Molecular Phylogenetics and Evolution 127 (2018) 376–386

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Ancient hybridization and admixture in macaques (genus *Macaca*) inferred from whole genome sequences

Zhenxin Fan[a], Anbo Zhou[b], Naoki Osada[c], Jianqiu Yu[d], Juan Jiang[a], Peng Li[a], Lianming Du[a], Lili Niu[d], Jiabo Deng[d], Huailiang Xu[e], Jinchuan Xing[b], Bisong Yue[a,*], Jing Li[a,*]

[a] *Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu 610065, Sichuan, People's Republic of China*
[b] *Department of Genetics, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA*
[c] *Graduate School of Information Science and Technology, Hokkaido University, Sapporo, Hokkaido 060-0814, Japan*
[d] *Chengdu Zoo, Institute of Chengdu Wildlife, Chengdu 610081, People's Republic of China*
[e] *College of Animal Science and Technology, Sichuan Agricultural University, Ya'an, People's Republic of China*

ARTICLE INFO

Keywords:
Primate
Genome evolution
Genetic diversity
Hybridization

ABSTRACT

The evolutionary history of the stump-tailed macaque (*Macaca arctoides*) and its genetic relationship to other macaques is a subject of continuing controversy. Here, we have reported the first genome sequences of two stump-tailed macaques and one Assamese macaque (*M. assamensis*). Additionally, we have investigated the genetic diversity between macaque species and analyzed ancient hybridization events. Genome-wide analyses demonstrated that the stump-tailed macaque is more closely related to *sinica* species than to *fascicularis/mulatta* species. This topology contradicts the mitochondrial sequence-based phylogeny that places the stump-tailed macaque into the *fascicularis/mulatta* group. However, our results further show that stump-tailed macaques have genetic backgrounds distinct from *sinica* species, and present evidence of gene flows with rhesus macaques. We suggest that an ancient introgression occurred after stump-tailed macaques diverged from *sinica* species. The distinct gene flow between *proto-arctoides* and *proto-mulatta* resulted in the transfer of rhesus macaque-type mitochondria into *proto-arctoides*. The rhesus macaque-type mitochondria remained in populations because of genetic drift during the bottleneck. The PSMC results and morphological and geographic evidence are consistent with the mitochondria capture pattern in the stump-tailed macaque. The molecular clock estimates suggest that the mitochondrial transference into stump-tailed macaques occurred 0.4–1.4 million years ago. Furthermore, we detected extensive admixtures between different macaque species, indicating that gene flow has played an important role in the evolutionary history of the genus *Macaca*.

1. Introduction

Natural hybridizations or gene flow between closely related species are important mechanisms for speciation and the evolutionary history of animals, and an appreciation of their roles in the evolutionary diversification has been growing over the past decade (Pinho and Hey, 2010; Zinner et al., 2011; Abbott et al., 2013). Through the introduction of new genetic variation and new allelic combinations, hybridization may influence the evolutionary trajectory of the hybrid population, the parental populations, or both (Arnold, 1992; Rieseberg, 1997; Rieseberg et al., 2003). Natural hybridizations were reported in almost all the major primate clades (Cortés-Ortiz et al., 2007; Rumpler et al., 2008; Zinner et al. (2009a); Ackermann and Bishop, 2010; Reich et al., 2010), including macaques. To date, there is no genome-wide report about the importance of ancient hybridization and admixture for macaque evolutionary history and macaque speciation.

The genus *Macaca* is one of the most successful primate radiations. Macaques are widely distributed in Central and Southeast Asia, with the exception of the Barbary macaque (*M. sylvanus*) in North Africa (Fooden, 1979). A new macaque species *M. leucogenys* was reported in southern Tibet recently (Li et al., 2015), which increases the number of macaque species to at least 23. Consequently, *Macaca* is one of the most species richness genera within the Old World Monkey (OWM) primate lineage. With such a variety of forms and species differing in

Abbreviations: CR1, Chinese rhesus macaque from Yunnan Province; CR2, Chinese rhesus macaque from Sichuan Province; CE1, crab-eating macaque from Vietnam; CE2, crab-eating macaque from Malaysia; TM, Tibetan macaque; XH, Assamese macaque; SM1, stump-tailed macaque 1; SM2, stump-tailed macaque 2

* Corresponding authors.
E-mail addresses: bsyue@scu.edu.cn (B. Yue), ljtjf@126.com (J. Li).

https://doi.org/10.1016/j.ympev.2018.03.038
Received 2 July 2017; Received in revised form 23 March 2018; Accepted 30 March 2018
Available online 31 March 2018

morphology, habitat and behavior, macaques represent an attractive species complex for studying species radiation and evolution.

Macaques are a monophyletic group in the tribe Papionini, diverging from other primates in northern Africa during the late Miocene from 7 to 8 million years ago (mya) (Delson, 1980). Macaques invaded Eurasia about 5.5 mya and then split into several phyletic lineages in Asia. Based on morphological (Fooden, 1976; Delson, 1980) and molecular evidence (Hoelzer and Melnick, 1996), the extant macaque species can be separated into four distinct species groups: *sylvanus*, *silenus*, *sinica*, and *fascicularis* (Zinner et al., 2013; Roos et al., 2014). Alternatively, extant macaque species have been delineated into seven species groups, where there are an additional three groups to the aforementioned four and these are the arctoides, mulatta and Sulawesi groups. In the seven species group definition, *M. arctoides* is separated from the *sinica* group creating the *arctoides* group. While, *M. mulatta*, *M. fuscata*, and *M. cyclopis* are separated from the *fascicularis* group forming the *mulatta* group. Sulawesi macaques are separated from the *silenus* group to form the Sulawesi group (Zinner et al., 2013; Roos et al., 2014).

The speciation and radiation of the Asian lineage occurred relatively recently at about 3 mya. This process was further influenced by natural events, such as climatic and eustatic changes during the Late Pliocene and Pleistocene, with potentially significant habitat disturbance, range expansion and/or declines (Zinner et al. (2009b); Abbott et al., 2013). Therefore, hybridization or secondary contact among different macaque lineages may have occurred and resulted in a complex phylogeny and evolutionary history of macaques (Eudey (1979); Delson, 1980). In fact, hybridizations or gene flow have been reported either by field observations or by phylogenetic analysis in the rhesus macaque (*M. mulatta*), crab-eating macaque (*M. fascicularis*) (Fooden, 1995; Hamada et al., 2006; Yan et al., 2011), Bornean pig-tailed macaques (*M. nemestrina*) (Ziegler et al., 2007), various macaque species on the island of Sulawesi (Ciani et al., 1989; Evans et al., 2001, Evans et al. (2003)), and the Arunachal macaque (*M. munzala*) (Sinha (2013)). Molecular evidence also suggests that ancient hybridizations occurred between different species groups (*e.g. sinica* and *fascicularis* group) (Tosi et al., 2000, 2003a, 2003b), the Tibetan macaque (*M. thibetana*) and the Chinese rhesus macaque (Fan et al., 2014).

In particular, the stump-tailed macaque (*M. arctoides*), named for its characteristically short tail, may also have a hybrid origin. The continental distribution of this species ranges from southern China and northeastern India to the northwestern tip of West Malaysia on the Malay Peninsula (Fooden, 1990; Groves, 2001). Compared to other macaque species, the stump-tailed macaque has the most distinctive reproductive anatomy (strikingly prolonged glans penis and exocervix) (Fooden, 1985) and unique reproductive behavior. Therefore, its evolutionary relationship to other macaques is debatable (Table S1). Moreover, the species was once confused with the Tibetan macaque, a member of the *sinica* group, and these two species were regarded as only subspecifically distinct in China (Bertrand, 1969). Fooden (1990) suggested that the stump-tailed macaque was probably derived from a Tibetan macaque-like ancestor. However, previous studies attempting to determine the phylogenetic position of the stump-tailed macaque generated conflicting results. After comparing the discrepancy between different molecular systems, Tosi et al. (2000, 2003a) and Li et al. (2009) suggested a hybrid origin of the stump-tailed macaque. However, these studies were based on a small number of markers. Genome-wide analysis of more closely related macaque species could provide compelling evidence for the hybridization hypothesis. Additionally, genome-wide analysis can investigate the pattern and underlying mechanism of the hybridization and speciation event in macaque evolutionary history.

Recently, the genomic approach has become an effective tool to resolve admixture issues and has been successfully applied to various taxa (*e.g.* wolves and dogs (Freedman et al., 2014; Fan et al., 2016); macaques (Yan et al., 2011; Fan et al., 2014); and Neandertals and humans (Green et al., 2010). The Indian rhesus macaque (*M. mulatta mulatta*) genome (Rhesus Macaque Genome Sequencing and Analysis Consortium, et al., 2007) provides a high-quality reference for other macaques (Fang et al., 2011; Higashino et al., 2012; Fan et al., 2014; Osada et al., 2015). In this study, we generated whole genome sequences of two stump-tailed macaques and one Assamese macaque (*M. assamensis*) at high coverage. We assembled a dataset of nine macaque genomes (including the reference genome) and conducted comprehensive analyses based on genome-wide data and mitochondrial genome data. We aimed to: (1) investigate the genetic background of the stump-tailed macaque and Assamese macaque, and their genetic relationship to other macaque species; (2) determine the hybridization mechanism in stump-tailed macaque's evolutionary history; and (3) investigate hybridization and gene flow between macaque species.

2. Materials and methods

2.1. Ethics statement

All samples were obtained following the regulations for the implementation of China on the protection of terrestrial wild animals (State Council Decree [1992] No. 13). All animals received humane care and study protocols complied with the institution's guidelines. All animal protocols involved in this study were reviewed and approved by the Ethics Committees of Sichuan University, China. All procedures employed in the experiments were examined and approved by the Ethics Committees of Sichuan University, China. Throughout the procedure, care was taken to avoid macaque suffering and to ensure animal welfare.

2.2. Samples and sequencing

Blood samples from two female stump-tailed macaques and one male Assamese macaque were collected from Chengdu Zoo (Sichuan Province, China). According to zoo records, the two stump-tailed macaques were rescued from the wild (Yunnan province) as they had been injured by poachers. The macaques were badly injured and it was determined that it was in their best interests to be housed in the zoo, rather than being released into the wild. The two individuals have since been captive in the Chengdu Zoo. The Assamese macaque was transferred from Guangxi Zoo (Guangxi Province, China), and was likely a wild-born individual. Genomic DNA was extracted from whole blood using the standard phenol–chloroform method (Sambrook and Russell, 2001). The whole genome sequencing was performed using Illumina HiSeq 2000 and 2500 at BGI and Novogene in China. Paired-end libraries with insert sizes of ~500 bp were generated for each sample. Library preparation and all sequencing runs were performed according to the manufacturer's protocols. The clean reads of the two stump-tailed macaques and the Assamese macaque have been deposited in the NCBI Short Read Archive under accession number PRJNA305009.

Together with published genomes, a dataset was assembled with full genome sequences of nine macaques and one Guinea baboon (*Papio papio*) as an outgroup (Table S2). Rhesus, crab-eating and Tibetan macaque genotypes were obtained from our previous study (Fan et al., 2014). The Guinea baboon raw data was downloaded from the Illumina HiSeq 2000 platform, and consisted of one male from NCBI (SRX652597 and SRX652598) and was genotyped with our pipeline.

2.3. Short reads mapping, genotyping, and post-genotype filters

The 100 bp (stump-tailed macaque 1) and 125 bp (stump-tailed macaque 2 and Assamese macaque) paired-end short reads were aligned to the Indian rhesus macaque genome (rheMac2) using Bowtie2 (Langmead and Salzberg, 2012) under the local alignment algorithm with the very sensitive model (-D 20 -R 3 -N 0 -L 20 -i S, 1, 0.50) and proper insert sizes. Default options were used for other parameters.

Picard and GATK toolsets (Depristo et al., 2011) were then applied to process the alignments to single nucleotide variant (SNV) calls. The whole pipeline converted the short reads to bam format alignment files, and then generated genotype calls in Variant Call Format (VCF). The improperly mapped reads and PCR duplicates were discarded during the genotyping. The pipeline is the same as used in our previous studies (see Fan et al., 2014, 2016; Freedman et al., 2014; Zhang et al., 2014), and the details of command lines are listed in the Supplementary Note.

Based on the raw SNV calls, a series of data quality filters were applied to improve the quality of genotype calls. These filters were designed to minimize the errors from sequencing and alignment, and to exclude the regions exhibiting accelerated evolutionary rates that are not caused by positive selection (Fan et al., 2014, 2016; Freedman et al., 2014; Zhang et al., 2014). Therefore, we used two levels of filters, the Genome Filters (GF) and Sample Filters (SF). The GFs were based on the features of the reference genome and polymorphism across all the samples, while SFs were based on the genotype results of each independent sample. We only used sites that were successfully passed through GFs and SFs. Details of these filters were described in our previous study (Fan et al., 2014).

2.4. Population genetic analysis

Pairwise allele-sharing genetic distance calculation and PCA were performed using MATLAB as previously described in Xing et al. (2009, Xing et al. (2010)). All variants in the genome were used for the genetic distance calculation between macaque species. To improve resolution, the Guinea baboon was removed and only the eight macaques (without reference) were included in the PCA. An Autosomal Neighbor-joining (NJ) tree was constructed based on allele-sharing distance between samples using the neighbor-joining method implemented by the Matlab® Bioinformatics Toolbox. Branch support was calculated by bootstrap analysis consisting of 1000 samplings with replacements, and a consensus tree was inferred using the consense program of Phylip 3.69 (Plotree and Plotgram, 1989). Nine sequences were used for mitochondrial genome phylogenetic analysis, including two *M. arctoides* mitochondrial genomes generated in this study and the following published genomes: *M. assamensis* (NC_023795), *M. fascicularis* (NC_012670), *M. mulatta* (NC_005943), *M. sylvanus* (NC_002764), *M. thibetana* (NC_011519) and *M. arctoides* (NC_025513). The Guinea baboon (NC_020009) was used as an outgroup. Sequences were aligned with MEGA7 (Kumar et al., 2016). MrBayes v3.1.2 (Ronquist and Huelsenbeck, 2003) was used for Bayesian inference (BI). The Akaike Information Criterion (AIC) was used to choose the optimal nucleotide substitution model. Phylogenetic analyses were performed using the parameters of best-fit GTR + G model and was run for 5,000,000 generations, sampling every 100 generations.

The program MCMCTree within the PAML package (Yang, 1997) was used for species divergence time estimates. MCMCTree generates a phylogenetic tree based on the topology of the autosomal NJ tree, incorporating Bayesian estimates of species divergence times, calculated using soft calibration point constraints. Three calibration points (LCA of *sinica-fascicularis* group (2.89–2.90 mya), *fascicularis* group (1.67–1.68 mya) and *sinica* group (1.39–1.40 mya)) were based on five autosomal genes (*ALB3, IRBP3, TNP2, TTR1* and *vWF11*), two Y-chromosomal loci (*SRY* and *TSPY*), and one Xq13.3 (Jiang et al., 2016). Correlation between mutation rates and branches were allowed, and likelihoods from all five possible mutation models (JC69, K80, F81, F84, HKY85) were compared to select the best mutation model for estimating mutation rates. The model HKY85 was used since it had the maximum likelihood ($\ln L = -122536446.3416$). Since the divergence time was estimated using calibration points in calendar years, it was not necessary to specify generation time.

To estimate the divergence between stump-tailed macaque and other macaques, genetic distance was calculated using the genetic distance metric described by Gronau et al. (2011). The genome was scanned with 50 kb non-overlapping window, and then pairwise comparisons were calculated between and within different species across the genome.

Genome-wide admixture estimates were obtained using a model-based algorithm implemented in ADMIXTURE (version 1.02) (Alexander et al., 2009). Single nucleotide variants (SNVs) were filtered to eliminate the effects of SNVs in linkage disequilibrium. Filtering involved the removal of SNVs that had a pairwise r2 > 0.2 within 50 SNV windows using PLINK (Purcell et al., 2007), as recommended by the authors of ADMIXTURE.

2.5. PSMC

The pairwise sequentially Markovian coalescent (PSMC) (Li and Durbin, 2011) method was used to infer demographic history. Briefly, the method uses the distribution of heterozygote sites across the genome and a Hidden Markov Model to reconstruct the history of effective population sizes. The key parameters have been explored and successfully applied in similar macaque genomic studies (Higashino et al., 2012; Fan et al., 2014). The following parameters were used in this study: numbers of iterations = 25, time interval = 1 * 6 + 58 * 1, mutation rate per generation = 2.5×10^{-8} and generation time = 6. To validate the confidence in PSMC findings, 100 bootstrap replicates were run for each genome. The genome was divided into segments of 5 Mb in length to sample a bootstrap replicate. The segments were then sampled with replacement to obtain a sequence with approximately the same length as the original genome defined by the "-b" option in the PSMC software.

2.6. ABBA-BABA tests

The ABBA-BABA test (*D*-statistics) was performed between closely related populations to test whether there was gene flow between the stump-tailed macaque and other macaques. The ABBA-BABA test detected asymmetries in allele sharing between either of two receiving lineages (P1 and P2), and a source lineage (P3), given an outgroup (O) (Durand et al., 2011). For each comparison, the *D* statistic was calculated in 1 Mb windows along the genome. While, only sites passing GF2 and SF filters were considered. One allele was randomly selected from each genotype for each site. Following Durand et al. (2011), the standard error of the statistic was calculated using a jackknife procedure and a Z-score was obtained by dividing the value of the *D* statistic by its standard error. Z-scores with absolute values ≥3 were considered significant, indicating evidence for gene flow between the P3 and one of the receiving lineages (P1 for negative Z-scores, P2 for positive values). Different macaques were assigned as P1, P2 and P3 and the Guinea baboon was used as an outgroup (O) in all tests. The Chinese rhesus macaque (Sichuan) was excluded in these analyses due to its low coverage. This test was performed with our own Python script, which has been successfully used in our previous studies (Fan et al., 2014, 2016).

3. Results

3.1. Genome data and genetic diversity

We sequenced two female stump-tailed macaques and one male Assamese macaque. Overall, we generated more than 2.7 billion clean reads in the three genomes (Table S2). In each sample, more than 94% of the reads were mapped to the reference genome, resulting in an effective coverage from 20.5 to 54.0 fold for the three samples (Table S2). Together with published genomes, we assembled the full genome sequences of nine macaques (including the reference genome rheMac2) and one outgroup (Guinea baboon) (Tables S2 and S3). All the macaque genomes in this study have higher than 20-fold coverage, except one (Chinese rhesus macaque from Sichuan) that had a low coverage of 10.52-fold (Table S3).

Table 1
The numbers of different sites, genome-wide heterozygosity and SNP rate in each sample.

Sample	Total useable sites	Non-variant sites	SNPs			Heterozygosity	SNP rate
			Heterozygous	Homozygous	Total		
Chinese rhesus macaque (Yunnan)	2,264,095,197	2,254,841,365	5,843,678	3,410,154	9,253,832	0.00258	0.00409
Chinese rhesus macaque (Sichuan)	1,637,245,909	1,631,075,721	4,219,597	1,950,591	6,170,188	0.00258	0.00377
Malaysian crab-eating macaque	2,245,361,069	2,233,769,107	6,654,349	4,937,613	11,591,962	0.00296	0.00516
Vietnamese crab-eating macaque	2,261,970,886	2,250,120,265	7,100,845	4,749,776	11,850,621	0.00314	0.00524
Tibetan macaque	2,281,356,163	2,269,585,173	2,019,479	9,751,511	11,770,990	0.00089	0.00516
Assamese macaque	2,011,347,545	1,999,354,817	5,364,861	6,627,867	11,992,728	0.00267	0.00596
Stump-tailed macaque 1	2,280,352,231	2,267,919,335	3,646,467	8,786,429	12,432,896	0.00160	0.00545
Stump-tailed macaque 2	2,079,812,789	2,069,015,844	3,106,273	7,690,672	10,796,945	0.00149	0.00519
Guinea baboon	2,118,205,588	2,094,916,416	1,109,430	22,179,742	23,289,172	0.00052	0.01099

We then identified SNVs in each sample (Table 1). All three new genomes had more than two billion total useable sites, which covered more than 80% of the reference genome (stump-tailed macaques: 91.04% and 83.03%; Assamese macaque: 80.3%). The SNV rates were consistent between the three genomes, with the highest rate in the Assamese macaque (0.00596; Table 1). All three new genomes had more than 10 million SNVs, similar to other macaques (Table 1). The Transitions/Transversions (Ti/Tv) ratios (2.17–2.21) are consistent with previous human and macaque genomics studies (2–2.2) (DePristo et al., 2011; Yan et al., 2011; Lachance et al., 2012; Fan et al., 2014). The two stump-tailed macaques had the lowest genome-wide autosomal SNV heterozygosity (0.0016 and 0.00149) among the macaques, with the exception of the Tibetan macaque (0.00089) (Table 1).

We further identified the species-specific homozygous SNVs (Fig. S1) and excluded the Chinese rhesus macaque (Sichuan) from this analysis due to its low coverage. The Tibetan macaque had the highest number of specific SNVs (1,026,502), whereas stump-tailed macaque 1 (69,587) and stump-tailed macaque 2 (62,722) had the lowest number of the seven macaques. The Assamese macaque had more specific SNVs (378,691) than the Chinese rhesus macaque (Yunnan; 349,420). Although the two crab-eating macaques had very similar numbers of sample-specific SNVs (476,120 and 472,016), the species-specific SNVs shared between them are only 142,166. This result is consistent with the two macaques being from very different populations and one of them had gene flow with the Chinese rhesus macaque. In contrast, the two stump-tailed macaques shared a very large number of species-specific SNVs (916,864), indicating a very similar genetic background.

To provide a measure of the divergence between the different macaques, we estimated the overall genetic distance (Table S4). The two stump-tailed macaques had the smallest genetic distance (0.0010), smaller than the distance between the two Chinese rhesus macaques (0.0024) and between the two crab-eating macaques (0.0033). The genetic distances of the two stump-tailed macaques to the Assamese macaque (0.0035) and to the Tibetan macaque (0.0036) were similar and substantially lower than the distances of stump-tailed macaques to either of the *fascicularis* group species. The Assamese macaque had the smallest genetic distance from the Tibetan macaque (0.0027), indicating the Assamese macaque and Tibetan macaque share a very close genetic relationship. We also calculated the pairwise genetic distance between the three new macaques with other macaques in the 50 kb non-overlapping window (Fig. 1 and Table S5). The results were consistent with the overall genetic distance (Table S4).

3.2. Phylogenetic tree and PCA

All autosomal SNVs were used to construct a genome-wide phylogenetic tree (Fig. 2a). The overall topology supported two major clades within sampled macaques with 100% bootstrap support for every branch. The first clade included the two Chinese rhesus macaques and two crab-eating macaques, whereas the second clade included the two stump-tailed macaques, Tibetan macaque, and Assamese macaque. Specifically, the two stump-tailed macaques grouped together, while the Tibetan macaque grouped with the Assamese macaque forming a sister clade to the stump-tailed macaques. This tree topology is consistent with the Y-chromosome tree proposed by Tosi et al. (2000), which clearly grouped the stump-tailed macaque into the *sinica* group.

However, the mitochondrial tree based on 12 protein-coding genes showed different topologies (Fig. 2b). Mitochondrial sequences of the two stump-tailed macaques in this study were first grouped with the third stump-tailed macaque sequence (GenBank No. NC_025513). In contrast to the autosome tree, the mitochondrial tree placed the three stump-tailed macaques in the *fascicularis/mulatta* clade, being more closely related to the rhesus macaque (NC_005943) than to the crab-eating macaque (NC_012670). The position of the Assamese macaque is consistent in both the autosome tree and the mitochondrial genome tree (Fig. 2a and b).

A principal component analysis (PCA) without outgroup showed that macaques were divided into three clusters: the stump-tailed macaques, the *sinica* group, and the *fascicularis/mulatta* macaques (Fig. 2c). The first principal component (PC1) accounted for ~52.8% of the variance in the dataset, where the stump-tailed macaques and the *sinica* group are separated from the *fascicularis/mulatta* macaques. PC2 (18.3% of the variance) separated stump-tailed macaques from the Tibetan macaque and the Assamese macaque. Consistent with the phylogenetic trees, the Assamese macaque and Tibetan macaque are closely related (Fig. 2c).

Next, we estimated the divergence time between the macaques using the whole-genome data (Table 2). Three calibration points were defined based on a previous study (Table 2) (Jiang et al., 2016). Using these calibration points, we estimated the divergence time of the Tibetan macaque and Assamese macaque to be 0.51 mya (CI: 0.48–0.54 mya). The two stump-tailed macaques separated from each other at 0.41 mya (CI: 0.39–0.45 mya). The two crab-eating macaques were from very different populations, and the Vietnamese crab-eating macaque had frequent gene flow with the Chinese rhesus macaque. This may contribute to their relative old divergence time of 1.32 mya (CI: 1.28–1.41 mya). We assumed no gene flow when estimating the divergence time with the MCMCTree, which may underestimate the divergence time if gene flow occurred between species.

3.3. PSMC

The pairwise sequentially Markovian coalescent model (PSMC) was applied to investigate ancestral demographic trajectories (Fig. 3). All the macaques exhibited similar demographic trajectories until about 700 thousand years ago (kya). Since then both stump-tailed macaques showed very different trajectories, especially when compared to *fascicularis/mulatta* group species (Fig. 3a). Although all macaque populations in this study began growing at 700 kya, the stump-tailed

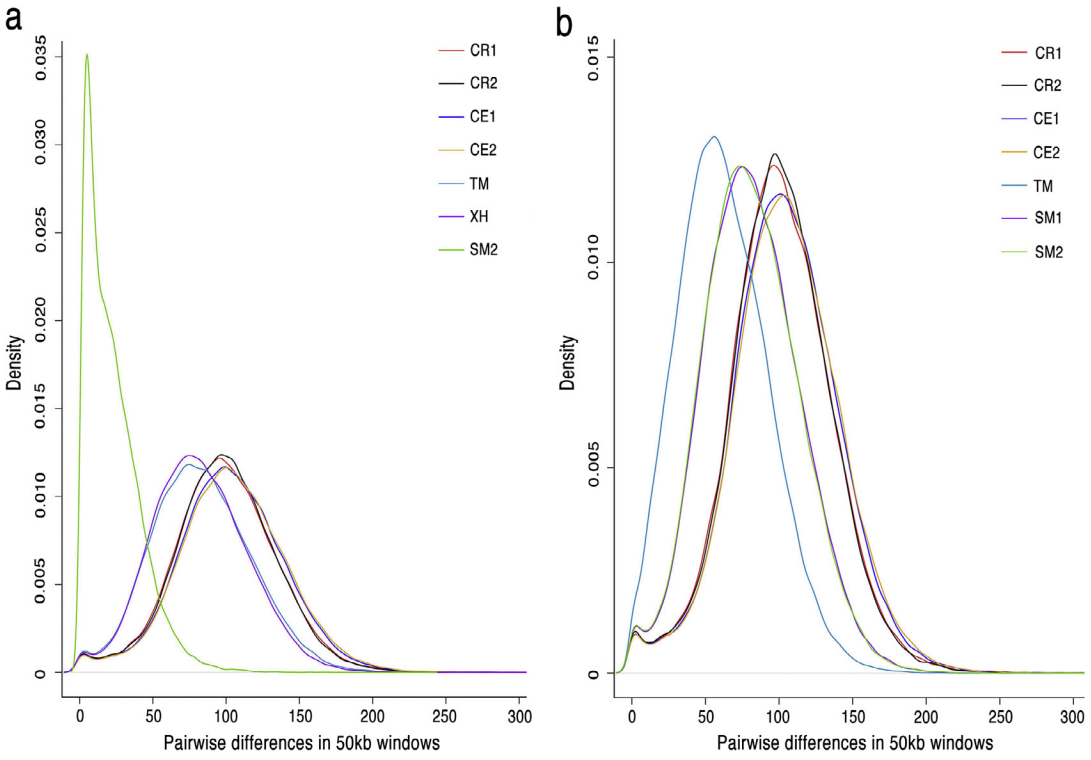

Fig. 1. Pairwise differences across the genome. Pairwise differences between macaques were estimated in 50 kb non-overlapping windows across the genome. (a) The divergence between stump-tailed macaque 1 and other macaques. (b) The divergence between Assamese macaque and other macaques.

macaques had the highest effective population size (N_e) between ~320 and 700 kya. When other macaques underwent population growth or stagnation, the two stump-tailed macaques began experiencing population decline at ~320 kya and did not experience further population growth until very recently (~30 kya; Fig. 3). The Assamese macaque had very similar demographic trajectories to the Tibetan macaque, but the Assamese macaque had higher N_e than the Tibetan macaque after 700 kya (Fig. 3b).

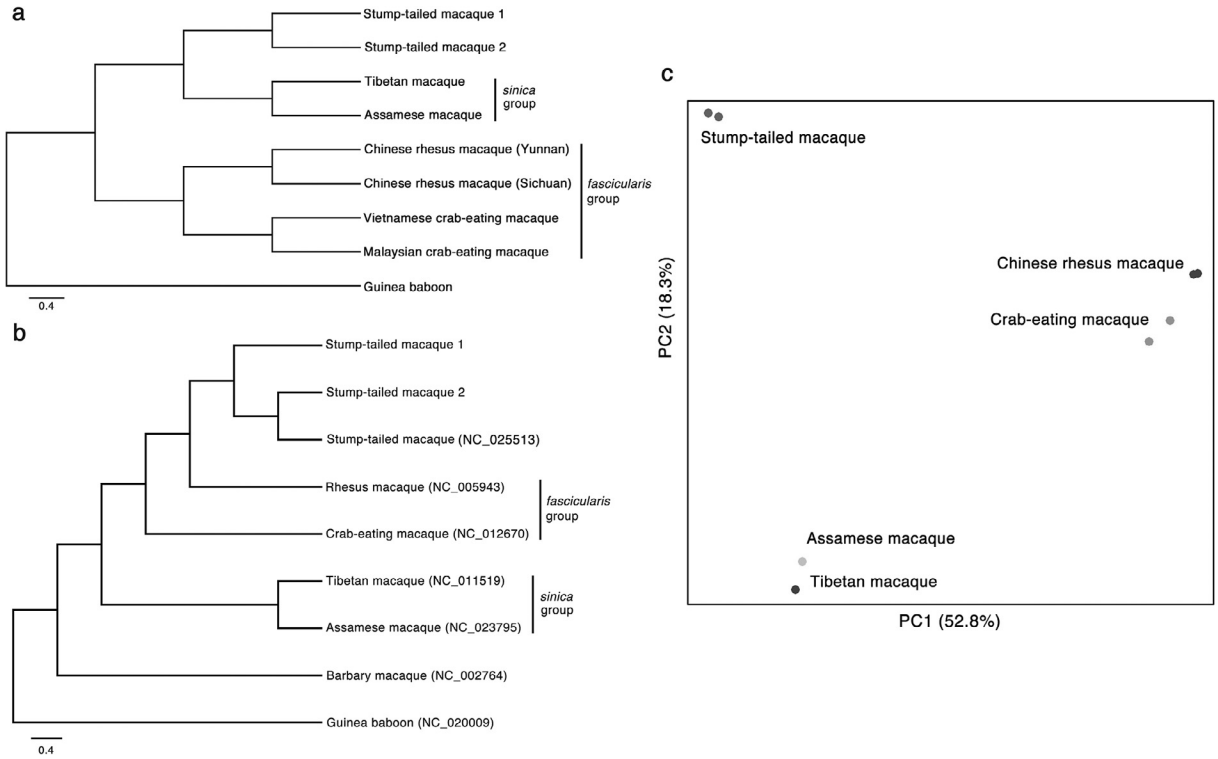

Fig. 2. Phylogenetic tree and Principle Component Analysis (PCA). (a) The genome-wide phylogenetic tree. All the branches are supported by 100% bootstrap runs. (b) The phylogenetic tree based on 12 protein-coding genes of mitochondrial genome. (c) The genome-wide PCA results. The results from PC1 to PC2 and the variance explained by each PC are shown.

Table 2
The divergence time between different macaques in this study.

Comparison	Divergence time (mya)	95% CI (mya)
Stump-tailed macaque 1/Stump-tailed macaque 2	0.41	0.39–0.45
Tibetan macaque/Assamese macaque	0.51	0.48–0.54
Stump-tailed macaques/Tibetan macaque-Assamese macaque	1.40	1.39–1.40
Chinese rhesus macaque (Sichuan)/Chinese rhesus macaque (Yunnan)	1.10	1.00–1.13
Malaysian crab-eating macaque/Vietnamese crab-eating macaque	1.32	1.28–1.41
Chinese rhesus macaques/Crab-eating macaques	1.68	1.67–1.68

3.4. Individual group membership inferred from ADMIXTURE

We used the program ADMIXTURE (Alexander et al., 2009) to assess the ancestry of each individual from 2 to 8 inferred ancestral populations (K). The likelihood value reached the first peak when K = 4, although the CV error was still high. The maximum likelihood was achieved at K = 6 (Table S6). We described the results from K = 2 to 6 due to the small number of samples (Figs. S3 and S4). When K = 2, the *sinica/arctoides* clade (including Tibetan macaque, Assamese macaque and stump-tailed macaques) and *fascicularis/mulatta* clade (including Chinese rhesus macaques and carb-eating macaques) were identified, which suggests that the stump-tailed macaque has an autosomal *sinica* group identity. At K = 3, the stump-tailed macaque and the other two *sinica* species separate into their own component. Rhesus macaques and crab-eating macaques within the *fascicularis/mulatta* group further separate into their own component at K = 4. When K = 5, stump-tailed macaques and the Tibetan macaque/Assamese macaque lineage have a near 100% of their own component, indicating a very close relationship between Tibetan macaque and Assamese macaque. However, *fascicularis* group macaques have a mixture of three components. This is consistent with our results of genome-wide genetic distances and phylogenetic tree (Fig. S3). At K = 6, the Tibetan macaque, Assamese macaque and the two crab-eating macaques all separated into their own components (Fig. S4).

3.5. Gene flow events

We used the ABBA-BABA test to detect gene flow between the stump-tailed macaque and other macaques (Tables S7–S9). Contrary to our expectations, the two individual stump-tailed macaques showed different admixture signals when tested for gene flow with *sinica* group species. All the tests generated significant results indicating that stump-tailed macaque 1 had gene flow with the Tibetan macaque, whereas stump-tailed macaque 2 had gene flow with the Assamese macaque (Table S7).

We then tested whether *fascicularis/mulatta* species (Indian rhesus macaque, Chinese rhesus macaque and crab-eating macaques) had gene

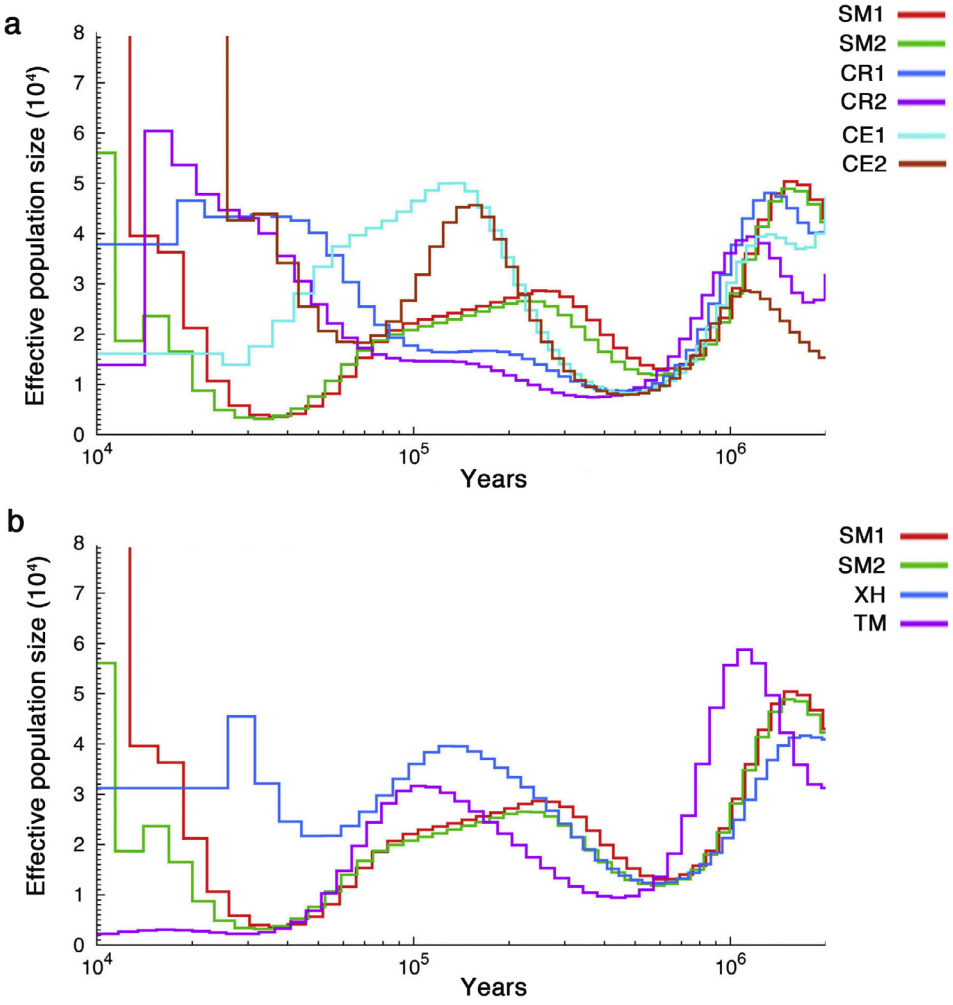

Fig. 3. Historical changes in effective population sizes of macaques. Reconstruction of historical patterns of effective population size for macaque genomes using the PSMC method. (a) stump-tailed macaques and the *fascicularis* group species. (b) stump-tailed macaques and the *sinica* group species.

flow with *sinica/arctoides* species (Table S8). All the tests showed significant gene flow between stump-tailed macaques and *fascicularis/mulatta* species. Finally, we examined which *fascicularis/mulatta* species had gene flow with the stump-tailed macaques (Table S9). Rhesus macaques (Indian rhesus macaque and Chinese rhesus macaque) showed the strongest gene flow signals with the stump-tailed macaque, whereas the Malaysian crab-eating macaque did not show any signal. The Vietnamese crab-eating macaque exhibited significant gene flow with the stump-tailed macaque in some tests, probably because ~30% of its genome was of Chinese rhesus macaque origin (Yan et al., 2011). It is notable that the two stump-tailed macaque individuals again showed opposite signals when tested for gene flow with Indian or Chinese rhesus macaques (Table S9). The results showed that stump-tailed macaque 2 had gene flow with the Indian rhesus macaque, whereas stump-tailed macaque 1 had gene flow with Chinese rhesus macaque or Vietnamese crab-eating macaque.

In addition to the gene flow involving stump-tailed macaques, we also detected complicated gene flows between different macaque species as well as between species groups. For instance, there was significant gene flow between the Tibetan macaque and the Chinese rhesus macaque, and between the Tibetan macaque and the two crab-eating macaques (Table S10). In addition, we detected that the Assamese macaque had gene flow with the rhesus macaques (Indian rhesus macaque and Chinese rhesus macaque) and the Vietnamese crab-eating macaque (Table S11 and S12).

4. Discussion

4.1. Genetic diversity

In this study, we generated the first high coverage genome sequences of two stump-tailed macaques and one Assamese macaque, increasing the number of genomes of *sinica* group species to four. More than 94% of the total reads of each sample could be mapped to the reference rhesus macaque genome, indicating a close relationship between macaque species. We identified more than 10 million high quality SNVs in each macaque and found a relatively low heterozygosity rate in both stump-tailed macaques compared to other macaque species (except the Tibetan macaque who had the lowest heterozygosity in all examined macaques). The number of stump-tailed macaque-specific SNVs was the second largest in all macaques (914,864), demonstrating its distinctive genetic background compared to other macaques. Despite the large number of stump-tailed macaque-specific SNVs, less than 10% were specific to each stump-tailed macaque, indicating a very close genetic background between the two individuals.

With multiple samples in different species groups, we could assess the genetic diversity within and between species groups (Tables S4 and S5). The overall diversity pattern is consistent with the phylogenetic relationship between species (Figs. 1b and S2). However, the Assamese macaque had the smallest divergence with the Tibetan macaque, which is smaller than the genetic difference between the two crab-eating macaques (Tables S4 and S5). There are two potential reasons that contribute to this unexpected observation: first, these two *sinica* species have a very close relationship and their divergence time was smaller than between the two crab-eating macaques (Table 2). Indeed, a close relationship between the two species was suggested by previous phylogenetic studies based on mtDNA sequences (Zhang and Shi, 1993; Hayasaka et al., 1996) as well as morphological characteristics (Delson, 1980). Secondly or alternatively, there are unusually high genetic differences between the two crab-eating macaque individuals. This possibility is supported by the Malaysian crab-eating macaque being from the Indonesian-Malaysian population that maintains the highest genetic diversity (Delson, 1980; Higashino et al., 2012; Fan et al., 2014), while the Vietnamese crab-eating macaque is from the Indochinese population that has frequent gene flow with the Chinese rhesus macaque (Stevison and Kohn, 2008; Bonhomme et al., 2009; Yan et al., 2011).

4.2. The phylogenetic position of stump-tailed macaque

Based on different taxonomic characteristics and phylogenetic markers, the stump-tailed macaque has been classified as being either the sole member of the *arctoides* group, belonging to the *sinica* group, or belonging to the *fascicularis* group (Hayasaka et al., 1996; Morales and Melnick, 1998; Li and Zhang, 2005). Our genome-wide genetic divergence, phylogenetic tree, PCA, as well as ADMIXTURE analysis robustly support that the stump-tailed macaque is closer to the *sinica* group than to the *fascicularis* group. This *sinica* grouping is in agreement with studies based on morphology, nuclear genes, and *Alu* elements (Cronin et al., 1980; Delson, 1980; Deinard and Smith, 2001; Tosi et al., 2003a; Li et al., 2009; Fooden and Lanyon, 2010). Furthermore, we found that the stump-tailed macaque is a distinct species that diverged before the split of the Tibetan macaque and Assamese macaque. However, due to the lack of genome data from other *sinica* species, our results cannot reject the placement of the stump-tailed macaque in its own species group, which has been suggested by several recent taxonomy studies on the genus *Macaca* (Zinner et al., 2013; Roos et al., 2014).

In contrast to the autosomal result, the phylogenetic tree based on mitochondrial genomes grouped both stump-tailed macaques in the *fascicularis* group, specifically with rhesus macaques, with high confidence (bootstrap value 100; Fig. 2b). The rhesus type mitochondria was also found in stump-tailed macaques from different distribution areas (Thailand, Malaysia, Vietnam and China) (Table S1). This is consistent with previous phylogenetic analyses based on one or several mitochondrial genes (Tosi et al., 2003b; Jiang et al., 2016). Such strong discrepancies between mitochondrial and nuclear phylogeny are usually due to two processes that can confound phylogenetic inference: large-scale introgressive hybridizations between species or incomplete lineage sorting (ILS) of ancestral polymorphisms. If it is ILS of mitochondrial DNA, two ancestral mitochondrial haplotypes have to be maintained through the internode of the *sinica/fascicularis* group ancestor, which is highly unlikely. Moreover, mtDNA is less likely to show ILS because it has a small effective population size (Ziegler et al., 2007), which argues strongly against an ILS hypothesis. Therefore, we believe the observed discrepancy between mitochondrial and nuclear genome is more likely to be driven by hybridization.

4.3. Ancient hybridizations pattern in stump-tailed macaque

Sex-specific hybridization often occurs in primate species, in this case, where a limited number of males or females from one taxon transferring into the population of another taxon (Zinner et al., 2011). Given that female philopatry and male dispersal are the most common behaviors in macaques, male-mediated asymmetric introgression is the most likely outcome. For instance, Yan et al. (2011) estimated that ~30% of the Vietnamese crab-eating macaque genome was of Chinese rhesus macaque origin. Yan et al. (2011) suggested that this was a result of ancient gene flow from Chinese rhesus macaque males to female crab-eating macaques. The hybridization pattern in the stump-tailed macaque is different from the Vietnamese crab-eating macaque, and shares high similarity with the nuclear genome of the Chinese rhesus macaque. A male-mediated introgression pattern has been suggested in the stump-tailed macaque by Li et al. (2009) and Jiang et al. (2016) based on nuclear markers, mitochondrial genes, and *Alu* elements. A male-mediated gene flow from a *sinica* species (probably the proto-*assamensis/thibetana*) into a population of rhesus macaque followed by a nuclear swamping would alter the nuclear genomes of the rhesus macaque population but retain the mtDNA from its rhesus lineage matrilineal ancestor. If the stump-tailed macaque originated from male-mediated hybridization, it should have a small genetic divergence from the Tibetan macaque and Assamese macaque, the gene flow rate between them should be very high, and the migration rate should be

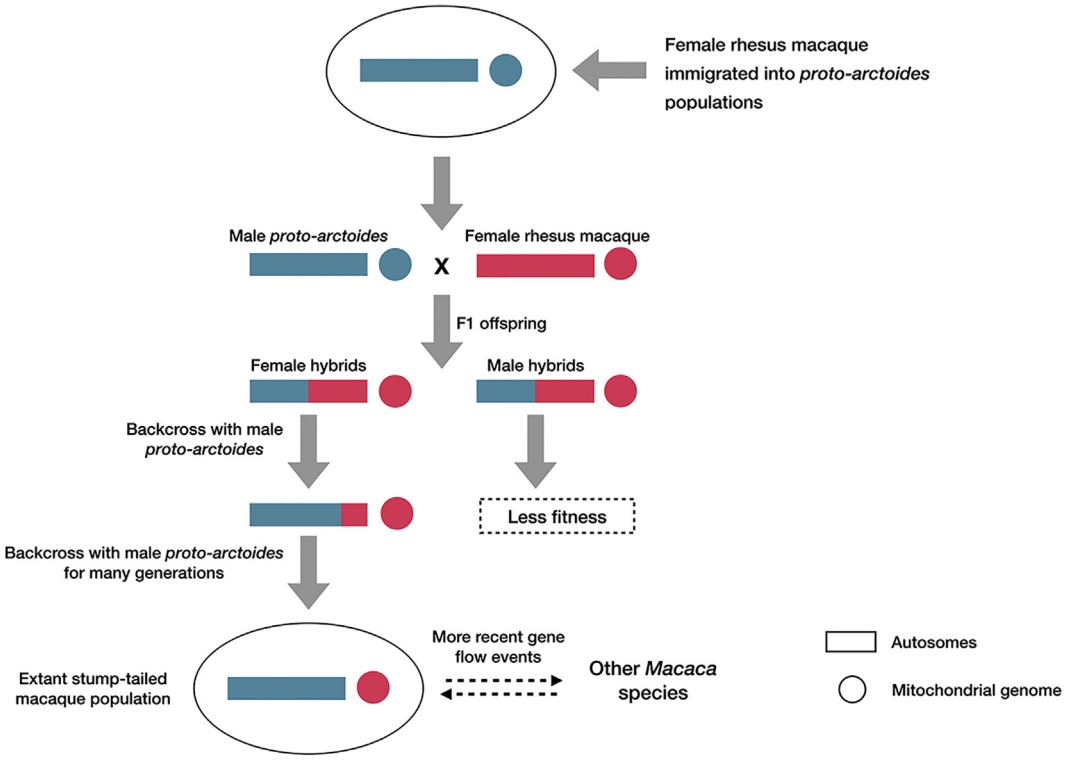

Fig. 4. Mitochondria capture hybridization scenario in stump-tailed macaques. Females of rhesus macaque immigrated into *proto-arctoides* populations producing hybrid offspring with rhesus–type mitochondria (red) and a chimeric nuclear genome of half rhesus (red) and half *proto-arctoides* (blue). Female offspring showed hybrid advantages so that male *proto-arctoides* preferred mating with hybrid offspring. If the back-crossing continued for generations, the nuclear genome of the *proto-arctoides* would barely change but the population would absorb mitochondrial genome of rhesus.

strongly biased to male. However, our genome-wide phylogenetic tree showed an early divergence of the stump-tailed macaque before the split of the Tibetan macaque and the Assamese macaque (Fig. 2a). In addition, both ADMIXTURE (Fig. S3) and PCA (Fig. 2c) indicated the stump-tailed macaque had a unique genetic background compared to other *sinica* species (K = 3), which argues against the male-mediated origin scenario. Instead, we infer an introgression in the stump-tailed macaque after it diverged from other *sinica* species, which transferred the rhesus macaque-type mitochondria into the ancestral stump-tailed macaque population (Fig. 4). Since then, the rhesus macaque-type mitochondria had risen to a high frequency or fixed in stump-tailed macaques either because of genetic drift, during the population bottleneck, or natural selection.

Fa (1989) concluded that the *sinica* group probably originated from an ancestral *sinica-radiata* stock. As the small-bodied, long-tailed *sinica-radiata* stock moved northwards, their body size may have increased and their tail length may have reduced due to adaptation to the cooler temperatures. This could have given rise to new stock including *proto-arctoides* and *proto-assamensis/thibetana*. Our results suggest the divergence time of *proto-arctoides* (~1.40 mya) is right after the divergence of rhesus macaque from *fascicularis* species (~1.68 mya). Soon after their divergences, the expansion of the *proto-arctoides* was influenced by a rapid northward expansion of the rhesus macaque resulting in overlapping geographic distributions of these two macaque populations. The expansion of the rhesus macaque may have also contributed to the hybridization with another sympatric species of the Tibetan macaque, which has been reported previously (Fan et al. 2014). Despite the primarily female philopatry system, female dispersal could have occurred, especially during significant habitat disturbance (Clutton-Brock, 1989; Swedell et al., 2011). It is possible that climate and habitat changes that occurred during the glacial period of the Pleistocene forced a rhesus population and *proto-arctoides* populations together. The rhesus males were outcompeted by *proto-arctoides* males due to the relative small body size of the rhesus macaque. Males of F1 offspring were either sterile or disadvantageous in selections, whereas female offspring showed hybrid advantages so that male *proto-arctoides* preferred mating with hybrid offspring. If the cross between female hybrids and male *proto-arctoides* continued for generations, the nuclear genome of the *proto-arctoides* would barely change but the population would absorb the mitochondrial genome of the rhesus macaque and give rise to the unique genetic entity of extant stump-tailed macaques. This extant form differed from the *proto-arctoides* by retaining a rhesus–type mitochondria but had *sinica* group autosomes, and X-, and Y- chromosomes (Fig. 4). Furthermore, we conclude that the mitochondria-transferring event must have occurred before the radiation of this species because the rhesus–type mitochondria were found in all stump-tailed macaques from different distribution ranges (Table S1). According to our divergence estimation, the hybridization event occurred about 0.4–1.4 mya. However, we cannot exclude another scenario of male introgression followed by nuclear swamping (Jiang et al., 2016). It was suggested that a small isolated rhesus macaque population was adjacent to a large *proto-arctoides* population. Male *proto-arctoides* invaded, and reproduced within the small rhesus population. After many generations, genomes of the small rhesus population were swamped by *proto-arctoides* genes and finally went cyto-nuclear extinct. However, the rhesus-type mitochondria were maintained because of maternal inheritance. It is difficult to decide which of the two scenarios is more likely. More individuals from different stump-tailed macaque populations are needed to determine the exact introgression pattern.

Although recent studies have highlighted the taxonomic breadth of natural hybridization in the primate order, information about hybridization patterns is still limited. Our results have implied distinct hybridization patterns occurred in the evolutionary history of *M. arctoides* and Vietnamese crab-eating macaque despite being sympatric species of southern Asia (Tosi et al., 2003b). The hybridization patterns are consistent with previous studies in colobine monkeys (Roos et al.,

2011) and baboons (Burrell et al., 2009; Keller et al., 2010), suggesting that sex-specific dispersal patterns, promoted by a respective social organization of the species involved, can result in different hybridization scenarios. The hybridization pattern shown in our genome-wide analyses could account for some of the unique characteristics in the stump-tailed macaque. For instance, the distinctive prolonged glans penis and exocervix in stump-tailed macaque reproductive organs, and the strikingly whitish neonatal pelage (Fooden, 1985) are probably results of hybridization between *sinica* and *fascicularis* groups. Stump-tailed macaques also display a mixture of copulatory behavior from the two groups. The protracted single-mount ejaculatory behavior in the stump-tailed macaque may have been derived from the bonnet macaque (a *sinica* group species). However, the copulatory relationship in the stump-tailed macaque resembles the multiple-mount ejaculators, such as the rhesus macaque (Fooden, 1990).

4.4. Gene flow among macaques

Hybridizations between different macaques have been previously reported (Fooden, 1995; Tosi et al., 2000, 2003a; Hamada et al., 2006; Yan et al., 2011; Fan et al., 2014). Based on a comparative genomic analysis, we also detected extensive admixtures between different macaques, some of which were reported previously and some of which were reported here for the first time. For example, in the ADMIXTURE analyses (K = 4 and 5), the Vietnamese crab-eating macaque showed a mixture genetic component with Chinese rhesus macaques. This result is congruous with the results of Yan et al. (2011), who indicated that the crab-eating macaque genome was shaped by introgression after hybridization with the Chinese rhesus macaque. We also detected significant gene flow between the Tibetan macaque and the Chinese rhesus macaque (Table S10), which had been reported in Fan et al. (2014). Notably, the ABBA-BABA test also detected significant gene flow between the Assamese macaque and the rhesus macaque (Table S11 and S12), suggesting extensive gene flow between different macaque species. However, as indicated by Stevison and Kohn (2009) the extensive gene flow inferred between macaque species could also be due to the recent divergences of macaque species (Table 2) and their shared ancestry. Whether the gene flow we detected between macaques is caused by contemporary hybridization or the result of incomplete lineage sorting from a polymorphic common ancestor warrants further investigation.

Some gene flow events appear to have occurred very recently, indicating gene flow between macaques is an ongoing process. The two stump-tailed macaque individuals were both from southwest China and we found that they shared a close genetic distance. However, we detected different gene flow signals in the two individuals. Within *sinica* group species, one individual had significant gene flow with the Tibetan macaque, whereas the other one with the Assamese macaque (Table S7). Similarly, within the *fascicularis* group, one individual had significant gene flow with the Indian rhesus macaque whereas the other with the Chinese rhesus macaque or Vietnamese crab-eating macaque (Table S9). The different patterns imply that the gene flow must have occurred after the two populations that the individuals came from formed, and therefore was very recent. Unfortunately, the ABBA-BABA test could not detect the direction of the gene flow, thus gene flow from more individuals will need to be examined to determine the exact pattern and mechanism of this recent gene flow. Nevertheless, by combining genome-wide analysis on different macaque species, the present study indicates that gene flow between macaque species occurred more commonly than originally believed and hybridization patterns might be more complicated. Therefore this study has provided new insights into the evolutionary history of primates and shed light on the evolutionary processes and the adaptation of species. Moreover, a better understanding of the ancient and ongoing hybridization in macaques is valuable for their applications in biomedical research and will help researchers to interpret differences between experimental groups and investigate species-specific effects.

Acknowledgments

This work was supported by the National Natural Science Foundation of China 31530068 (JL), 31270431 (JL) and 31501871 (ZF), the Outstanding Youth Scholarship Foundation of Sichuan University (JL). JX is supported by NHGRI grant R00HG005846. We thank Dr. Megan Price at Sichuan University read the whole manuscript and help to modify the language.

Authors' contributions

ZF, BY, and JL contributed to the design of this research. JY, LN, JD and HX collected the samples. JJ and PL performed the experiments. ZF, AZ, NO, LD and JX contributed to data analysis. ZF, AZ, NO, JX, BY and JL wrote the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ympev.2018.03.038.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., Butlin, R.K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S.H., Hermansen, J.S., Hewitt, G., Hudson, A.G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P., Möst, M., Mullen, S., Nichols, R., Nolte, A.W., Parisod, C., Pfennig, K., Rice, A.M., Ritchie, M.G., Seifert, B., Smadja, C.M., Stelkens, R., Szymura, J.M., Väinölä, R., Wolf, J.B., Zinner, D., 2013. Hybridization and speciation. J. Evol. Biol. 26, 229–246.

Ackermann, R.R., Bishop, J.M., 2010. Morphological and molecular evidence reveals recent hybridization between gorilla taxa. Evolution 64, 271–290.

Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19, 1655–1664.

Arnold, M.L., 1992. Natural hybridization as an evolutionary process. Annu Rev Ecol Syst. 23, 237–261.

Bertrand, M., 1969. The behavioral repertoire of the stump-tail macaque: A descriptive and comparative study. Bibliotheca Primatologica No. 1. S. Karger, Basel, Switzerland.

Bonhomme, M., Cuartero, S., Blancher, A., Crouau-Roy, B., 2009. Assessing natural introgression in 2 biomedical model species, the rhesus macaque (Macaca mulatta) and the long-tailed macaque (Macaca fascicularis). J. Hered. 100, 158–169.

Burrell, A.S., Jolly, C.J., Tosi, A.J., Disotell, T.R., 2009. Mitochondrial evidence for the hybrid origin of the kipunji, Rungwecebus kipunji (Primates: Papionini). Mol. Phylogenet. Evol. 51 (2), 340–348.

Ciani, A.C., Stanyon, R., Scheffrahn, W., Sampurno, B., 1989. Evidence of gene flow between sulawesi macaques. Am. J. Primatol. 17 (4), 257–270.

Clutton-Brock, T.H., 1989. Female transfer and inbreeding avoidance in social mammals. Nature 337 (6202), 70–72.

Cortés-Ortiz, L., Duda Jr., T.F., Canales-Espinosa, D., García-Orduña, F., Rodríguez-Luna, E., Bermingham, E., 2007. Hybridization in large-bodied New World primates. Genetics 176, 2421–2425.

Cronin, J.E.R., et al., 1980. Molecular evolution and systematics of the genus Macaca. In: Lindburg, D.G. (Ed.), The Macaques: Studies in Ecology Behavior and Evolution. Van Nostrand Reinhold Co., New York, pp. 31–51.

Deinard, A., Smith, D.G., 2001. Phylogenetic relationships among the macaques: evidence from the nuclear locus NRAMP1. J. Hum. Evol. 41, 45–59.

Delson, E., 1980. Fossil macaques, phyletic relationships and a scenario of deployment. In: Lindburg, D.G. (Ed.), The Macaques: Studies in Ecology, Behavior, and Evolution. Van Nostrand Reinhold Co, New York, pp. 10–30.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., McKenna, A., Fennell, T.J., Kernytsky, A.M., Sivachenko, A.Y., Cibulskis, K., Gabriel, S.B., Altshuler, D., Daly, M.J., 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 43, 491–498.

Durand, E.Y., Patterson, N., Reich, D., Slatkin, M., 2011. Testing for ancient admixture between closely related populations. Mol. Biol. Evol. 28, 2239–2252.

Eudey, A.A., 1979. Differentiation and dispersal of macaques (Macaca spp.) in Asia. Doctoral dissertation. Dep Anthropol, Univ California, Davis.

Evans, B.J., Supriatna, J., Andayani, N., Setiadi, M.I., Cannatella, D.C., Melnick, D.J., 2003. Monkeys and toads define areas of endemism on Sulawesi. Evolution 57, 1436–1443.
Evans, B.J., Supriatna, J., Melnick, D.J., 2001. Hybridization and population genetics of two macaque species in Sulawesi, Indonesia. Evolution 55 (8), 1686–1702.
Fa, J.E., 1989. The genus Macaca: a review of taxonomy and evolution. Mammal Rev. 19, 45–81.
Fan, Z., Zhao, G., Li, P., Osada, N., Xing, J., Yi, Y., Du, L., Silva, P., Wang, H., Sakate, R., Zhang, X., Xu, H., Yue, B., Li, J., 2014. Whole-genome sequencing of Tibetan macaque (Macaca thibetana) provides new insight into the macaque evolutionary history. Mol. Biol. Evol. 31 (6), 1475–1489.
Fan, Z., Silva, P., Gronau, I., Wang, S., Armero, A.S., Schweizer, R.M., Ramirez, O., Pollinger, J., Galaverni, M., Ortega Del-Vecchyo, D., Du, L., Zhang, W., Zhang, Z., Xing, J., Vilà, C., Marques-Bonet, T., Godinho, R., Yue, B., Wayne, R.K., 2016. Worldwide patterns of genomic variation and admixture in gray wolves. Genome Res. 26 (2), 163–173.
Fang, X., Zhang, Y., Zhang, R., Yang, L., Li, M., Ye, K., Guo, X., Wang, J., Su, B., 2011. Genome sequence and global sequence variation map with 5.5 million SNPs in Chinese rhesus macaque. Genome Biol. 12, R63.
Fooden, J., 1976. Provisional classification and key to living species of macaques (Primates: Macaca). Folia Primatol. 25, 225–236.
Fooden, J., 1979. Taxonomy and evolution of the sinica group of macaques: I. Species and subspecies accounts of Macaca sinica. Primates 20, 109–140.
Fooden, J., 1985. The stump-tail macaques of China. Am. J. Primatol. 30, 8–11.
Fooden, J., 1990. The bear macaque, Macaca arctoides: a systematic review. J. Hum. Evol. 19, 607–686.
Fooden, J., 1995. Systematic review of southeast Asian longtail macaques, Macaca fascicularis (Raffles, 1821). Fieldiana Zool. 81, 1–206.
Fooden, J., Lanyon, S.M., 2010. Blood-protein allele frequencies and phylogenetic relationships in Macaca: a review. Am. J. Primatol. 17, 209–241.
Freedman, A.H., Gronau, I., Schweizer, R.M., Ortega-Del Vecchyo, D., Han, E., Silva, P.M., Galaverni, M., Fan, Z., Marx, P., Lorente-Galdos, B., Beale, H., Ramirez, O., Hormozdiari, F., Alkan, C., Vilà, C., Squire, K., Geffen, E., Kusak, J., Boyko, A.R., Parker, H.G., Lee, C., Tadigotla, V., Wilton, A., Siepel, A., Bustamante, C.D., Harkins, T.T., Nelson, S.F., Ostrander, E.A., Marques-Bonet, T., Wayne, R.K., Novembre, J., 2014. Genome sequencing highlights the dynamic early history of dogs. PLoS Genet. 10, e1004016.
Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H., Hansen, N.F., Durand, E.Y., Malaspinas, A.S., Jensen, J.D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H.A., Good, J.M., Schultz, R., Aximu-Petri, A., Butthof, A., Höber, B., Höffner, B., Siegemund, M., Weihmann, A., Nusbaum, C., Lander, E.S., Russ, C., Novod, N., Affourtit, J., Egholm, M., Verna, C., Rudan, P., Brajkovic, D., Kucan, Ž., Gušic, I., Doronichev, V.B., Golovanova, L.V., Lalueza-Fox, C., de la Rasilla, M., Fortea, J., Rosas, A., Schmitz, R.W., Johnson, P.L.F., Eichler, E.E., Falush, D., Birney, E., Mullikin, J.C., Slatkin, M., Nielsen, R., Kelso, J., Lachmann, M., Reich, D., Pääbo, S., 2010. A draft sequence of the Neandertal genome. Science 328, 710–722.
Gronau, I., Hubisz, M.J., Gulko, B., Danko, C.G., Siepel, A., 2011. Bayesian inference of ancient human demography from individual genome sequences. Nature Genet. 43, 1031–1034.
Groves, C.P., 2001. Primate Taxonomy. Smithsonian Institution Press, Washington DC.
Hamada, Y., Urasopon, N., Hadi, I., Malaivijitnond, S., 2006. Body size and proportions and pelage color of free-ranging *Macaca mulatta* from a zone of hybridization in Northeastern Thailand. Int J Primatol. 27 (2), 497–513.
Hayasaka, K., Fujii, K., Horai, S., 1996. Molecular phylogeny of macaques: implications of nucleotide sequences from an 896-base pair region of mitochondrial DNA. Mol. Biol. Evol. 13, 1044–1053.
Higashino, A., Sakate, R., Kameoka, Y., Takahashi, I., Hirata, M., Tanuma, R., Masui, T., Yasutomi, Y., Osada, N., 2012. Whole-genome sequencing and analysis of the Malaysian cynomolgus macaque (Macaca fascicularis) genome. Genome Biol. 13, R58.
Hoelzer, G.A., Melnick, D.J., 1996. Evolutionary relationships of the macaques. In: Evolution and Ecology of Macaque Societies. Publisher Cambridge University Press, pp. 3–19.
Jiang, J., Yu, J., Li, J., Li, P., Fan, Z., Niu, L., Deng, J., Yue, B., Li, J., 2016. Mitochondrial genome and nuclear markers provide new insight into the evolutionary history of macaques. PLoS One 11 (5), e0154665.
Keller, C., Roos, C., Groeneveld, L.F., Fischer, J., Zinner, D., 2010. Introgressive hybridization in southern African baboons shapes patterns of mtDNA variation. Am. J. Phys. Anthropol. 142 (1), 125–136.
Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33 (7), 1870–1874.
Lachance, J., Vernot, B., Elbers, C.C., Ferwerda, B., Froment, A., Bodo, J.M., Lema, G., Fu, W., Nyambo, T.B., Rebbeck, T.R., Zhang, K., Akey, J.M., Tishkoff, S.A., 2012. Evolutionary history and adaptation from high-coverage whole-genome sequences of diverse African hunter-gatherers. Cell 150, 457–469.
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.
Li, C., Zhao, C., Fan, P.F., 2015. White cheeked macaque (Macaca leucogenys): a new macaque species from Medog, southeastern Tibet. Am. J. Primatol. 77, 753–766.
Li, H., Durbin, R., 2011. Inference of human population history from individual whole-genome sequences. Nature 475, 493–496.
Li, J., Han, K., Xing, J., Kim, H.S., Rogers, J., Ryder, O.A., Disotell, T., Yue, B., Batzer, M.A., 2009. Phylogeny of the macaques (Cercopithecidae: Macaca) based on Alu elements. Gene 448, 242–249.
Li, Q.Q., Zhang, Y.P., 2005. Phylogenetic relationships of the macaques (Cercopithecidae: Macaca), inferred from mitochondrial DNA sequences. Biochem. Genet. 43, 375–386.
Morales, J.C., Melnick, D.J., 1998. Phylogenetic relationships of the macaques (Cercopithecidae: Macaca), as revealed by high resolution restriction site mapping of mitochondrial ribosomal genes. J. Hum. Evol. 34, 1–23.
Osada, N., Hettiarachchi, N., Babarinde, I.A., Saitou, N., Blancher, A., 2015. Whole-genome sequencing of six mauritian cynomolgus macaques (Macaca fascicularis) Reveals a genome-wide pattern of polymorphisms under extreme population bottleneck. Genome Biol. Evol. 7, 821–830.
Pinho, C., Hey, J., 2010. Divergence with gene flow: models and data. Annu. Rev. Ecol. Evol. Syst. 41, 215–230.
Plotree, D., Plotgram, D., 1989. PHYLIP-phylogeny inference package (version 3.2). Cladistics 5, 163–166.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575.
Reich, D., Green, R.E., Kircher, M., Krause, J., Patterson, N., Durand, E.Y., Viola, B., Briggs, A.W., Stenzel, U., Johnson, P.L., Maricic, T., Good, J.M., Marques-Bonet, T., Alkan, C., Fu, Q., Mallick, S., Li, H., Meyer, M., Eichler, E.E., Stoneking, M., Richards, M., Talamo, S., Shunkov, M.V., Derevianko, A.P., Hublin, J.J., Kelso, J., Slatkin, M., Pääbo, S., 2010. Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature 468, 1053–1060.
Rhesus Macaque Genome Sequencing and Analysis Consortium, Gibbs, R.A., Rogers, J., Katze, M.G., Bumgarner, R., Weinstock, G.M., Mardis, E.R., Remington, K.A., Strausberg, R.L., Venter, J.C., Wilson, R.K., Batzer, M.A., Bustamante, C.D., Eichler, E.E., Hahn, M.W., Hardison, R.C., Makova, K.D., Miller, W., Milosavljevic, A., Palermo, R.E., Siepel, A., Sikela, J.M., Attaway, T., Bell, S., Bernard, K.E., Buhay, C.J., Chandrabose, M.N., Dao, M., Davis, C., Delehaunty, K.D., Ding, Y., Dinh, H.H., Dugan-Rocha, S., Fulton, L.A., Gabisi, R.A., Garner, T.T., Godfrey, J., Hawes, A.C., Hernandez, J., Hines, S., Holder, M., Hume, J., Jhangiani, S.N., Joshi, V., Khan, Z.M., Kirkness, E.F., Cree, A., Fowler, R.G., Lee, S., Lewis, L.R., Li, Z., Liu, Y.S., Moore, S.M., Muzny, D., Nazareth, L.V., Ngo, D.N., Okwuonu, G.O., Pai, G., Parker, D., Paul, H.A., Pfannkoch, C., Pohl, C.S., Rogers, Y.H., Ruiz, S.J., Sabo, A., Santibanez, J., Schneider, B.W., Smith, S.M., Sodergren, E., Svatek, A.F., Utterback, T.R., Vattathil, S., Warren, W., White, C.S., Chinwalla, A.T., Feng, Y., Halpern, A.L., Hillier, L.W., Huang, X., Minx, P., Nelson, J.O., Pepin, K.H., Qin, X., Sutton, G.G., Venter, E., Walenz, B.P., Wallis, J.W., Worley, K.C., Yang, S.P., Jones, S.M., Marra, M.A., Rocchi, M., Schein, J.E., Baertsch, R., Clarke, L., Csürös, M., Glasscock, J., Harris, R.A., Havlak, P., Jackson, A.R., Jiang, H., Liu, Y., Messina, D.N., Shen, Y., Song, H.X., Wylie, T., Zhang, L., Birney, E., Han, K., Konkel, M.K., Lee, J., Smit, A.F., Ullmer, B., Wang, H., Xing, J., Burhans, R., Cheng, Z., Karro, J.E., Ma, J., Raney, B., She, X., Cox, M.J., Demuth, J.P., Dumas, L.J., Han, S.G., Hopkins, J., Karimpour-Fard, A., Kim, Y.H., Pollack, J.R., Vinar, T., Addo-Quaye, C., Degenhardt, J., Denby, A., Hubisz, M.J., Indap, A., Kosiol, C., Lahn, B.T., Lawson, H.A., Marklein, A., Nielsen, R., Vallender, E.J., Clark, A.G., Ferguson, B., Hernandez, R.D., Hirani, K., Kehrer-Sawatzki, H., Kolb, J., Patil, S., Pu, L.L., Ren, Y., Smith, D.G., Wheeler, D.A., Schenck, I., Ball, E.V., Chen, R., Cooper, D.N., Giardine, B., Hsu, F., Kent, W.J., Lesk, A., Nelson, D.L., O'brien, W.E., Prüfer, K., Stenson, P.D., Wallace, J.C., Ke, H., Liu, X.M., Wang, P., Xiang, A.P., Yang, F., Barber, G.P., Haussler, D., Karolchik, D., Kern, A.D., Kuhn, R.M., Smith, K.E., Zwieg, A.S., 2007. Evolutionary and biomedical insights from the rhesus macaque genome. Science 316, 222–234.
Rieseberg, L.H., 1997. Hybrid origins of plant species. Annu. Rev. Ecol. Syst. 28, 359–389.
Rieseberg, L.H., Raymond, O., Rosenthal, D.M., et al., 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. Science 301, 1211–1216.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Roos, C., Boonratana, R., Supriatna, J., Fellowes, J., Groves, C., Nash, S., Rylands, A., Mittermeier, R., 2014. An updated taxonomy and conservation status review of Asian primates. Asian Primates J. 4, 2–38.
Roos, C., Zinner, D., Kubatko, L.S., Schwarz, C., Yang, M., Meyer, D., Nash, S.D., Xing, J., Batzer, M.A., Brameier, M., Leendertz, F.H., Ziegler, T., Perwitasari-Farajallah, D., Nadler, T., Walter, L., Osterholz, M., 2011. Nuclear versus mitochondrial DNA: evidence for hybridization in colobine monkeys. BMC Evol. Biol. 11, 77.
Rumpler, Y., Warter, S., Hauwy, M., Fausser, J.L., Roos, C., Zinner, D., 2008. Comparing chromosomal and mitochondrial phylogenies of sportive lemurs (Genus Lepilemur, Primates). Chromos. Res. 16, 1143–1158.
Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Press, New York.
Stevison, L., Kohn, M., 2008. Determining genetic background in captive stocks of cynomolgus macaques (Macaca fascicularis). J. Med. Primatol. 37, 311–317.
Swedell, L., Saunders, J., Schreier, A., Davis, B., Tesfaye, T., Pines, M., 2011. Female "dispersal" in hamadryas baboons: transfer among social units in a multilevel society. Am. J. Phys. Anthropol. 145, 360–370.
Stevison, L.S., Kohn, M.H., 2009. Divergence population genetic analysis of hybridization between rhesus and cynomolgus macaques. Mol. Ecol. 18 (11), 2457–2475.
Sinha A, C.D., Datta, A, Gama, N, Kumar, S., Madhusudan, M.D., 2013. The Arunachal Macaque. In: Johnsingh, A.J.T. (Ed.), Mammals of South Asia – Volume 1, Chapter: The Arunachal Macaque, Universities Press, Hyderabad, India, Nima Manjrekar, pp. 198–210.
Tosi, A.J., Morales, J.C., Melnick, D.J., 2000. Comparison of Y chromosome and mtDNA phylogenies leads to unique inferences of macaque evolutionary history. Mol. Phylogenet. Evol. 17, 133–144.
Tosi, A.J., Disotell, T.R., Morales, J.C., Melnick, D.J., 2003a. Cercopithecine Y-chromosome data provide a test of competing morphological evolutionary hypotheses. Mol. Phylogenet. Evol. 27, 510–521.
Tosi, A.J., Morales, J.C., Melnick, D.J., 2003b. Paternal, maternal, and biparental

molecular markers provide unique windows onto the evolutionary history of macaque monkeys. Evolution 57, 1419–1435.
Xing, J., Watkins, W.S., Witherspoon, D.J., Zhang, Y., Guthery, S.L., Thara, R., Mowry, B.J., Bulayeva, K., Weiss, R.B., Jorde, L.B., 2009. Fine-scaled human genetic structure revealed by SNP microarrays. Genome Res. 19, 815–825.
Xing, J., Watkins, W.S., Shlien, A., Walker, E., Huff, C.D., Witherspoon, D.J., Zhang, Y., Simonson, T.S., Weiss, R.B., Schiffman, J.D., Malkin, D., Woodward, S.R., Jorde, L.B., 2010. Toward a more uniform sampling of human genetic diversity: a survey of worldwide populations by high-density genotyping. Genomics 96, 199–210.
Yan, G., Zhang, G., Fang, X., Zhang, Y., Li, C., Ling, F., Cooper, D.N., Li, Q., Li, Y., van Gool, A.J., Du, H., Chen, J., Chen, R., Zhang, P., Huang, Z., Thompson, J.R., Meng, Y., Bai, Y., Wang, J., Zhuo, M., Wang, T., Huang, Y., Wei, L., Li, J., Wang, Z., Hu, H., Yang, P., Le, L., Stenson, P.D., Li, B., Liu, X., Ball, E.V., An, N., Huang, Q., Zhang, Y., Fan, W., Zhang, X., Li, Y., Wang, W., Katze, M.G., Su, B., Nielsen, R., Yang, H., Wang, J., Wang, X., Wang, J., 2011. Genome sequencing and comparison of two nonhuman primate animal models, the cynomolgus and Chinese rhesus macaques. Nat. Biotechnol. 29, 1019–1023.
Yang, Z., 1997. PAML : a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13, 555–556.
Zhang, W., Fan, Z., Han, E., Hou, R., Zhang, L., Galaverni, M., Huang, J., Liu, H., Silva, P., Li, P., Pollinger, J.P., Du, L., Zhang, X., Yue, B., Wayne, R.K., Zhang, Z., 2014. Hypoxia adaptations in the grey wolf (Canis lupus chanco) from Qinghai-Tibet Plateau. PLoS Genet. 10, e1004466.
Zhang, Y.P., Shi, L.M., 1993. Phylogenetic relationships of macaques as inferred from restriction endonuclease analysis of mitochondrial DNA. Folia Primatol. 60, 7–17.
Ziegler, T., Abegg, C., Meijaard, E., Perwitasari-Farajallah, D., Walter, L., Hodges, J.K., Roos, C., 2007. Molecular phylogeny and evolutionary history of Southeast Asian macaques forming the M. silenus group. Mol. Phylogenet. Evol. 42, 807–816.
Zinner, D., Groeneveld, L.F., Keller, C., Roos, C., 2009a. Mitochondrial phylogeography of baboons (Papio spp.): indication for introgressive hybridization? BMC Evol. Biol. 9, 83.
Zinner, D., Arnold, M.L., Roos, C., 2009b. Is the new primate genus Rungwecebus a baboon? PLoS One 4, e4859.
Zinner, D., Arnold, M.L., Roos, C., 2011. The strange blood: natural hybridization in primates. Evol. Anthropol. 20, 96–103.
Zinner, D., Fickenscher, G.H., Roos, C., 2013. Family Cercopithecidae (Old World Monkeys). In: Mittermeier, R.A., Rylands, A.B., Wilson, D.E. (Eds.), Handbook of the Mammals of the World, vol. 3, Primates, Barcelona, Lynx Edicions, pp. 550–627.