantly over the past 5 years as the human density has increased substantially, resulting in nearly complete loss of any large trees. Dokhola, Bormi, Dhamrai and Narayanganj are all located northeast of the Padma River. This river is a formidable geographic barrier to natural macaque migration.

Charmaguria is located southwest of the Padma River. This urban site still contains a large number of forested areas that can and do support a robust population of macaques. Karamjal, also southwest of the Padma River, is a forestry station located at the northern tip of the Sundarbans, the world's largest mangrove forest. This station is one of the tourist entry points into the Sundarbans and a small troop of rhesus macaques has become habituated to visitors at this site. The Sundarbans are prime habitat for rhesus macaques, though there are no comprehensive census data for macaques in this area estimates range to the thousands.²⁸

The nomadic Bade with their performing macaques were found as we traveled between sampling sites and during a visit to a village in western Bangladesh that is used as an intermittent base for the Bade. The Bade are readily identifiable by the distinctive design of their tents, usually erected by major roadsides on the edges of urban areas,

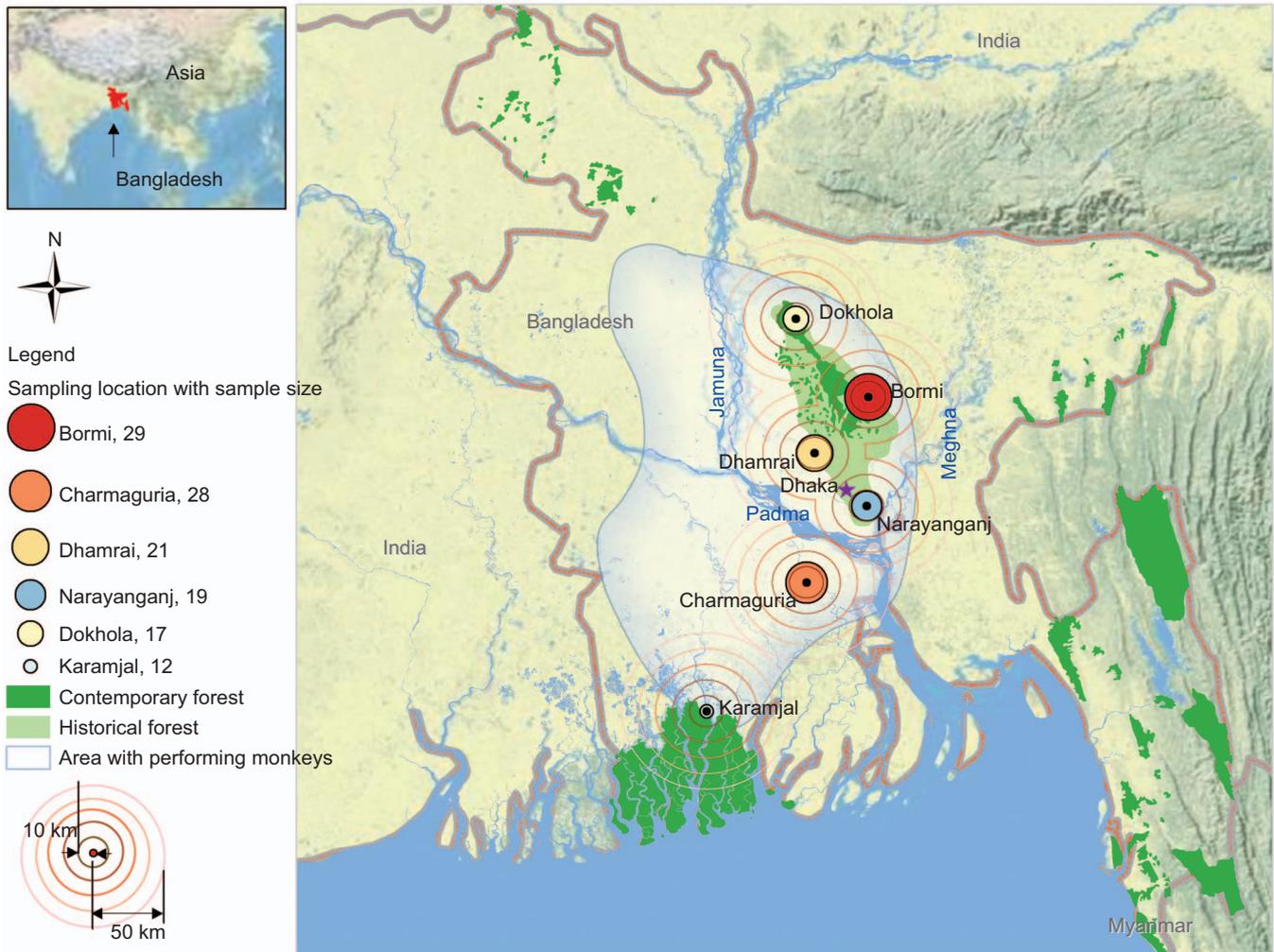


Figure 1 Rhesus macaque sampling site locations including relative geographic proximity and historical and extant forest patches. Rhesus macaques, *Macaca mulatta*, were sampled from six geographically distinct sites in Bangladesh ($n=126$). Concentric circles in 10 km intervals from the epicenter of the sites are shown. Existing forest areas are noted in dark green and historical forest cover is shown in lighter green. Natural macaque migration would have been possible between sites that were connected by forested areas. Nomadic Bade people travel throughout the subcontinent with their performing macaques. We sampled 38 of these performing monkeys from multiple locations within the shaded light blue area.

together with the specially crafted wooden boxes that contain their performing monkeys. Figure 1 indicates the range in which we encountered the performing monkeys and their nomadic owners.

Sample collection from free-ranging monkeys in Bangladesh

Our group has trapped, darted, sampled and released free-ranging rhesus macaques (*M. mulatta*) in Bangladesh since 2006. Capture and sampling methods are detailed in Jones-Engel *et al.*²⁹ Rhesus trapping, darting and sampling only took place after we had met with village stakeholders and explained our proposed project and long-term objectives. We only sampled at sites where we had support of the local residents. At sites where we could identify more than one group of rhesus macaques, we attempted to obtain samples from individuals from each group. Briefly, free-ranging macaques were trapped in portable cages (measuring approximately 2.5 m³) fabricated at the Washington National Primate Research Center. Captured macaques were hand-injected intramuscularly with 3 mg of Telazol/kg of estimated body weight. All anesthetized macaques were given a physical exam. While under anesthesia, we performed

morphometric analysis and collected biological samples. Using sterile technique and universal precautions, 5–10 mL of whole blood was collected via venipuncture from the femoral vein and stored in ethylenediaminetetraacetic acid tubes (VWR, Radnor, PA, USA) on cold packs. A portion of the whole blood was aliquoted and the remaining sample centrifuged to separate plasma from the cells and then aliquoted. Blood samples were kept on cold packs and then transferred to long term -80°C storage. Each macaque's dental formula was recorded for age assessment and animals were classified as 'adult' versus 'juvenile' based on the observed tooth eruption sequence. Each animal was given a unique identifier via a subcutaneous Avid microchip (Norco, CA, USA), to allow future longitudinal sampling. Following sample collection, the anesthetized animals were placed in recovery cages and allowed to recover fully before being released as a group back into their home range.

The same anesthetic and biological sampling protocols were used for animals that were darted and for the performing monkeys. We darted animals using the Telinject (Agua Dolce, CA, USA) blowpipe system. To minimize risk of injury, young animals and pregnant

females were not darted. Prior to anesthetizing and sampling the performing monkeys, we discussed the procedures and our research objectives with the monkeys' owners and collected information on the age of the animal and where it was originally trapped.

Microsatellite analyses

Microsatellite assay. Genomic DNA was extracted from 200 μ L aliquots of whole blood collected in ethylenediaminetetraacetic acid using the QIAamp DNA Blood Mini kit (Qiagen GmbH, Hilden, Germany). Animals that had tested positive for SFV were assayed using a 13-autosomal microsatellites panel (D01S548, D03S1768, D05S1457, D06S501, D07S794, D08S116, D09S921, D04S2365, D10S611, D11S2002, D13S765, D15S823 and D18S537).³⁰ Polymerase chain reaction (PCR) was carried out in a 15 μ L volume reaction using 25–35 ng of total genomic DNA, 0.5 mM of each set of primers and 7.5 μ L of 2X master mix (3 mM MgCl₂, 400 μ M of each dNTP and 50 U/mL Taq DNA polymerase; Promega, Madison, WI, USA). PCR conditions were: a partial denaturation at 95 °C for 10 min and 30 cycles with 45 sec at 95 °C, 30 sec at 52 °C–58 °C and 30 sec extension at 72 °C, and a final extension of 30 min at 72 °C was added in the last cycle. Fluorescently labeled³⁰ PCR products were separated on an Applied Biosystems 3730 capillary sequencer and scored using Gene Marker v1.95 (SoftGenetics LLC, State College, PA, USA).

Population genetic analysis. DNA fragment sizes, allele frequencies, and observed (H_o) and expected heterozygosity (H_e) assuming equilibrium conditions were determined using Cervus v3.0.3.³¹ Missing data (no amplifications) were reported by loci but not considered in the analyses. We used a model-based clustering method for multilocus genotype data (13 microsatellites) to infer population structure and assign individuals to populations as implemented in Structure v2.3.3.^{32,33} The model assumes K populations (where K is inferred from the data) and each of which is characterized by a set of allele frequencies at each locus. Individuals are probabilistically assigned to populations, or jointly to two or more populations if their microsatellite data indicate that they are admixed. We evaluated the observed genetic diversity at different K values ($K=2-10$), and each K value was run independently 10 times with a burn-in period of 10 000 iterations followed by 10 000 iterations. Because individuals may have mixed ancestry, we applied the admixture model that assumes correlated allele frequency, in which each individual draws some fraction of his/her genome from each of the K populations. Structure Harvester Web v0.6.92³⁴ was used to compute Delta K values. R 2.15.1 (<http://www.R-project.org>) and ggplot2 (v0.9.1)³⁵ were used to graphically display the results.

Testing for foamy virus antibodies

Western blot (WB). All antibody testing was done using plasma samples. Rhesus macaque fibroblast Telo-RF (Tf) cell line was infected with SFV *M. mulatta* or *M. fascicularis*²¹ or left uninfected. After 48 h, cells were lysed in 1 \times sodium disulfide sample buffer (12.5% 50 mM Tris-HCl (pH 6.8), 10% glycerol, 2% sodium disulfide, 5% 2-mercaptoethanol, 0.01% bromophenol blue) and samples electrophoresed through 10% polyacrylamide gels. After transfer of proteins to polyvinylidene difluoride membranes (Millipore, Billerica, MA, USA), the membranes were blocked in phosphate-buffered saline+10.05% Tween-20 (PBST) containing 2% nonfat dry milk, and rinsed with water and left to dry. Strips were made, each containing three lanes (molecular weight marker, uninfected Tf and infected Tf cell lysates) and primary antibody (monkey plasma) was added at a 1:200 dilution and incubated

overnight at 4 °C. Strips were washed three times in PBST before adding secondary goat anti-monkey horseradish peroxidase (Santa Cruz Biotechnology, Santa Cruz, CA, USA) at a 1:5000 dilution for 1 h at room temperature (RT). Strips were washed three more times before rinsing with water and applying 3,3',5,5'-tetramethylbenzidine stabilized substrate for horseradish peroxidase (Promega) to visualize bands.

Enzyme-linked immunosorbent assay (ELISA). The general ELISA method has been described previously.³⁶ Ninety-six-well ELISA plates (Nalge Nunc, Rochester, NY, USA) were coated with 50 ng/well GST-Gag diluted in carbonate buffer (10 mM Na₂CO₃, 40 mM NaHCO₃, pH 9.4) and incubated overnight at 4 °C. Heat-inactivated plasma were serially diluted from 1:50 to 1:1600 in blocking buffer and placed on the plates for 1 h at RT. Plates were washed three times with PBST. The plates were then incubated with a 1:5000 dilution of goat anti-monkey horseradish peroxidase secondary antibody for 1 h at RT. Plates were washed three times with PBST and 100 μ L 3,3',5,5'-tetramethylbenzidine liquid substrate system (Sigma-Aldrich, St Louis, MO, USA) was added to each well and incubated at RT for 8 min. One hundred microliter 1 N H₂SO₄ was added to stop the reaction. Plates were read at 450 nm.

Testing for SRV-D antibodies

We and others have successfully used multiplex microbead immunoassays which require only a small volume (less than 5 μ L) for the detection of antibodies to panels of viruses including SRV-D in non-human primates.³⁷⁻³⁹ Using commercially available reagents from Charles River Laboratories (Wilmington, MA, USA), plasma samples were reacted with a liquid suspension array of microbeads coated with an SRV-D or other test or control antigens. After washing, the beads were reacted with biotinylated detector antibodies and then, after additional washing, Streptavidin-R-Phycoerythrin E. After a final washing to remove all unbound material, the beads were analyzed in a Bioplex reader (BioRad, Hercules, CA, USA). By monitoring the spectral properties of the beads and amount of Streptavidin-R-Phycoerythrin E fluorescence associated with each, the presence or absence of antibody to SRV-D antigens was determined.

FV PCR

DNA isolation from whole blood. Total genomic DNA was extracted from the whole blood of macaques using the QIAamp DNA blood mini kit (Qiagen, Valencia, CA USA) according to manufacturer's protocol. Briefly, 200 μ L of whole blood was mixed with proteases and equal volume of lysis buffer AL and 100% ethanol. The mixture was loaded onto columns and centrifuged. After two washes with AW1 and AW2 buffers, DNA was eluted twice with 200 μ L of elution buffer AE.

PCR amplification of gag fragments from genomic DNA. A total of 10 μ L of the isolated DNA was used for PCR amplification of gag fragments by a two round PCR. First round PCR with gag primers primer sequence: G1 (+) (nucleotide (nt) 114–134, taken from SFVmac 21010 sequence NC_001364 NC_001736) 5'-AGG ATGG TGG GGA CCA GCT A-3'; G2 (-) (nt 1451–1433) 5'-CAG GAA GAG GTA CTC GGG G-3', was previously described.⁴⁰ The reactions were denatured at 95 °C for 3 min, followed by two cycles of 95 °C for 30 sec, 38 °C for 1 min, 72 °C for 1 min, then followed by 25 cycles of 95 °C for 30 sec, 52 °C for 1 min and 72 °C for 1 min. Water was used as a negative control. 2 μ L of the first round PCR product was used as

template for the second round PCR. We used *gag* primers G3 (+) (nt 190–212) 5'-CAA CCT AGA TGG AGA GCT GAA GG-3'; G4 (-) (nt 1425–1406) 5'-GGG GAC AAA CTA CGG CTG GG-3', to amplify the *gag* fragment. The reactions were denatured at 95 °C for 3 min, followed by 35 cycles of 95 °C for 30 sec, 52 °C for 1 min and 72 °C for 1 min. The expected size of the PCR product for *gag* is 1235 bp.

Cloning and sequencing of SFV *gag*

The FV PCR products were gel purified using a QIAquick gel extraction kit (Qiagen) according to manufacturer's protocol. The purified fragments were ligated into the TOPO-TA cloning vector (Invitrogen, Carlsbad, CA, USA) and transformed into *Escherichia coli* TOP10 competent cells (Invitrogen). Multiple colonies were individually selected from each *gag* sample and plasmid DNA was purified and sequenced. Sequencing was carried out using M13 forward and reverse primers, as well as internal forward and reverse primers for *gag* by GeneWiz Inc. (Seattle, WA, USA). SFV *gag* sequences generated in the current study have been deposited in GenBank with the following accession numbers KC991378 - KC992185.

SRV-D DNA testing

Real-time PCR was used to amplify the DNA extracted from whole blood to detect the presence of SRV-D (serotype 1) proviral DNA. SRV-D primers and probes were derived from highly conserved sequences within the envelope gene of SRV-D serotype 1, 2, 3, 4 and 5 viral genomes.⁴¹ Primers: 5'-CTG GWC AGC CAA TGA CGG G-3', 5'-CGC CTG TCT TAG GTT GGA GTG-3'; Probes: 6FAM—TCA CTA ACC TAA GAC AGG AGG GTC GTC A—TAMRA, 6FAM—TCC TAA ACC TAA GAC AGG AGG GCT GTC A—TAMRA. In addition, primers and probes designed to amplify the housekeeping gene oncostatin M were multiplexed into the reaction as a control for the presence of amplifiable DNA. 15 ul of isolated DNA was used in the following cycling conditions: 1 cycle for 2 min at 50 °C, 1 cycle for 10 min at 95 °C, 55 cycles for 15 s at 95 °C and 1 cycle for 1 min at 60 °C.

FV biocomputational analysis

Sequences were aligned using MUSCLE v3.8.31;⁴² the relatively small number of apparent insertion/deletion events led to an alignment that was <0.3% gap. Amino-acid translation was performed using EMBOSS v6.4.0.0.⁴³

A large phylogenetic tree was built from the entire *gag* nucleic acid sequence alignment using FastTree v2.1.3 with a Jukes–Cantor model.⁴⁴ The tips of this tree were colored by sampling location using a python script together with the Bio.Phylo⁴⁵ Biopython⁴⁶ library. This colored tree was rendered using Archaeopteryx v0.957 beta (<http://www.phylosoft.org/archaeopteryx/>). Additionally, a NeighborNet splits network was constructed using SplitsTree⁴⁷ using LogDet distances and the NeighborNet algorithm.

Nucleic acid sequences were clustered using UCLUST v1.1⁴⁸ at a 97.5% identity, producing a total of 32 clusters. Three of these clusters (containing five sequences from animals MBG132, MBG189 and MBG110) clustered separately from other larger clusters as a result of G to A hypermutation (a more complete analysis in preparation for publication). When the putative hypermutation sites within these clusters were removed from the alignments, these sequences no longer clustered separately. Thus, these sequences were considered part of the clusters with which they clustered when these hypermutation sites were excluded; we define a viral 'strain' in the context of this study to be one of these clusters. For each macaque, the number of unique

strains to which that animal's sequences had been assigned was computed. Based on these counts, macaques were identified as either singly or multiply infected.

When the resulting clusters were ordered by size, there were six that were substantially larger than the rest. Each of these six was uniquely associated with a geographic region, except for Bormi, which had two clusters associated with it. We named each of these strains by the geographic location where it dominated; in the case of Bormi, the two clusters were arbitrarily designated Bormi1 and Bormi2. This assignment of viral strains is used in the rest of the present paper, furnishing the labeling in Figure 2.

An iterative method was employed to detect recombinant strains among these major viral strains. Two representative sequences were selected from each strain (except four for the Dhamrai strain) for recombination analysis with GARD,⁴⁹ using the HKY+ Γ model. The trees generated from each non-recombinant alignment segment were inspected, and the strain with highest observed discord between trees was removed. This process was repeated until no recombination was detected within the alignment using GARD. This analysis removed both strains from the Bormi region, and left four strains found in the Charmaguria, Dhamrai, Dokhalo and Karamjal monkeys, which were identified as non-recombinant. Each step of the GARD results was confirmed by a PHI test analysis⁵⁰ as implemented in SplitsTree;⁴⁷ this test showed identical results with a *P*-value of less than 0.05 for each step. These four identified non-recombinant strains will be referred to as the 'parental' strains, which forms a subset of the above-defined 'core' strains.

The parental strains were used as genotype representatives for cBrother v2.0⁵¹ analyses, where a representative from each cluster was evaluated for recombinant relationships to these potential parental strains. Two independent cBrother runs of 1.1 million generations were run for each cluster's representative sequence, with the initial 10% discarded as burn-in, and sampling every 1000 generations. Convergence was assessed using the Gelman–Rubin diagnostic included with the cBrother distribution. Portions of evaluated sequences which were found to be descendant from a particular parental strain at 90% posterior probability or higher for at least 200 contiguous base pairs were identified as being partially descended from that parental strain.

For each of Bormi1 and Bormi2, which we had identified as recombinant strains in our GARD and PHI test analysis, two representative sequences were aligned with two representatives from each of the parental strains. These alignments were split at the corresponding breakpoints identified from the above cBrother analysis (304 and 496, respectively). Each of the resulting four alignments was analyzed using PhyML 3.0⁵² under a HKY85 model with 100 bootstrap replicates and trees were generated using FigTree v1.3.1 (<http://tree.bio.ed.ac.uk/software/figtree>). The resulting trees put the recombinant sequences in the same positions as those inferred by cBrother.

Within-host pairwise distances were computed using R v2.15.1 (<http://www.R-project.org>) and the ape⁵³ package's *dist.dna* function with the 'raw' model (normalized Hamming distance). Figure 3 was also created using R and ape.

Construction of a 6 nucleotide deletion mutant of *gag* in SFV-1

To create a *gag* gene containing the 6 nucleotide deletion in pSFV-1,³⁶ a two-step PCR protocol was used. The two primer sets used to PCR amplify overlapping fragments were: F1 (5'-GAG AGC TGA AGG TCG GGC TG-3') R1 (5'-GGC CTG GTG ATG GTA AGG GTG AAG CCA-3') and F2 (5'-GGC CTG GTG ATG GTA AGG GTG

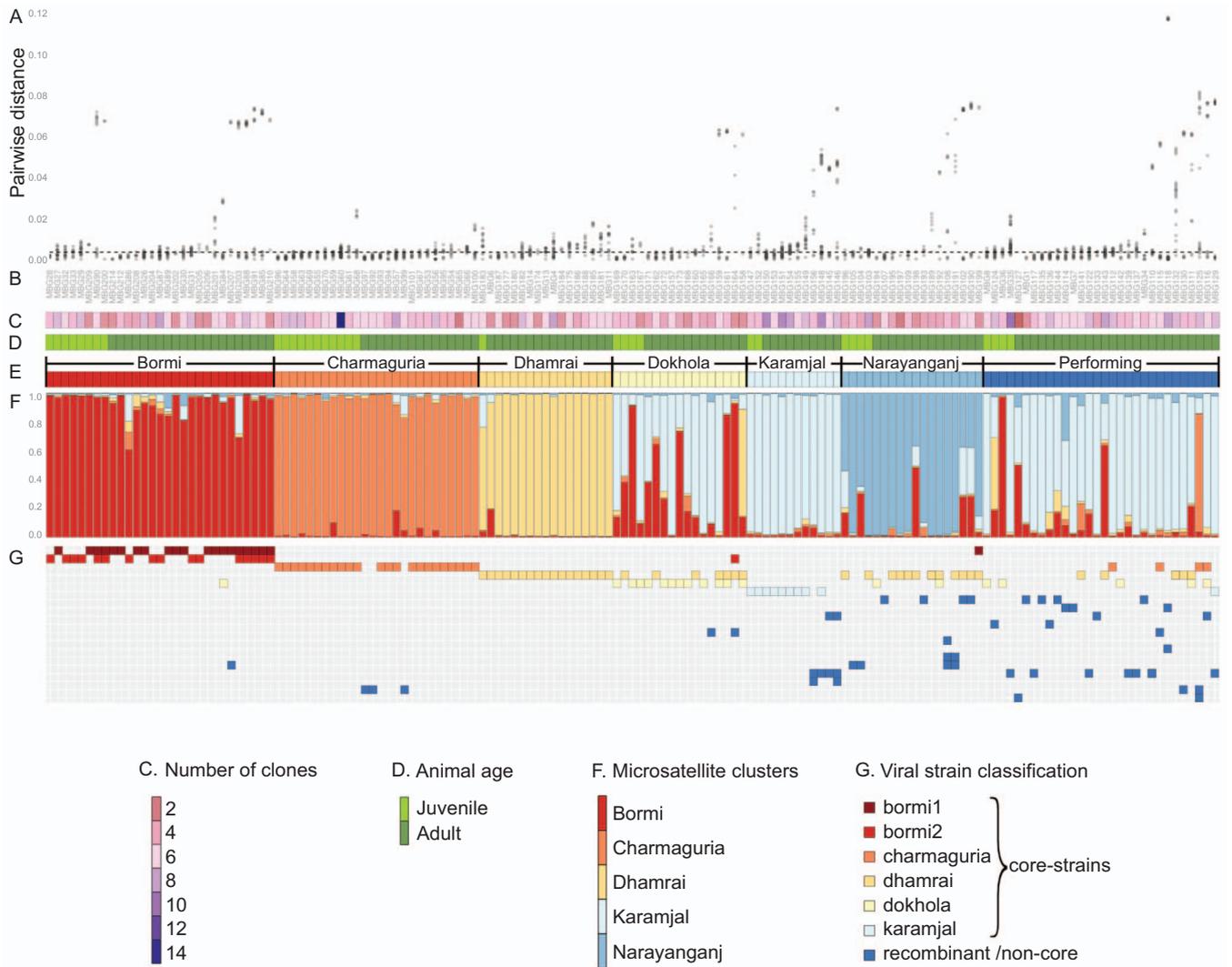


Figure 2 Relationship between monkey and viral population structures. Each animal is represented by a single column spanning the figure's sections. Columns are sorted by viral diversity, animal age, and sampling population. **(A)** Normalized Hamming distance for each pair of clones from each animal, representing intra-host viral diversity. The dotted line represents expected distance (due to PCR error) between distinct PCR products of a single DNA molecule. **(B)** The identifier of the animal to which each column corresponds. Respectively, **(C)**, **(D)** and **(E)** represent number of viral clones, animal age class and sampling population for each monkey. **(F)** Microsatellite clusters as inferred by STRUCTURE; admixture between genotype clusters is represented by bar height. Microsatellite colors are assigned according to an observed correlation with sampling location. **(G)** Viral strains cloned from each animal, classified as described in the text. All sampled monkeys are represented except for one FV negative infant monkey (MBG211).

AAG CCA-3') R2 (5'-CAC GAG CCT CTG ACT GCG GTT-3'). The outer primers F1 and R2 were then used to PCR amplify a 1.4 kb fragment (containing the deletion) using the overlapping fragments as template. The resulting PCR fragment and pSFV-1 vector were digested with BsaI and BsaBI, gel purified and ligated together. Several individual clones were sequenced to confirm the 6 nucleotide deletion.

Assay of SFV-1 and deletion mutants

Transient transfection of Tf cells with pSFV-1 or the *gag* deletion mutant clones described above were performed following Qiagen's Polyfect transfection protocol. Cells were harvested after 48 h and lysed for WB analysis, to detect Gag proteins. At the same time, extracellular supernatants from transfections were collected for an simian foamy-activated beta-galactosidase assay and clarified of cell debris by

low-speed centrifugation.^{21,54} Simian foamy-activated beta-galactosidase cells are baby hamster kidney cells containing a β -galactosidase gene under the control of the SFV-1 long terminal repeat. When cells are infected with SFV, the viral *tas* gene activates the β -gal gene. To measure infectivity, a β -galactosidase expression assay was performed as previously described.⁵⁴

RESULTS

M. mulatta populations used for sampling

Figure 1 is a map of Bangladesh with all the sampling sites indicated. A total of 164 rhesus macaques were sampled for the study, including 126 free ranging animals at six distinct sites, representing between 4% and 35% of the estimated population at each site (Table 1). Additionally, 38 performing macaques were sampled at five locations. Estimates of the total number of performing macaques in Bangladesh

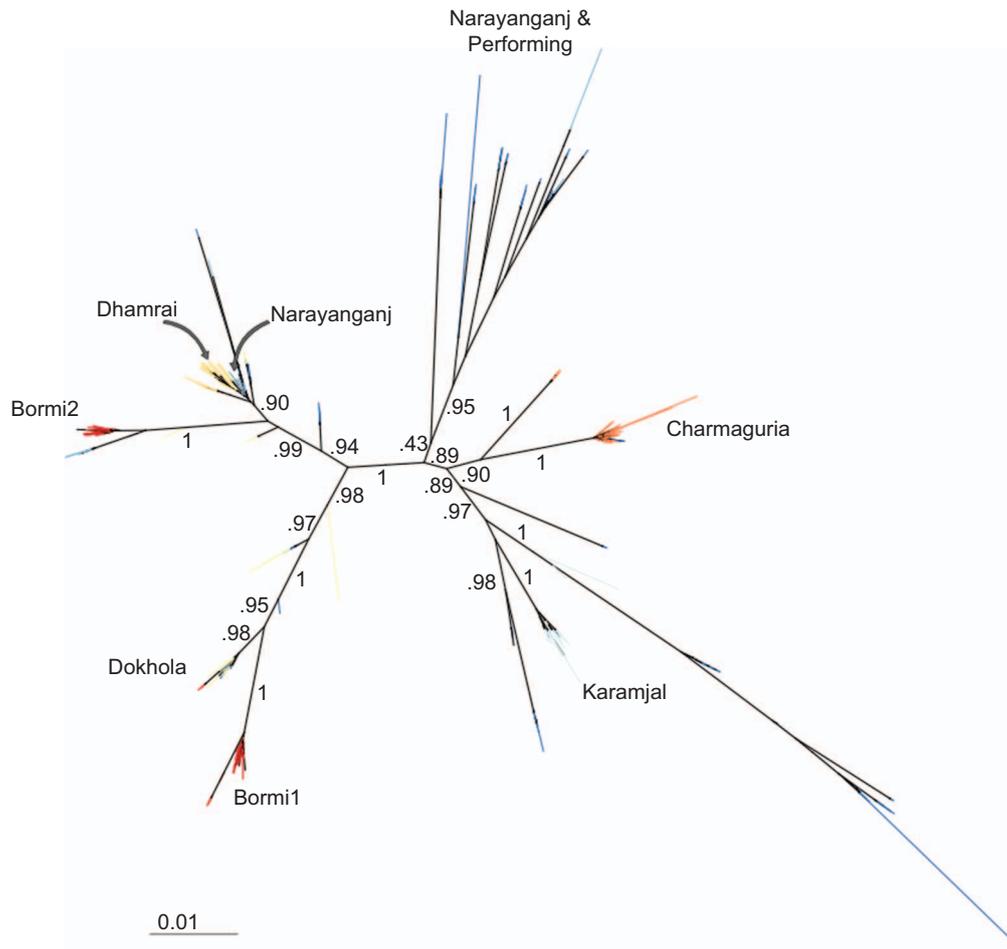


Figure 3 Foamy virus sequence diversity overview. Phylogenetic tree built from all sequenced *gag* nucleotide clones using FastTree using the Jukes–Cantor sequence evolution model. Due to recombination, this tree should not be interpreted as an evolutionary history, but rather as an indication of the clustering seen in the sequence data. The edges of the tree are labeled with SH-like local supports.

vary widely, from 500 to 5000. Because performing monkeys and their owners are nomadic and we have no *a priori* knowledge of the precise origins of the monkeys; for the purposes of our SFV *gag* diversity analyses we treat all 38 performing monkeys in this dataset as a single population.

Population structure of macaques using microsatellite DNA

Analysis of microsatellite DNA revealed five distinct genetic clusters, explained in part by geographic separation (Figures 2E and 2F). Supplementary Table S1 shows the number of alleles found for the 13 loci in different locations in Bangladesh and the observed heterozygosity (H_o) and expected heterozygosity (H_e) assuming equilibrium

conditions. All the loci assayed were polymorphic with varying numbers of alleles (from 4 to 19, see Supplementary Table S1 for details). Out of the 13 loci, four were not in Hardy–Weinberg equilibrium. Using a Bayesian clustering algorithm,⁵⁵ the posterior probability for different numbers of clusters ($K=2-10$) was computed and the clustering patterns obtained with $K=5$ were associated with the highest delta K values ($\Delta K = \text{mean}(|L'(K)|)/\text{sd}(L(K))$) (structure harvester, available from: http://taylor0.biology.ucla.edu/struct_harvest/).³⁴

Each of the five microsatellite clusters identified in this analysis was found to be most closely associated with a particular macaque population, these being the Bormi, Charmaguria, Dhamrai, Karamjal and Narayanganj populations. Of these populations, Narayanganj stood

Table 1 Each area, estimated population size and number of monkeys sampled by age class and sex

Location	Dokhola	Bormi	Dhamrai	Narayanganj	Charmaguria	Karamjal	Performing
Years sampled	2011	2007; 2009; 2011	2006; 2011	2010; 2011	2008; 2009	2011	2006; 2008; 2010; 2011
Context	Forest reserve	Urban	Urban	Urban	Urban	Forest reserve	Performing
Estimated population size ^a	>300	102	102	55	210	>300	500–5000 ^a
Sample size	17	29	21	19	28	12	38
Male; female	8; 9	17; 12	9; 12	12; 7	9; 19	6; 6	15; 23
Juveniles; adults ^b	0; 17	3; 26	4; 17	1; 18	3; 25	0; 12	5; 33

^a Population surveys in the urban and forested areas as reported by Hasan *et al.*²⁸ Estimates of performing monkey numbers are based on interviews with community leaders.

^b Animals were considered juveniles if based on their dental eruption sequence they were less than 24 months old.

out as having the greatest degree of admixture (i.e., combination of different lineages). The Dokhola monkeys seemed to be mostly admixtures of the Bormi and Karamjal genotypes. The reason for this apparent admixture is unclear. Most of the performing monkeys also bore a strong relationship to the Karamjal genotype. This relationship with Karamjal suggests a possible source for many of these performing monkeys.

Viral antibodies to FV Gag proteins

Initially we tested for FV antibodies using WB; 17 animals were tested in this way. The positive monkeys showed a clear doublet corresponding to the major gag proteins (data not shown). It became apparent that performing WB assays on all 164 animals sampled would not be feasible; therefore, we developed an ELISA assay using GST-Gag protein as the antigen. We performed ELISA on many of the WB-assayed animals and in all cases the WB and ELISA results matched. We therefore screened the remaining animals by ELISA. Negative animals showed no reactivity to GST-Gag at a 1:50 plasma dilution, whereas positive animals showed clear reactivity at plasma dilutions of 1:400 or greater (Table 2). We found that 149 of the 164 animals sampled have antibodies to SFV Gag. Details of the locations and ages of the FV antibody-positive and -negative animals are summarized in Table 3. SFV antibody-negative adult animals were only found in the performing monkey population. It is important to note that all the adults ($n=89$) in Dharmai, Bormi, Charmaguria, Narayanganj, Karamjal and Dokhola, were antibody-positive. Among these free-ranging populations, only 7 of the 36 younger animals, those estimated to be less than 24 months of age based on their dental eruption sequence, lacked SFV antibodies. Four out of five of the younger performing monkeys were SFV antibody positive as were 26/33 of the adult performing monkeys. Interestingly, all seven of the adult performing monkeys that were SFV negative were female.

Absence of SRV-D in sampled population

Asian macaques are the predominant natural hosts for SRV-D. Infections are not uncommon in captive macaque colonies, with prevalence varying widely from none to over 80%, determined at least partially by geographic origin and husbandry/management practices.⁵⁶ However, neither antibodies to SRV-D nor SRV-D viral DNA sequences were detected in any of the 164 animals sampled in the present study. While this finding is in sharp contrast to numerous reports of prevalence in captive colonies, there have been only limited reports of natural infections in the wild. These include less than 5% prevalence in *Pongo pygmaeus* (orangutan) from Indonesia⁵⁷ and isolation of SRV-D of a novel type from *Macaca mulatta* and *Semnopithecus entellus* from Rajasthan, India.⁵⁸

The diagnostic antibody and PCR assays used in this study were developed and validated against the original known exogenous SRV-D 1, 2, 3, 4 and 5 serotypes.^{39,41} As reported previously, they are highly sensitive and specific and their cross-reactivity to more divergent serotypes that may be emerging is unknown. The SRV4 subtypes found in captive colony monkeys by investigators in Tsukuba, Japan and San

Table 2 Assay of anti-Gag antibodies using ELISA

Dilution	1:50 ^a	1:100 ^b	1:200 ^b	1:400 ^b	1:800 ^b	Total
Reactivity	14	1	1	5	125	146

^a Plasma are considered negative if they had an OD450 nm reading less than one at this dilution.

^b The highest dilution where an individual MBG plasma was reactive to Gag antigen.

Table 3 Geographic and age distribution of SFV seropositive rhesus macaques

	Infant ^a	Juveniles ^b	Adults ^c	Total
Dhamrai	NA	1/5	16/16	17/21
Bormi	NA	8/8	21/21	29/29
Charmaguria	0/1	11/12	15/15	26/28
Narayanganj	NA	4/5	14/14	18/19
Karamjal	NA	2/2	10/10	12/12
Dokhola	NA	4/4	13/13	17/17
Performing	NA	4/5	26/33	30/38

^a Infants: animals aged less than 9 months based on dental formula.

^b Juveniles: animals aged at less than 24 months based on dental eruption sequence.

^c Adult: animals with erupted first and second molars and permanent anterior dentition.

NA: no animal in this group were sampled.

Antonio, Texas were detected in the assays used for this study.^{59,60} However, other newer isolates such as SRV6 and SRV7 found in India have not been available for testing.⁶¹

Sequencing of SFV proviruses from peripheral blood

We initially screened antibody positive animals by PCR using a highly conserved portion of the integrase (IN) region of *pol* used by others for phylogenetic analyses.⁶² However, for sequence analyses, we used primers to a portion of the *gag* gene which molecular genetics studies suggest would allow the most sequence diversity to be identified, since it only contains a small region with known function in viral replication.⁶³⁻⁶⁵ We were able to amplify *gag* fragments from all of the SFV antibody-positive macaques. We obtained *gag* sequences from individual PCR products for each animal, with a median number of 6 *gag* clones per animal (Figure 2C).

Phylogenetic analyses of SFV *gag*

An initial phylogenetic analysis to gain an overview of SFV *gag* diversity in our samples (Figure 3) analysis showed that the viruses formed a number of tight clusters in the tree. As this initial analysis did not take viral recombination or other possible confounding effects into account, it only serves to visualize our concept of strains and suggests that there may be a correlation between viral genotype and sampling location. This observation was confirmed by our SplitsTree⁴⁷ analysis (Figure 4). SplitsTree produces a 'splits network' capable of displaying non-treelike sequence difference information in a graphical form. In the figure, each set of parallel edges represents a 'split', which is a collection of sequence differences that separates one group of sequences from another. In this case, non-treelike evolution is expected to be from recombination; putative recombinant sequences are found at the end of the parallelograms formed by edges. This tree also shows clustering by region, and abundant recombinant sequences.

These observations were formalized by a clustering analysis using UCLUST v1.1⁴⁸ at a 97.5% identity threshold. The resulting clusters were used to define strains, as described in the section on 'MATERIALS AND METHODS'. The six most populous strains were recognized as particularly representative of the bulk of the viral population diversity, and bore significant structural congruence with monkey population structure. These six strains will be called the 'core' strains and are noted as such in Figure 2G. Two of the strains, Bormi1 and Bormi2, were found to be recombinants of other core strains. The other, non-recombinant strains will be called 'parental strains'. Differences between these strains are reflected on both nucleotide and amino-acid levels (Figure 5).

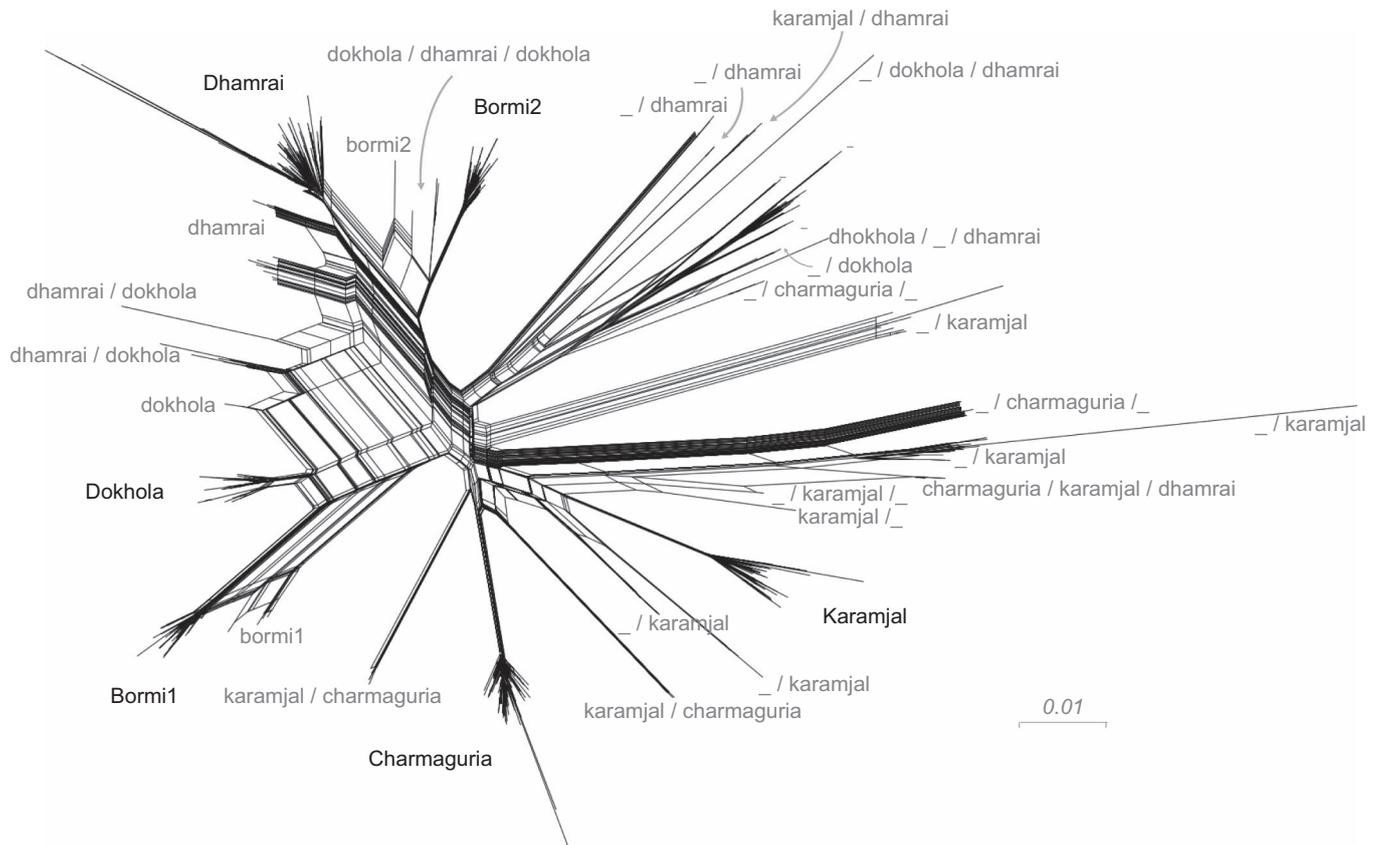


Figure 4 Splits network giving an overview of sequence diversity, and built by SplitsTree with strain labels. SplitsTree produces a ‘splits network’ capable of displaying non-tree-like sequence difference information in a graphical form. In the figure, each set of parallel edges represents a ‘split’, which is a collection of sequence differences that separates one group of sequences from another. In this case, non-tree-like evolution is expected to be from recombination; putative recombinant sequences are found at the end of the parallelograms formed by edges. Strain labels have multiple identifiers if they were found to be recombinant by cBrother; e.g., a/b indicates a recombinant between core sequences a and b. An underscore (_) indicates the presence of a >200 bp region in which no one core strain was clearly found as the source.

Of the monkey populations, three were strongly associated with a single core strain primarily found in that region: Charmaguria, Dhamrai and Karamjal. In these three areas, we only found the corresponding core strain, and not core strains from other areas. For convenience, we named these strains according to the sampling population with which they were closely associated.

Two other populations, Bormi and Dokhola, stood out as having very structured viral populations. The Bormi monkey viruses separated into two distinct strains, almost entirely unique to this monkey population. We named these strains Bormi1 and Bormi2. These strains are not closely related (Figure 3). This separation of the virus population into two strains is not paralleled by a separation of the macaque population into two microsatellite groups (Figure 2F), suggesting two strains co-existing within a homogenous host population. The Dokhola monkey viruses also separated into two strains; one was identified as the Dhamrai strain, and another, seen predominantly in this population, was named the Dokhola strain. Recombination analysis using GARD,⁴⁹ cBrother⁵¹ and the PHI test⁵⁰ demonstrated that recombination played a role in shaping these viral populations. Charmaguria, Dokhola, Dhamrai and Karamjal core strains showed no evidence of recombination according to GARD or the PHI test. On the other hand, the two Bormi core strains each appeared to be derived from a recombinant lineage that has become established in the population. Bormi1 *gag* was found to be a recombinant form of some

otherwise unsampled strain and the Dokhola strain, with a breakpoint at about nt 304 (Figures 6A and 6B). We were unable to identify a source for the first portion of this strain via a BLAST v2.2.25 search of either known Bangladeshi SFV viral strains or the NCBI nr (complete non-redundant) database, suggesting that this first portion is unique to the Bormi1 strain. The Bormi2 strain was found to be a recombinant form of the Dokhola and Dhamrai strains, with a breakpoint at about nt 496 in the 1125 nucleotide *gag* fragment we sequenced (Figures 6C and 6D).

The monkey populations with the least phylogenetically structured and most diverse viral populations were the Narayanganj and performing monkeys. Aside from a strong prevalence of the Dhamrai core strain within the Narayanganj monkeys, the viruses found in these populations were exceptionally divergent. Many of the very diverse viruses seen in the Narayanganj and performing populations were identified as recombinant in our cBrother analysis, while some appeared to be entirely distinct from the rest of the viral population. Additionally, all of the non-core strain viruses sampled from within Charmaguria, Dokhola and Karamjal were identified as having a recombinant relationship to the core strain associated with the sampled population in question.

We considered PCR-generated recombination unlikely given the very low copy number of FV in the samples (one copy per 10⁴–10⁶ cells); however, in order to eliminate the possibility that the observed

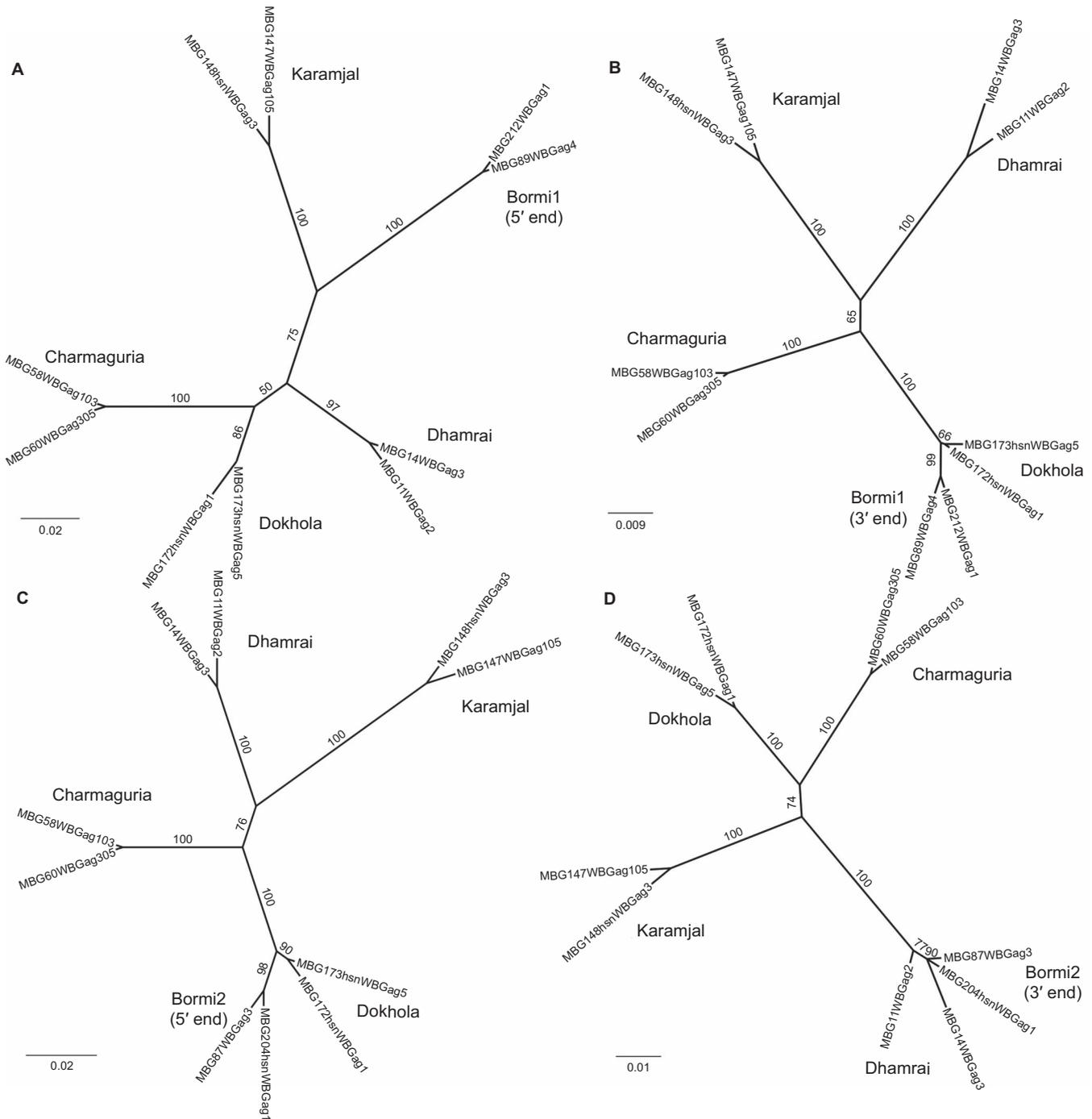


Figure 6 Recombination between Bormi strains and non-recombinant parental strains. Phylogenetic trees constructed by running PhyML on representative sequences which have been partitioned into separate alignments along recombination breakpoints inferred by cBrother. **(A)** and **(B)** show the relationship between Bormi1 and the core strains; **(C)** and **(D)** the relationship between Bormi2 and the core strains. **(A)** is built from sequence data between nucleotides 1 and 304, **(B)** from 305 on, **(C)** from 1 to 496 and **(D)** from 497 on (nucleotide (nt) number refers to position in the 1125 nt fragment of gag we sequenced, not to the nt position of the full length gag gene). Internal branches are labeled with their bootstrap values to indicate confidence.

DISCUSSION

Synanthropy and disease risk

Rhesus macaques are the consummate synanthropic primate. Their ecological and behavioral flexibility allows them to thrive in human-altered environments. Among the Order Primates, only humans are more widely geographically distributed than the rhesus. Over centuries, facilitated by cultural traditions that encourage tolerance of NHP,

rhesus macaques have come to inhabit niches in densely populated Bangladesh, where they come into frequent contact with humans. This long-term, complex interspecies association presents an opportunity to study how NHP-borne viruses originate, change and spread within and between NHP and human populations, and to better appreciate how the interplay between host ecology and viral evolution can potentially lead to the emergence of viruses of increased pathogenicity.

Table 4 Percentage of amino-acid sequence similarity^a in the region of *gag* analyzed between strains and SFV-1

	Bormi1	Bormi2	Charmaguria	Dhamrai	Dokhola	Karamjal	SFV-1
Bormi1		99	99	99	99	97	95
Bormi2			99	99	99	97	95
Charmaguria				99	99	95	95
Dhamrai					99	96	95
Dokhola						97	95
Karamjal							93

^a Similarity between two strains is expressed as percentage of identical amino acids and conservative substitutions. The amino-acid sequence is derived from a consensus of all clones for each strain, using the portion of *gag* that we sequenced (1235 nt). SFV-1 is the cloned macaque SFV described by Mergia *et al.*³⁶

Foamy viruses naturally infect many mammalian species, including all NHP, cats and cows, as reviewed in ref. 14. Infants receive maternal antibodies against FV which protect them from infection. Young juveniles are not FV-infected. However, as NHP, cats and cows age, they become infected from older animals, which transmit the virus in saliva, through bites, licks or sharing of cud. Virtually, all adult animals that are not kept isolated from others of their species are infected. Infection is life-long, but there is no pathology associated with FV in natural hosts. DNA from latent proviruses is present in many tissues including peripheral blood, but virus and viral RNA indicative of replication is only detected in the oral mucosa and saliva. One study showed that in NHP, it is the superficial differentiating epithelial cells of the oral mucosal tissues that are infected.¹⁷

Here we present results describing the prevalence and geographic distribution of the NHP-borne retrovirus, SFV, in Bangladeshi rhesus macaques that were sampled from multiple contexts. SFV infection was highly prevalent in free-ranging macaques, with all adult animals having been infected and 24% of the adults infected with more than one strain of SFV. This is consistent with rates of SFV infection measured in other free-ranging macaque populations;²¹ however co-infection of single animals with more than one strain has not been well documented. Because our analyses only included 6–10 proviral clones per macaque, it is likely that the diversity we describe is an underestimate, and that we have missed some viral strains. Only 75% of the performing macaques we examined were infected by SFV. Performing macaques are often acquired by their owners as young animals, and subsequently, the animals have reduced contact with other NHP and hence, fewer opportunities to become infected with SFV.⁶⁷

Our analyses indicated that, for the majority of Bangladeshi macaques, SFV strains co-segregate with macaque populations, with one predominant strain associated with the majority of animals in a given area (Figure 2G). There were two exceptions to this finding. First, in Bormi, we found two distinct SFV core strains circulating in the macaque population, which we have called Bormi1 and Bormi2. Bormi samples from each of the 3 years under study contained both the Bormi1 and Bormi2 strains. Further longitudinal data are required to determine if these two strains will continue to persist or if one will

eventually predominate in Bormi. Second, in the port city of Narayanganj, we observed a wide diversity of viruses. We found that several of these closely resembled viruses that predominate in other macaque populations, while others appeared to be novel. Additionally, several viruses from Narayanganj macaques appeared to be recombinants of core strains from other sites.

Biogeographic effects

Geographic factors may explain some of the observed patterns of microsatellite and SFV variation. Patterns of historical and contemporary forest cover and the country's major rivers are likely to have influenced the geographic distribution of macaques and SFV. Like many countries in South and Southeast Asia, large areas of Bangladesh have been deforested over the past decades. As recently as 50 years ago, the forests of the Modhupur National Park covered much of central Bangladesh between the Padma, Jamuna and Meghna rivers. This extensive forest would have allowed the free movement of fauna, including subadult rhesus macaques, which normally leave their home groups and range in search of new groups. Phylogenetic analysis of SFV *gag* (Figure 3) indicated that the Bormi2 strain is a recombinant virus containing portions of the parental Dokhola and Dhamrai strains. Prior to widespread deforestation, Bormi, Dokhola and Dhamrai would all have fallen within the historical forests of the Modhupur Park (Figure 1), allowing macaques carrying the two parental strains to come into contact, setting the stage for viral recombination. In contrast, Charmaguria is located >30 km south of the Padma River and in a region of Bangladesh that has for decades lacked any significant, connected forest patches. The rhesus macaques in Charmaguria are effectively an isolated population. This isolation is reflected in the homogeneity seen in their microsatellites and SFV *gag* sequences.

Anthropogenic effects

Two groups stood out from the rest in terms of increased viral phylogenetic diversity: the performing macaques and the macaques from Narayanganj. Of the sites analyzed in this study, Narayanganj is associated with the highest human population density and the greatest commerce. It is located in the center of the country, close to the capital Dhaka and on a major river transit route. The diversity of viruses from the macaques in Narayanganj could be due to its use as a commercial hub. As with many Asian countries containing NHP habitats, pet NHP are common in Bangladesh and, as elsewhere, NHP owners face a dilemma when confronted with a monkey acquired as an infant that has grown into an aggressive adult. The solution is sometimes to release pets in an area with an existing NHP population.⁶⁸ It was not uncommon for us to find macaques with collars still about their necks in a free-ranging group. Narayanganj may simply be Bangladesh's 'melting pot' for released and/or escaped pet macaques. Natural migration of

Table 5 Infectivity of SFV-1 with a 6 nucleotide deletion in *gag*

Virus ^a	Titer ^b (units/mL)
WT SFV-1	2.0×10 ⁵
pSFVΔ6nt-1	9.5×10 ⁴
pSFVΔ6nt-2	1.2×10 ⁴
pSFVΔ6nt-3	3.0×10 ⁶

^a WT SFV-1 or 3 randomly picked molecular clones containing the 6 nucleotide deletion in *gag* were transfected into Tf cells. After 48 h, supernatants were collected.

^b Supernatants were titered on simian foamy-activated beta-galactosidase cells.

macaques is unlikely to be a factor in Narayanganj, as there has been no forest within 50 km of Narayanganj for decades. Thus, the genetically and SFV heterogeneous monkey population that we observed in Narayanganj today is most likely a byproduct of human-assisted migration.

The observed genetic heterogeneity among the performing monkeys may be linked to the sociogeography and habits of nomadic people who travel throughout South Asia with their performing animals.^{10,69,70} The Bade in Bangladesh have a centuries old tradition of capturing macaques from forested or urban areas and training them to perform for audiences. Performing monkeys are traded amongst nomadic groups, sometimes across national borders. Bade performing monkey owners stated that the Sundarban mangrove forest, near Karamjal, is a prime source for their monkeys, which is consistent with both our microsatellite and viral data.

Retroviral recombination

Our data show that individual animals can harbor more than one viral strain (Figures 2A and 2G). Whether infection with two viruses results from simultaneous infection from a single bite, transferring saliva containing multiple viruses, or from multiple bites is not known. However, our data showing a higher prevalence of multiple infections in older macaques are more consistent with the latter interpretation. This might imply that, SFV infection does not always protect from infection by other strains.

In this paper, we report the amplification and sequencing of DNA from blood samples and not from saliva. Analysis of SFV from the saliva of animals infected with multiple viral strains is needed to determine whether these animals also shed more than one viral strain, and this work is in progress. The existence of more than one viral strain in the monkey populations under study raises the possibility of historical and contemporary retroviral recombination. Indeed, our phylogenetic data suggest that recombination does occur in SFV. However, it is surprising that in Bormi, where two core strains are found, there is very little recombination detected so far between these strains. Further studies are needed to determine whether infection by a recombinant virus has biological consequences for the host.

Based on our results, we hypothesize that the observed patterns of genetic variation in SFV from these Bangladeshi rhesus macaques are explained, at least in part, by the human translocation of macaques. Translocation of infectious agents between populations of macaques sets the stage for simultaneous infection with more than one viral strain, as we detected in 17% of all macaques sampled regardless of age, and for the generation of recombinant viral genomes, which we also detected. Genetic recombination can produce retroviruses of altered phenotypes, as has been demonstrated for lentiviruses.⁷¹ We have shown that frequent human/macaque contacts provide the contexts for NHP-to-human viral transmission,^{26,27} (companion manuscript in review). Future work will expand the scope and detail of our understanding of the interplay between host–pathogen–environment, as we are only beginning to appreciate how such complex dynamics at the human/NHP interface in Asia could affect the emergence of new diseases in humans.

ACKNOWLEDGMENTS

The authors wish to thank our student assistants and faculty in the Wildlife Branch of the Department of Zoology at Jahangirnagar University, Bangladesh. We also thank the Bangladesh Forest Department for their permission and constant support of our research program. Drs John Heidrich (Ventana Animal Clinic, Albuquerque, NM, USA) and Michael Schillaci (University of Toronto Scarborough, Toronto, Ont., Canada) as well as Hanna and Leah Engel (Seattle,

WA, USA) and Lynn Johnson (National Geographic Magazine, Washington, DC, USA) were outstanding assistants during multiple field seasons. We thank Robin Watanabe (University of Washington, Seattle, WA, USA), Ann Rosenthal, Shannon Flynn (University of California, Davis, CA, USA), Connor McCoy, Leah Engel and Brandon Duffy (Fred Hutchinson Cancer Research Center, Seattle, WA, USA) for their laboratory technical assistance and discussions and Amanda Zeller and Regina Liszanckie (University of Washington, Seattle, WA, USA) for administrative support. We are grateful to Drs Harmit Malik and Michael Emerman (Fred Hutchinson Cancer Research Center, Seattle, WA, USA) for their insightful suggestions and comments. This research was supported by funding from NIH-NIAID grants R01 AI078229, R01AI078229-03S1, R03 AI064865, NIH-NCI grant CA18282, NIH-NCCR grant P51 RR000166 and New Development Institutional Support from the Fred Hutchinson Cancer Research Center.

Supplementary Information accompanies the paper on *Emerging Microbes & Infections's* website (<http://www.nature.com/EMI/>).

- 1 Bailes E, Gao F, Bibollet-Ruche F *et al*. Hybrid origin of SIV in chimpanzees. *Science* 2003; **300**: 1713–1716.
- 2 Daszak P, Cunningham AA, Hyatt AD. Anthropogenic environmental change and the emergence of infectious diseases in wildlife. *Acta Trop* 2001; **78**: 103–116.
- 3 Jones KE, Patel NG, Levy MA *et al*. Global trends in emerging infectious diseases. *Nature* 2008; **451**: 990–993.
- 4 Weiss RA, McMichael AJ. Social and environmental risk factors in the emergence of infectious diseases. *Nat Med* 2004; **10**(12 Suppl): S70–S76.
- 5 Plowright RK, Foley P, Field HE *et al*. Urban habituation, ecological connectivity and epidemic dampening: the emergence of Hendra virus from flying foxes (*Pteropus* spp.). *Proc Biol Sci* 2011; **278**: 3703–3712.
- 6 Engering A, Hogerwerf L, Slingenbergh J. Pathogen–host–environment interplay and disease emergence. *Emerg Microbes Infect* 2013; **2**: e5.
- 7 Khatun UH, Ahsan MF, Røskaft E. Attitudes of the local community towards the conservation of the common langur (*Semnopithecus entellus*) in Keshabpur, Bangladesh. *Int J Biodivers Conserv* 2012; **4**: 385–399.
- 8 Feeroz MM. Species diversity and population density of non-human primates in north-east and south-east of Bangladesh. *Ecoprint* 2011; **8**: 53–57.
- 9 Hasan MK, Feeroz MM. Distribution and status of long-tailed macaque (*Macaca fascicularis aurea* I. Geoffroy Saint-Hilaire, 1830) in Bangladesh. *J Threat Taxa* 2010; **2**: 1342–1344.
- 10 Mack TM, Noble J Jr. Natural transmission of smallpox from man to performing monkeys. An ecological curiosity. *Lancet* 1970; **1**: 752–754.
- 11 Oberste MS, Feeroz MM, Maher K *et al*. Characterizing the picornavirus landscape among synanthropic nonhuman primates in Bangladesh, 2007 to 2008. *J Virol* 2013; **87**: 558–571.
- 12 Oberste MS, Feeroz MM, Maher K *et al*. Naturally acquired picornavirus infections in nonhuman primates at the Dhaka Zoo. *J Virol* 2013; **87**: 572–580.
- 13 Karlsson EA, Engel GA, Feeroz MM *et al*. Influenza virus infection in nonhuman primates. *Emerg Infect Dis* 2012; **18**: 1672–1675.
- 14 Linal ML. Foamy viruses. In: Knipe DM, Howley PM (ed.) *Fields virology*. 5th ed. Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins, 2007: 2245–2263.
- 15 Han GZ, Worobey M. An endogenous foamy-like viral element in the coelacanth genome. *PLoS Pathog* 2012; **8**: e1002790.
- 16 Linal ML. Why aren't foamy viruses pathogenic? *Trends Microbiol* 2000; **8**: 284–289.
- 17 Murray SM, Picker LJ, Axthelm MK, Hudkins K, Alpers CE, Linal ML. Replication in a superficial epithelial cell niche explains the lack of pathogenicity of primate foamy virus infections. *J Virol* 2008; **82**: 5981–5985.
- 18 Falcone V, Leupold J, Clotten J *et al*. Sites of simian foamy virus persistence in naturally infected African green monkeys: latent provirus ubiquitous, whereas viral replication is restricted to the oral mucosa. *Virology* 1999; **257**: 7–14.
- 19 Voevodin AF, Marx PA. Spumaviruses. In: *Simian virology*. London: Blackwell Publishing Ltd, 2009: 217–236.
- 20 Lerche NW. Simian retroviruses: infection and disease—implications for immunotoxicology research in primates. *J Immunotoxicol* 2010; **7**: 93–101.
- 21 Jones-Engel L, Steinkraus KA, Murray SM *et al*. Sensitive assays for simian foamy viruses reveal a high prevalence of infection in commensal, free-ranging, Asian monkeys. *J Virol* 2007; **81**: 7330–7337.
- 22 Boneva RS, Switzer WM, Spira TJ *et al*. Clinical and virological characterization of persistent human infection with simian foamy viruses. *AIDS Res Hum Retroviruses* 2007; **23**: 1330–1337.
- 23 Switzer WM, Bhullar V, Shanmugam V *et al*. Frequent simian foamy virus infection in persons occupationally exposed to nonhuman primates. *J Virol* 2004; **78**: 2780–2789.
- 24 Switzer WM, Garcia AD, Yang C *et al*. Coinfection with HIV-1 and simian foamy virus in west Central Africans. *J Infect Dis* 2008; **197**: 1389–1393.
- 25 Sandstrom PA, Phan KO, Switzer WM *et al*. Simian foamy virus infection among zoo keepers. *Lancet* 2000; **355**: 551–552.

- 26 Jones-Engel L, Engel GA, Schillaci MA *et al*. Primate-to-human retroviral transmission in Asia. *Emerg Infect Dis* 2005; **11**: 1028–1035.
- 27 Jones-Engel L, May CC, Engel GA *et al*. Diverse contexts of zoonotic transmission of simian foamy viruses in Asia. *Emerg Infect Dis* 2008; **14**: 1200–1208.
- 28 Hasan MK, Aziz MA, Alam SMR *et al*. Distribution of rhesus macaques (*Macaca mulatta*) in Bangladesh: inter-population variation in group size and composition. *Primate Conserv* 2013; **26**: 125–132.
- 29 Jones-Engel L, Engel GA, Heidrich J *et al*. Temple monkeys and health implications of commensalism, Kathmandu, Nepal. *Emerg Infect Dis* 2006; **12**: 900–906.
- 30 Kanthaswamy S, von Dollen A, Kurushima JD *et al*. Microsatellite markers for standardized genetic management of captive colonies of rhesus macaques (*Macaca mulatta*). *Am J Primatol* 2006; **68**: 73–95.
- 31 Kalinowski ST, Taper ML, Marshall TC. Revising how the computer program cervus accommodates genotyping error increases success in paternity assignment. *Mol Ecol* 2007; **16**: 1099–106.
- 32 Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 2000; **155**: 945–959.
- 33 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.
- 34 Earl DA, vonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing structure output and implementing the Evanno method. *Conservation Genetics Resources* 2012; **4**: 359–361.
- 35 Wickham H. *ggplot2: elegant graphics for data analysis*. New York: Springer, 2009.
- 36 Mergia A, Wu M. Characterization of provirus clones of simian foamy virus type 1. *J Virol* 1998; **72**: 817–822.
- 37 Kuller L, Watanabe R, Anderson D, Grant R. Development of a whole-virus multiplex flow cytometric assay for antibody screening of a specific pathogen-free primate colony. *Diagn Microb Infect Dis* 2005; **53**: 185–193.
- 38 Khan IH, Mendoza S, Yee J *et al*. Simultaneous detection of antibodies to six nonhuman-primate viruses by multiplex microbead immunoassay. *Clin Vaccine Immunol* 2006; **13**: 45–52.
- 39 Liao Q, Guo H, Tang M *et al*. Simultaneous detection of antibodies to five simian viruses in nonhuman primates using recombinant viral protein based multiplex microbead immunoassays. *J Virol Methods* 2011; **178**: 143–152.
- 40 Schweizer M, Neumann-Haefelin D. Phylogenetic analysis of primate foamy viruses by comparison of pol sequences. *Virology* 1995; **207**: 577–582.
- 41 White JA, Todd PA, Rosenthal AN *et al*. Development of a generic real-time PCR assay for simultaneous detection of proviral DNA of simian Betaretrovirus serotypes 1, 2, 3, 4 and 5 and secondary uniplex assays for specific serotype identification. *J Virol Methods* 2009; **162**: 148–154.
- 42 Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 2004; **5**: 113.
- 43 Rice P LIBA. EMBOS: the European Molecular Biology Open Software Suite. *Trends Genet* 2000; **16**: 276–277.
- 44 Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009; **26**: 1641–1650.
- 45 Talevich E, Invergo BM, Cock PJ, Chapman BA. Bio.Phylo: a unified toolkit for processing, analyzing and visualizing phylogenetic trees in Biopython. *BMC Bioinformatics* 2012; **13**: 209.
- 46 Cock PJ, Antao T, Chang JT *et al*. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 2009; **25**: 1422–1423.
- 47 Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 2006; **23**: 254–267.
- 48 Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010; **26**: 2460–2461.
- 49 Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD. GARD: a genetic algorithm for recombination detection. *Bioinformatics* 2006; **22**: 3096–3098.
- 50 Bruen TC, Philippe H, Bryant D. A simple and robust statistical test for detecting the presence of recombination. *Genetics* 2006; **172**: 2665–2681.
- 51 Fang F, Ding J, Minin VN, Suchard MA, Dorman KS. cBrother: relaxing parental tree assumptions for Bayesian recombination detection. *Bioinformatics* 2007; **23**: 507–508.
- 52 Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biol* 2010; **59**: 307–321.
- 53 Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 2004; **20**: 289–290.
- 54 Yu SF, Linial ML. Analysis of the role of the *bel* and *bet* open reading frames of human foamy virus by using a new quantitative assay. *J Virol* 1993; **67**: 6618–6624.
- 55 Hubisz MJ, Falush D, Stephens M, Pritchard JK. Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour* 2009; **9**: 1322–1332.
- 56 Nandi JS, Bhavalkar-Potdar V, Tikute S, Raut CG. A novel type D simian retrovirus naturally infecting the Indian Hanuman langur (*Semnopithecus entellus*). *Virology* 2000; **277**: 6–13.
- 57 Warren KS, Niphuis H, Heriyanto, Verschoor EJ, Swan RA, Heeney JL. Seroprevalence of specific viral infections in confiscated orangutans (*Pongo pygmaeus*). *J Med Primatol* 1998; **27**: 33–37.
- 58 Nandi JS, Van DS, Chhangani AK, Mohnot SM. New simian beta retroviruses from rhesus monkeys (*Macaca mulatta*) and langurs (*Semnopithecus entellus*) from Rajasthan, India. *Virus Genes* 2006; **33**: 107–116.
- 59 Zao CL, Armstrong K, Tomanek L *et al*. The complete genome and genetic characteristics of SRV-4 isolated from cynomolgus monkeys (*Macaca fascicularis*). *Virology* 2010; **405**: 390–396.
- 60 Hara M, Sata T, Kikuchi T *et al*. Isolation and characterization of a new simian retrovirus type D subtype from monkeys at the Tsukuba Primate Center, Japan. *Microbes Infect* 2005; **7**: 126–131.
- 61 Nandi JS, Tikute SA, Chhangani AK *et al*. Natural infection by simian retrovirus-6 (SRV-6) in Hanuman langurs (*Semnopithecus entellus*) from two different geographical regions of India. *Virology* 2003; **311**: 192–201.
- 62 Hussain AI, Shanmugam V, Bhullar VB *et al*. Screening for simian foamy virus infection by using a combined antigen Western blot assay: evidence for a wide distribution among Old World primates and identification of four new divergent viruses. *Virology* 2003; **309**: 248–257.
- 63 Stange A, Mannigel I, Peters K *et al*. Characterization of prototype foamy virus gag late assembly domain motifs and their role in particle egress and infectivity. *J Virol* 2005; **79**: 5466–5476.
- 64 Patton GS, Morris SA, Chung W, Bieniasz PD, McClure MO. Identification of domains in gag important for prototypic foamy virus egress. *J Virol* 2005; **79**: 6392–6399.
- 65 Rethwilm A. Molecular biology of foamy viruses. *Med Microbiol Immunol* 2010; **199**: 197–207.
- 66 Switzer WM, Salemi M, Shanmugam V *et al*. Ancient co-speciation of simian foamy viruses and primates. *Nature* 2005; **434**: 376–380.
- 67 Blewett EL, Black DH, Lerche NW, White G, Eberle R. Simian foamy virus infections in a baboon breeding colony. *Virology* 2000; **278**: 183–193.
- 68 Jones-Engel L, Engel GA, Schillaci MA, Babo R, Froehlich J. Detection of antibodies to selected human pathogens among wild and pet macaques (*Macaca tonkeana*) in Sulawesi, Indonesia. *Am J Primatol* 2001; **54**: 171–178.
- 69 Schillaci MA, Jones-Engel L, Engel GA *et al*. Prevalence of enzootic simian viruses among urban performance monkeys in Indonesia. *Trop Med Int Health* 2005; **10**: 1305–1314.
- 70 Berland JC. *No five fingers are alike*. Cambridge, MA: Harvard University Press, 1982.
- 71 Gao F, Bailes E, Robertson DL *et al*. Origin of HIV-1 in the chimpanzee *Pan troglodytes* troglodytes. *Nature* 1999; **397**: 436–441.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivative Works 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0>

Supplementary Information accompanies the paper on *Emerging Microbes & Infections*'s website (<http://www.nature.com/EMI/>).