American Journal of Primatology

WILEY

AMERICAN JOURNAL OF PRIMATOLOGY

RESEARCH ARTICLE OPEN ACCESS

Re-Evaluating the Demographic History of, and Inferring the Fine-Scale Recombination Landscape For, Wild Chinese Rhesus Macaques (*Macaca mulatta*)

John W. Terbot II | Adriana Calahorra-Oliart | Cyril J. Versoza | Devangana Shah | Vivak Soni | Susanne P. Pfeifer | Jeffrey D. Jensen

Center for Evolution and Medicine, School of Life Sciences, Arizona State University, Tempe, AZ, USA

Correspondence: Susanne P. Pfeifer (susanne@spfeiferlab.org) | Jeffrey D. Jensen (jeffrey.d.jensen@asu.edu)

Received: 5 July 2025 | **Revised:** 14 October 2025 | **Accepted:** 22 October 2025

Funding: This study was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award numbers R35GM139383 to J.D.J. and R35GM151008 to SPP. C.J.V. was additionally supported by the National Science Foundation CAREER Award DEB-2045343 to S.P.P. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Science Foundation.

Keywords: catarrhine | population genomics | population history | primate | recombination

ABSTRACT

As a major model for biomedical research, the rhesus macaque (*Macaca mulatta*) is one of the most important and heavily studied nonhuman primates. Despite this importance, the level of population structure and subspecific division in this species remains relatively unclear; for example, the number of proposed subspecies in the literature ranges from one to six within China, with additional populations found across India. Motivated by an interest in comparing recombination rate landscapes between rhesus macaque subspecies, we re-evaluated the demographic history of this group using a previously published data set from 79 wild-born individuals sampled across 17 regions in China. In so doing, we found that previously published demographic models utilizing five subspecies did not well reproduce empirical levels or patterns of genomic variation. Thus, we re-performed demographic inference, finding instead multiple lines of support for a single, interbreeding population (i.e., an absence of population structuring), as well as a population size-change history linking periods of population growth and contraction to historical patterns of glaciation. Finally, utilizing this well-fitting population history, we inferred a genome-wide, fine-scale recombination rate map for this population, finding mean rates consistent with those estimated in other closely related populations and species. However, we also observed notable difference in the fine-scale landscape between rhesus macaques of Chinese and Indian origin – two populations widely used as models in biomedical research – highlighting the importance of accounting for population-specific demographic history and recombination rate variation in future population genomic studies of this species.

1 | Introduction

The rhesus macaque (*Macaca mulatta*) is one of the most well-studied of the nonhuman primates. This partly owes to the fact that, because of its biological similarities to humans and general abundance, it is frequently used in biomedical research (Gibbs et al. 2007; Wiseman et al. 2013; Warren et al. 2020; and see the reviews of Chiou et al. 2020; Rogers 2022). As the most geographically widespread of four member species in the *fascicularis* group of the Macaca genus, it occurs across much of

John W. Terbot II, Adriana Calahorra-Oliart and Cyril J. Versoza contributed equally.

Southern, Eastern, and Southeastern Asia. Other species in this group have more restricted ranges within Southeastern Asia (*M. fascicularis*), Taiwan (*M. cyclopis*), and Japan (*M. fuscata*) (Thierry 2017). However, despite this biomedical importance and impressive geographic distribution, the demographic history of this group – including both historical size changes and population structuring – remains poorly characterized.

Complicating this issue further, speciation within the *fascicularis* group may be ongoing as there is evidence for current and ancient hybridization between *M. mulatta* and *M. fascicularis* (Osada et al. 2010; Phadphon et al. 2022), including potentially current reinforcement speciation (Bailey et al. 2023). Genetic differentiation between *M. mulatta* and *M. cyclopis* is similarly weak (Chu et al. 2007; Zhou et al. 2024b); however, the Japanese macaque (*M. fuscata*) appears to be comparatively distinct from both *M. mulatta* and *M. cyclopis* (Ito et al. 2021; Zhou et al. 2024b). It is perhaps unsurprising that incomplete speciation and hybridization within this group appears pervasive, considering that it is believed that the *fascicularis* group itself has a hybrid origin between the *sinica* and *silenus* groups dating to 3.45–3.56 million years ago (Zhang et al. 2023), indicating long-standing, porous species boundaries in the genus.

Despite species delineation within this group being unclear, there have been numerous attempts to define various subpopulations of *M. mulatta* as subspecies given its frequent use as a nonhuman primate model system and thus, the inherent implications of these characterizations for biomedical research. For example, such designations have been proposed based on differences between the Indian and Chinese subpopulations observed in morphometric data (Clarke and O'Neil 1999), mitochondrial DNA (Smith and McDonough 2005), and levels of exonic diversity (Ferguson et al. 2007). Interestingly, previous work examining genetic distinctions between these two subpopulations have described greater differentiation in coding regions relative to putatively neutral, noncoding regions (Trask et al. 2011). Differentiation between Indian and Chinese rhesus macaques may even exceed that observed amongst the *M. mulatta*, *M. cyclopis*, and *M. fuscata* species (Zhou et al. 2024b).

Within the Chinese subpopulation, various authors have defined between one and six subspecies (Jiang et al. 1991; Zhang and Shi 1993; Liu et al. 2018), including *M. m. mulatta* in the Southwestern section of the Chinese range, *M. m. brevicaudus* found on islands near Hong Kong, *M. m. littoralis* in the Eastern section of the range, *M. m. lasiotis* in the Northwestern section of the range, *M. m. tcheliensis* in the Northeastern-most portion of the range, and *M. m. vestita* in the Tibetan portion of the range. The rationale for these subspecific divisions has largely relied on morphological differences (Jiang et al. 1991; Zhang et al. 2021), mitochondrial DNA (Zhang and Shi 1993; Zhou et al. 2023), and microsatellite data (Zhang et al. 2023). Recent work by Liu et al. (2018) used full genome sequencing of 81 individuals (79 wild-born individuals sequenced for that study, and 2 sequenced previously) across five of these putative subspecies (including all of those listed above, aside from *M. m. vestita*). In that study, they presented evidence for the rigor of the five subspecies classifications, and correspondingly estimated the demographic history of each. As it is appropriate to account for the impacts of both direct selection and selection at linked sites when performing such inference (see the discussions of Charlesworth and Jensen 2021; Soni and Jensen 2025), they filtered out exonic and exonic-adjacent regions to only include putatively neutral sites in their analyses. Additionally, they pruned those sites to reduce linkage effects between variants, and down-sampled to produce even sampling across the five putative subspecies.

Motivated by an interest in quantifying the fine-scale recombination landscape of these populations using this data, we evaluated the fit of this estimated demographic history to the empirical data. This step is necessary, given previous evidence that the demographic history of the population may bias inferred recombination rates (e.g., Dapper and Payseur 2018). In simulating neutral genomic regions under the previously inferred 5-subspecies demographic history for the individuals in the Liu et al. data set, we found a poor fit between the resulting site frequency spectra (SFS) of the model to that observed in the empirical data. As such, we re-evaluated the population structure and population history of this sample and found the greatest statistical support for all samples in fact belonging to a single population. Utilizing this result, we further estimated the history of population size change; the resulting models were generally consistent in suggesting population growth following the end of the Xixiabangma Glaciation and population contraction associated with the timing of the last glacial period – a result corroborated by estimates previously observed with this data set (Liu et al. 2018; Zhou et al. 2024a). Notably, unlike the previous work, the resulting fit of the empirical data to this single-deme, population-size change model was found to be highly consistent.

Utilizing this newly inferred population history, we further inferred the fine-scale recombination landscape for this Chinese population, finding rates consistent with those estimated in other closely related primate populations and species (see the review of Stapley et al. 2017 and references therein). However, we also observed notable differences in the fine-scale landscape between rhesus macaques of Chinese origin and that previously reported in the Indian subpopulation (Versoza et al. 2024), highlighting the importance of accounting for population-specific demographic history and recombination rate variation in future population genomic studies of this species.

2 | Materials and Methods

2.1 | Ethics Statement

Not applicable: no new sequencing data was generated for this study.

2.2 | Whole-Genome Sequencing Data of Chinese Rhesus Macaques

To infer the demographic history and fine-scale recombination landscape of Chinese rhesus macaques (*M. mulatta*), we downloaded publicly available whole-genome sequencing data of 79 wild-born individuals from the NCBI Sequence Read Archive

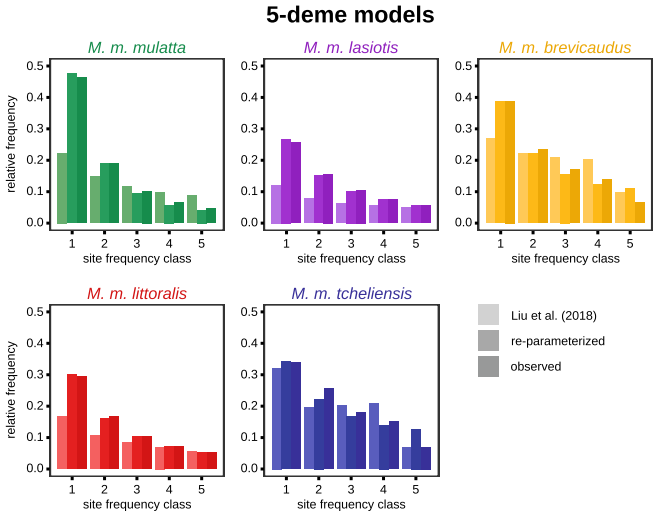

FIGURE 1 | 5-deme models. Simulated, mean, folded, relative site frequency spectra (SFS) for each of the five previously designated subspecies alongside the SFS observed from the empirical data: *Macaca mulatta mulatta* (shown in green), *M. m. lasiotis* (purple), *M. m. brevicaudus* (yellow), *M. m. littoralis* (red), and *M. m. tcheliensis* (blue). Each SFS bin contains bars for the relative frequency of the site frequency class for the data simulated under the parameterized model inferred by Liu et al. (2018) (left, lightly shaded), the re-parameterization of this model performed in this study (middle, moderately shaded), and the empirically observed SFS (right, darkly shaded). Only the first five bins for each deme are depicted here; the full SFS for all bins can be found in Supporting Figure 2. Note that simulated results from both model parameterizations are visually distinct from the empirical SFS for all five demes.

(BioProject: PRJNA345528; Liu et al. 2018). This data set consists of five *M. m. tcheliensis* (individual IDs: *C_rhe_1–C_rhe_5*), 28 *M. m. littoralis* (*C_rhe_6– C_rhe_21* and *C_rhe_27– C_rhe_38*), five *M. m. brevicaudus* (*C_rhe_22–C_rhe_26*), 31 *M. m. lasiotus* (*C_rhe_39–C_rhe_69*), and 10 *M. m. mulatta* (*C_rhe_70–C_rhe_79*) individuals sampled across 17 regions in China (see Figure 1a in Liu et al. 2018).

Following best practices in the field (Pfeifer 2017; van der Auwera and O'Connor 2020), we pre-processed the raw sequencing data by converting the files downloaded in FASTQ format to unmapped BAMs and adding read group information using the *FastqToSam* tool implemented in the Genome Analysis Toolkit (GATK; version 4.1.8.1, unless noted otherwise). We then marked adapter sequences using GATK's *MarkIlluminaAdapters* and converted the files back to FASTQ format, discarding any identified adapter sequences using GATK's *SamToFastq*. Next, we mapped the pre-processed reads against the Indian-origin rhesus macaque genome assembly (rheMac10; GenBank assembly: GCA_003339765.3; Warren et al. 2020) using the Burrows-Wheeler Aligner (BWA-MEM) v.0.7.17 (Li and Durbin 2009), marking split reads as secondary using the '*-M*' option. We adjusted available meta-information using GATK's *MergeBamAlignment*, marked duplicate reads using GATK's *MarkDuplicates*, aggregated data for each individual using GATK's *MergeSamFiles*, and indexed files using GATK's *BuildBamIndex*. As additional duplicates may have arisen after data aggregation, we carried out a second round of duplicated marking as described above. To improve base quality scores, we first performed a local realignment around small insertions and deletions using the *RealignerTargetCreator* and *IndelRealigner* functions implemented in GATK v.3.7.0 before adjusting base qualities using GATK's *BaseRecalibrator* and *ApplyBQSR* based on information available from a variant catalog of 526 captive rhesus macaque individuals (Harris 2019). We then called germline variants from high-quality read mappings ('*--minimum-mapping-quality* 40') via local reassembly of haplotypes using GATK's *HaplotypeCaller* with the species-specific heterozygosity rate set to 0.0024 (as previously reported by Warren et al. 2020) and the '*-ERC*' flag set to BP_RESOLUTION to obtain information about all sites (i.e., variable and invariable) accessible to this study, and jointly genotyped all individuals at each site in the genome ('*-allSites*').

To obtain high-confidence sites, we used GATK's *SelectVariants* to limit the call set to sites called on the autosomes (i.e., chromosomes 1–20) that were genotyped in all individuals ('$AN = 158$', for the 79 diploids in the study). In the absence of an experimentally validated data set for the species that could be used as an input for the Gaussian mixture model implemented in GATK, we filtered variant (here biallelic single nucleotide polymorphisms; SNPs) and invariant sites following the recommendations of the developer team (i.e., $QD < 2.0$, $QUAL < 30.0$, $SOR > 3.0$, $FS > 60.0$, $MQ < 40.0$, $MQRankSum < -12.5$, $ReadPosRankSum < -8.0$, with acronyms defined by the GATK package) using GATK's *VariantFiltration* tool. This data set contained 42.1 million SNPs with a transition-transversion (Ts/Tv) ratio of 2.27 in the accessible genome (Supporting Table 1).

Lastly, for the sake of comparison, we created a second version of the data set by applying the non-standard filter criteria of Liu et al. (2018) (i.e., using GATK's *VariantFiltration* with '*--filterExpression*"$QD < 5.0 \,||\, FS > 60.0 \,||\, MQ < 40.0 \,||\, ReadPosRankSum < -8.0 \,||\, MQRankSum < -12.5$" *--genotype FilterExpression* "$DP < 4.0$"' and filtering out any segregating sites where the reference genome contains an *N* or genotype information is missing for > 20% of the individuals) and compared the folded SFS from our study to that obtained from the data set published by the authors (Supporting Figure 1; note that although the authors also called invariant sites in their study, only information regarding the variable sites is available via the study's data repository [10.5524/100484]).

2.3 | High-Quality Datasets of Chinese Rhesus Macaques for Demographic and Recombination Rate Inference

To exclude regions prone to sequencing, alignment and variant calling errors, we applied minimum and maximum cutoffs on the depth of sequencing coverage to the GATK Best Practice filtered call set, following recommendations by Wu et al. (2020) (i.e., only considering regions of the genome where the depth of coverage passed a two-sided Poisson test with a *P*-value $> 2 \times 10^{-4}$ in each individual, assuming that λ is equal to the individual mean read depth).

To obtain a high-quality data set for demographic inference, we restricted the depth-filtered call set to putatively neutral, non-repetitive genomic regions to circumvent the biasing effects of

selection (Ewing and Jensen 2014, 2016; Johri et al. 2020, 2021; and see the reviews of Charlesworth and Jensen 2021, 2024). Specifically, following the recommendations by Johri et al. (2020, 2023), we removed both sites constrained across primates (based upon phyloP and phasCons algorithms; Kuderna et al. 2024) as well as those within 10 kb of exonic regions to avoid the effects of purifying and background selection, respectively. The final data set contained 7.9 million putatively neutral SNPs with a Ts/Tv of 2.20 in the accessible genome (Supporting Table 1).

To obtain a high-quality data set for recombination rate inference, we followed the recommendations outlined in earlier studies of recombination in nonhuman primates (Auton et al. 2012; Stevison et al. 2016; Pfeifer 2020a; Soni et al. 2025a, 2025b) and restricted the depth-filtered call set to sites outside of both SNP clusters (using GATK's *VariantFiltration* tool with the options '*--cluster-size* 3' and '*--cluster-window-size* 10' to exclude any regions containing ≥ 3 SNPs within any 10bp-window) and regions blacklisted by the ENCODE project (Amemiya et al. 2019) that showed no signs of an excess in levels of heterozygosity (assessed using VCFtools v.0.1.16 with a Hardy-Weinberg equilibrium p-value < 0.01; Danecek et al. 2011). We then phased the filtered sites using BEAGLE v.5.5 (Browning et al. 2021) to attain haplotypes. The final data set contained 27.6 million phased SNPs with a Ts/Tv of 2.26 in the accessible genome (Supporting Table 1).

2.4 | Evaluating the Demographic Model of Chinese Rhesus Macaques Inferred by Liu et al. (2018)

Our initial step was to evaluate the ability of the parameterized model inferred by Liu et al. (2018) to produce SFS in line with the observed empirical data. To that end, we calculated the multi-dimensional and pairwise folded SFS for each of the five previously designated subspecies – *M. m. brevicaudus*, *M. m. lasiotus*, *M. m. littoralis*, *M. m. mulatta*, and *M. m. tcheliensis* – using easySFS v.0.0.1 (https://github.com/isaacovercast/easySFS). Notably, unlike the Liu et al. (2018) study which, in addition to filtering for putatively neutral sites, pruned sites based on patterns of linkage disequilibrium (LD) and down-sampled to emulate even-sampling across demes, these SFS were calculated without down-sampling individuals or any further filtering other than that noted above. This step thus serves to evaluate the impact of alternate filtering on downstream estimation. Next, we ran fastsimcoal2 v.2.8.0.0 (Excoffier et al. 2013, 2021; Marchi et al. 2024) to produce 100 replicate coalescent simulations of 1Mb-long genomic segments under the published model, assuming a neutral mutation rate of 1.08e-8 mutations/site/generation and a recombination rate of 1e-8 events/site/generation (i.e., the rates used by Liu et al. 2018). Following this, we compared the resulting simulated SFS to the empirical SFS. Observing discordance, we then re-estimated the parameters of Liu et al.'s branching and migration model using fastsimcoal2. This program conducts parameter optimization via a maximum (composite) likelihood search calculated from simulated SFS in comparison to the composite likelihood of the empirical SFS. Specifically, we ran this parameter estimation for 100 replicates each with 150,000 simulations performed per parameter, a total of 50 maximization cycles, and a re-setting of the parameter search after 10 consecutive failed cycles ('*-n*150000 *-L* 50 *-y* 10'), and then simulated 100 replicates of 1MB-long genomic segments with the newly estimated model parameters.

2.5 | Estimating the Population Structure of Chinese Rhesus Macaques

Following our re-parameterization of Liu et al's model, we next evaluated whether simpler models consisting of fewer parameters may be able to fit the observed data similarly well or better. To that end, we estimated population structure using two alternative methods, ADMIXTURE v.1.3.0 (Alexander et al. 2009) and fastSTRUCTURE v.1.0 (Raj et al. 2014), both implementing the ancestry optimization algorithms of STRUCTURE (Pritchard et al. 2000; Falush et al. 2003) under either a maximum likelihood (ADMIXTURE) or variational Bayesian (fastSTRUCTURE) framework. We applied these methods on putatively neutral sites, both per-chromosome (autosomal, $n = 20$) and on a full genome sample consisting of all autosomes concatenated together using BCFtools v.1.9 (Danecek et al. 2021). We generated the appropriate input files using PLINK v.1.9.0-b.7.7 (Purcell et al. 2007), and prepared versions of these input files for comparison that pruned SNPs based on LD using a window size of 50 SNPs, a step size of 10 SNPs, and pruning SNPs with r^2 values greater than 0.1 ('*--indep-pairwise* 50 10 0.1'). We ran both ADMIXTURE and fastSTRUCTURE on these input files with evaluated values of K (i.e., the number of subpopulations or demes) ranging from 1 to 8, and determined the optimal splitting by minimizing cross-validation errors for ADMIXTURE and maximizing marginal likelihoods for fastSTRUCTURE.

Finally, we calculated pairwise F_{ST} values for each chromosome under the subspecific assignment of Liu et al. (2018) using VCFtools (Danecek et al. 2011) via the process outlined by Weir and Cockerham (1984). For each calculation of pairwise F_{ST} values, we used either the full sample set of individuals, or five randomly chosen individuals from each subspecies to produce a uniform, sub-sampled comparison. In both cases, we included all putatively neutral SNPs in the calculations without performing any LD pruning. To obtain pairwise deme comparisons, we then averaged across per-chromosome values.

2.6 | Inferring the Demographic History of Chinese Rhesus Macaques Using Fastsimcoal2 and δaδi

The results from both ADMIXTURE and fastSTRUCTURE strongly favored a single deme (see the "Results and Discussion" section below); however, F_{ST} values indicated possible structure between *M. m. brevicaudus* and *M. m. tcheliensis*. Therefore, we parameterized and simulated both single-deme and 3-deme models using fastsimcoal2. Specifically, we used the empirical data to prepare multi-dimensional and pairwise, folded SFS for the 3-deme models alongside a single-dimension, folded SFS for the single-deme models, using easySFS for fastsimcoal2 and δaδi's Spectrum library. For the 3-deme models,

we used one of two branching patterns – either *M. m. tcheliensis* splitting first with *M. m. brevicaudus* and all other samples as sister groups, or *M. m. brevicaudus* splitting first with *M. m. tcheliensis* and all other samples as sister groups. For the single deme models, we tested models using 0 to 4 size changes, running 100 replicates of parameterizations for each model using fastsimcoal2, with the same parameters as detailed above (i.e., '*-n*150000 *-L* 50 *-y* 10'), and selected the best parameterization based on log maximum likelihood scores. Following this, we used the best parameterizations of each model to produce 100 simulated SFS for 1Mb-long genomic segments using fastsimcoal2, and performed further model selection based on log maximum likelihood scores and comparison of these simulated SFS to the empirical SFS.

To further confirm the parameterized models produced by fastsimcoal2, we parameterized single-deme models with 0 to 4 size change events using an alternative inference approach as implemented in δaδi v.2.1.0 (Gutenkunst et al. 2009). In contrast to fastsimcoal2, δaδi uses a diffusion approximation approach to conduct a likelihood search for optimized parameters by calculating a continuous approximation of the expected SFS under a particular demographic model and comparing that to an empirical SFS. For each model, we ran five replicates of optimization with randomly different initial values. We performed model selection based on maximum likelihood; additionally, we simulated the parameterized models for a 1Mb-long genomic segment using fastsimcoal2 to produce simulated SFS for comparison to the empirical SFS.

As a final step to differentiate between the three best-fitting models (i.e., the 3-event and 4-event models inferred by fastsimcoal2 and the 3-event model inferred by δaδi), we calculated the per-site Watterson's theta (θ_W; Watterson 1975) for the simulated data and compared it to θ_W obtained from the empirical data using pylibseq v.0.2.3 (Thornton 2003). We additionally compared the mean and variance of these values using the *t. test* function implemented in the R stats package v.4.2.2 (R Core Team 2023) with a Bonferroni corrected α of 0.00833 (from an α of 0.05, corrected for six total comparisons).

2.7 | Inferring the Recombination Landscape of Chinese Rhesus Macaques Using Pyrho

To infer fine-scale rates of recombination, we implemented pyrho v.0.1.7 (Spence and Song 2019), a demography-aware estimator that relies on patterns of LD observed in sequencing data as an indirect signal to estimate underlying recombination rates. Briefly, we first generated a likelihood lookup table using pyrho's *make_table* function, with the inferred 4-event population size change history ('*--popsizes* 160646,640374,275692,393753,60170 *--epochtimes* 1707,9494,42956,67421') assuming a species-specific, per site, per generation mutation rate of 5.8e-9 ('*--mu* 5.8e-9', as estimated by Wang et al. 2020). Adopting the recommendations of the developers for large sample sizes (Spence and Song 2019), we calculated an approximate lookup table ('*--approx*') for a population that is 50% larger ('*-N*237') and down-sampled it to the actual sample size ('*-n* 158', as there are 79 diploid individuals in the study). Next, we ran pyrho's *hyperparam* function with the per-generation mutation rate set as 5.8e-9 ('*--mu* 5.8e-9') to determine the optimal parameter settings for window size and block penalty. Lastly, we utilized pyrho's *optimize* function with the optimal parameters – a window size of 30 ('*--windowsize* 30') and a block penalty of 50 ('*--blockpenalty* 50') – to estimate recombination rates per site per generation across the genome.

To benchmark the performance of pyrho under the population history inferred for Chinese rhesus macaques (for details, see "Inferring the demographic history using fastsimcoal2 and δaδi"), we used msprime v.1.3.3 (Baumdicker et al. 2022) to simulate 10 replicates of the longest chromosome (i.e., chromosome 1; length: 223.6 Mb) with multiple parameter combinations under the best-fitting model (i.e., the 4-event model) assuming a neutral mutation rate of 1.08e-8 mutations/site/generation (i.e., the rate used by Liu et al. 2018) and sampling 79 individuals to mirror the empirical call set. We set the recombination rate for the simulated models as 1×10^{-8} events/site/generation (following Liu et al. 2018), to estimate any inference bias potentially generated by the population history itself. Finally, with these simulations on hand, we ran pyrho's inference procedure under the 4-event population size change history inferred by fastsimcoal2 as described above.

2.8 | Comparing the Recombination Landscapes of Chinese and Indian Rhesus Macaques

To gain insights into the similarities of, and differences between, the recombination landscapes of Chinese and Indian rhesus macaques – two populations widely used in biomedical research – we compared the recombination rates inferred here for wild-born Chinese individuals to previous estimates obtained from captive-born individuals of Indian origin (see Versoza et al. 2024 for details) by calculating Pearson's correlation coefficient at the 1 kb, 10 kb, 100 kb, 1 Mb, 3 Mb, and 5 Mb scales in R and plotting correlations using ggplot2 v.3.5.1 (Wickham 2016).

3 | Results and Discussion

3.1 | Whole-Genome Sequencing Data of Chinese Rhesus Macaques

Population genomic data from 79 wild-born rhesus macaques (*M. mulatta*) sampled across 17 geographic regions allowed for the detection of genetic variation from five putative subspecies, with *M. m. brevicaudus* (five individuals) found on the Hainan Island in the South China Sea, and *M. m. lasiotus* (31 individuals), *M. m. mulatta* (ten individuals), *M. m. littoralis* (28 individuals), and *M. m. tcheliensis* (five individuals) distributed across mainland China, inhabiting the western, western-central, southern, and northern parts of the country, respectively (see Figure 1a in Liu et al. 2018). Specifically, to infer the demographic history and fine-scale recombination landscape of Chinese rhesus macaques, two high-quality datasets were generated by mapping the quality-controlled reads to the species-specific genome assembly, and calling variant and invariant sites according to standard best practices in the field (Pfeifer 2017; van der Auwera and O'Connor 2020; and see "Materials and Methods" for details). Subsequently, this call set

was limited to (1) putatively neutral, non-repetitive genomic regions located a sufficient distance away from constrained and functional elements to circumvent the biasing effects of direct selection as well as selection at linked sites during demographic inference (Ewing and Jensen 2014, 2016; Johri et al. 2020, 2021; and see the reviews of Charlesworth and Jensen 2021, 2024) and (2) haplotypes inferred in a manner similar to those of previous studies of recombination in nonhuman primates (Auton et al. 2012; Stevison et al. 2016; Pfeifer 2020a; Soni et al. 2025b). The final datasets contained 7.9 million putatively neutral SNPs (Ts/Tv: 2.20) and 27.6 million phased SNPs (Ts/Tv: 2.26) in the accessible genome, respectively (Supporting Table 1).

3.2 | Evaluating the Demographic Model of Chinese Rhesus Macaques Inferred by Liu et al. (2018)

Simulations of the demographic model of Chinese rhesus macaques inferred by Liu et al. (2018) using the empirical genomes were performed using fastsimcoal2, and the fit of the model to the data was found to be unsatisfactory for all five suggested demes (see Figure 1 for the first five bins for each deme and Supporting Figure 2 for the full SFS for all bins). Utilizing the same 5-deme structure but re-estimating the parameters of this model produced a SFS with a closer overall shape to the empirical SFS (Figure 1 and Supporting Figure 2), though the fit remained poor. This re-parameterized model generally produced larger population sizes, aside from particular ancestral populations, and the timing of the ancestral merger events were also generally older than the estimates presented by Liu et al. (2018) (Supporting Table 2; and see Supporting Figures 3 and 4). Notably however, estimates of contemporary migration rates varied considerably from this previously published model, generally being an order of magnitude greater. Nonetheless, the continuing poor fit of the 5-deme model, along with the corresponding high migration rates inferred under this model implying a lack of underlying population structure, suggests that the five subspecies assumption of Liu et al. (2018) must be reconsidered.

3.3 | Estimating the Population Structure of Chinese Rhesus Macaques

The results from both ADMIXTURE and fastSTRUCTURE were overwhelmingly supportive of all 79 individuals belonging to a single deme, with a single deme minimizing the cross-validation error in ADMIXTURE and maximizing the marginal likelihood in fastSTRUCTURE for each chromosome as well as the whole genome (Table 1). These results were consistent for both the unpruned and LD-pruned datasets (see Supporting Tables 3 and 4 for the results of the ADMIXTURE and fastSTRUCTURE runs, respectively).

To avoid the impact of uneven sampling when calculating F_{ST}, five individuals were selected from each putative subspecies, sampling individuals at random from the subspecies with larger sample sizes, i.e., *M. m. lasiotis* (*C_rhe_40*, *C_rhe_48*, *C_rhe_59*, *C_rhe_62*, and *C_rhe_66*), *M. m. littoralis* (*C_rhe_9*, *C_rhe_17*, *C_rhe_19*, *C_rhe_35*, and *C_rhe_36*), and *M. m. mulatta* (*C_rhe_70*, *C_rhe_72*, *C_rhe_74*, *C_rhe_75*, and *C_rhe_79*). The mean F_{ST} values averaged across chromosomes amongst these previously suggested demes were generally modest, ranging from 0.0099 (*M. m. littoralis* and *M. m. mulatta*) to 0.0606 (*M. m. lasiotis* and *M. m. tcheliensis*) (Table 2). The only pair of subspecies outside of this range was *M. m. brevicaudus* and *M. m. tcheliensis*, with a mean F_{ST} value of 0.0975 (Supporting Table 5). There were no individual chromosomes that deviated from this overall pattern; F_{ST} values for a given pair of subspecies were similar for each chromosome (Supporting Table 6). Calculating F_{ST} using the data from all 79 individuals included in the study reduced the mean F_{ST} values for each pair of subspecies (ranging from 0.0169 to 0.0373), with the exception of the *M. m. brevicaudus* and *M. m. tcheliensis* pair which retained the same value as no down-sampling occurred in these groups (Supporting Tables 5 and 7). Notably, the *M. m. brevicaudus* and *M. m. tcheliensis* sampling locations are the most geographically isolated from one other, particularly when considering distance and potential barriers including rivers and islands, with the former inhabiting the Hainan island off the

TABLE 1 | Results of the population structure analyses.

K	ADMIXTURE CV error	fastSTRUCTURE Marginal likelihood
1	0.29189	−0.4067
2	0.30701	−0.4067
3	0.32117	−0.4067
4	0.33838	−0.4067
5	0.35614	−0.4067
6	0.38166	−0.4067
7	0.38579	−0.4067
8	0.39342	−0.4067

Note: ADMIXTURE and fastSTRUCTURE results using unpruned data from the entire genome. The second column contains the cross-validation (CV) error from this data analyzed using ADMIXTURE; note that the lowest value is for $K = 1$, with K being the number of sub-populations or demes. The third column contains the marginal likelihood values from this data analyzed using fastSTRUCTURE; note that this value is flat for all value of K as, even when more than a single deme was available, the most-likely model only used a single demographic grouping.

TABLE 2 | F_{ST} values.

deme A	deme B	Mean F_{ST}
M. m. tcheliensis	*M. m. brevicaudus*	0.09747
M. m. tcheliensis	*M. m. lasiotis*	0.06059
M. m. brevicaudus	*M. m. lasiotis*	0.05899
M. m. tcheliensis	*M. m. mulatta*	0.05242
M. m. brevicaudus	*M. m. mulatta*	0.04782
M. m. littoralis	*M. m. brevicaudus*	0.04402
M. m. tcheliensis	*M. m. littoralis*	0.04119
M. m. lasiotis	*M. m. mulatta*	0.01517
M. m. littoralis	*M. m. lasiotis*	0.01471
M. m. littoralis	*M. m. mulatta*	0.00989

Note: Mean, pairwise F_{ST} values of chromosomes between previously designated Chinese rhesus macaque subspecies using five representative individuals (randomly chosen without replacement) per subspecies.

southern coast of China and the latter living nearly 2,000 km apart, in the north of Henan (Brandon-Jones et al. 2004 and references therein). Thus, while the elevated F_{ST} values between this pair may indeed suggest a level of isolation by distance, taken together these results strongly suggest that all samples – previously assumed to compose five demes matching the geographical sampling locations – were in fact collected from a single deme genetically.

3.4 | Inferring the Demographic History of Chinese Rhesus Macaques

Given the lack of observed population structure, we used fastsimcoal2 to estimate parameters for a total of five single-deme models (consisting of 0 to 4 population re-sizing events) as well as two 3-deme models (one with *M. m. brevicaudus* branching first, and another with *M. m. tcheliensis* branching first, with all other samples combined into a single deme) to account for the possibility of structuring between the two geographically most isolated populations, *M. m. brevicaudus* and *M. m. tcheliensis*. Estimated parameters for the two 3-deme models were rather comparable, with the largest differences being in the estimated current day population sizes of the demes and in the timing of the split between the *M. m. lasiotis/M. m. littoralis/M. m. mulatta* deme with either *M. m. brevicaudus* or *M. m. tcheliensis* (Supporting Table 8; and see Supporting Figures 5 and 6). However, the simulated SFS for the estimated 3-deme models were visibly distinct from the observed SFS (Supporting Figure 7). Taken together with the fastSTRUCTURE and ADMIXTURE results, this provided further evidence for the lack of meaningful population structuring in this data, and thus suggested the need to evaluate single-deme estimation.

Model comparison of the single-deme models was first performed by comparing the log likelihood values for the best performing replicate of each model. From this, the greatest value was for the 4-event model (4-event: $L = -26{,}265{,}297.379$; 3-event: $L = -26{,}265{,}468.903$; 2-event: $L = -26{,}265{,}512.238$; 1-event: $L = -26{,}269{,}065.621$; 0-event: $L = -26{,}670{,}470.593$; Table 3). This result was confirmed through comparison of simulated SFS under these model parameters to the empirical SFS, with both the 3-event and 4-event models having fairly concordant fits (Figure 2a). Additionally, the fit of these simulated SFS to the empirical SFS was much stronger than that observed in any of the evaluated multiple-deme models (Supporting Figure 8). Moreover, estimated parameters between the 2-event, 3-event, and 4-event models were reasonably similar; differences were primarily found in estimated historical population sizes. Specifically, the current population size was estimated to be between ~125,000 to 160,000 haploid genomes, the ancestral population size was estimated to be between ~60,000 to 68,000 haploid genomes, and the initial population size change was inferred to have occurred between ~60,000 to 68,000 generations ago (Table 3 and Figure 2b,c; all parameters are available in Supporting Table 9).

Of the models analyzed by δaδi for comparison (for details, see the "Materials and Methods" section, and for a similar demographic comparison between fastsimcoal2 and δaδi, see Terbot et al. 2025), the best-fitting demographic history consisted of three population size change events (maximum likelihood = −726.45). The next best-fitting model was the four-size change event model (maximum likelihood = − 1,244.23). The inferred demographic history consists of an ancestral population size of ~62,000 haploid genomes, an expansion to ~300,000 haploid genomes ~58,000 generations ago, followed by a rapid expansion to ~3,300,000 haploid genomes ~4,650 generations ago, and finally a contraction to ~170,000 haploid genomes ~3,000 generations ago (Figure 2d, and see Supporting Table 9 for the precise parameter values). While the maximum population size estimated by δaδi is substantially larger than those estimated by fastsimcoal2, the current and ancestral population sizes, as well as the timing of the first population size change event, are comparable to those estimated by fastsimcoal2 (Table 3). Additionally, the fit of the simulated SFS under this model relative to the empirical data was similar to that of the best-performing fastsimcoal2 models (Figure 2a), and as such was also found to fit the observed data well.

Notably, the demographic model estimated for a single deme in this study shares similarities with the demography previously estimated using PSMC (Li and Durbin 2011) applied to this data set (Liu et al. 2018; Zhou et al. 2024a, 2024b) – which was additionally found to be highly similar for all five previously assumed demes, consistent with our improved single-deme model fit. Specifically, when using a generation time of 11 years (as assumed by Liu et al. 2018), the initial population growth in the model reported here occurred around 742 kya, roughly corresponding to the conclusion of the Xixiabangma Glaciation (~1,200-800 kya). Additionally, the most recent population contraction occurred around 19 kya, corresponding to the timing of the Last Glaciation (~70-10 kya). Both of these major population size changes and their association with the occurrence or ending

TABLE 3 | Single deme models.

	N_{cur}	N_{anc}	T_{max}	Log likelihood	Δ *logL*
0 events	190,135	190,135	na	−26,670,470.593	−405,173.214
1 event	315,178	60,772	67,228	−26,269,065.621	−3,768.242
2 events	153,697	65,437	60,773	−26,265,512.238	−214.859
3 events	126,626	65,196	68,368	−26,265,468.903	−171.524
4 events	160,646	60,170	67,421	−26,265,297.379	0.000

Note: Summary of results from single-deme models (consisting of 0 to 4 population re-sizing events) parameterized by fastsimcoal2. For each model only partial parameters are presented here (N_{cur} is the current population size in terms of haploid genomes, N_{anc} is the ancestral population size in terms of haploid genomes before any size changes, and T_{max} is the time in terms of "generations ago" of the first estimated size change from the ancestral population size). The log likelihood scores for each of these models is presented and the Δ*logL* is the difference between each model's log likelihood and the maximum log likelihood across all models.

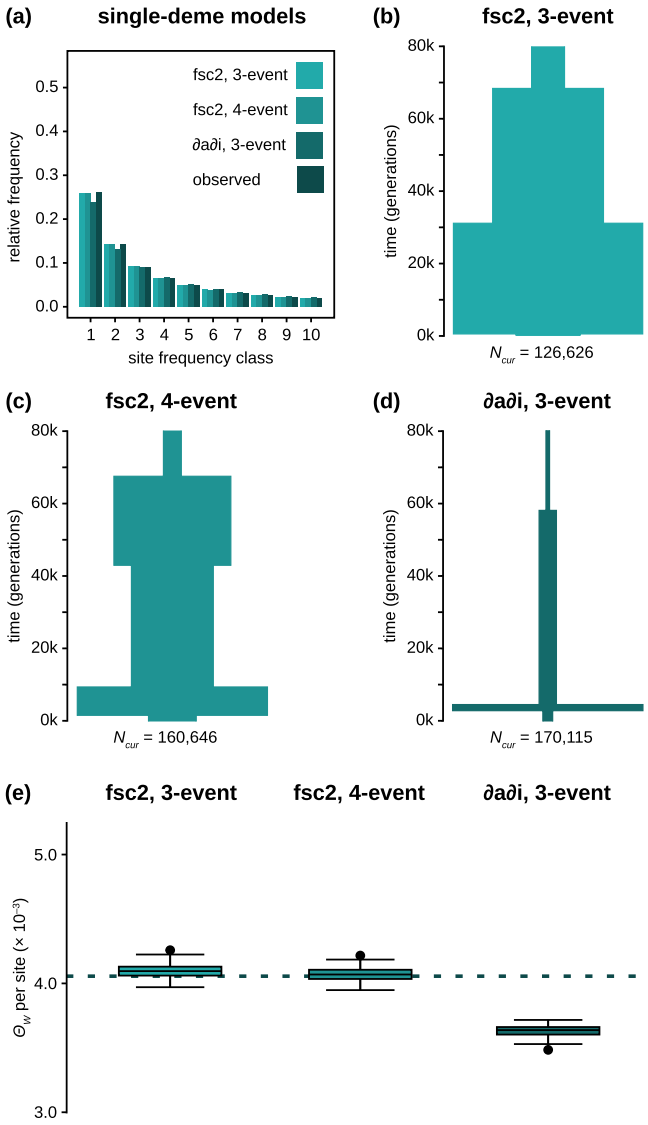

FIGURE 2 | Single deme models. (a) Simulated, mean, folded, relative site frequency spectra (SFS) for the best fitting single-deme models alongside the SFS observed from the empirical data. Each SFS bin contains bars for the relative frequency of the site frequency class for the data simulated under the 3-event model parameterized by fastsimcoal2 (fsc2) (leftmost, very lightly shaded), the 4-event model parameterized by fastsimcoal2 (second from the left, lightly shaded), the 3-event model parameterized by δaδi (second from the right, shaded), and the empirically observed SFS (rightmost, darkly shaded). Only the first ten bins for each deme are depicted here; the full SFS for all bins can be found in Supporting Figure 8. Note that all single-deme models correspond well to the observed SFS, though the overall fit of the fastsimcoal2 models are stronger. (b–d) Diagrams of the demographic history inferred by the three best-fitting, single-deme models: fastsimcoal2's 3-event model, fastsimcoal2's 4-event model, and δaδi's 3-event model. Note that population sizes between diagrams are not equivalently scaled. (e) Box-and-whisker plots of the per-site θ_W values for 1Mb-long genomic segments simulated using fastsimcoal2 under the three best-performing models: fastsimcoal2's 3-event model, fastsimcoal2's 4-event model, and δaδi's 3-event model. The dashed horizontal line represents the mean, empirically observed per-site θ_W value based on 1Mb-long sliding windows (step size: 0.5 Mb).

of glaciation periods were previously reported in Liu et al. (2018) and Zhou et al. (2024a, 2024b). Population growth was also estimated to the same timeframe using PSMC with an independent data set (Xue et al. 2016). Our results thus further confirm the likely importance of these changes in glaciation to the historical population dynamics of rhesus macaques in China.

While the above results quantified multiple single-deme population histories that well match the empirical data in terms of the SFS, they produced differences in overall levels of variation. Specifically, based on comparisons for θ_W, each parameterized model was significantly different from one another based on a Bonferroni corrected *t*-test (Figure 2e; fastsimcoal2, 3-event vs. fastsimcoal2, 4-event, p-value = 0.00097; fastsimcoal2, 3-event vs. δaδi, 3-event, p-value ≪ 0.00001; fastsimcoal2, 4-event vs. δaδi, 3-event, p-value ≪ 0.00001). Importantly, only the 4-event model parameterized by fastsimcoal2 was not significantly different from the observed empirical distribution of θ_W across the entire putatively neutral portion of the genome (Figure 2e; observed mean $\theta_W = 0.004056$; fastsimcoal2, 3-event vs. observed, p-value = 0.00104; fastsimcoal2, 4-event vs. observed, p-value = 0.2489; δaδi, 3-event vs. observed, p-value ≪ 0.00001). Taken together, this combination of results thus suggests with confidence that the single-deme 4-population-size-change-event model parameterized with fastsimcoal2 is the most capable of recapitulating both observed levels and patterns of genetic variation and thus, is the preferred demographic model for Chinese rhesus macaques.

3.5 | Inferring the Recombination Landscape of Chinese Rhesus Macaques

Indirect inference of the recombination landscape of Chinese rhesus macaques using patterns of LD under this best-fitting demographic model (i.e., fastsimcoal2's 4-event model) yielded an average genome-wide rate of 0.79 cM/Mb across 1Mb-windows. To determine the performance of the demography-aware recombination rate estimator pyrho under this specific demographic history, coalescent simulations were performed, simulating a region of ~220 Mb (i.e., the longest chromosome) under fastsimcoal2's 4-event model, assuming a neutral mutation rate (1.08e-8 per site per generation, i.e., the rate used by Liu et al. 2018) and a primate-like recombination rate (1 cM/Mb), sampling 79 individuals (i.e., the number of individuals in the empirical data set) in each replicate. Notably, these simulations demonstrated that, although pyrho's estimates are precise, they consistently underestimated the simulated recombination rate under the species' demographic history characterized by a severe population contraction ~3,000 generations ago (Supporting Figure 9). Recent research by Dutheil (2024) has shown that pyrho's inference is relatively robust to population declines at large sample sizes (> 50 diploid individuals) even when recombination landscapes are heterogeneous, though some degree of underestimation is to be expected (see figure 2 in Dutheil 2024). Moreover, even in their simple demographic models tested, population recombination rates were frequently underestimated in the presence of gene conversion, particularly in declining populations such as the one inferred here (see Figure 4 in Dutheil 2024). As gene conversion

tends to occur at a much higher rates than crossovers in many primates (Williams et al. 2015; Halldorsson et al. 2016; Wall et al. 2022; Versoza et al. 2024; Palsson et al. 2025; Versoza et al. 2025), patterns of gene conversion, such as those previously reported in the species (Versoza et al. 2024), likely contribute to the observed underestimation of recombination rate.

Taking the simulation results at face value, the actual average genome-wide recombination rate of Chinese rhesus macaques might thus be nearer to 1.21 cM/Mb (Figure 3a). Interestingly, the recombination rate estimate obtained here from wild-born individuals is considerably higher than previous estimates inferred from captive individuals housed at research colonies, which displayed a stark reduction in overall recombination rate compared to humans (average genome-wide recombination rates for colony-born rhesus macaques have been inferred to range from 0.433 ± 0.333 cM/Mb to 0.448 ± 0.286 cM/Mb [Xue et al. 2016, 2020], whereas humans exhibit average rates around 1.322 ± 1.399 cM/Mb [International HapMap Consortium 2007]). Notably, most earlier estimates available for rhesus macaques failed to account for the demographic history of the species (Xue et al. 2016, 2020) – a neglect that is particularly problematic given that captive individuals frequently originate from severely bottlenecked populations. Population declines and bottlenecks result in increased LD across the genome (as summarized by Walsh and Lynch 2019); consequently, unaccounted for population size changes can reduce the power and increase the false positive rate of indirect recombination rate methods relying on these patterns (for a discussion, see Dapper and Payseur 2018; Johri et al. 2022). More recent estimates have partially alleviated this issue by inferring, and accounting for, the size history of the population in rhesus macaques of Indian origin (Versoza et al. 2024); though it should be noted that the sequentially Markovian coalescence approach employed in the study assumed constant rates of mutation and recombination across the genome. The violation of this assumption – as is the case in primate genomes (see the reviews of Stapley et al. 2017; Pfeifer 2020b) – will bias estimates of coalescence times and is thus expected to result in both inaccurate estimations of demographic history as well as an underestimation of recombination rates (Terhorst et al. 2017; Sellinger et al. 2020; and see Sellinger et al. 2021 for a discussion). Nevertheless, despite differences in the overall rates between this previous demography-aware estimate for Indian-origin rhesus macaques (0.88 cM/Mb; Versoza et al. 2024) and those reported here for Chinese-origin rhesus macaques, recombination landscapes display a strong conservation at the broad-scale (Pearson's $r = 0.765$ at the 1 Mb-scale; Figure 3b). At the finer scales, however, differences between the two populations become more prominent, as expected from previous work in other primates (see the reviews of Stapley et al. 2017; Johnston 2024 and references therein).

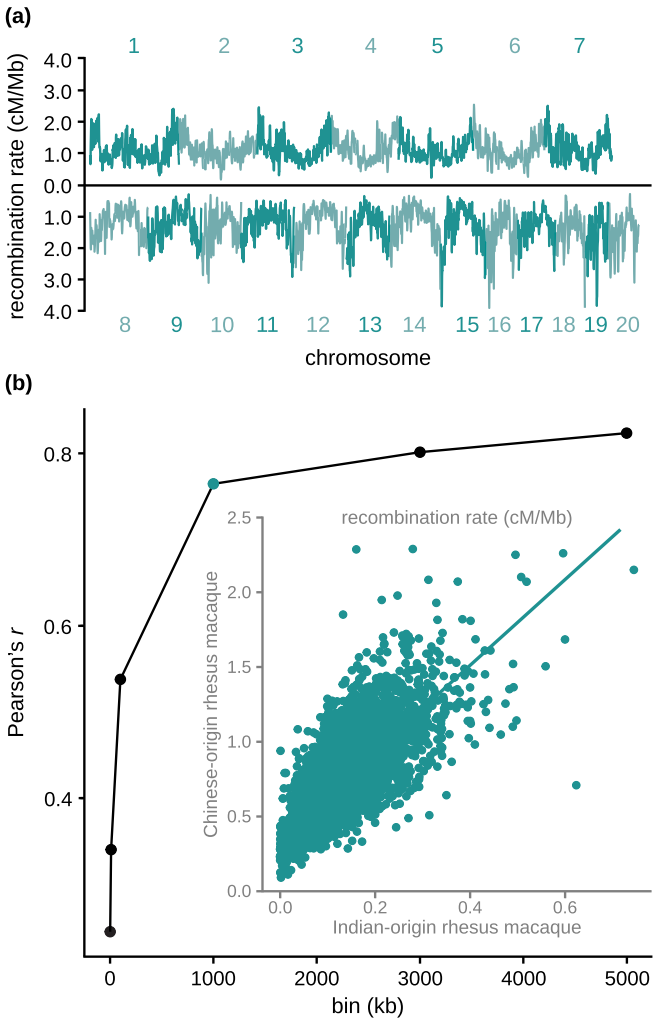

FIGURE 3 | Recombination landscape inferred for Chinese rhesus macaques. (a) Genome-wide recombination rate across each autosome (chromosomes 1-20) indirectly inferred using patterns of linkage disequilibrium and accounting for the best-fitting demographic model (i.e., fastsimcoal2's 4-event model). (b) Pearson correlation between the recombination rates inferred here for wild-born Chinese individuals and those previously obtained from captive-born individuals of Indian origin (see Versoza et al. 2024 for details) at different scales. The inlay shows the correlation at the broad (1 Mb) scale.

4 | Conclusions

We found the genomic evidence for the designation of five subspecies of rhesus macaque in China to be statistically unsupported. Specifically, the poor fit of the SFS of multi-deme models to the empirical data, combined with the tendency of those models to gravitate towards extensive gene flow to improve the fit, both suggested the need to re-evaluate population structure. Based on multiple structure analyses, as well as the strong fit of the SFS of a parameterized single-deme model to the empirical data, these results taken together suggest that the individuals from this collection of sampling locations are better described as a single deme.

Notably, these samples neither include individuals from the Tibetan subspecies (*M. m. vestita*) nor do they include samples from the more Western, Indian subpopulation of rhesus macaques (and there is indeed strong support for population structure generally between Chinese and Indian rhesus populations as assessed by F_{ST}, with a previously suggested split time ~162 kya; Hernandez et al. 2007). Thus, while these results support a single deme across the 17 sampling regions in China discussed here, they do not speak more generally to other possible subspecific divisions within the species. Additionally, the morphological variation between these subspecies previously reported

(Jiang et al. 1991; Zhang et al. 2021) may still be relevant for studies focusing on that aspect of their biology or using those characteristics to assess geographic regions of origin. It may be possible that these morphological differences are the result of environmental variation and phenotypic plasticity rather than purely heritable genetic differences more commonly associated with subspecific divisions. Alternatively, it may be possible that genetic structure between these subpopulations does exist to some degree within the exonic regions excluded from the demographic estimation performed here, as previously argued (Trask et al. 2011). This may be similar to the differentiation between *M. mulatta* and *M. fascicularis* either through reported reinforcement speciation ongoing between these two species (Bailey et al. 2023) or through differences in gene regulatory regions (Zhang et al. 2025). However, it is unclear how such incipient speciation could result in divergence limited only to exonic sites and not present in some capacity within the putatively neutral regions of the genome.

Utilizing this proposed well-fitting population history, we additionally characterized the fine-scale recombination rate map across these genomes, finding genome-wide recombination rates comparable with those observed in other primate species (see the reviews of Stapley et al. 2017; Johnston 2024). However, despite the broad-scale recombination rates being well conserved between different populations of rhesus macaques, we observed notable differences in the fine-scale landscapes between Chinese-origin and Indian-origin macaques – two populations widely used as models in biomedical research.

Taken together, given the unique demographic histories and recombination landscapes of Chinese-origin and Indian-origin rhesus macaques, it will thus be important for future biomedical and genomic studies to account for the differences observed between these populations to improve upon the characterization of adaptive and pathogenic variation in the species (see the discussion in Johri et al. 2022).

Acknowledgments

We would like to thank all members of the Jensen Lab and Pfeifer Lab for helpful discussion. Computations were performed on the Sol supercomputer at Arizona State University (Jennewein et al. 2023) and on the Open Science Grid (OSG 2006, 2015; Pordes et al. 2007; Sfiligoi et al. 2009).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

This study was based on whole-genome sequencing data publicly available from the NCBI SRA (BioProject: PRJNA345528). All scripts used to generate simulations, and to analyze both empirical and simulated data, are available at the GitHub repository: https://github.com/jwterbot2/Chinese_rhesus. The LD-based recombination map generated in this study is available at http://spfeiferlab.org/data.

References

Alexander, D. H., J. Novembre, and K. Lange. 2009. "Fast Model-Based Estimation of Ancestry in Unrelated Individuals." *Genome Research* 19, no. 9: 1655–1664. https://doi.org/10.1101/gr.094052.109.

Amemiya, H. M., A. Kundaje, and A. P. Boyle. 2019. "The ENCODE Blacklist: Identification of Problematic Regions of the Genome." *Scientific Reports* 9, no. 1: 9354.

Auton, A., A. Fledel-Alon, S. Pfeifer, et al. 2012. "A Fine-Scale Chimpanzee Genetic Map From Population Sequencing." *Science* 336, no. 6078: 193–198. https://doi.org/10.1126/science.1216872.

van der Auwera, G., and B. O'Connor. 2020. *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra* (1st ed.). Sebastopol, CA: O'Reilly Media.

Bailey, N., C. Ruiz, A. Tosi, and L. Stevison. 2023. "Genomic Analysis of the Rhesus Macaque (*Macaca mulatta*) and the Cynomolgus Macaque (*Macaca fascicularis*) Uncover Polygenic Signatures of Reinforcement Speciation." *Ecology and Evolution* 13, no. 10: 1–19. https://doi.org/10.1002/ece3.10571.

Baumdicker, F., G. Bisschop, D. Goldstein, et al. 2022. "Efficient Ancestry and Mutation Simulation With msprime 1.0." *Genetics* 220, no. 3: iyab229. https://doi.org/10.1093/genetics/iyab229.

Brandon-Jones, D., A. A. Eudey, T. Geissmann, et al. 2004. "Asian Primate Classification." *International Journal of Primatology* 25: 97–164. https://doi.org/10.1023/B:IJOP.0000014647.18720.32.

Browning, B. L., X. Tian, Y. Zhou, and S. R. Browning. 2021. "Fast Two-Stage Phasing of Large-Scale Sequence Data." *American Journal of Human Genetics* 108, no. 10: 1880–1890. https://doi.org/10.1016/j.ajhg.2021.08.005.

Charlesworth, B., and J. D. Jensen. 2021. "Effects of Selection at Linked Sites on Patterns of Genetic Variability." *Annual Review of Ecology, Evolution, and Systematics* 52, no. 1: 177–197. https://doi.org/10.1146/annurev-ecolsys-010621-044528.

Charlesworth, B., and J. D. Jensen. 2024. "Population Genetics." In *Encyclopedia of Biodiversity*, edited by S. M. Scheiner, (3rd ed., 467–483. Amsterdam: Elsevier.

Chiou, K. L., M. J. Montague, E. A. Goldman, et al. 2020. "Rhesus Macaques as a Tractable Physiological Model of Human Ageing." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 375, no. 1811: 20190612. https://doi.org/10.1098/rstb.2019.0612.

Chu, J. H., Y. S. Lin, and H. Y. Wu. 2007. "Evolution and Dispersal of Three Closely Related Macaque Species, *Macaca mulatta*, *M. cyclopis*, and *M. fuscata*, in the Eastern Asia." *Molecular Phylogenetics and Evolution* 43, no. 2: 418–429. https://doi.org/10.1016/j.ympev.2006.11.022.

Clarke, M. R., and J. A. S. O'Neil. 1999. "Morphometric Comparison of Chinese-Origin and Indian-Derived Rhesus Monkeys (*Macaca mulatta*)." *American Journal of Primatology* 47, no. 4: 335–346. https://doi.org/10.1002/(SICI)1098-2345(1999)47:4<335::AID-AJP5>3.0.CO;2-Y.

Danecek, P., A. Auton, G. Abecasis, et al. 2011. "The Variant Call Format and Vcftools." *Bioinformatics* 27, no. 15: 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

Danecek, P., J. K. Bonfield, J. Liddle, et al. 2021. "Twelve Years of SAMtools and BCFtools." *GigaScience* 10, no. 2: 1–4. https://doi.org/10.1093/gigascience/giab008.

Dapper, A. L., and B. A. Payseur. 2018. "Effects of Demographic History on the Detection of Recombination Hotspots From Linkage Disequilibrium." *Molecular Biology and Evolution* 35, no. 2: 335–353. https://doi.org/10.1093/molbev/msx272.

Dutheil, J. Y. 2024. "on the Estimation of Genome-Average Recombination Rates." *Genetics* 227, no. 2: iyae051. https://doi.org/10.1093/genetics/iyae051.

Ewing, G. B., and J. D. Jensen. 2014. "Distinguishing Neutral From Deleterious Mutations in Growing Populations." *Frontiers in Genetics* 5: 7. https://doi.org/10.3389/fgene.2014.00007.

Ewing, G. B., and J. D. Jensen. 2016. "The Consequences of Not Accounting for Background Selection in Demographic Inference." *Molecular Ecology* 25, no. 1: 135–141. https://doi.org/10.1111/mec.13390.

Excoffier, L., I. Dupanloup, E. Huerta-Sánchez, V. C. Sousa, and M. Foll. 2013. "Robust Demographic Inference From Genomic and SNP Data." *PLoS Genetics* 9, no. 10: e1003905. https://doi.org/10.1371/journal.pgen.1003905.

Excoffier, L., N. Marchi, D. A. Marques, R. Matthey-Doret, A. Gouy, and V. C. Sousa. 2021. "fastsimcoal2: Demographic Inference Under Complex Evolutionary Scenarios." *Bioinformatics* 37, no. 24: 4882–4885. https://doi.org/10.1093/bioinformatics/btab468.

Falush, D., M. Stephens, and J. K. Pritchard. 2003. "Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies." *Genetics* 164, no. 4: 1567–1587. https://doi.org/10.1093/genetics/164.4.1567.

Ferguson, B., S. L. Street, H. Wright, et al. 2007. "Single Nucleotide Polymorphisms (SNPs) Distinguish Indian-Origin and Chinese-Origin Rhesus Macaques (*Macaca mulatta*)." *BMC Genomics* 8: 43. https://doi.org/10.1186/1471-2164-8-43.

Gibbs, R. A., J. Rogers, M. G. Katze, et al. 2007. "Evolutionary and Biomedical Insights From the Rhesus Macaque Genome." *Science* 316, no. 5822: 222–234. https://doi.org/10.1126/science.1139247.

Gutenkunst, R. N., R. D. Hernandez, S. H. Williamson, and C. D. Bustamante. 2009. "Inferring the Joint Demographic History of Multiple Populations From Multidimensional SNP Frequency Data." *PLoS Genetics* 5, no. 10: e1000695. https://doi.org/10.1371/journal.pgen.1000695.

Halldorsson, B. V., M. T. Hardarson, B. Kehr, et al. 2016. "The Rate of Meiotic Gene Conversion Varies by Sex and Age." *Nature Genetics* 48, no. 11: 1377–1384. https://doi.org/10.1038/ng.3669.

Harris, A. R. 2019. Rhesus macaque SNP calls VCF. Zenodo. Dataset. Available from: https://doi.org/10.5281/zenodo.3515522.

Hernandez, R. D., M. J. Hubisz, D. A. Wheeler, et al. 2007. "Demographic Histories and Patterns of Linkage Disequilibrium in Chinese and Indian Rhesus Macaques." *Science* 316, no. 5822: 240–243. https://doi.org/10.1126/science.1140462.

International HapMap Consortium, K. A. Frazer, D. G. Ballinger, D. R. Cox, et al. 2007. "A Second Generation Human Haplotype Map of Over 3.1 Million SNPs." *Nature* 449, no. 7164: 851–861. https://doi.org/10.1038/nature06258.

Ito, T., T. Hayakawa, N. Suzuki–Hashido, et al. 2021. "Phylogeographic History of Japanese Macaques." *Journal of Biogeography* 48, no. 6: 1420–1431. https://doi.org/10.1111/jbi.14087.

Jennewein, D. M., J. Lee, C. Kurtz, et al. 2023. "The Sol Supercomputer at Arizona State University." *Pract and Exp in Adv Res Comp* 2023: 296–301.

Jiang, X., Y. Wang, and S. Ma. 1991. "Taxonomic Revision and Distribution of Subspecies of Rhesus Monkey (*Macaca mulatta*) in China." *Zoological Research* 12, no. 3: 241–247.

Johnston, S. E. 2024. "Understanding the Genetic Basis of Variation in Meiotic Recombination: Past, Present, and Future." *Molecular Biology and Evolution* 41, no. 7: msae112. https://doi.org/10.1093/molbev/msae112.

Johri, P., C. F. Aquadro, M. Beaumont, et al. 2022. "Recommendations for Improving Statistical Inference in Population Genomics." *PLoS Biology* 20, no. 5: e3001669. https://doi.org/10.1371/journal.pbio.3001669.

Johri, P., B. Charlesworth, and J. D. Jensen. 2020. "Toward An Evolutionarily Appropriate Null Model: Jointly Inferring Demography and Purifying Selection." *Genetics* 215, no. 1: 173–192. https://doi.org/10.1534/genetics.119.303002.

Johri, P., S. P. Pfeifer, and J. D. Jensen. 2023. "Developing an Evolutionary Baseline Model for Humans: Jointly Inferring Purifying Selection With Population History." *Molecular Biology and Evolution* 40, no. 5: msad100. https://doi.org/10.1093/molbev/msad100.

Johri, P., K. Riall, H. Becher, L. Excoffier, B. Charlesworth, and J. D. Jensen. 2021. "The Impact of Purifying and Background Selection on the Inference of Population History: Problems and Prospects." *Molecular Biology and Evolution* 38, no. 7: 2986–3003. https://doi.org/10.1093/molbev/msab050.

Kuderna, L. F. K., J. C. Ulirsch, S. Rashid, et al. 2024. "Identification of Constrained Sequence Elements Across 239 Primate Genomes." *Nature* 625, no. 7996: 735–742. https://doi.org/10.1038/s41586-023-06798-8.

Li, H., and R. Durbin. 2009. "Fast and Accurate Short Read Alignment With Burrows-Wheeler Transform." *Bioinformatics* 25, no. 14: 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

Li, H., and R. Durbin. 2011. "Inference of Human Population History From Individual Whole-Genome Sequences." *Nature* 475, no. 7357: 493–496. https://doi.org/10.1038/nature10231.

Liu, Z., X. Tan, P. Orozco-TerWengel, et al. 2018. "Population Genomics of Wild Chinese Rhesus Macaques Reveals a Dynamic Demographic History and Local Adaptation, With Implications for Biomedical Research." *GigaScience* 7, no. 9: 1–14. https://doi.org/10.1093/gigascience/giy106.

Marchi, N., A. Kapopoulou, and L. Excoffier. 2024. "Demogenomic Inference From Spatially and Temporally Heterogeneous Samples." *Molecular Ecology Resources* 24, no. 1: 1–11. https://doi.org/10.1111/1755-0998.13877.

Osada, N., Y. Uno, K. Mineta, Y. Kameoka, I. Takahashi, and K. Terao. 2010. "Ancient Genome-Wide Admixture Extends Beyond the Current Hybrid Zone Between *Macaca fascicularis* and *M. mulatta*." *Molecular Ecology* 19, no. 14: 2884–2895. https://doi.org/10.1111/j.1365-294X.2010.04687.x.

OSG. 2006. "OSPool." *OSG*. https://doi.org/10.21231/906P-4D78.

OSG. 2015. "Open Science Data Federation." *OSG*. https://doi.org/10.21231/0KVZ-VE57.

Palsson, G., M. T. Hardarson, H. Jonsson, et al. 2025. "Complete Human Recombination Maps." *Nature* 639, no. 8055: 700–707. https://doi.org/10.1038/s41586-024-08450-5.

Pfeifer, S. P. 2017. "From Next-Generation Resequencing Reads to a High-Quality Variant Data Set." *Heredity* 118, no. 2: 111–124. https://doi.org/10.1038/hdy.2016.102.

Pfeifer, S. P. 2020a. "A Fine-Scale Genetic Map for Vervet Monkeys." *Molecular Biology and Evolution* 37, no. 7: 1855–1865. https://doi.org/10.1093/molbev/msaa079.

Pfeifer, S. P. 2020b. "Spontaneous Mutation Rates." In *The Molecular Evolutionary Clock. Theory and Practice*, edited by S. Y. W. Ho, 35–44. Springer Nature.

Phadphon, P., S. Kanthaswamy, R. F. Oldt, Y. Hamada, and S. Malaivijitnond. 2022. "Population Structure of *Macaca fascicularis aurea*, and Their Genetic Relationships With *M. f. fascicularis* and *M. mulatta* Determined by 868 RADseq-Derived Autosomal SNPs—A Consideration for Biomedical Research." *Journal of Medical Primatology* 51, no. 1: 33–44. https://doi.org/10.1111/jmp.12554.

Pordes, R., D. Petravick, B. Kramer, et al. 2007. "The Open Science Grid." *Journal of Physics: Conference Series* 78: 012057. https://doi.org/10.1088/1742-6596/78/1/012057.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. "Inference of Population Structure Using Multilocus Genotype Data." *Genetics* 155, no. 2: 945–959. https://doi.org/10.1093/genetics/155.2.945.

Purcell, S., B. Neale, K. Todd-Brown, et al. 2007. "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *American Journal of Human Genetics* 81, no. 3: 559–575. https://doi.org/10.1086/519795.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raj, A., M. Stephens, and J. K. Pritchard. 2014. "fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets." *Genetics* 197, no. 2: 573–589. https://doi.org/10.1534/genetics.114.164350.

Rogers, J. 2022. "Genomic Resources for Rhesus Macaques (*Macaca mulatta*)." *Mammalian Genome* 33, no. 1: 91–99. https://doi.org/10.1007/s00335-021-09922-z.

Sellinger, T. P. P., D. Abu Awad, M. Moest, and A. Tellier. 2020. "Inference of Past Demography, Dormancy and Self-Fertilization Rates From Whole Genome Sequence Data." *PLoS Genetics* 16, no. 4: e1008698. https://doi.org/10.1371/journal.pgen.1008698.

Sellinger, T. P. P., D. Abu-Awad, and A. Tellier. 2021. "Limits and Convergence Properties of the Sequentially Markovian Coalescent." *Molecular Ecology Resources* 21, no. 7: 2231–2248. https://doi.org/10.1111/1755-0998.13416.

Sfiligoi, I., D. C. Bradley, B. Holzman, P. Mhashilkar, S. Padhi, and F. Wurthwein 2009. The Pilot Way to Grid Resources Using GlideinWMS. In: 2009 WRI World Congress on cOmputer Science and Information Engineering. Washington, DC: IEEE. p. 428–432.

Smith, D. G., and J. McDonough. 2005. "Mitochondrial DNA Variation in Chinese and Indian Rhesus Macaques (*Macaca mulatta*)." *American Journal of Primatology* 65, no. 1: 1–25. https://doi.org/10.1002/ajp.20094.

Soni, V., and J. D. Jensen. 2025. "Inferring Demographic and Selective Histories From Population Genomic Data Using a 2-step Approach in Species With Coding-Sparse Genomes: An Application to Human Data." *G3: Genes, Genomes, Genetics* 15, no. 4: jkaf019. https://doi.org/10.1093/g3journal/jkaf019.

Soni, V., C. Versoza, J. D. Jensen, and S. P. Pfeifer. 2025a. "Inferring the Landscape of Mutation and Recombination in the Common Marmoset (*Callithrix jacchus*) in the Presence of Twinning and Hematopoietic Chimerism." *bioRxiv [pre-print]*. https://doi.org/10.1101/2025.07.01.662565.

Soni, V., C. J. Versoza, J. W. Terbot, J. D. Jensen, and S. P. Pfeifer. 2025b. "Inferring Fine-Scale Mutation and Recombination Rate Maps in Aye-Ayes (*Daubentonia madagascariensis*)." *Ecology and Evolution*. in press. https://doi.org/10.1101/2024.12.28.630620.

Spence, J. P., and Y. S. Song. 2019. "Inference and Analysis of Population-Specific Fine-Scale Recombination Maps Across 26 Diverse Human Populations." *Science Advances* 5, no. 10: eaaw9206. https://doi.org/10.1126/sciadv.aaw9206.

Stapley, J., P. G. D. Feulner, S. E. Johnston, A. W. Santure, and C. M. Smadja. 2017. "Variation in Recombination Frequency and Distribution Across Eukaryotes: Patterns and Processes." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 372, no. 1736: 20160455.

Stevison, L. S., A. E. Woerner, J. M. Kidd, et al. 2016. "The Time Scale of Recombination Rate Evolution in Great Apes." *Molecular Biology and Evolution* 33, no. 4: 928–945. https://doi.org/10.1093/molbev/msv331.

Terbot, J. W., V. Soni, C. J. Versoza, S. P. Pfeifer, and J. D. Jensen. 2025. "Inferring the Demographic History of Aye-Ayes (*Daubentonia madagascariensis*) From High-Quality, Whole-Genome, Population-Level Data." *Genome Biology and Evolution* 17, no. 1: evae281. https://doi.org/10.1093/gbe/evae281.

Terhorst, J., J. A. Kamm, and Y. S. Song. 2017. "Robust and Scalable Inference of Population History From Hundreds of Unphased Whole Genomes." *Nature Genetics* 49, no. 2: 303–309. https://doi.org/10.1038/ng.3748.

Thierry, B. 2017. "Macaque (Macaca)." *Int Encycl Primatol.*: 1–4. https://doi.org/10.1002/9781119179313.wbprim0040.

Thornton, K. 2003. "Libsequence: A C++ Class Library for Evolutionary Genetic Analysis." *Bioinformatics* 19, no. 17: 2325–2327. https://doi.org/10.1093/bioinformatics/btg316.

Trask, J. A. S., R. S. Malhi, S. Kanthaswamy, et al. 2011. "The Effect of SNP Discovery Method and Sample Size on Estimation of Population Genetic Data for Chinese and Indian Rhesus Macaques (*Macaca mulatta*)." *Primates* 52, no. 2: 129–138. https://doi.org/10.1007/s10329-010-0232-4.

Versoza, C. J., A. Lloret-Villas, J. D. Jensen, and S. P. Pfeifer. 2025. "A Pedigree-Based Map of Crossovers and Non-Crossovers in Aye-Ayes (*Daubentonia madagascariensis*)." *Genome Biology and Evolution* 17: evaf072. Epub ahead of print. https://doi.org/10.1093/gbe/evaf072.

Versoza, C. J., S. Weiss, R. Johal, B. La Rosa, J. D. Jensen, and S. P. Pfeifer. 2024. "Novel Insights Into the Landscape of Crossover and Noncrossover Events in Rhesus Macaques (*Macaca mulatta*)." *Genome Biology and Evolution* 16, no. 1: evad223. https://doi.org/10.1093/gbe/evad223.

Wall, J. D., J. A. Robinson, and L. A. Cox. 2022. "High-Resolution Estimates of Crossover and Noncrossover Recombination From a Captive Baboon Colony." *Genome Biology and Evolution* 14, no. 4: evac040. https://doi.org/10.1093/gbe/evac040.

Walsh, B., and M. Lynch. 2019. *Evolution and Selection of Quantitative Traits*. Oxford: Oxford University Press.

Wang, R. J., G. W. C. Thomas, M. Raveendran, et al. 2020. "Paternal Age in Rhesus Macaques Is Positively Associated With Germline Mutation Accumulation but Not With Measures of Offspring Sociability." *Genome Research* 30, no. 6: 826–834. https://doi.org/10.1101/gr.255174.119.

Warren, W. C., R. A. Harris, M. Haukness, et al. 2020. "Sequence Diversity Analyses of an Improved Rhesus Macaque Genome Enhance Its Biomedical Utility." *Science* 370, no. 6523: abc6617. https://doi.org/10.1126/science.abc6617.

Watterson, G. A. 1975. "On the Number of Segregating Sites in Genetical Models Without Recombination." *Theoretical Population Biology* 7, no. 2: 256–276. https://doi.org/10.1016/0040-5809(75)90020-9.

Weir, B. S., and C. C. Cockerham. 1984. "Estimating F-Statistics for the Analysis of Population Structure." *Evolution* 38, no. 6: 1358. https://doi.org/10.2307/2408641.

Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York. ISBN 978-3-319-24277-4: Springer-Verlag.

Williams, A. L., G. Genovese, T. Dyer, et al. 2015. "Non-Crossover Gene Conversions Show Strong GC Bias and Unexpected Clustering in Humans." *eLife* 4: e04637. https://doi.org/10.7554/eLife.04637.

Wiseman, R. W., J. A. Karl, P. S. Bohn, F. A. Nimityongskul, G. J. Starrett, and D. H. O'Connor. 2013. "Haplessly Hoping: Macaque Major Histocompatibility Complex Made Easy." *ILAR Journal* 54, no. 2: 196–210. https://doi.org/10.1093/ilar/ilt036.

Wu, F. L., A. I. Strand, L. A. Cox, et al. 2020. "A Comparison of Humans and Baboons Suggests Germline Mutation Rates Do Not Track Cell Divisions." *PLoS Biology* 18, no. 8: e3000838. https://doi.org/10.1371/journal.pbio.3000838.

Xue, C., M. Raveendran, R. A. Harris, et al. 2016. "The Population Genomics of Rhesus Macaques (*Macaca mulatta*) Based on Whole-Genome Sequences." *Genome Research* 26, no. 12: 1651–1662. https://doi.org/10.1101/gr.204255.116.

Xue, C., N. Rustagi, X. Liu, et al. 2020. "Reduced Meiotic Recombination in Rhesus Macaques and the Origin of the Human Recombination Landscape." *PLoS One* 15, no. 8: e0236285. https://doi.org/10.1371/journal.pone.0236285.

Zhang, B. L., W. Chen, Z. Wang, et al. 2023. "Comparative Genomics Reveals the Hybrid Origin of a Macaque Group." *Science Advances* 9, no. 22: 1–13. https://doi.org/10.1126/sciadv.add3580.

Zhang, R., C. Wu, T. Chen, J. Zhang, and P. Zhang. 2021. "Morphological Characteristics of *Macaca Mulatta brevicaudus*." *Acta Anthropol Sin* 40, no. 1: 97–108. https://doi.org/10.16359/j.cnki.cn11-1963/q.2018.0040.

Zhang, S., N. Xu, L. Fu, et al. 2025. "Integrated Analysis of the Complete Sequence of a Macaque Genome." *Nature* 640, no. 8059: 714–721. https://doi.org/10.1038/s41586-025-08596-w.

Zhang, Y., and L. Shi. 1993. "Phylogeny of Rheusus Monkeys (*Macaca mulatta*) as Revealed by Mitochondrial DNA Restriction Enzyme

Analysis." *International Journal of Primatology* 14, no. 4: 587–605. https://doi.org/10.1007/BF02215449.

Zhou, Y., J. Tian, M. Han, and J. Lu. 2024a. "The Phylogenetic Relationship and Demographic History of Rhesus Macaques (*Macaca mulatta*) in Subtropical and Temperate Regions, China." *Ecology and Evolution* 14, no. 5: 1–13. https://doi.org/10.1002/ece3.11429.

Zhou, Y., J. Tian, H. Jiang, M. Han, Y. Wang, and J. Lu. 2024b. "Phylogeography and Demographic History of Macaques, *Fascicularis* Species Group, in East Asia: Inferred From Multiple Genomic Markers." *Molecular Phylogenetics and Evolution* 194, no. April 2023: 108042. https://doi.org/10.1016/j.ympev.2024.108042.

Zhou, Y., J. Tian, and J. Lu. 2023. "Genetic Structure and Recent Population Demographic History of Taihangshan Macaque (*Macaca mulatta tcheliensis*), North China." *Integrative Zoology* 18, no. 3: 530–542. https://doi.org/10.1111/1749-4877.12677.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supplementary Figure 1: Folded site frequency spectra for the dataset obtained by applying the nonstandard filter criteria of Liu et al. **Supplementary Figure 2:** Simulated, mean, folded, relative site frequency spectra (SFS) for each of the five previously designated subspecies alongside the SFS observed from the empirical data: Macaca mulatta mulatta (shown in green), M. m. lasiotis (purple), M. m. brevicaudus (yellow), M. m. littoralis (red), and M. m. tcheliensis (blue). Each SFS bin contains bars for the relative frequency of the site frequency class for the data simulated under the parameterized model inferred by Liu et al. (2018) (left, lightly shaded), the re-parameterization of this model performed in this study (middle, moderately shaded), and the empirically observed SFS (right, darkly shaded). **Supplementary Figure 3:** Diagram of the 5-deme demographic model inferred by Liu et al. **Supplementary Figure 4:** Diagram of the 5-deme demographic model inferred by Liu et al. (2018) as reparameterization in this study, with the coloring of each putative subspecies following the scheme described in Supplementary Figure 2. **Supplementary Figure 5:** Diagram of 3-deme model with M. m. brevicaudus branching first, with the coloring of each group following the scheme described in Supplementary Figure 2. **Supplementary Figure 6:** Diagram of 3-deme model with M. m. tcheliensis branching first, with the coloring of each group following the scheme described in Supplementary Figure 2. **Supplementary Figure 7:** Simulated, mean, folded, relative site frequency spectra (SFS) for simulations using the best, parameterized 3-deme models alongside the SFS observed from the empirical data. **Supplementary Figure 8:** Simulated, mean, folded, relative site frequency spectra (SFS) for the best fitting single-deme models alongside the SFS observed from the empirical data. **Supplementary Figure 9:** Performance of pyrho under the best-fitting demographic model (i.e., fastsimcoal2's 4-event model), assuming a neutral mutation rate (1.08e-8 per base pair per generation, i.e., the rate used by Liu et al. 2018) and a primate-like recombination rate (1cM/Mb), sampling 79 individuals (i.e., the number of individuals in the empirical dataset) in each replicate. **Supplementary Table 1:** Summary of the datasets. The final dataset filtered according to the Genome Analysis Toolkit Best Practices (GATK BP) contained 42.1 million SNPs with a transition-transversion (Ts/Tv) ratio of 2.27 in the accessible genome. **Supplementary Table 2:** Best parameter values for the 5-deme model presented by Liu et al. **Supplementary Table 3:** Results from ADMIXTURE for the complete dataset (i.e., unpruned) and the dataset pruned for linkage. **Supplementary Table 4:** Results from fastSTRUCTURE for the complete dataset (i.e., unpruned) and the dataset pruned for linkage. **Supplementary Table 5:** Mean, pairwise FST values between previously designated Chinese rhesus macaque subspecies. **Supplementary Table 6:** Mean, pairwise FST by chromosome using five randomly sampled individuals from each deme. **Supplementary Table 7:** Mean, pairwise FST by chromosome using all 79 individuals. **Supplementary Table 8:** Best parameterizations for the 3-deme models. Population sizes are given in terms of haploid genomes and timing is provided in terms of "generations ago". **Supplementary Table 9:** Best parameterizations for the single-deme models estimated using fastsimcoal2 (fsc2) and δaδi. Population sizes are given in terms of haploid genomes and timing is provided in terms of "generations ago".