2017

Alu insertion polymorphisms as evidence for population structure in baboons

Cody J. Steely

Jane Phillips-Conroy

et al

Alu Insertion Polymorphisms as Evidence for Population Structure in Baboons

Cody J. Steely[1], Jerilyn A. Walker[1], Vallmer E. Jordan[1], Thomas O. Beckstrom[1], Cullen L. McDaniel[1], Corey P. St. Romain[1], Emily C. Bennett[1], Arianna Robichaux[1,2], Brooke N. Clement[1,3], Muthuswamy Raveendran[4], The Baboon Genome Analysis Consortium[†], Kim C. Worley[4,5], Jane Phillips-Conroy[6], Clifford J. Jolly[7], Jeff Rogers[4,5], Miriam K. Konkel[1], and Mark A. Batzer[1,*]

[1]Department of Biological Sciences, Louisiana State University

[2]Department of Biological and Physical Sciences, Northwestern State University of Louisiana

[3]School of Veterinary Medicine, Louisiana State University

[4]Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas

[5]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas

[6]Department of Neuroscience, Washington University, St. Louis

[7]Department of Anthropology, New York University

*Corresponding author: E-mail: mbatzer@lsu.edu.

Accepted: September 7, 2017

[†]Membership in the Baboon Genome Analysis Consortium is listed in the supplementary file S1, Supplementary Material online.

Abstract

Male dispersal from the natal group at or near maturity is a feature of most baboon (*Papio*) species. It potentially has profound effects upon population structure and evolutionary processes, but dispersal, especially for unusually long distances, is not readily documented by direct field observation. In this pilot study, we investigate the possibility of retrieving baboon population structure in yellow (*Papio cynocephalus)* and kinda (*Papio kindae)* baboons from the distribution of variation in a genome-wide set of 494 *Alu* insertion polymorphisms, made available via the recently completed Baboon Genome Analysis Consortium. *Alu* insertion variation in a mixed population derived from yellow and olive (*Papio anubis*) baboons identified each individual's proportion of heritage from either parental species. In an unmixed yellow baboon population, our analysis showed greater similarity between neighboring than between more distantly situated groups, suggesting structuring of the population by male dispersal distance. Finally (and very provisionally), an unexpectedly sharp difference in *Alu* insertion frequencies between members of neighboring social groups of kinda baboons suggests that intergroup migration may be more rare than predicted in this little known species.

Key words: *Alu*, population genetics, population structure, retrotransposon.

Introduction

Baboons (genus *Papio*) are distributed throughout most of sub-Saharan Africa and southwestern Arabia. Six major forms, now generally recognized as species, have broad, contiguous but nonoverlapping ranges, and are all interfertile, hybridizing where their ranges meet. The many studies that have been carried out on baboons include analyses of phylogenetic diversity, behavior, and ecology, and they have been widely recognized as useful analogs for understanding human evolution, as well as in biomedical studies of disease processes such as cardiovascular disease and obesity (Jolly 2001; Premawardhana et al. 2001; Cox et al. 2013; Yeung et al. 2016). Numerous studies of the social behavior of wild baboons have documented common features, including the basic unit of social organization, a permanent social group (usually called a troop) that includes individuals of all ages and both sexes. Other important features of social behavior vary among species, one of which concerns patterns of sex-specific dispersal. In hamadryas (*Papio hamadryas*) (Swedell et al. 2011) and, apparently, Guinea baboons

(*Papio papio*) (Fischer et al. 2017) most males are philopatric, remaining to breed in their troop of birth. Among olive baboons (*Papio anubis)*, yellow baboons (*Papio cynocephalus)*, and chacma babons (*Papio ursinus*), nearly all males, when they reach sexual maturity, emigrate from their natal troop and join another, where they breed. Some males change groups more than once. The dispersal regime of kinda baboons (*Papio kindae*) has yet to be reported (though recent work has been enlightening, see Jolly et al. 2011), but anecdotal evidence suggests that it too is a male-dispersal species. Near-universal male dispersal is clearly a major determinant of the genetic structure of the wider population, and thus, potentially, an influence on evolutionary processes such as adaptive population divergence and differentiation due to isolation by distance (Wright 1943).

The occurrence of male dispersal in baboons has been firmly established by naturalistic, observation-based studies. Such studies, however, are rarely able to recognize and monitor known, individual members of more than a single troop, or at most a cluster of neighboring troops. Animals are often seen joining or leaving the focal group(s), but their origin or destination is usually unknown. Longer range dispersal, in particular, can rarely be documented by tracking known, migrant individuals, and its frequency and impact are likely to be underestimated.

Molecular genetic approaches offer an alternative approach to the problem, first, by documenting the population genetic structure directly from the distribution of quantifiable variation among social groups across the landscape, and, also, potentially identifying individuals whose genetic distinctiveness marks them as possible long-distance migrants. Past studies have attempted to retrieve broader-scale population structure in baboons by surveying the distribution of genetic markers. These include isozymes (Olivier et al. 1974; Shotake et al. 1977; Rogers and Kidd 1993), RFLPs (Newman et al. 2004), blood type antigens (Socha et al. 1977), microsatellites (St. George et al. 1998; Woolley-Barker 1999), and single-locus *Alu* insertions (Szmulewicz et al. 1999). Such studies are typically restricted to a few genetic loci, so that many individuals must to be sampled to reveal population structure. In recent years, the development of whole-genome, DNA-level techniques has allowed access to a much greater density of genetic information per individual (Bergey et al. 2013).

One developing source of genetic markers is *Alu* insertion polymorphism. *Alu* elements are a class of primate-specific retrotransposons, derived from 7SL RNA, that are present in high copy number throughout the genome (Reviewed in Deininger et al. 2003). *Alu* elements are fairly short (~300 bp), making them easy to amplify and genotype by locus specific polymerase chain reaction (PCR) assays. They mobilize through RNA intermediates and insert new copies in novel locations in the genome (reviewed in Batzer and Deininger 2002; Cordaux and Batzer 2009; Konkel et al. 2010; Levin and Moran 2011). *Alu* elements are non-long terminal repeat, non-autonomous retrotransposons that require the proteins encoded by L1 elements to mobilize (Dewannieux et al. 2003). The process of *Alu* mobilization in primates has created a series of distinct subfamilies or clades of elements that share common diagnostic or subfamily specific mutations (Slagel et al. 1987; Willard et al. 1987; Britten et al. 1988; Jurka and Smith 1988; Deininger et al. 1992). These subfamilies of *Alu* elements have dispersed throughout primate genomes at varying points through evolutionary time, giving rise to a varying number of elements per subfamily. This has led to different rates of *Alu* distribution throughout the Primate order (Konkel et al. 2010; Walker et al. 2012; Rogers et al. under revision). Not all these subfamilies are mobilization competent in the same time interval. In fact, a relatively small number of all *Alu* elements present in a genome are mobilization competent (Deininger et al. 1992; Cordaux et al. 2004; Han et al. 2005; Konkel et al. 2010).

Mobile elements are valuable tools for determining phylogenetic relationships among species and genetic structure within populations, as the ancestral state of any candidate locus is the absence of the element (Batzer and Deininger 1991, 2002; Ray et al. 2006). Another attribute of *Alu* elements that makes them valuable for such studies is that they are identical by descent and nearly homoplasy free, reducing the risk of homoplasy-induced sources of error (Ray et al. 2006). *Alu* elements have been used in a number of recent population genetic and phylogenetic studies throughout the primate order (Batzer et al. 1994; Hamdi et al. 1999; Szmulewicz et al. 1999; Schmitz et al. 2001; Salem et al. 2003; Roos et al. 2004; Ray et al. 2005a; Ray et al. 2005b; Witherspoon et al. 2006; Kriegs et al. 2007; Li et al. 2009; Meyer et al. 2012; Hartig et al. 2013; McLain et al. 2013). In this pilot study, we investigate the potential for using a panel of *Alu* insertion polymorphisms to retrieve the structure of natural baboon populations from comparatively small, representative samples of animals. This approach is made possible by the newly available, uniquely dense database of *Alu* insertion polymorphisms generated as a part of the Baboon Genome Analysis Consortium (Rogers et al. under revision). Along with the previously sequenced olive baboon (Panu_2.0, GenBank accession GCA_000264685.1), recently genomes from 15 baboons have been sequenced, including multiple individuals from each of the six recognized species. These new data have made available a plethora of new mobile element insertion polymorphisms.

Materials and Methods

This study included a total of 42 yellow baboons (*P. cynocephalus*), 15 kinda baboons (*P. kindae*), and 3 olive baboons (*P. anubis*). Twelve of the yellow baboons sampled were captive animals from the Southwest Foundation for Biomedical Research (SFBR), probably all descended from baboons captured in the early 1960s near Amboseli National Park, Kenya, (~2.6°S, 37.0°E).

The remaining 30 yellow baboons were captured, sampled, and released by J.P.-C. and J.R. from a wild population living under natural conditions in Mikumi National Park, Tanzania (~7.3°S, 37.0°E) (Rogers 1989). The yellow baboons of Mikumi, especially the Viramba troops, have been the subject of many studies, some spanning many years (Norton et al. 1987; Rhine et al. 1988, 1992; Wasser and Starling 1988; Rogers and Kidd 1993, 1996). These have documented aspects of ecology, social structure, and demography, including female philopatry and male dispersal. Previous genetic work at Mikumi has shown that this male prereproductive dispersal maintains a large effective population size, and hence high levels of genetic variation within troops (Rogers and Kidd 1996). The Mikumi animals sampled for our study belonged to seven different troops, each represented by four or five individuals. Troop foraging ranges were extensive and seasonally variable, so that the distances between them reported here are approximate, but serve to illustrate their dispersal. Troops 1, 2, and 5 (Nyeusi, Barabara, and Punk, respectively) had distinct but overlapping ranges. All three visited, and were captured, at the same, central, trapping site; near the Headquarters of the Animal Behavior Research Unit (ABRU). Troops 3 and 4 (Viramba 1 and Viramba 2) were derived from a single troop that had divided. Their overlapping ranges were centered ~6 km northeast of the ABRU Headquarters. Group 6 (Ikoya) was centered ~4 km southwest of the ABRU HQ, and Group 7 (Kisorobi) lived ~13 km north of it.

Kinda baboons have been less extensively studied than other baboon species, and have often been classified as a subspecies of *P. cynocephalus* (Jolly 1993; Grubb et al. 2003). In recent years, however, studies based on their morphology, genetics, and behavior have documented their distinctiveness and led to their widespread recognition as a "major form" (Frost et al. 2003), and more recently a full species (Jolly et al. 2011; Zinner et al. 2013; Weyher et al. 2014; Rogers et al. under revision). The 15 kinda baboon samples in this study were collected by J.P.-C., C.J.J., and J.R. from wild animals captured, sampled, and released at Chunga, the northern headquarters of Kafue National Park, Zambia (15.05°S, 26.00°E). The animals were trapped at two different sites, ~1.5 km apart. Five came from the "Chunga School" site, and ten from the "Chunga Headquarters" site. Observations suggested that each of these trapping sites was mostly frequented by a different troop of baboons, but that the two troops had closely adjacent and overlapping ranges.

While tranquilized, all animals in the Mikumi and Chunga samples were weighed and sexed, and assigned to an age-class on the basis of dental eruption (Phillips-Conroy and Jolly 1988).

Alu Ascertainment and Oligonucleotide Primer Design

Alu elements for this study were ascertained in two ways. In the first method, elements that were found in the reference genome of the olive baboon, *Papio anubis* (Panu_2.0) (GenBank accession GCA_000264685.1), were compared with the genome of the rhesus macaque (rheMac3) to ensure that they resulted from insertion events that occurred after the genera *Macaca* and *Papio* had diverged. The second method utilized *Alu* elements located by interrogating bam files (binary format used for storing sequence data) of sequences from the genomes of sequenced *Papio* individuals using an in house pipeline (Jordan et al. In preparation). Briefly, the Burrows–Wheeler Aligner (Li and Durbin 2009) was used to align *Alu* consensus sequences to the reads located in various bam files. These locations for potentially novel insertions were compared with known *Alu* element locations in the olive baboon reference genome (Panu_2.0). The starting point of potential *Alu* insertions was then estimated. Nucleotide sequences adjacent to these breakpoints were extracted from the olive baboon reference genome and aligned to the orthologous location of Rhesus Macaque (rheMac8), Chimpanzee (panTro4), and Human (hg19) genomes. MUSCLE (Edgar 2004) was then used to align each of the orthologous locations and a modified version of Primer 3 (Untergasser et al. 2012) was used to design all primers.

Polymerase Chain Reaction

Locus specific PCR amplification was performed under the following conditions: 15–50 ng of DNA template, 200 nM of each forward and reverse primers, 200 μM dNTPs in 1× PCR buffer (50 mM KCl/10 mM Tris–HCl), 1.5 mM $MgCl_2$, and 1–2 units of *Taq* DNA polymerase; for a final volume per reaction of 25 μl. The conditions for the reactions were as follows: an initial denaturation at 94 °C for 1 min, followed by 32 cycles of denaturation at 94 °C, an annealing step at optimal annealing temperatures for each primer pair, and extension at 72 °C for 30 s. The reactions were terminated with a final extension step at 72 °C for 2 min. PCR products were run out using gel electrophoresis on 2% agarose, which was stained with 0.2 μg/ml ethidium bromide. DNA fragments were visualized using UV fluorescence and genotyped from the resulting images.

STRUCTURE Analysis

Analyses of population structure were performed using STRUCTURE 2.3.4 software (Falush et al. 2003). Genotype data were entered into an Excel spreadsheet, with "1, 1" indicating an insertion that is fixed present in an individual, "1, 0" indicating an insertion that is heterozygous in an individual, and "0, 0" indicating that an insertion is absent in an individual. Genotype data for all 494 polymorphic *Alu* insertions were uploaded into STRUCTURE and analyses were performed on kinda and yellow baboons separately. No information about geographic location, or origin of the samples was incorporated in any of these analyses. All analyses were performed under the admixture model, which assumes

that individuals could be of mixed ancestry. To determine the number of population clusters (K), initial analyses were run on each population (all yellow baboons, only Mikumi yellow baboons, and all kinda baboons) with 20,000 burn-in, followed by 200,000 MCMC iterations with five replicates for each K value (from one to seven). When the most likely value of K was found, as determined by the likelihood values produced by STRUCTURE, the procedure was repeated with 100,000 burn-in, followed by 1,000,000 MCMC iterations with five replicates of the most likely K value. These five replicates were averaged to generate the final data set, which was then graphed in Microsoft Excel. For instances where there were multiple K values that shared similar likelihood values, each of these K values was put through the second procedure of 100,000 burn-in with 1,000,000 MCMC iterations with five replicates to ensure that the most likely value of *K* was selected.

Principal Components Analysis

Genotype data for each individual were entered into an Excel spreadsheet. The data were uploaded into R (version 3.2.5) (R Development Core Team 2016). All missing data for each individual were omitted from the analysis, using the "na.omit" command, so as to not skew results for any of the loci or individuals. This process removed 21 loci, leaving 473 loci to be analyzed. A principal component analysis was run using the "prcomp" library and the resulting values were imported into Microsoft Excel, which was used to create figures.

Where appropriate, we applied simple nonparametric tests in the SPSS Statistics package (IBM Corp 2016), treating the first population cluster score of each individual as a variable, to test for significant differences among populations.

Results

Our data set contained 494 *Alu* insertion polymorphisms that were genotyped on a full panel of 79 *Papio* individuals, including representatives of all six known species (Rogers et al. under revision). Of these 494 loci, 115 loci were ascertained from the genome of a *P. cynocephalus* individual (Mikumi 5026). The binomial data for these 494 loci can also be found on the Batzer Lab website (https://biosci-batzerlab.biology.lsu.edu/, last accessed July 14, 2017) for the Baboon Genome Analysis Consortium manuscript.

Regional Diversity in Yellow Baboons

Though the naturalistic behavior of the actual ancestors of the SFBR baboons was not recorded, the capture sites are very close to the Amboseli National Park, where baboons have been extensively studied for many years, and are well known to show near universal male dispersal (Alberts and Altmann 2001; Charpentier et al. 2008). Charpentier et al. (2008), found that the average dispersal age for male yellow baboons at Amboseli was ~8 years (96 months), though the timing of this dispersal event varies from individual to individual. A contributing factor to the variance in male dispersal age at Amboseli may be natural hybridization between olive and yellow baboons, with males that have substantial olive baboon ancestry tending to disperse at a younger age (Alberts and Altmann 2001).

Natural hybridization was active in the Amboseli region when the ancestors of the SFBR baboons were captured (Maples and McKern 1967) and has continued sporadically since that time (Alberts and Altmann 2001; Charpentier et al. 2012). Hybridization between yellow and olive baboons is also known to have occurred in captivity at the SFBR (Ackermann et al. 2006, 2014). The SFBR individuals examined here were reported to be unmixed yellow baboons, but it was not possible to confirm this identification by examining their external appearance. To check for olive baboon admixture, we included in the analysis *Alu* data from three olive baboons of known Kenyan ancestry (Rogers et al. under revision).

Within the diversity panel of all yellow baboons, plus the three *P. anubis* samples, 411 of our 494 loci were polymorphic, and the STRUCTURE analysis, which uses a Bayesian approach to determine the number of population clusters present in a data set and assign individuals to a specific cluster, revealed two distinct population clusters (fig. 1). The first cluster, shown in green, is at or nearly at 100% in the three olive baboons, ranging from 99.9% to 100% (the percentage shown for each cluster shows how well an individual identifies with a given population cluster). By contrast, individuals that were sampled in Mikumi National Park carry genomes composed almost exclusively (99.18–99.94%) of population Cluster 2, shown in yellow (fig. 1). The SFBR individuals show genomes with a much greater range in the percentage of each population cluster, with Cluster 1 spanning from 26.8% to 71.46% in these individuals.

Variation among Social Groups at Mikumi

To elucidate the population structure of the baboons from Mikumi National Park (table 1), a second STRUCTURE analysis was carried out (fig. 2*A*). Of our 494 loci, 295 were polymorphic in Mikumi yellow baboons. Two distinct population clusters were found, with a wide range of both clusters being found throughout the seven troops (fig. 2*A*).

In a Kruskal–Wallis one-way ANOVA test, the diversity among all seven troops failed to reach statistical significance ($\chi^2 = 11.18$, D.F. = 6, $P = 0.08$). A similar Kruskal–Wallis one-way ANOVA with troops grouped by the four trapsites (ABRU HQ, Viramba, Ikoya, Kisorobi), however, did find statistically significant diversity among sites ($\chi^2 = 11.10$, D.F. = 3,

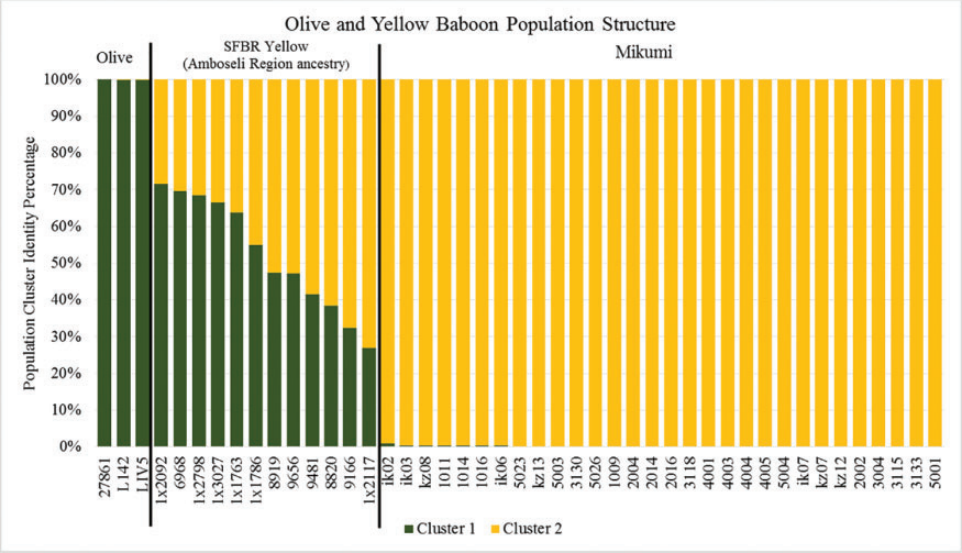

Fig. 1.—Population structure analysis using 494 *Alu* elements. Yellow individuals from Mikumi National Park and SFBR are included, along with three olive individuals. The yellow baboons from the SFBR show varying levels of admixture between the olive baboons and the yellow baboons from Mikumi. The olive baboons that were part of the analysis included the reference individual and two diversity samples, and were of Kenyan origin.

$P = 0.011$), and the difference between Viramba and ABRU, the two multi-troop sites, was still more marked (Mann–Whitney $U = 13.0$, 2-tailed $P = 0.002$) (fig. 2*B*). There was, however, no significant difference between Viramba 1 and Viramba 2 ($U = 10.0$, $P = 1.000$), or among the troops (Nyeusi, Barabara, Punk) sampled at ABRU HQ ($\chi^2 = 0.030$; D.F = 2; $P = 0.98$) (fig. 2*B*). Individual 5001 from the "Punk" troop was the only individual that was found to be a significant outlier from the troop of origin (fig. 2*B*).

Kinda Baboons

Of the 494 *Alu* elements examined in our study, 296 were polymorphic in the 15 kinda baboons on our panel. Further information for these individuals can be found in table 2. Of these, 76 were ascertained from the genome of individual BZ11050, and the another 21 elements were ascertained from the genome of individual BZ11047. The STRUCTURE analysis found two population clusters with varying levels of admixture in each individual (fig. 3*A*).

All animals trapped at Chunga HQ were assigned to population cluster 2, with variable amounts of cluster 1 admixture. Four of the five animals trapped at the Chunga School site were very similar to each other, and identified strongly with population cluster 1, with the lowest cluster identification of those four individuals being 95.2%. However, the fifth individual trapped at this site (BZ11047) showed 96.44% identity with cluster 2, thus closely resembling the animals from Chunga HQ. These relationships can also be seen in the principal component analysis (PCA) (fig. 3*B*). PC 1 makes up 13.62% of the total variation, and PC 2 makes up 12.82% of the total

Table 1
Information for Mikumi Yellow Baboon Samples

Mikumi ID	Sex	Group/Troop	Estimated Age Group	Mirgant Status
1009	Female	1/Nyeusi	Adult	Natal
1011	Female	1/Nyeusi	Adult	Natal
1014	Female	1/Nyeusi	Adult	Natal
1016	Female	1/Nyeusi	Adult	Natal
2002	Female	2/Barabara	Adult	Natal
2004	Male	2/Barabara	Adult	Migrant
2014	Female	2/Barabara	Adult	Natal
2016	Female	2/Barabara	Adult	Natal
3004	Female	3/Viramba 1	Adult	Natal
3115	Female	3/Viramba 1	Adult	Natal
3118	Male	3/Viramba 1	Subadult	Natal
3130	Female	3/Viramba 1	Juvenile	Natal
3133	Female	3/Viramba 1	Adult	Natal
4001	Female	4/Viramba 2	Adult	Natal
4003	Male	4/Viramba 2	Subadult	Natal
4004	Male	4/Viramba 2	Juvenile	Natal
4005	Male	4/Viramba 2	Adult	Migrant
5001	Male	5/Punk	Adult	Migrant
5003	Male	5/Punk	Adult	Migrant
5004	Male	5/Punk	Adult	Migrant
5023	Male	5/Punk	Juvenile	Natal
5026	Female	5/Punk	Adult	Natal
IK02	Male	6/Ikoya	Subadult	Natal
IK03	Male	6/Ikoya	Adult	Migrant
IK06	Male	6/Ikoya	Subadult	Natal
IK07	Male	6/Ikoya	Adult	Migrant
KZ07	Male	7/Kizorobi	Subadult	Natal
KZ08	Male	7/Kizorobi	Subadult	Natal
KZ12	Male	7/Kizorobi	Adult	Migrant
KZ13	Male	7/Kizorobi	Adult	Migrant

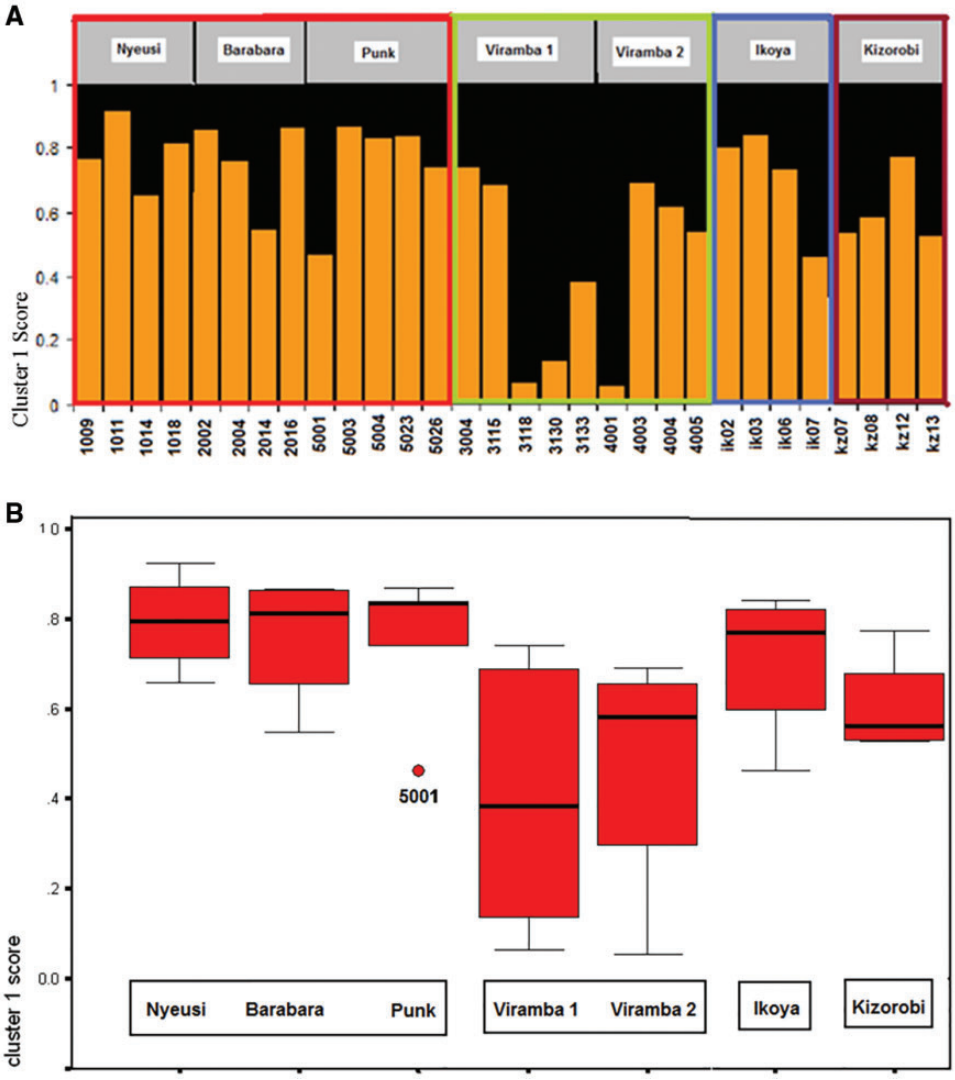

Fig. 2.—(*A*) Population structure analysis for Mikumi yellow baboons using 494 *Alu* insertions. Individuals are separated by capture site. (*B*) Box plot of distribution (median and interquartile range) of Cluster 1 scores in Mikumi animals. Results are shown for each troop, showing individual 5001 as the sole outlier. Group names contained in boxes along the *X* axis are groups collected from the same trapsite.

variation. Although the two social groups are generally separated, individual BZ11047 (a male individual), which was an outlier in the structure analysis, is also an outlier in the PCA.

Discussion

The SFBR yellow baboons were clearly distinguished from Mikumi animals by the presence, in varying amounts, of a second component, shared with olive baboons. It appears very likely that this component was derived by admixture with *P. anubis*, but more genomic analysis would be required to determine whether the interbreeding occurred in the wild over many generations of hybridization, or during the few generations of captivity at the SFBR.

Mikumi National Park is distant from any known, currently active baboon hybrid zone, likely preventing any observable levels of recent hybridization and admixture. Delving into the seven different Mikumi yellow troops from which we had samples, there was a considerable range of variation among individuals, but also wide overlap between troops in *Alu* admixture score (Cluster scores from STRUCTURE), and where more than one troop was sampled at a single trapping location, there was no significant difference in their *Alu* admixture scores.

All nine adult males in the Mikumi sample can be presumed to be immigrants to the troop in which they were living when sampled, but only one (5001, in "Punk" troop 5) was flagged by his admixture score as an outlier from his troop (fig. 2*B*). This lack of strong genetic divergence

Table 2
Kinda Baboon Sample Information

Kinda ID	Sex	Social Group	Estimated Age (months)	Weight (kg)
BZ11001	Male	Chunga HQ	188	15.95
BZ11002	Male	Chunga HQ	112	16.5
BZ11004	Male	Chunga HQ	68	9.25
BZ11005	Male	Chunga HQ	205	14.2
BZ11011	Male	Chunga HQ	50	6.15
BZ11012	Female	Chunga HQ	35	4.4
BZ11024	Female	Chunga HQ	153	10.4
BZ11030	Female	Chunga HQ	92	8.4
BZ11031	Male	Chunga HQ	130	15.6
BZ11032	Female	Chunga HQ	155	9.3
BZ11033	Male	Chunga school	76	14.9
BZ11045	Female	Chunga school	210	11.8
BZ11046	Male	Chunga school	15	2.55
BZ11047	Male	Chunga school	130	14.3
BZ11050	Female	Chunga school	76	12.3

among troops is not unexpected for a population in which male dispersal is the rule. In this situation, there may be a tendency for intertroop divergence to occur by genetic drift, especially if reproductive skew among males is pronounced, but this will be offset by intertroop gene flow, as every individual is the outbred offspring of parents born in different troops.

Our results do, however, suggest some population structuring by distance at Mikumi. The pooled membership of troops that ranged around the ABRU HQ (Nyeusi, Barabara, and Punk) differed strongly from those sampled at Viramba, living ~6 km away (Viramba 1 and 2 troops). Studies of the patterns of male dispersal in baboons (Packer 1979) and other cercopithecine species with similar male dispersal behavior (Cheney and Seyfarth 1983) suggest that males may preferentially disperse to neighboring troops, and in some cases male siblings disperse to the same nonnatal troop (Cheney and Seyfarth 1983). This may create local networks of troops that exchange males frequently and hence retain similar allele frequencies. Animal 5001, who was an outlier from his troop of residence (Punk), but resembled members of the Viramba groups, may exemplify unusually long-distance migration.

The findings from the kinda baboons at Chunga appear surprisingly different. Structure analysis showed that most individuals clustered strongly with others trapped at the same site, with little indication of admixture (fig. 3*A*). However, individual BZ11047, an adult male trapped at the Chunga School site, much more closely matched the animals trapped at Chunga HQ, This observation was reinforced by PCA, which also showed BZ11047 falling with the Chunga HQ social group. BZ11047 might have previously migrated from the HQ and joined the School group, but it is also very possible that he was an HQ troop member visiting the School baiting and trapping site, where he was captured. What we find interesting and potentially significant, however, is that members of the two social groups at Chunga, even though their ranges overlapped, were very distinct in their *Alu* admixture scores. If this finding is confirmed with larger samples, it suggests that kinda baboons, at least those at Chunga, differ from yellow, olive, and chacma baboons in patterns of male dispersal. A major difference in nuclear genetic markers between groups with closely adjacent and overlapping ranges is contrary to expectation if male dispersal were nearly universal, as in most other species of the genus. It also contrasts with the general similarity of *Alu* admixture values across the much more geographically scattered yellow baboon troops at Mikumi. This suggests that male Chunga baboons are more likely than yellow baboons at Mikumi to stay to breed within their troop of birth. Much more work will be needed, however, to determine whether this apparent contrast is a chance artifact of small samples, a peculiarity of the Chunga population, or a universal difference between the two baboon species in patterns of male dispersal.

More generally, our findings illustrate the potential of multi-locus, whole genome *Alu* insert polymorphism to document population structure, even if comparatively few individuals are sampled from each constituent subpopulation. In a context where interspecies hybridization is known or suspected, *Alu* insert polymorphisms clearly identify individuals of mixed heritage, and provide an estimate of the contribution of each parental population to the genome of such hybrid individuals. Within a widespread population of a single species, *Alu* elements can document subpopulation structure, and hence help to infer patterns of dispersal. To the extent that local clusters of troops form genetically differentiated subpopulations, *Alu* insertion polymorphism profiles may also distinguish rare, long distance migrant individuals and suggest their origin—a valuable resource for researchers in the field. Although there is extensive evidence that male migration impacts baboon population structure, future studies could investigate fission/fusion events and reproductive skew in baboons, as these factors have been found to play a role in the population genetics of other primates (Ober et al. 1984; Dittus 1988; Widdig et al. 2004). Other future studies on baboon population genetics should include more individuals from these same social groups, include new social groups for these two species of baboons, and study social groups from the other species of baboons to ensure that this approach is effective in determining genetic differences in all *Papio* species.

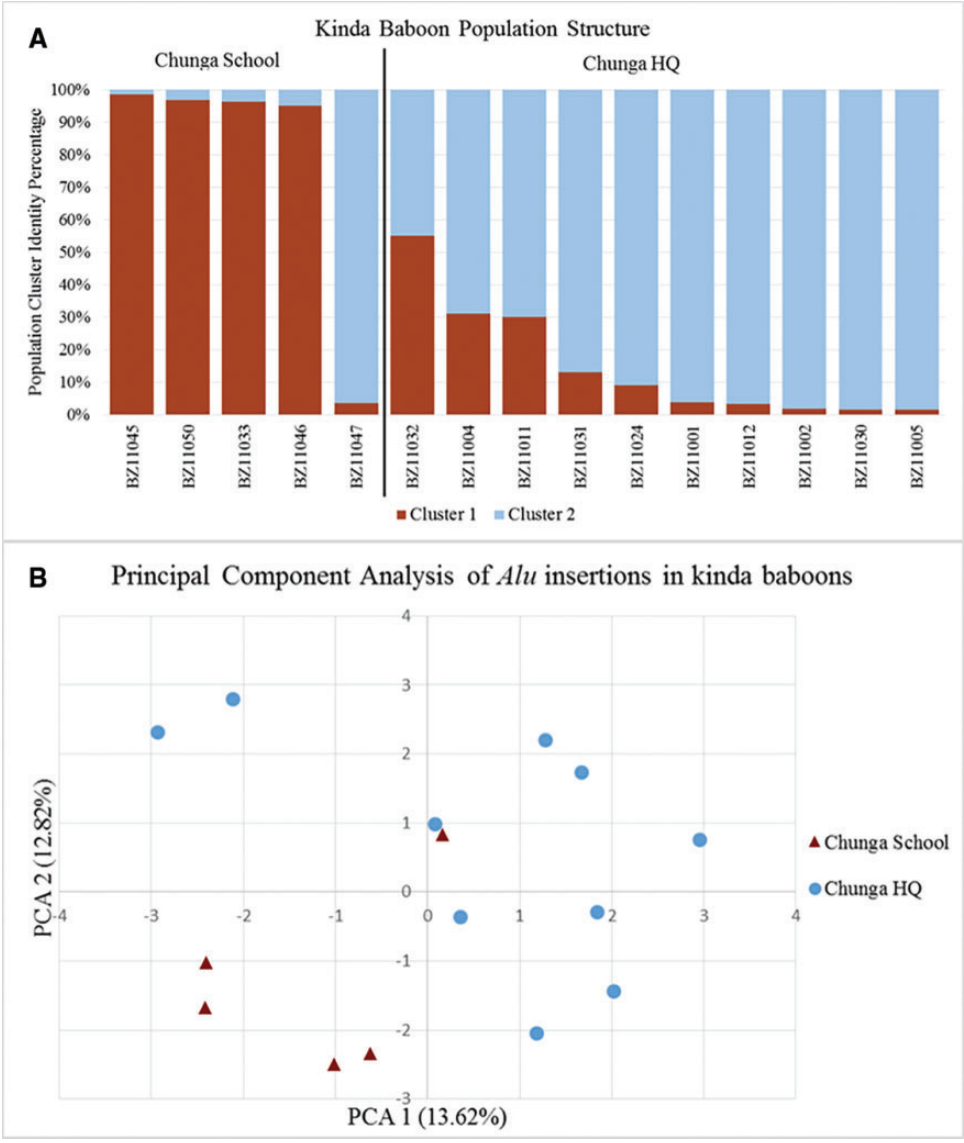

Fig. 3.—(*A*) *Alu*-based population structure analysis for 15 kinda baboons. There are two inferred population clusters, showing varying degrees of admixture. The social groups are separated by a black line. The percentage with which an individual identifies with a population cluster is shown on the *Y* axis. (*B*) Principal component analysis of 473 *Alu* insertions in our two social groups of kinda baboons. With the exception of one male individual, the Chunga School social group clusters closely together, whereas the Chunga HQ social group shows slightly more variability.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

The authors would like to thank all members of the Batzer Lab and the Baboon Genome Analysis Consortium for all of their advice and helpful suggestions. The samples received from the Southwest Foundation for Biomedical Research were sincerely appreciated. Fieldwork in Tanzania and Zambia was kindly facilitated by the Serengeti Research Institute and the Zambia Wildlife Authority, respectively. The laboratory research was supported by National Institutes of Health (RO1 GM59290) to M.A.B. A.R. was supported in part by the Louisiana Biomedical Research Network (LBRN) with funding from the National Institute of General Medical Sciences of the National Institutes of Health under Award Number P20GM103424 and by the Louisiana Board of Regents Support Fund. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or Louisiana Board of Regents. B.N.C. was funded in part by the National

Institutes of Health Award Number 5T35OD011151-12 and the LSU School of Veterinary Medicine. Fieldwork at Mikumi was funded by BNS 83-03506 to J.P.-C. and J.R., and in Zambia by NSF1029302 to J.P.-C., J.R., and C.J.J.

Author Contributions

J.A.W. and M.A.B. designed the research; C.J.J., J.P.-C., and J.R. collected kinda baboon and Mikumi yellow baboon samples, and supplied information about kinda and Mikumi yellow baboon samples; C.J.S., J.A.W., V.E.J., T.O.B., C.L.M., C.P.S.-R., E.C.B., A.R., and B.N.C. conducted the experiments; M.K.K. and T.O.B. performed the *Alu* element analysis of the reference olive baboon genome (Panu_2.0); C.J.S., V.E.J., and T.O.B. performed computational analyses of the whole genome sequence data from the Baboon Genome Consortium project at Baylor College of Medicine and designed primers; C.J.S. performed the Structure analyses and PCA; C.J.J. performed Kruskal–Wallis one-way ANOVA test and created the associated figure. C.J.S. wrote the first draft which was revised by J.A.W., M.R., K.C.W., C.J.J., J.P.-C., J.R., M.K.K., and M.A.B.

Author Information

A.R. conducted experiments for this project in the Department of Biological Sciences, LSU-Baton Rouge as a participant in the Louisiana Biomedical Research Network (LBRN) while completing a degree in the Department of Biological and Physical Sciences at Northwestern State University of Louisiana, Natchitoches, LA. B.N.C. conducted experiments for this project in the Department of Biological Sciences, LSU-Baton Rouge as a member of the NIH Biomedical Research Experience for Veterinary Scholars.

Literature Cited

Ackermann RR, Rogers J, Cheverud JM. 2006. Identifying the morphological signatures of hybridization in primate and human evolution. J Hum Evol. 51(6):632–645.

Ackermann RR, Schroeder L, Rogers J, Cheverud JM. 2014. Further evidence for phenotypic signatures of hybridization in descendant baboon populations. J Hum Evol. 76:54–62.

Alberts SC, Altmann J. 2001. Immigration and hybridization patterns of yellow and anubis baboons in and around Amboseli, Kenya. Am J Primatol. 53(4):139.

Batzer MA, Deininger PL. 2002. Alu repeats and human genomic diversity. Nat Rev Genet. 3(5):370–379.

Batzer MA, Deininger PL. 1991. A human-specific subfamily of Alu sequences. Genomics 9(3):481–487.

Batzer MA, et al. 1994. African origin of human-specific polymorphic Alu insertions. Proc Natl Acad Sci U S A. 91(25):12288–12292.

Bergey CM, Pozzi L, Disotell TR, Burrel AS. 2013. A new method for genome-wide marker development and genotyping holds great promise for molecular primatology. Int J Primatol. 34(2):303–314.

Britten RJ, Baron WF, Stout DB, Davidson EH. 1988. Sources and evolution of human Alu repeated sequences. Proc Natl Acad Sci U S A. 85(13):4770–4774.

Charpentier MJ, et al. 2012. Genetic structure in a dynamic baboon hybrid zone corroborates behavioural observations in a hybrid population. Mol Ecol. 21(3):715–731.

Charpentier MJ, Tung J, Altmann J, Alberts SC. 2008. Age at maturity in wild baboons: genetic, environmental and demographic influences. Mol Ecol. 17(8):2026–2040.

Cheney DL, Seyfarth RM. 1983. Nonrandom dispersal in free-ranging vervet monkeys: social and genetic consequences. Am Nat. 122(3):392–412.

Cordaux R, Batzer MA. 2009. The impact of retrotransposons on human genome evolution. Nat Rev Genet. 10(10):691–703.

Cordaux R, Hedges DJ, Batzer MA. 2004. Retrotransposition of Alu elements: how many sources? Trends Genet. 20(10):464–467.

Cox LA, et al. 2013. Baboons as a model to study genetics and epigenetics of human disease. Ilar J. 54(2):106–121.

Deininger PL, Batzer MA, Hutchison CA III, Edgell MH. 1992. Master genes in mammalian repetitive DNA amplification. Trends Genet. 8(9):307–311.

Deininger PL, Moran JV, Batzer MA, Kazazian HH Jr. 2003. Mobile elements and mammalian genome evolution. Curr Opin Genet Dev. 13(6):651–658.

Dewannieux M, Esnault C, Heidmann T. 2003. LINE-mediated retrotransposition of marked Alu sequences. Nat Genet. 35(1):41–48.

Dittus WPJ. 1988. Group fission among wild toque macaques as a consequence of female resource competition and environmental stress. Anim Behav. 36(6):1626–1645.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32(5):1792–1797.

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164(4):1567–1587.

Fischer J, et al. 2017. Charting the neglected West: the social system of Guinea baboons. Am J Phys Anthropol. 162:15–31.

Frost SR, Marcus LF, Bookstein FL, Reddy DP, Delson E. 2003. Cranial allometry, phylogeography, and systematics of large-bodied papionins (primates: Cercopithecinae) inferred from geometric morphometric analysis of landmark data. Anat Rec A Discov Mol Cell Evol Biol. 275(2):1048–1072.

Grubb P, et al. 2003. Assessment of the diversity of African primates. Int J Primatol. 24(6):1301–1357.

Hamdi H, Nishio H, Zielinski R, Dugaiczyk A. 1999. Origin and phylogenetic distribution of Alu DNA repeats: irreversible events in the evolution of primates. J Mol Biol. 289(4):861–871.

Han K, et al. 2005. Under the genomic radar: the stealth model of Alu amplification. Genome Res. 15(5):655–664.

Hartig G, et al. 2013. Retrophylogenomics place tarsiers on the evolutionary branch of anthropoids. Sci Rep. 3:1756.

IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk (NY): IBM Corp.

Jolly CJ. 1993. Species, subspecies and baboon systematics. In: Kimbel WH, Martin LB, editors. Species, species concepts, and primate evolution. New York: Plenum Press.

Jolly CJ. 2001. A proper study for mankind: analogies from the Papionin monkeys and their implications for human evolution. Am J Phys Anthropol. 116(Suppl 33):177–204.

Jolly CJ, Burrell AS, Phillips-Conroy JE, Bergey C, Rogers J. 2011. Kinda baboons (*Papio kindae*) and grayfoot chacma baboons (*P. ursinus griseipes*) hybridize in the Kafue river valley, Zambia. Am J Primatol. 73(3):291–303.

Jurka J, Smith T. 1988. A fundamental division in the Alu family of repeated sequences. Proc Natl Acad Sci U S A. 85(13):4775–4778.

Konkel MK, Walker JA, Batzer MA. 2010. LINEs and SINEs of primate evolution. Evol Anthropol. 19(6):236–249.

Kriegs JO, Churakov G, Jurka J, Brosius J, Schmitz J. 2007. Evolutionary history of 7SL RNA-derived SINEs in Supraprimates. Trends Genet. 23(4):158–161.
Levin HL, Moran JV. 2011. Dynamic interactions between transposable elements and their hosts. Nat Rev Genet. 12(9):615–627.
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14):1754–1760.
Li J, et al. 2009. Phylogeny of the macaques (Cercopithecidae: Macaca) based on Alu elements. Gene 448(2):242–249.
Maples WR, McKern TW. 1967. A preliminary report on classification of the Kenya baboon. Baboon Med Res. 2:13–22.
McLain AT, et al. 2013. Analysis of western lowland gorilla (*Gorilla gorilla gorilla*) specific Alu repeats. Mob DNA 4(1):26.
Meyer TJ, et al. 2012. An Alu-based phylogeny of gibbons (hylobatidae). Mol Biol Evol. 29(11):3441–3450.
Newman TK, Jolly CJ, Rogers J. 2004. Mitochondrial phylogeny and systematics of baboons (Papio). Am J Phys Anthropol. 124(1):17–27.
Norton GW, Rhine RJ, Wynn GW, Wynn RD. 1987. Baboon diet: a five-year study of stability and variability in the plant feeding and habitat of the yellow baboons (*Papio cynocephalus*) of Mikumi National Park, Tanzania. Folia Primatol (Basel) 48(1–2):78–120.
Ober C, et al. 1984. Demographic components of gene frequency change in free-ranging macaques on Cayo Santiago. Am J Phys Anthropol. 64(3):223–231.
Olivier TJ, Buettner-Janusch J, Buettner-Janusch V. 1974. Carbonic anhydrase isoenzymes in nine troops of Kenya baboons, *Papio cynocephalus* (Linnaeus 1766). Am J Phys Anthropol. 41(2):175–189.
Packer C. 1979. Inter-troop transfer and inbreeding avoidance in *Papio anubis*. Anim Behav. 27:1–36.
Phillips-Conroy JE, Jolly CJ. 1988. Dental eruption schedules of wild and captive baboons. Am J Primatol. 15(1):17–29.
Premawardhana U, Adams MR, Birrell A, Yue DK, Celermajer DS. 2001. Cardiovascular structure and function in baboons with Type 1 diabetes – a transvenous ultrasound study. J Diabetes Complications 15(4):174–180.
R Development Core Team. 2016. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
Ray DA, et al. 2005a. Inference of human geographic origins using Alu insertion polymorphisms. Forensic Sci Int. 153(2–3):117–124.
Ray DA, et al. 2005b. Alu insertion loci and platyrrhine primate phylogeny. Mol Phylogenet Evol. 35(1):117–126.
Ray DA, Xing J, Salem AH, Batzer MA. 2006. SINEs of a nearly perfect character. Syst Biol. 55(6):928–935.
Rhine RJ, Norton GW, Rogers J, Wasser SK. 1992. Secondary sex ratio and maternal dominance rank among wild yellow baboons (*Papio cynocephalus*) of Mikumi National Park, Tanzania. Am J Primatol. 27(4):261–273.
Rhine RJ, Wasser SK, Norton GW. 1988. Eight-year study of social and ecological correlates of mortality among immature baboons of Mikumi National Park, Tanzania. Am J Primatol. 16(3):199–212.
Rogers J, et al. 1989. Genetic structure and microevolution in a population of Tanzanian yellow baboons (*Papio hamadryas cynocephalus*) [PhD thesis]. [Princeton, NJ]: Yale University.
Rogers J, Kidd KK. 1993. Nuclear DNA polymorphisms in a wild population of yellow baboons (*Papio hamadryas cynocephalus*) from Mikumi National Park, Tanzania. Am J Phys Anthropol. 90(4):477–486.
Rogers J, Kidd KK. 1996. Nucleotide polymorphism, effective population size, and dispersal distances in the yellow baboons (*Papio hamadryas cynocephalus*) of Mikumi National Park, Tanzania. Am J Primatol. 38(2):157–168.
Rogers J, et al. Under revision. The comparative genomics, epigenomics and complex population history of *Papio* baboons. Nature.
Roos C, Schmitz J, Zischler H. 2004. Primate jumping genes elucidate strepsirrhine phylogeny. Proc Natl Acad Sci U S A. 101(29):10650–10654.
Salem AH, et al. 2003. Alu elements and hominid phylogenetics. Proc Natl Acad Sci U S A. 100(22):12787–12791.
Schmitz J, Ohme M, Zischler H. 2001. SINE insertions in cladistic analyses and the phylogenetic affiliations of Tarsius bancanus to other primates. Genetics 157(2):777–784.
Shotake T, Nozawa K, Tanabe Y. 1977. Blood protein variations in baboons. I. Gene exchange and genetic distance between Papio anubis, Papio hamadryas and their hybrid. Jpn J Genet. 52(3):223.
Slagel V, Flemington E, Traina-Dorge V, Bradshaw H, Deininger P. 1987. Clustering and subfamily relationships of the Alu family in the human genome. Mol Biol Evol. 4(1):19–29.
Socha WW, Wiener AS, Moor-jankowski J, Jolly CJ. 1977. Blood groups of baboons. Population genetics of feral animals. Am J Phys Anthropol. 47(3):435–442.
St. George D, et al. 1998. Microsatellite variation in two populations of free-ranging yellow baboons (*Papio hamadryas cynocephalus*). Int J Primatol. 19:273–285.
Swedell L, et al. 2011. Female "dispersal" in hamadryas baboons: transfer among social units in a multilevel society. Am J Phys Anthropol. 145(3):360–370.
Szmulewicz MN, et al. 1999. An Alu insertion polymorphism in a baboon hybrid zone. Am J Phys Anthropol. 109(1):1–8.
Untergasser A, et al. 2012. Primer3–new capabilities and interfaces. Nucleic Acids Res. 40(15):e115.
Walker JA, et al. 2012. Orangutan Alu quiescence reveals possible source element: support for ancient backseat drivers. Mob DNA 3:8.
Wasser SK, Starling AK. 1988. Proximate and ultimate causes of reproductive suppression among female yellow baboons at Mikumi National Park, Tanzania. Am J Primatol. 16(2):97–121.
Weyher AH, Phillips-Conroy JE, Fourrier MS, Jolly CJ. 2014. Male-driven grooming bouts in mixed-sex dyads of Kinda baboons (*Papio kindae*). Folia Primatol (Basel) 85(3):178–191.
Widdig A, et al. 2004. A longitudinal analysis of reproductive skew in male rhesus macaques. Proc R Soc B Biol Sci. 271(1541):819.
Willard C, Nguyen HT, Schmid CW. 1987. Existence of at least three distinct Alu subfamilies. J Mol Evol. 26(3):180–186.
Witherspoon DJ, et al. 2006. Human population genetic structure and diversity inferred from polymorphic L1(LINE-1) and Alu insertions. Hum Hered. 62(1):30–46.
Woolley-Barker T. 1999. Social organization and genetic structure in a baboon hybrid zone [PhD thesis]. [ProQuest NYU]: New York University.
Wright S. 1943. Isolation by distance. Genetics 28(2):114–138.
Yeung KR, et al. 2016. A cross-sectional study of ageing and cardiovascular function over the baboon lifespan. PLoS One 11(7):e0159576.
Zinner D, Wertheimer J, Liedigk R, Groeneveld LF, Roos C. 2013. Baboon phylogeny as inferred from complete mitochondrial genomes. Am J Phys Anthropol. 150(1):133–140.

Associate editor: Emmanuelle Lerat