

RESEARCH ARTICLE OPEN ACCESS

Inferring the Landscapes of Mutation and Recombination in the Common Marmoset (*Callithrix jacchus*) in the Presence of Twinning and Hematopoietic Chimerism

Vivak Soni  | Cyril J. Versoza | Jeffrey D. Jensen  | Susanne P. Pfeifer 

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

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

Received: 17 January 2026 | **Revised:** 2 August 2026 | **Accepted:** 2 September 2026

Funding: National Institute of General Medical Sciences, Grant/Award Numbers: R35GM151008, R35GM139383; National Science Foundation, Grant/Award Number: DEB-2045343

Keywords: Callitrichidae | mutation | neutral divergence | platyrrhine | primate | recombination

ABSTRACT

The common marmoset (*Callithrix jacchus*) is one of the most commonly used nonhuman primate models in biomedical research. Yet, the population genomics of the species remains relatively poorly characterized, despite the critical importance of this knowledge in many areas of research including genome-wide association studies, models of polygenic risk scores, and scans for the targets of selection. This neglect owes, at least in part, to two biological peculiarities related to the reproductive mode of the species—frequent twinning and sibling chimerism—which are likely to affect standard population genetic approaches relying on assumptions underlying the Wright–Fisher model. Using high-quality population genomic data, we here infer the rates and landscapes of mutation and recombination—two fundamental processes dictating the levels and patterns of genetic variability—in the presence of these biological features, and discuss our findings in light of recent work in primates. Our results suggest that, while the species exhibits relatively low neutral mutation rates, rates of recombination are in the range of those observed in other anthropoids. Moreover, the recombination landscape of common marmosets, like that of many vertebrates, appears dominated by PRDM9-mediated hotspots, with artificial intelligence-based models predicting an intricate 3D-structure of the species-specific PRDM9-DNA binding complex in silico. Apart from providing novel insights into the population genetics of common marmosets, given the importance of the availability of fine-scale maps of mutation and recombination for evolutionary inference, this work will also serve as a valuable resource to aid future genomic research in this widely studied system.

1 | Introduction

The introduction of new genetic variation through the process of mutation and its reorganization through crossover and non-crossover events—the two possible outcomes of meiotic recombination—are key evolutionary mechanisms that influence genomic diversity. Both processes are known to be highly variable across species, with rates differing between populations,

individuals, and sites within a genome (see the reviews by Baer et al. 2007; Lynch 2010; Hodgkinson and Eyre-Walker 2011; Pfeifer 2020b for an overview of mutation rate variation and a discussion of its genetic determinants, and the reviews by Ritz et al. 2017; Stapley et al. 2017; Johnston 2024 for an overview of recombination rate variation). Importantly, the uneven distribution of mutation and recombination rates across genomes can

Jeffrey D. Jensen and Susanne P. Pfeifer are co-corresponding authors and jointly supervised the project.

Vivak Soni and Cyril J. Versoza contributed equally to this study.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2026 The Author(s). *American Journal of Primatology* published by Wiley Periodicals LLC.

profoundly influence interactions amongst other evolutionary processes; for instance, heterogeneity in these underlying rates may substantially modify the effects of selection at linked sites, and thereby modify expectations of both levels and patterns of genetic variation (reviewed by Charlesworth and Jensen 2021, 2022). Additionally, relying on simplified, species-averaged rates of mutation and recombination—as is common practice in many evolutionary inference applications—has been shown to potentially lead to mis-inference in downstream analyses, including those estimating population history and the distribution of fitness effects (Dapper and Payseur 2018; Samuk and Noor 2022; Ghafoor et al. 2023; Soni et al. 2024; Soni and Jensen 2025). Yet, despite their crucial importance, both processes remain relatively poorly characterized in many vertebrates.

There are two primary classes of approach for estimating both mutation and recombination rates in primates and other large organisms. Direct approaches utilize genomic data from parent-offspring trios or multigeneration pedigrees in order to detect de novo mutations (reviewed by Pfeifer 2020b, and see also Pfeifer 2021), or contemporary crossover and noncrossover events occurring between generations (reviewed by Clark et al. 2010). The resolution of these pedigree-based inference approaches at the genome-scale remains relatively coarse due to the limited number of de novo mutations and meiotic exchanges that can be observed in the small number of generations characterizing the sample. Hence, given the hundreds to thousands of pedigreed individuals required to obtain high-resolution maps, most studies typically yield only a genome-wide estimate over the limited generational span studied, rather than providing detailed information about the finer scale landscapes. Moreover, due to the extensive sample and sequencing requirements, direct approaches tend to be both labor-intensive and costly, limiting their application, particularly in organisms for which resources remain limited. Indirect approaches, on the other hand, rely on population genetic theory and information about the genealogy of the sample over longer evolutionary timescales. Indirect estimates of mutation rate, for example, rely on species-level divergence data, based on the observation that the neutral mutation rate is equivalent to the neutral divergence rate (Kimura 1968). This allows for the inference of historically averaged mutation rates from phylogenetic sequence data in neutrally evolving genomic regions. Although in principle straightforward, this approach is limited by the availability of high-quality genome annotations necessary to identify neutrally evolving regions—which are lacking for many nonmodel organisms—and is often accompanied by significant uncertainties related to divergence and generation times, typically resulting in a range of possible mutation rates (see e.g., Soni, Versoza, Terbot, et al. 2025). Indirect recombination rate inference relies on polymorphism rather than divergence data, analyzing unrelated individuals in order to estimate historical sex-averaged rates of recombination based on the extent of observed linkage disequilibrium (LD) in the genome (reviewed by Stumpf and McVean 2003; Peñalba and Wolf 2020). For this reason, these methods are sensitive to other evolutionary forces shaping LD (Dapper and Payseur 2018; Samuk and Noor 2022) and it is thus crucial to account for the underlying population history when performing such analyses (Johri et al. 2020, 2022; Jensen 2023). Yet, despite these caveats, these approaches are also uniquely capable of

providing a fine-scale, genome-wide mapping of underlying sex-averaged rates over many generations necessary for evolutionary analyses (see the discussion in Johri et al. 2022).

Within primates specifically, high-quality, fine-scale rate estimation has generally been focused upon humans and their closest relatives, the great apes, as well as a handful of species of biomedical or conservation interest (e.g., Kong et al. 2002; Auton et al. 2012; Stevison et al. 2016; Pfeifer 2020a; Xue et al. 2020; Wall et al. 2022; Versoza et al. 2024, 2026; Versoza, Ehmke, et al. 2025; Versoza, Lloret-Villas, et al. 2025; Soni, Versoza, Terbot, et al. 2025; Spatola et al. 2026; and see discussion of Soni, Pfeifer, et al. 2025). In the primates studied to date, as well as in numerous other organisms, meiotic recombination has been found to be concentrated in hotspots, the location of which is primarily determined by the zinc-finger protein PRDM9 (Baudat et al. 2010; Myers et al. 2010; Parvanov et al. 2010). Notably, PRDM9 has evolved rapidly across primates, showing a high variability in the number of zinc-fingers as well as their nucleotide contact residues, resulting in differences in the predicted nucleotide binding sequence and, consequently, in the hotspot positioning even between closely related species (reviewed by Stapley et al. 2017; Lorenz and Mpaulo 2022; Johnston 2024; and see Schwartz et al. 2014 for a characterization of the allelic diversity in PRDM9 zinc finger domains of different primate species). As many evolutionary and biomedical studies require knowledge of the fine-scale patterns of recombination across the genome—such as rare disease gene mapping and genome-wide association studies that rely on patterns of LD to detect associations between genetic variants and phenotypic traits—and given the extensive variation in the distribution of recombination observed, it is thus important to obtain detailed maps characterizing the recombination frequency specific to the organism of interest. Yet, deeper in the primate clade, high-quality estimates of fine-scale mutation and recombination rates, or hotspot mapping, remain sparse—particularly across platyrrhines and prosimians (strepsirrhines)—limiting both our understanding of primate evolution in general and the accuracy and resolution of clinically relevant genomic studies in primate model systems specifically.

In order to extend this inference, we here examine the common (or white-tufted-ear) marmoset, *Callithrix jacchus*, a relatively abundant platyrrhine monkey of the Callitrichidae family. As one of the smallest statured anthropoids, common marmosets have risen in biomedical prominence over the past two decades particularly as a commonly used model for the study of aging, neurodegeneration, and neurodevelopment disorders, as well as autoimmune and infectious disease dynamics (see e.g., the reviews by Miller et al. 2016; Philippens and Langermans 2021; Han et al. 2022), owing largely to their early sexual maturity (~15–18 months of age), short gestation period (~145 days), and high fecundity (birthing up to four offspring in a single pregnancy, with pregnancies occurring biannually). Interestingly, a notable peculiarity of marmoset biology, and that of several other callitrichines, is the frequent ovulation of two (or more) ova per cycle (Tardif and Jaquish 1997), resulting in common dizygotic (fraternal) twin births (Ward et al. 2014), and the exchange of hematopoietic stem cells during embryonic development via anastomoses in a single shared placenta, causing sibling chimerism (Hill 1932; Wislocki 1939; Benirschke

et al. 1962; Gengozian et al. 1969). As a result, samples of blood and many other tissues obtained from marmosets have been observed to contain a mixture of genetic material originating from both the sampled individual and their unsampled littermate (Ross et al. 2007; Sweeney et al. 2012; del Rosario et al. 2024; and see the commentary by Chiou and Snyder-Mackler 2024). Notably, recent genomic analyses have demonstrated that the high frequency of twinning together with chimeric sampling shape both levels and patterns of observed genetic variation—including LD—in this species, with Soni, Versoza, Vallender, et al. (2025) finding that a neglect of these biological factors can result in significant mis-inference of population history. Yet, the impact of these unusual reproductive dynamics on other types of population genomic inference remains unexplored to date.

Although a direct mutation rate point estimate exists for common marmosets (0.43×10^{-8} per base pair per generation; Yang et al. 2021), this estimate was based on a single trio and, although the sampled tissues (muscle, liver, and spleen) are expected to exhibit lower levels of chimerism than blood, the authors did not correct for the effects of chimerism. In this instance, the potential “counting” of twin-shared variants in a single individual sampled from a triplet may be expected to confound mutation counts and thus calculated rates. To lower the potential impact of chimerism on their estimate, the authors applied highly stringent filter criteria; however, this in turn may be expected to impact the overall number of genuine de novo mutations identifiable in their study. More recently, Mao et al. (2024) inferred a much-increased human–marmoset neutral divergence relative to the closely related owl monkey–human divergence. In this case, the long-term neutral divergence rate would not be expected to be impacted by twinning and chimeric sampling (i.e., neutral divergence will be driven by the neutral mutation rate, μ). In contrast to these studies of mutation, to the best of our knowledge, no fine-scale genetic map yet exists for the species. Helpfully, utilizing novel population genomic data from individuals sequenced to high-coverage and annotated at the gene-level—and directly accounting for twinning and chimerism—Soni, Versoza, Vallender, et al. (2025) recently described a well-fitting population history consisting of a rapid reduction in population size roughly 7000 years ago, followed by a modest recovery—a necessary prerequisite to interpret patterns of genetic variation and LD. Here, we utilize this modeling framework accounting both for this unique reproductive biology as well as population history, combined with high-quality whole-genome sequencing data, in order to quantify the rates and characterize the landscape of mutation and recombination in the common marmoset. Beyond providing valuable biological insights into the location and frequency of mutation and recombination in this platyrrhine, these fine-scale maps will play a crucial role in enhancing future biomedical and evolutionary analyses integrating the significant nuance of accurate patterns of underlying genomic rate heterogeneity, thus improving our ability to study heritable genetic disorders in this widely used model system. In this effort, we have also uniquely characterized the effects of twinning and chimeric sampling on indirect rate estimation more generally and describe the underlying impact on patterns of LD—insights which will be beneficial to future research focused on organisms characterized by chimerism.

2 | Materials and Methods

2.1 | Animal Subjects

Animals were maintained in accordance with the guidelines of the Harvard Medical School Standing Committee on Animals and the Guide for Care and Use of Laboratory Animals of the Institute of Laboratory Animal Resources, National Research Council, as well as the American Society of Primatologists Principles for the Ethical Treatment of Nonhuman Primates. All samples were collected during routine veterinary care under approved protocols at the New England Primate Research Center (NEPRC Base Colony Protocol # 8601361-R98; Animal Welfare Assurance # A3431-01).

2.2 | Whole Genome Sequencing

We sequenced the genomes of 15 captive common marmosets (*C. jacchus*) from twin litters of the colony previously housed at the New England Primate Research Center using DNA derived from archived whole blood samples. For each DNA sample, we generated 150 bp paired-end reads from a PCR-free library sequenced on the DNBseq platform at the Beijing Genomics Institute (BGI Group, Shenzhen, China), targeting a genome-wide average coverage of $\sim 35\times$. Importantly, genetic analysis confirmed that the selected individuals were largely unrelated, sharing no more than 1/16 of their DNA with any other individual from the colony included in this study. This partly owes to the fact that these samples originated from a DNA bank including individuals sampled during earlier phases of the colony. Additionally, colony pedigree records indicate that these individuals were the product of twin births. Helpfully, recent work studying the evolutionary dynamics on the common marmoset X chromosome in these same individuals (Soni et al. 2026) was able to utilize expected X/Y ratios in order to further assess chimeric twin contributions. Namely, for individuals with different-sex twins one may expect an observed X/Y ratio of ~ 0.75 when sequencing either twin, if blood samples indeed represent roughly evenly mixed contributions from both siblings (i.e., with 2 X chromosomes contributed from the female twin, and 1 X and 1 Y chromosome contributed from the male twin). As shown in Table 1 of Soni et al. (2026), this is indeed near the empirical observation when sampling individuals with different-sex twins (and, also as expected, same-sex female twins were found to have a ratio of ~ 1 and same-sex male twins were found to have a ratio of ~ 0.5 , with some inherent sampling variance during sequencing).

2.3 | Read Mapping, Variant Calling, and Filtering

Following best practices in the field (Pfeifer 2017), we trimmed adapters, polyX tails, and low-quality ends from the raw reads using SOAPnuke v.1.5.6 with BGI-recommended parameter settings (i.e., “-n 0.01 -l 20 -q 0.3 -A 0.25 --cutAdaptor -Q 2 -G --polyX --minLen 150”; Chen et al. 2018) before mapping them to the species-specific reference genome (mCalJa1.2.pat.X; GenBank Accession Number: GCA_011100555.2; Yang et al. 2021) using BWA-MEM v.0.7.17 (Li and Durbin 2009). To avoid potential biases during variant calling and genotyping, we marked duplicates

TABLE 1 | Inferred rates of neutral mutation.

	Divergence time	Generation time	
		1.5 years	2.0 years
<i>C. jacchus</i> – <i>C. kuhlii</i>	0.59 mya	2.50E – 09	3.34E – 09
	0.82 mya	1.80E – 09	2.40E – 09
	1.09 mya	1.36E – 09	1.81E – 09
<i>C. jacchus</i> – <i>H. sapiens</i>	32 mya	2.80E – 09	3.74E – 09
	33 mya	2.72E – 09	3.62E – 09
	36 mya	2.49E – 09	3.32E – 09

Note: Comparison of indirectly inferred mean neutral mutation rate estimates based on *C. jacchus*–*C. kuhlii* divergence (mean neutral divergence rate: 0.0029) and *C. jacchus*–*H. sapiens* divergence (0.06) for three possible divergence times (0.59 million years ago [mya], 0.82 mya, and 1.09 mya for *C. jacchus*–*C. kuhlii* [Malukiewicz et al. 2021]; 32 mya, 33 mya, and 36 mya for *C. jacchus*–*H. sapiens* [Glazko 2003]) and two possible generation times (1.5 years and 2.0 years; Tardif et al. 2003; Okano et al. 2012; Schultz-Darken et al. 2016; Han et al. 2022).

in the mapped reads using the Genome Analysis Toolkit (GATK) v.4.2.6.1 *MarkDuplicates* (van der Auwera and O'Connor 2020). From these high-quality, duplicate-marked reads (“--minimum-mapping-quality 40”), we called variants for each individual using the GATK *HaplotypeCaller* in the “--pcr-indel-model NONE” mode as our data were obtained from PCR-free libraries. We then combined individual-level variant call sets and jointly genotyped them using GATK’s *CombineGVCFs* and *GenotypeGVCFs*, respectively. In order to obtain information regarding the genomic regions accessible to our study, we emitted reference confidence scores at each locus (“--emit-ref-confidence BP_RESOLUTION”) and included genotypes of both variant and invariant loci in the output (“--include-non-variant-sites”). Lastly, we separated the jointly genotyped call set into biallelic single-nucleotide polymorphisms (SNPs) and invariant loci with genotype information available in all individuals (“AN=30”), with downstream analyses focusing on the autosomes (i.e., chromosomes 1–22) only.

Due to the lack of a well-curated, experimentally validated data set that could be harnessed for filtering variants via a machine-learning framework, we used the GATK *VariantFiltration* tool to quality-control the discovered loci following the developers’ hard filtering recommendations (i.e., $QD < 2.0$, $QUAL < 30.0$, $MQ < 40.0$, $FS > 60.0$, $SOR > 3.0$, $MQRankSum < -12.5$, $ReadPosRankSum < -8.0$). As spurious variation can complicate the estimation of recombination rates from patterns of LD observed in population genomic data, we applied several additional stringent filtering criteria to limit the number of false positives in our data set, following the methodologies established in prior research (e.g., Auton et al. 2012; Stevison et al. 2016; Pfeifer 2020a; Soni, Versoza, Terbot, et al. 2025). In brief, as a skewed read coverage often results in poorly supported genotypes, we first used GATK’s *SelectVariants* to remove loci located within genomic regions exhibiting less than half, or more than twice, the individual’s genome-wide average coverage. Second, we excluded variants that were tightly clustered (using GATK’s *VariantFiltration* with “--cluster-size 3” and “--cluster-window-size 10”) as well as those exhibiting high levels of heterozygosity (using VCFtools v.0.1.14 to filter out variants that have a Hardy–Weinberg equilibrium exact test $p < 0.01$; Danecek et al. 2011), a common indicator of local genome mis-assembly. Lastly, we used BEDTools v.2.30.0 (Quinlan 2014) to eliminate regions blacklisted by the ENCODE Project Consortium due to their poor performance in next-generation sequencing (wgEncodeDacMapabilityConsensusExcludable.bed and wgEncodeDukeMapabilityRegionsExcludable.bed; Amemiya et al. 2019) by

converting genomic coordinates between the human (hg38) and the common marmoset genome assemblies using the UCSC liftOver tool (Raney et al. 2024).

After filtering, the population-level data set contained 7,198,428 autosomal biallelic SNPs in the accessible genome (Table S1).

2.4 | Phasing

We estimated haplotypes from the population-level genotype data using BEAGLE v.5.5 (Browning et al. 2021), a progressive phasing algorithm that has been shown to be highly accurate (Williams et al. 2012). During phasing, a total of 16,701 loci (0.23%) became nonpolymorphic due to a change in genotype and were thus excluded from further analyses.

2.5 | Inferring Fine-Scale Rates of Neutral Divergence and Mutation

In order to infer fine-scale rates of neutral divergence and mutation, we first used Cactus v.2.9.2 (Armstrong et al. 2020) to extract the 239 primate genomes (Kuderna et al. 2024) from the 447-way multiple species alignment (Zoonomia Consortium 2020) and update the common marmoset genome with the most recent reference genome for the species utilized in this study. To this end, we removed the common marmoset genome included in the multiple species alignment using the *halRemoveGenome* function and then extracted the alignment block consisting of the ancestral PrimateAnc232 genome and the genome of the closely related Wied’s black-tufted-ear marmoset (*Callithrix kuhlii*) using the *hal2fasta* function. We then realigned the current reference genome for the common marmoset with the genomes of PrimateAnc232 and *C. kuhlii* using the previously inferred branch lengths and updated the alignment using the *halReplaceGenome* function. With this updated multispecies alignment on hand, we identified fixed single nucleotide differences along the marmoset lineage (i.e., between the common marmoset and the ancestral primate PrimateAnc232 genomes) using the *halSummarizeMutations* function. In contrast to the common marmoset genome, the genome of Wied’s marmoset is highly fragmented (293,512 scaffolds with a scaffold N50 of 15.1 kb totaling 2.6 Gb compared to 1233 scaffolds with a scaffold N50 of 137 Mb totaling 2.9 Gb), thus we removed

alignments less than 10 kb in length in order to avoid spurious and/or incomplete alignments that might artificially inflate estimates of neutral divergence and mutation. Additionally, in order to obtain fixed differences along the human-marmoset branch, we obtained the subalignment consisting of the common marmoset, human, and *PrimatesAnc003* genomes from the multispecies alignment using the *cactus-hal2maf* function, converted this alignment back into HAL format (Hickey et al. 2013) using the *maf2hal* function, and then retrieved point mutations along the branch from *C. jacchus* to the ancestral *PrimatesAnc003* using the *halBranchMutations* function. In order to obtain neutral substitutions between *C. jacchus* and *C. kuhlii* as well as between *C. jacchus* and humans, we excluded both variants known to segregate in any of the species (based on published population-level polymorphism data available for humans of Yoruban ancestry [1000 Genomes Project Consortium 2015] and common marmosets [Soni, Versoza, Vallender, et al. 2025]; note that no such population-level polymorphism data were available for Wied's marmoset) as well as sites within 10 kb of functional regions (based on the protein-coding genes annotated in the common marmoset genome). Though neutral divergence (and thus neutral mutation rate inference) is unaffected by selection at linked sites (Birky and Walsh 1988), previous work has shown that 10 kb is a sufficient distance from functional regions to avoid the effects of selection at linked sites in shaping levels and patterns of variation in primate-like genomes (Johri et al. 2020, 2021, 2023). Using these data sets, we then calculated neutral divergence at both the broad (genome-wide) scale and the fine scale (using 1 kb, 10 kb, 100 kb, and 1 Mb sliding windows with a step size of 500 bp, 5 kb, 50 kb, and 500 kb, respectively) by dividing the number of substitutions by the number of sites accessible to our study. To derive mutation rates, we divided the rates of neutral divergence between *C. jacchus* and *C. kuhlii* by the divergence time in generations. For this, we used three possible divergence time estimates—0.59 million years ago (mya), 0.82 mya, and 1.09 mya (based on the range of divergence times between *C. jacchus* and *C. kuhlii* inferred by Malukiewicz et al. 2021)—and two possible generation time estimates—1.5 years and 2.0 years (Tardif et al. 2003; Okano et al. 2012; Schultz-Darken et al. 2016; Han et al. 2022).

2.6 | Inferring Fine-Scale Rates of Recombination

We used two different approaches to infer fine-scale rates of recombination: the demography-unaware estimator LDhat (McVean et al. 2002, 2004; Auton and McVean 2007) and the demography-aware estimator pyrro (Spence and Song 2019).

- **LDhat:** We inferred genome-wide fine-scale recombination rates in the common marmoset using LDhat v.2.2 (McVean et al. 2002, 2004; Auton and McVean 2007). To this end, we first generated a lookup table containing the coalescent likelihoods for every two-locus haplotype configuration possible in our sample of 15 diploid individuals (i.e., 30 haplotypes: $-n$ 30), using LDhat *complete* with the suggested maximum population-scaled recombination rate ρ of 100 ($-rhomax$ 100) and a grid size of 201 ($-n_pts$ 201) to improve accuracy. Next, we obtained region-based population-scaled estimates by running the *interval* function of LDhat with a block penalty of 5 ($-bpen$ 5) for 60 million iterations ($-its$ 60000000) using a

sampling scheme of 40,000 iterations ($-samp$ 40000) across windows of 4000 SNPs with a 200 SNP step size. After discarding the burn-in of the Monte Carlo Markov Chain using LDhat's *stat* function ($-burn$ 500), we then combined the region-based estimates at the midpoint of each overlapping window to obtain estimates at the chromosome-scale. Following standard practices in the field (e.g., Auton et al. 2012; Pfeifer 2020a), we excluded large localized peaks in recombination rate (with $\rho > 100$ between adjacent SNPs) to minimize the impact of genome assembly errors leading to artificial breaks in LD. In total, we identified 1352 such regions and masked them together with the 50 SNPs adjacent on each side by setting the recombination rate to 0 (masking a total of 42,348 SNPs, or 0.59%). Finally, we used the effective population size (N_e) based on the mean nucleotide diversity (θ_w) observed in the empirical data to convert the population-scaled recombination rate to a per-generation recombination rate, assuming a per-site per-generation mutation rate of 0.81×10^{-8} as per Soni, Versoza, Vallender, et al. (2025).

- **pyrho:** We also inferred genome-wide recombination rates using pyrro v.0.1.7 (Spence and Song 2019). In brief, as recommended by the developers, we first used pyrro's *make_table* function to compute an approximate likelihood lookup table ($--approx$) for a 50% larger sample size ($-N$ 45) under the demographic model previously inferred by Soni, Versoza, Vallender, et al. (2025) to account for historical population size changes in the species (i.e., a population size reduction approximately 3500 generations ago, followed by an exponential population recovery) and then downsampled this table to 15 diploid individuals ($-n$ 30) to match our empirical sample size. Next, we used the *hyperparam* function to determine suitable hyperparameter settings for pyrro and then ran the *optimize* function with the recommended window size ($--windowsize$ 30) and smoothness penalty ($--blockpenalty$ 50) to estimate genome-wide recombination rates. In all steps, we assumed a per-site per-generation mutation rate of 0.81×10^{-8} ($--mu$ 0.81e-8) as per Soni, Versoza, Vallender, et al. (2025) to internally convert the population-scaled recombination rate to a per-generation recombination rate.

To account for the impact of twinning and chimerism, we re-scaled the LDhat and pyrro estimates by multiplying the rates by a factor of 0.574 and 0.761, respectively (i.e., the misinference caused by twinning and chimerism under the population-specific demographic history as quantified via simulation; see Section 2.7).

2.7 | Evaluating the Impact of Twinning and Chimerism on Recombination Rate Inference

Both LDhat and pyrro rely on coalescent theory and the theoretical foundation provided by the Wright–Fisher (WF) model to infer fine-scale recombination rates from population genomic data. To evaluate the impact of twinning and chimerism—two model violations inherent to the biology of marmosets—on the recombination rate inference with these two approaches, we used SLiM 4.0.1 (Haller and Messer 2023) to simulate a population of marmoset individuals using the modeling framework

recently described in Soni, Versoza, Vallender, et al. (2025). Specifically, we simulated 10 replicates of a 1 Mb genomic region in a population of marmosets under the demographic model of the species recently inferred by Soni, Versoza, Vallender, et al. (2025)—consisting of an ancestral population of 61,198 individuals that collapsed to 17,931 individuals 3513 generations ago before recovering via exponential growth to a current day size of 33,830 individuals—assuming a mutation rate of 0.81×10^{-8} per base pair per generation and a recombination rate of 1.0×10^{-8} per base pair per generation (as per the rates used in Soni, Versoza, Vallender, et al. 2025). From each replicate, we sampled 15 chimeric individuals and estimated recombination rates using LDhat and pyrho (as described in Section 2.6).

2.8 | Inferring Recombination Hotspots

To infer the positions of recombination hotspots, we used the software package LDhot v.0.4 (Auton et al. 2014). In brief, we first ran the *ldhot* function to identify putative recombination hotspots in 3 kb windows with a step size of 1 kb and a background window of 100 kb centered around the hotspot, using 1000 simulations as recommended by the developers. Afterward, we combined significant windows to merge adjacent candidates using the *ldhot_summary* function with significance thresholds of 0.001 and 0.01 for calling and merging hotspots, respectively. To limit the number of spurious hotspots, we implemented a three-step filtering approach by combining the recommendations from the Great Ape Recombination Project (Stevenson et al. 2016) with those of Brazier and Glémin (2024), filtering out any hotspot candidates with a width longer than 10 kb, an intensity lower than 4 or higher than 200, or a rate less than five times the chromosome-wide average rate.

2.9 | In Silico Prediction of PRDM9 Binding Motifs

We used the ZOOPS model implemented in MEME v.5.5.7 (Bailey and Elkan 1994) to identify 10–15 bp motifs present not more than once in each of the 1000 hotspots with the highest intensity (including the flanking 500 bp to ensure that the entire hotspot is captured) while accounting for genomic background using cold spot regions matched for sequence length and GC-content. With these putative motifs on hand, we used FIMO v.5.5.7 to scan the complete set of hotspot windows for any occurrences and compared the frequency of each motif in the hotspot regions with that observed in 25,000 randomly sampled background regions using MOODS v.1.9.4.1 (Korhonen et al. 2009). We assessed the statistical significance by performing a Fisher's exact test in R v.4.2.2 (R Core Team 2022).

2.10 | In Silico Characterization of PRDM9 Binding

We used AlphaFold3 to predict the sequence-specific binding between PRDM9 and the putative PRDM9 binding motifs. To this end, we first retrieved the nucleotide sequence of PRDM9 from the common marmoset genome and aligned it against the PRDM9 protein sequence annotated in the human telomere-to-telomere assembly (T2T-CHM13v2.0; Nurk et al. 2022) using

GeneWise v.2.4.1 (Birney et al. 2004) to visually inspect the sequence for completeness. Next, we used the Expasy web server (Duvaud et al. 2021) to translate the nucleotide sequence into an amino acid sequence. We then used this translated amino acid sequence as input for a protein–protein BLAST (Altschul et al. 1990) search against the NCBI nonredundant protein sequences database (Figure S1) and visualized the resulting phylogenetic tree using EMBL's interactive Tree of Life (Letunic and Bork 2021), noting a high similarity of the query to PRDM9 sequences previously annotated in haplorrhines (Figure S2). Next, we used InterPro (Blum et al. 2025) to predict protein domains within this sequence, confirming the presence of a Krueppel-associated box (KRAB) domain, a SSX repressor domain (SSXRD), a PR/SET domain, and a C2H2-type zinc finger array (Figure S3). Finally, we ran AlphaFold3 (Abramson et al. 2024) by providing both the marmoset PRDM9 amino acid sequence and the putative PRDM9 binding motifs as input.

2.11 | Assessing the Correlation of Fine-Scale Rates of Recombination With Genomic Features

To assess the correlation of fine-scale rates of recombination with different genomic features, we calculated a number of summary statistics across 100 kb windows along the 22 autosomes of the common marmoset genome—including nucleotide diversity (based on our marmoset population genomic data), divergence (based on the updated 447-way mammalian multispecies alignment as described in Section 2.5), GC-content and exon-content (both based on the annotations of the common marmoset assembly; Yang et al. 2021)—and calculated partial Kendall's rank correlations across windows in which at least 50% of sites were accessible using R v.4.2.2 (R Core Team 2022).

3 | Results and Discussion

3.1 | Population Genomic Data

In order to infer fine-scale rates of neutral divergence, mutation, and recombination, we sequenced the genomes of 15 captive common marmoset (*C. jacchus*) individuals (7 females and 8 males) to an average 35-fold coverage. Using a mapping-based approach, we identified ~7.2 million SNPs across the autosomal genome with a transition-transversion ratio of 2.2 (Table S1). To facilitate rate inference, we estimated haplotypes from this population-level genotype data analogously to the 1000 Genomes (1000 Genomes Project Consortium 2015), PanMap (Auton et al. 2012), and Great Ape Recombination Maps (Stevenson et al. 2016) projects, which previously generated fine-scale genetic maps for humans (*Homo sapiens*), Western and Nigerian chimpanzees (*Pan troglodytes verus* and *P. t. ellioti*), bonobos (*P. paniscus*), and Western gorillas (*Gorilla gorilla gorilla*).

3.2 | The Landscape of Neutral Divergence and Mutation in the Common Marmoset Genome

We first extracted the 239 primate genomes (Kuderna et al. 2024) from the 447-way multispecies alignment (Zoonomia Consortium 2020) and updated the common marmoset genome to the current reference assembly available for the species (Yang et al. 2021). With this updated whole-genome alignment on

hand, we identified fixed single nucleotide differences along the marmoset lineage, that is, between the genome of the common marmoset and the genome of the closely related Wied's black-tufted-ear marmoset (*C. kuhlii*) (see Section 2). Using this data set, we then calculated neutral divergence across accessible sites at both the broad (genome-wide) scale and the fine scale (for additional details, see Section 2).

At the 1 Mb-scale, we observed a neutral divergence rate of 9.85×10^{-4} along the marmoset lineage relative to the reconstructed ancestor (see Figure S4 for the distributions of neutral divergence across a range of window sizes). To calculate the neutral mutation rate, we then drew the point estimate, upper and lower bounds of common marmoset divergence times relative to *C. kuhlii* (0.59 mya, 0.82 mya, and 1.09 mya; Malukiewicz et al. 2021), and generation times of 1.5 and 2.0 years (Tardif et al. 2003; Okano et al. 2012; Schultz-Darken et al. 2016; Han et al. 2022). Depending on the underlying assumptions, the mean neutral mutation rate varied from 0.14×10^{-8} mutations per base pair per generation (under a divergence time of 1.09 mya and a generation time of 1.5 years) to 0.33×10^{-8} mutations per base pair per generation (under a divergence time of 0.59 mya and a generation time of 2.0 years; Table 1 and see Figure 1a for density plots of neutral mutation rate estimates across this range of possible divergence and generation times and Figure 1b for the inferred genome-wide neutral mutation rates). Notably, these indirectly inferred neutral mutation rates are lower than the direct estimate of 0.43×10^{-8} mutations per base pair per generation obtained by

Yang et al. (2021) from a single parent-offspring trio. While indirect phylogenetic approaches are naturally unable to observe strongly deleterious/lethal mutations that are purged from the population, the higher pedigree-based estimate observed by Yang et al. (2021) might also partially be attributed to unaccounted for chimerism present in their genomic data (with nongermline tissues sampled for their study).

Given that the divergence time between *C. jacchus* and *C. kuhlii* is relatively short, we also generated fine-scale divergence estimates along the human-marmoset branch to obtain information over longer evolutionary timescales. Notably, comparisons between the mutation rate estimates indirectly obtained from the *C. jacchus*–*C. kuhlii* divergence and the neutral divergence estimates based on *C. jacchus*–*H. sapiens* alignments demonstrated a significant positive correlation at the fine scale ($r = 0.235$, $p = 1.71 \times 10^{-59}$; and see Figure S5 for the heterogeneity in neutral divergence and mutation rates across each autosome), providing additional confidence in our estimates. Moreover, mutation rates inferred using the divergence along the *C. jacchus*–*H. sapiens* branch are highly similar to those obtained using *C. jacchus*–*C. kuhlii* divergence, ranging from 0.25×10^{-8} mutations per base pair per generation (under a divergence time of 36 mya and a generation time of 1.5 years) to 0.37×10^{-8} mutations per base pair per generation (under a divergence time of 32 mya and a generation time of 2.0 years; Table 1). Although it is unlikely that the generation time of common marmosets has remained constant over the > 30 million years separating the two species (Glazko 2003), it is nevertheless

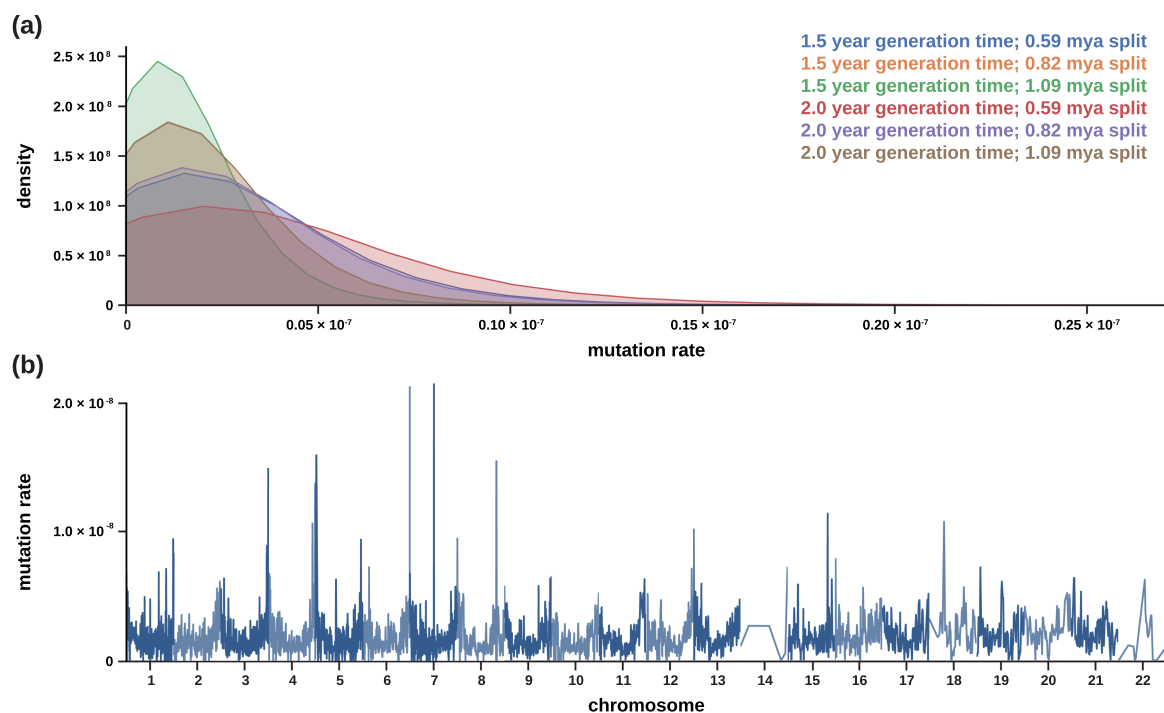


FIGURE 1 | Fine-scale rates of neutral mutation. (a) Density plots of the per-site per-generation mutation rate implied by the neutral divergence for two possible generation times (1.5 years and 2.0 years; Tardif et al. 2003; Okano et al. 2012; Schultz-Darken et al. 2016; Han et al. 2022) and three possible divergence times between the common marmoset (*C. jacchus*) and the closely related Wied's black-tufted-ear marmoset (*C. kuhlii*) (0.59 million years ago [mya], 0.82 mya, and 1.09 mya; Malukiewicz et al. 2021). (b) Genome-wide per-site per-generation neutral mutation rates for genomic windows of size 1 Mb, with a 500 kb step size (and see Figure S5 for the heterogeneity in neutral mutation rates across all autosomes). Neutral mutation rates were estimated from the rates of neutral divergence observed in > 10 kb-long alignments between *C. jacchus* and *C. kuhlii* (note that due to the limited number of such alignments on chromosomes 14 and 22, neutral mutation rate estimates for these two autosomes are relatively coarse).

encouraging that the marmoset-based and human–marmoset-based estimates are in such close correspondence.

Taking the opposite approach, we inferred marmoset divergence times based on the mean neutral divergence rate observed in the empirical data and the previously published pedigree-based mutation rate estimates for both common marmosets (0.43×10^{-8} mutations per base pair per generation; Yang et al. 2021) as well as the closely related (but nonchimeric) owl monkeys (0.81×10^{-8} mutations per base pair per generation; Thomas et al. 2018) for which a larger number of trios were available. The estimated divergence times ranged from 0.18 mya (under a per-site per-generation mutation rate of 0.81×10^{-8} and a generation time of 1.5 years) to 0.49 mya (under a per-site per-generation mutation rate of 0.43×10^{-8} and a generation time of 2.0 years) (Table S2). Notably, these whole-genome divergence times based on the marmoset and owl monkey mutation rates are considerably lower than those inferred by Malukiewicz et al. (2021) from mitochondrial DNA, due to the mutation rates being considerably higher than those inferred in this study.

3.3 | The Landscape of Recombination in the Common Marmoset Genome

We used two different approaches to infer fine-scale rates of recombination—the demography-unaware estimator LDhat (McVean et al. 2002, 2004; Auton and McVean 2007) and the demography-aware estimator pyrro (Spence and Song 2019)—both of which rely on coalescent theory and the theoretical foundation provided by the WF model to infer fine-scale recombination rates from population genomic data.

Unlike most primates, twinning and chimerism are the norm rather than the exception in marmosets (Hill 1932; Wislocki 1939; Benirschke et al. 1962; Ward et al. 2014). To evaluate the impact of these two violations of the WF model inherent to the biology of marmosets, we first assessed the performance of the two recombination rate estimators on simulated data. Using the framework recently described in Soni, Versoza, Vallender, et al. (2025), we modeled twinning and chimerism from hematopoietic stem cells (as observed in blood samples) by first simulating a non-WF model in SLiM (Haller and Messer 2023) in which pairs of marmoset individuals reproduce each generation to give birth to nonidentical twins and then combining their genotypes post-simulation to mimic a single chimeric individual (for additional details, see Section 2). More specifically, we simulated genomic regions of 1 Mb under the demographic model of the population inferred by Soni, Versoza, Vallender, et al. (2025)—consisting of an ancestral population of ~60 k individuals that experienced a population decline before recovering to about half of its original size—using a non-WF model with twinning only as well as with both twinning and chimeric sampling, assuming a constant per-site per-generation mutation rate of 0.81×10^{-8} and recombination rate of 1 cM/Mb (as per the rates used in Soni, Versoza, Vallender, et al. 2025), and sampling 15 chimeric individuals to match our empirical data. To aid the interpretation of the inference results, we additionally performed simulations under a standard WF model (i.e., without twinning or chimerism) for comparison. We then

inferred the genome-wide recombination rates in the simulated data using LDhat and pyrro. In contrast to pyrro, LDhat outputs the population-scaled recombination rate (ρ) and we thus used the effective population size (N_e) based on the mean nucleotide diversity (θ_w) observed in each simulation scenario to obtain a per-generation recombination rate.

Our simulations highlighted that both LDhat and pyrro tend to underestimate genome-wide recombination rates inferred under the marmoset demographic model (Figure 2a), likely due to the recent population contraction and subsequent exponential expansion having resulted in a ρ at the time of sampling that is substantially different from that during most of the population history. This observation is consistent with the recent findings of a simulation study conducted by Dutheil (2024) which highlighted that classical LD-based approaches tend to underestimate recombination rates in populations characterized by population size declines and recent growth, particularly when sample sizes are moderate as is the case here (see Figures 1 and 2 in Dutheil 2024; and see their Figure 4 for the increased mis-inference in the presence of gene conversion). In contrast, both approaches overestimate the per-generation recombination rate when twinning is common, as populations in which nonsingleton births are the predominant mode of reproduction are characterized by both an increased LD and decreased N_e relative to the standard expectations. Chimerism, on the other hand, has the opposite effect, with chimeric sampling expected to both break up haplotypes—as evidenced by the reduced mean r^2 (a measure of LD)—and increase haplotype diversity due to the generation of novel haplotypes relative to the standard expectations (Figure 2b).

Re-scaling the empirically observed recombination rate estimates to take into account the mis-inference caused by twinning and chimerism yielded an average sex-averaged genome-wide recombination rate of 0.91 ± 0.83 cM/Mb (100 kb windows) in common marmosets (Figure 3a), with sex-averaged crossover rates ranging from 0.65 cM/Mb for one of the longest autosomes (chromosome 4) to 1.29 cM/Mb for one of the shortest autosomes (chromosome 22) (Figure S6). Although the rates inferred here for common marmosets are in the same range as those previously reported in the great apes (e.g., 1.32 ± 1.40 cM/Mb in humans [International HapMap Consortium 2007], with an average rate of 0.945 cM/Mb in males and 1.518 cM/Mb in females [Halldorsson et al. 2019], and ~1.19 cM/Mb in chimpanzees, bonobos and gorillas [Stevenson et al. 2016]), rates previously estimated for biomedically relevant Cercopithecidae are substantially lower (e.g., 0.43 ± 0.33 cM/Mb in rhesus macaques [Xue et al. 2016] and 0.43 ± 0.44 cM/Mb in vervet monkeys [Pfeifer 2020a]) — however, it should be noted that many of the earlier studies did not account for the potentially confounding effects of population demography during inference (see discussions of Dapper and Payseur 2018; Johri et al. 2022), thus hindering the interpretation of the biological differences in observed rates.

In addition to the nearly two-fold variation between autosomes, the recombination landscape in common marmosets—like that of many other primates (e.g., Spencer et al. 2006; Auton et al. 2012; Pfeifer and Jensen 2016; Stevenson et al. 2016; Pfeifer 2020a; Wall et al. 2022; Versoza et al. 2024; Soni, Versoza,

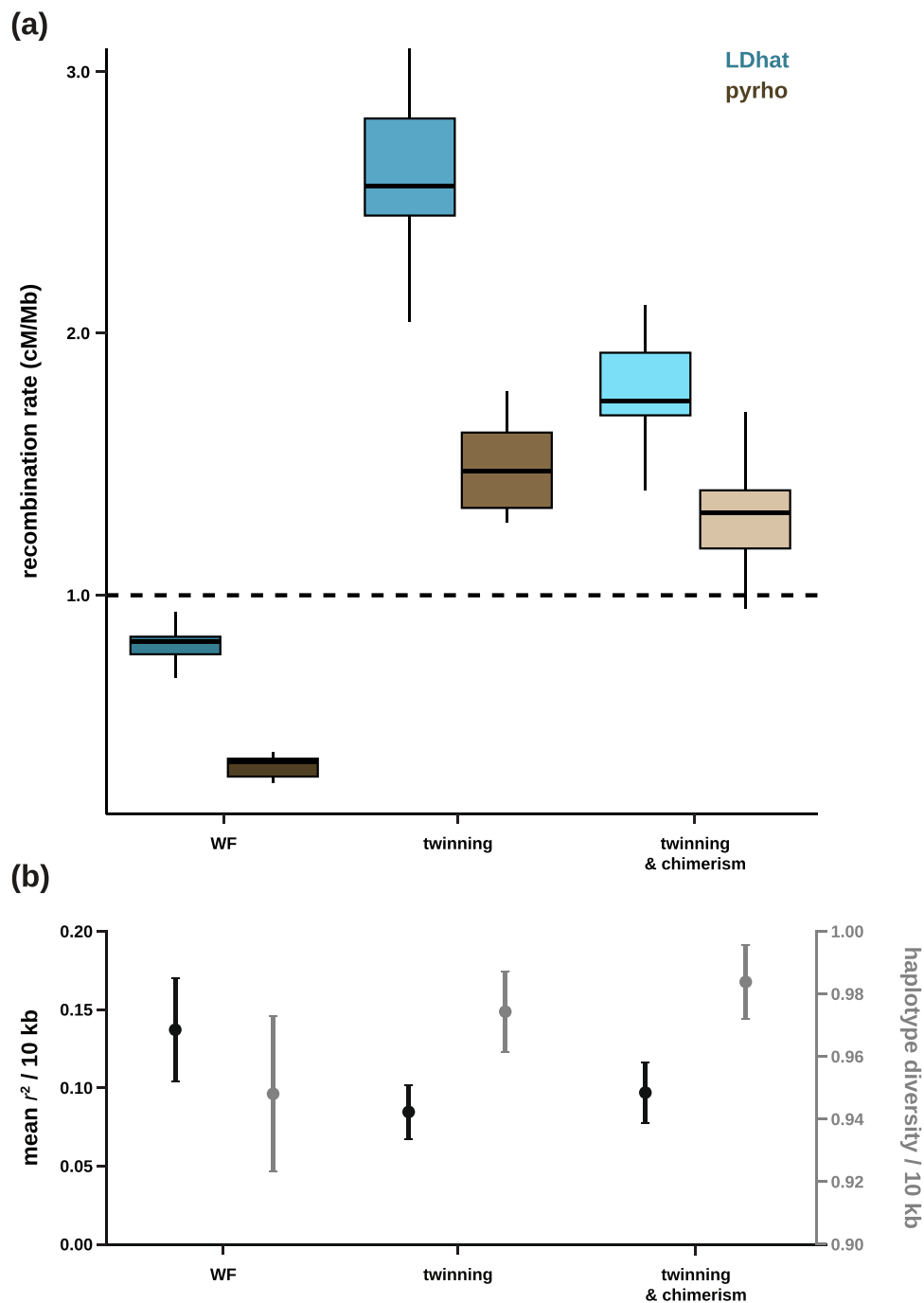


FIGURE 2 | Performance of recombination rate estimators under twinning and chimerism. (a) Performance of the demography-unaware recombination rate estimator LDhat (shown in blues) and the demography-aware recombination rate estimator pyrho (shown in browns) under the demographic history inferred for the population by Soni, Versoza, Vallender, et al. (2025) using a Wright–Fisher (WF) model and a non-WF framework that models twinning and chimerism—two model violations inherent to the biology of marmosets. The dashed line depicts the recombination rate that was used in the simulations (1 cM/Mb). (b) Mean r^2 (shown in black; left y-axis) and haplotype diversity (gray; right y-axis) calculated over 10 kb windows (with a 5 kb step size) across 10 simulation replicates under the demographic history inferred for the population by Soni, Versoza, Vallender, et al. (2025) using a WF model and a non-WF framework that models twinning and chimerism. Data points represent the mean value, whilst confidence intervals represent the standard deviation.

Terbot, et al. 2025; Spatola et al. 2026)—is highly heterogenous, with a pronounced enrichment in the distal (sub)telomeric regions and reduced rates within centromeric and pericentromeric regions (Figure 3a and see Figure S7 for recombination rate heterogeneity across each individual autosome). Moreover, genome-wide recombination rates exhibit a strong positive

association with nucleotide diversity ($r_t = 0.49$, $p = 0$)—highlighting the likely importance of selection at linked sites in shaping the landscape of genomic variation in the species (Soni, Versoza, Pfeifer, et al. 2025)—and, to a lesser extent, divergence ($r_t = 0.06$, $p = 2.07E - 17$) and GC-content ($r_t = 0.03$, $p = 6.90E - 06$), as may be expected if recombination in common

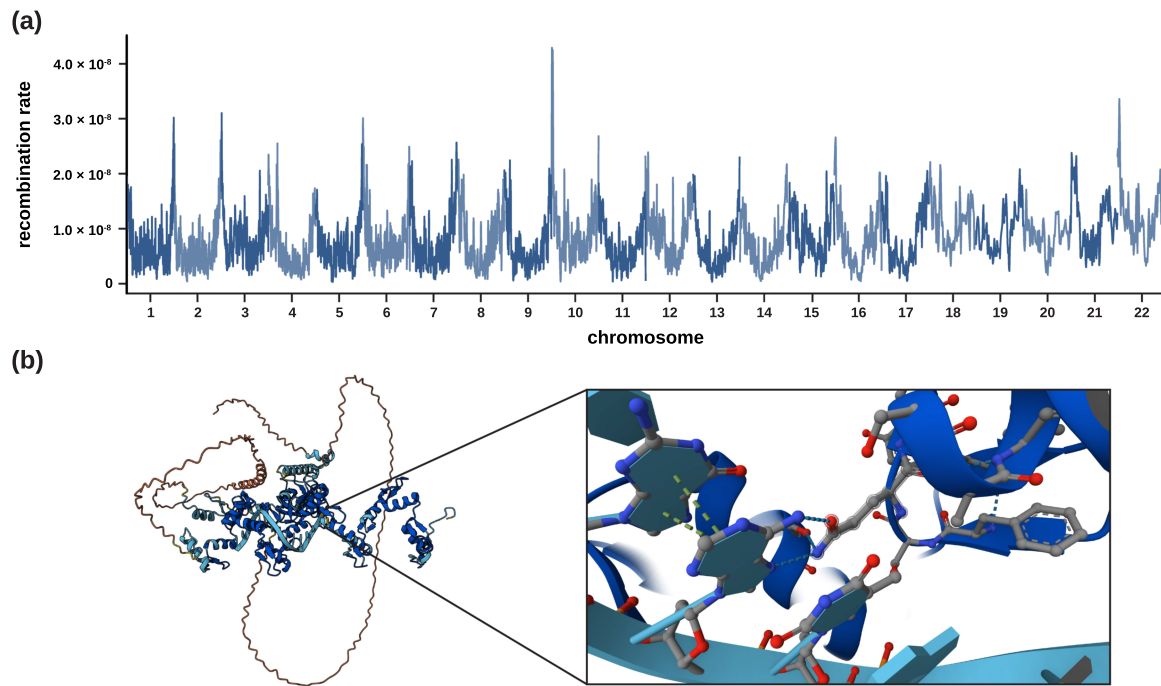


FIGURE 3 | Fine-scale rates of recombination. (a) Genome-wide per-site per-generation recombination rates as inferred by LDhat for genomic windows of size 1 Mb, with a 500 kb step size (and see Figure S7 for recombination rate heterogeneity across all autosomes and Figure S10 for the genome-wide recombination rates inferred by pyrho). (b) In silico binding prediction between PRDM9 and the putative binding sites. Predictions were performed using AlphaFold3 (predictions are provided for noncommercial use only, under and subject to AlphaFold Server Output Terms of Use found under alphafoldserver.com/output-terms). Per-atom confidence scores are color-coded with very high confidence predicted structures shown in dark blue (pLDDT > 90), confidently predicted structures in turquoise (90 > pLDDT > 70), low confidence predicted structures in yellow (70 > pLDDT > 50), and very low confidence predicted structures in orange (pLDDT < 50).

marmosets exhibits mutagenic effects (Halldorsson et al. 2019; Hinch et al. 2023) and is prone to GC-biased gene conversion (reviewed by Duret and Galtier 2009) (Figure S8). In concordance with PRDM9 directing recombination away from functional elements (Brick et al. 2012; Pratto et al. 2014), recombination rates significantly decrease within, and around the boundaries of, protein-coding genes ($r_r = -0.03$, $p = 1.09E - 04$) compared to neighboring regions (Figure S9). Notably, while these inferred patterns of recombination were highly similar between the two approaches ($r = 0.57$, $p < 2.2E - 16$ at the 1 kb-scale), the mean inferred rates of pyrho were much lower (0.25 ± 0.27 cM/Mb) as, despite accounting for historical fluctuations in population size, its internal conversion to the per-generation recombination rate does not account for the lower N_e caused by twinning (Figure S10).

3.4 | Characterization of Recombination Hotspots in the Common Marmoset Genome

To gain further insights into the landscape of fine-scale recombination, we inferred recombination hotspots using LDhot (Auton et al. 2014)—a statistical method implemented to directly process the recombination rate estimates obtained from LDhat. To obtain a high-quality data set, we implemented a three-step filtering approach by combining the recommendations from the Great Ape Recombination Project (Stevenson et al. 2016) with those of Brazier and Glémin (2024). First, as genuine recombination hotspots tend to be narrow (with ~70%–80% of recombination occurring in hotspots with a mean width of 2.3 kb in humans; 1000

Genomes Project Consortium 2010), we excluded hotspot candidates with a width larger than 10 kb. Second, as artificial breaks in LD resulting, for example, from mis-assembly and/or mis-alignment, can lead to localized peaks in recombination rate, we calculated the intensity of each hotspot candidate by dividing the recombination rate at its peak by the rate estimated for the 100 kb surrounding region and removed candidates with an intensity higher than 200. Additionally, due to the difficulty of distinguishing genuine low-intensity hotspots from background variation, we also applied a low-intensity filter, retaining only candidates with a minimum intensity of 4. Third, accounting for recombination rate variation across chromosomes to further narrow down the location of recombination hotspots, we split each remaining candidate region into windows of 2 kb with a step size of 1 kb, keeping only the hotspot windows exhibiting rates of at least five times the chromosome-wide average rate. This three-step filtering approach yielded 26,831 candidate hotspots—similar to the number of recombination hotspots initially identified in human populations (International HapMap Consortium 2005, 2007; Myers et al. 2005; 1000 Genomes Project Consortium 2010; Kong et al. 2002). We note, however, that a larger number of hotspots was identified in subsequent, larger-scale studies in humans (e.g., Halldorsson et al. 2019) and thus, it is expected that the number of hotspots identified here is likely an underestimate for the species given the available sample size.

In order to identify sequences that predict potential PRDM9 binding, we searched through the 1000 hotspots with the highest intensity for consistent sequence motifs. To this end, we used the ZOOPS model implemented in MEME (Bailey and

Elkan 1994) to identify degenerate motifs present not more than once in each hotspot region, while accounting for genomic background using cold spot regions matched for sequence length and GC-content. As sequence motifs underlying recombination hotspots in other primates tend to be short (e.g., 10–15 bp in humans, chimpanzees, and gorillas; Auton et al. 2012; Schwartz et al. 2014; Stevison et al. 2016), we limited the search to motifs with a minimum width of 10 bp and maximum width of 15 bp, keeping only motifs below a discovery threshold of $1e - 05$. Next, we used FIMO to scan the complete set of hotspot windows for occurrences of each of these motifs. As the predicted motifs are highly repetitive, we also performed permutation tests to compare the frequency of each motif in hotspot regions with that observed in 25,000 background regions randomly sampled from the common marmoset genome by matching the position weight matrix of each motif against the genomic sequences of each hotspot and cold spot region using the MOTif Occurrence Detection Suite (MOODS; Korhonen et al. 2009), requiring a $p < 0.01$ for a match (see Table S3 for the complete list of candidate motifs for the PRDM9 binding sequence). To further narrow down the list of candidates, we used AlphaFold3 (Abramson et al. 2024) to predict the sequence-specific binding between the marmoset PRDM9 sequence and the putative PRDM9 binding motifs. Interestingly, for the best candidate DNA binding motifs—the 15-mer GCTGGGATTACAGGC (e -value: $6.80E - 54$), present in 83.5% of the candidate hotspots and only 2.0% of the matching cold spots, and its degenerate generalization, the 15-mer GCTGGGAKYASWGGC (e -value: $5.40E - 17$)—AlphaFold3's machine-learning guided model predicted a protein-DNA complex with high-confidence (ipTM = 0.94 and 0.96, respectively; with 1.0 being the maximum value of the ipTM performance metric), indicating a very well-defined protein-DNA interface and supporting PRDM9-DNA binding in silico (see Figure 3b for a visualization of the predicted binding complex). Notably, while this predicted degenerate PRDM9 binding motif shares several key features with motifs previously observed in other primates (including the C/G-rich terminal regions and its overall length), there are also species-specific differences (such as the A/T-rich internal region) as might be anticipated from the rapid evolution of the PRDM9 zinc finger array across the primate lineage.

4 | Conclusion

By investigating, and ultimately correcting for, the effects of twinning and sibling chimerism on the inference of fine-scale mutation and recombination rate maps, we have here described these landscapes in one of the most commonly used nonhuman primate models in research, *C. jacchus*. Using high-quality population genomic data, we found that the species exhibits relatively low neutral mutation rates, and rates of recombination within the range observed amongst other anthropoids. Moreover, like many vertebrates, the recombination landscape in common marmosets is dominated by PRDM9-mediated hotspots, and we have described a putative 3D-structure of the species-specific PRDM9-DNA binding complex in silico.

While this resource was built from largely unrelated individuals previously housed at one of the more than 40 marmoset research colonies, and while there are likely fine-scale (e.g., hotspot) differences between these different populations/colonies, broad-scale

recombination rate patterns are likely conserved and consistent amongst different populations (as has been observed in other primates, including humans; e.g., Serre et al. 2005; Hinch et al. 2011). Importantly, there are also two distinct advantages provided in this data set owing to these colony records. First, pedigree records indicate that these individuals were themselves the product of twin births. While this observation cannot rule out an in utero reduction in litter size (i.e., an additional chimeric contribution from an unborn sibling may be present), it nonetheless provides an unusual amount of insight with regards to twin chimerism in confirming that these were indeed not singleton births. Second, while this is the first effort to our knowledge to model genome-wide fine-scale mutation and recombination rate maps whilst accounting for twinning and chimerism, we also have additional insight on these individuals specifically regarding the likely even chimeric mixing between twins in these blood samples. Specifically, in investigating X chromosome evolutionary dynamics utilizing these same samples, Soni et al. (2026) found that empirical observations closely matched expectations under a model in which blood between siblings is evenly mixed (i.e., a X/Y ratio around 0.75 for individuals with different-sex twins). While this sibling chimerism has been described for nearly a century (Hill 1932; Wislocki 1939), these observations hence further support the more recent results of del Rosario et al. (2024) in suggesting that an individual blood sample contains a roughly even mixture of genetic material from both twins. Thus, while an evenly mixed sibling chimeric contribution is supported for this data set, there is no a priori expectation that this ratio will be identical in other litter sizes, other tissues types, or in other chimeric species, and it would further be expected that alternative chimeric ratios would result in alternative levels and patterns of genomic variation. Importantly, however, the simulation and inference framework provided here can accommodate the modeling of any arbitrary chimeric ratio, and as such can be utilized in a variety of future applications across samples, tissues, and species. More generally, future empirical research as well as evolutionary modeling would benefit from a greater characterization of the underlying biological variance in twin contributions to blood samples as well as to other tissue types.

Apart from providing new insights into the population genetic processes shaping variation in common marmosets, and the novelty of extending these inference approaches for this relatively peculiar aspect of marmoset biology, these maps will also serve as a valuable resource for future studies in this biomedically important species—including in genome-wide association studies, polygenic risk score modeling, and genomic scans for targets of selection—with implications ranging from the improved study of neurodevelopment disorders to infectious disease dynamics. This is particularly the case as twinning and chimerism have been shown to modify expectations of LD and thus bias, for example, the inference of recombination rates if neglected, potentially producing genomic correlations that may be mistaken for the action of selection.

Author Contributions

Vivak Soni, Cyril J. Versoza, Jeffrey D. Jensen, and Susanne P. Pfeifer conceived of and conducted the study, interpreted results, and wrote the manuscript. Jeffrey D. Jensen and Susanne P. Pfeifer obtained research funding.

Acknowledgments

The authors would like to thank Eric J. Vallender for providing the marmoset samples used in this study, and the members of the Jensen and Pfeifer Labs for helpful comments and discussion. Library preparation and sequencing was performed at the Beijing Genomics Institute (BGI Group, Shenzhen, China). Computations were performed on the Sol supercomputer at Arizona State University (Jennewein et al. 2023). This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R35GM151008 to S.P.P. V.S. and J.D.J. were supported by National Institutes of Health Award Number R35GM139383 to J.D.J. C.J.V. was 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.

Ethics Statement

Animals were maintained in accordance with the guidelines of the Harvard Medical School Standing Committee on Animals and the Guide for Care and Use of Laboratory Animals of the Institute of Laboratory Animal Resources, National Research Council, as well as the American Society of Primatologists Principles for the Ethical Treatment of Nonhuman Primates. All samples were collected during routine veterinary care under approved protocols at the New England Primate Research Center (NEPRC Base Colony Protocol # 8601361-R98; Animal Welfare Assurance # A3431-01).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

All sequence data have been deposited under NCBI BioProject PRJNA1215741. The mutation and recombination rate maps generated in this study are available at https://github.com/vivaksoni/marmoset_mutation_recombination and <http://spfeiferlab.org/data>, respectively.

References

- 1000 Genomes Project Consortium. 2010. "A Map of Human Genome Variation From Population-Scale Sequencing." *Nature* 467, no. 7319: 1061–1073.
- 1000 Genomes Project Consortium. 2015. "A Global Reference for Human Genetic Variation." *Nature* 526, no. 7571: 68–74.
- Abramson, J., J. Adler, J. Dunger, et al. 2024. "Accurate Structure Prediction of Biomolecular Interactions With AlphaFold 3." *Nature* 630, no. 8016: 493–500.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. "Basic Local Alignment Search Tool." *Journal of Molecular Biology* 215, no. 3: 403–410.
- 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.
- Armstrong, J., G. Hickey, M. Diekhans, et al. 2020. "Progressive Cactus Is a Multiple-Genome Aligner for the Thousand-Genome Era." *Nature* 587, no. 7833: 246–251.
- Auton, A., A. Fedel-Alon, S. Pfeifer, et al. 2012. "A Fine-Scale Chimpanzee Genetic Map From Population Sequencing." *Science* 336, no. 6078: 193–198.
- Auton, A., and G. McVean. 2007. "Recombination Rate Estimation in the Presence of Hotspots." *Genome Research* 17, no. 8: 1219–1227.
- Auton, A., S. Myers, and G. McVean. 2014. "Identifying Recombination Hotspots Using Population Genetic Data." Preprint, arXiv. <https://doi.org/10.48550/arXiv.1403.4264>.
- Baer, C. F., M. M. Miyamoto, and D. R. Denver. 2007. "Mutation Rate Variation in Multicellular Eukaryotes: Causes and Consequences." *Nature Reviews Genetics* 8, no. 8: 619–631.
- Bailey, T. L., and C. Elkan. 1994. "Fitting a Mixture Model by Expectation Maximization to Discover Motifs in Biopolymers." In *Proceedings of the International Conference on Intelligent Systems for Molecular Biology*, Vol. 2: 28–36.
- Baudat, F., J. Buard, C. Grey, et al. 2010. "PRDM9 Is a Major Determinant of Meiotic Recombination Hotspots in Humans and Mice." *Science* 327, no. 5967: 836–840.
- Benirschke, K., J. M. Anderson, and L. E. Brownhill. 1962. "Marrow Chimerism in Marmosets." *Science* 138, no. 3539: 513–515.
- Blum, M., A. Andreeva, L. C. Florentino, et al. 2025. "InterPro: The Protein Sequence Classification Resource in 2025." *Nucleic Acids Research* 53, no. D1: D444–D456.
- Birky, C. W., and J. B. Walsh. 1988. "Effects of Linkage on Rates of Molecular Evolution." *Proceedings of the National Academy of Sciences of the United States of America* 85, no. 17: 6414–6418.
- Birney, E., M. Clamp, and R. Durbin. 2004. "GeneWise and Genome-wide." *Genome Research* 14, no. 5: 988–995.
- Brazier, T., and S. Glémin. 2024. "Diversity in Recombination Hotspot Characteristics and Gene Structure Shape Fine-Scale Recombination Patterns in Plant Genomes." *Molecular Biology and Evolution* 41, no. 9: msae183.
- Brick, K., F. Smagulova, P. Khil, R. D. Camerini-Otero, and G. V. Petukhova. 2012. "Genetic Recombination Is Directed Away From Functional Genomic Elements in Mice." *Nature* 485, no. 7400: 642–645.
- 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.
- 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: 177–197.
- Charlesworth, B., and J. D. Jensen. 2022. "How Can We Resolve Lewontin's Paradox?" *Genome Biology and Evolution* 14, no. 7: evac096.
- Chen, Y., Y. Chen, C. Shi, et al. 2018. "SOAPnuke: A MapReduce Acceleration-Supported Software for Integrated Quality Control and Preprocessing of High-Throughput Sequencing Data." *GigaScience* 7, no. 1: 1–6.
- Chiou, K., and N. Snyder-Mackler. 2024. "Marmosets Contain Multitudes." *eLife* 13: e97866.
- Clark, A. G., X. Wang, and T. Matise. 2010. "Contrasting Methods of Quantifying Fine Structure of Human Recombination." *Annual Review of Genomics and Human Genetics* 11: 45–64.
- Danecek, P., A. Auton, G. Abecasis, et al. 2011. "The Variant Call Format and VCFtools." *Bioinformatics* 27, no. 15: 2156–2158.
- 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.
- del Rosario, R. C. H., F. M. Krienen, Q. Zhang, et al. 2024. "Sibling Chimerism Among Microglia in Marmosets." *eLife* 13: RP93640.
- Duret, L., and N. Galtier. 2009. "Biased Gene Conversion and the Evolution of Mammalian Genomic Landscapes." *Annual Review of Genomics and Human Genetics* 10: 285–311.
- Dutheil, J. Y. 2024. "On the Estimation of Genome-Average Recombination Rates." *Genetics* 227, no. 2: iyae051.
- Duvaud, S., C. Gabella, F. Lisacek, H. Stockinger, V. Ioannidis, and C. Durinx. 2021. "Expasy, the Swiss Bioinformatics Resource Portal, as Designed by Its Users." *Nucleic Acids Research* 49, no. W1: W216–W227.

- Ghafoor, S., J. Santos, C. J. Versoza, J. D. Jensen, and S. P. Pfeifer. 2023. "The Impact of Sample Size and Population History on Observed Mutational Spectra: A Case Study in Human and Chimpanzee Populations." *Genome Biology and Evolution* 15, no. 3: evad019.
- Glazko, G. V. 2003. "Estimation of Divergence Times for Major Lineages of Primate Species." *Molecular Biology and Evolution* 20, no. 3: 424–434.
- Gengozian, N., J. S. Batson, C. T. Greene, and D. G. Gosslee. 1969. "Hemopoietic Chimerism in Imported and Laboratory-Bred Marmosets." *Transplantation* 8, no. 5: 633–652.
- Halldorsson, B. V., G. Palsson, O. A. Stefansson, et al. 2019. "Characterizing Mutagenic Effects of Recombination Through a Sequence-Level Genetic Map." *Science* 363, no. 6425: eaau1043.
- Haller, B. C., and P. W. Messer. 2023. "SLiM 4: Multispecies Eco-Evolutionary Modeling." *American Naturalist* 201: E127–E139.
- Han, H. J., S. J. Powers, and K. L. Gabrielson. 2022. "The Common Marmoset – Biomedical Research Animal Model Applications and Common Spontaneous Diseases." *Toxicologic Pathology* 50, no. 5: 628–637.
- Hickey, G., B. Paten, D. Earl, D. Zerbino, and D. Haussler. 2013. "HAL: A Hierarchical Format for Storing and Analyzing Multiple Genome Alignments." *Bioinformatics* 29, no. 10: 1341–1342.
- Hill, J. P. 1932. "II. Croonian Lecture – The Developmental History of the Primates." *Philosophical Transactions of the Royal Society of London. Series B* 221, no. 474–482: 45–178.
- Hinch, R., P. Donnelly, and A. G. Hinch. 2023. "Meiotic DNA Breaks Drive Multifaceted Mutagenesis in the Human Germ Line." *Science* 382, no. 6674: eadh2531.
- Hinch, A. G., A. Tandon, N. Patterson, et al. 2011. "The Landscape of Recombination in African Americans." *Nature* 476, no. 7359: 170–175.
- Hodgkinson, A., and A. Eyre-Walker. 2011. "Variation in the Mutation Rate Across Mammalian Genomes." *Nature Reviews Genetics* 12, no. 11: 756–766.
- International HapMap Consortium. 2005. "A Haplotype Map of the Human Genome." *Nature* 437, no. 7063: 1299–1320.
- International HapMap Consortium. 2007. "A Second Generation Human Haplotype Map of Over 3.1 Million SNPs." *Nature* 449, no. 7164: 851–861.
- Jennewein, D. M., J. Lee, C. Kurtz, et al. 2023. "The Sol Supercomputer at Arizona State University." In *Practice and Experience in Advanced Research Computing 2023: Computing for the Common Good (PEARC '23)*, 296–301. Association for Computing Machinery.
- Jensen, J. D. 2023. "Population Genetic Concerns Related to the Interpretation of Empirical Outliers and the Neglect of Common Evolutionary Processes." *Heredity* 130, no. 3: 109–110.
- 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.
- Johri, P., C. F. Aquadro, M. Beaumont, et al. 2022. "Recommendations for Improving Statistical Inference in Population Genomics." *PLoS Biology* 20, no. 5: e3001669.
- 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.
- 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.
- 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.
- Kimura, M. 1968. "Evolutionary Rate at the Molecular Level." *Nature* 217, no. 5129: 624–626.
- Kong, A., D. F. Gudbjartsson, J. Sainz, et al. 2002. "A High-Resolution Recombination Map of the Human Genome." *Nature Genetics* 31, no. 3: 241–247.
- Korhonen, J., P. Martinmäki, C. Pizzi, P. Rastas, and E. Ukkonen. 2009. "MOODS: Fast Search for Position Weight Matrix Matches in DNA Sequences." *Bioinformatics* 25, no. 23: 3181–3182.
- 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.
- Letunic, I., and P. Bork. 2021. "Interactive Tree Of Life (iTOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation." *Nucleic Acids Research* 49, no. W1: W293–W296.
- Li, H., and R. Durbin. 2009. "Fast and Accurate Short Read Alignment With Burrows-Wheeler Transform." *Bioinformatics* 25, no. 14: 1754–1760.
- Lorenz, A., and S. J. Mpaulo. 2022. "Gene Conversion: A Non-Mendelian Process Integral to Meiotic Recombination." *Heredity* 129, no. 1: 56–63.
- Lynch, M. 2010. "Evolution of the Mutation Rate." *Trends in Genetics* 26, no. 8: 345–352.
- Malukiewicz, J., R. A. Cartwright, N. H. A. Curi, et al. 2021. "Mitogenomic Phylogeny of *Callithrix* With Special Focus on Human Transferred Taxa." *BMC Genomics* 22, no. 1: 239.
- Mao, Y., W. T. Harvey, D. Porubsky, et al. 2024. "Structurally Divergent and Recurrently Mutated Regions of Primate Genomes." *Cell* 187, no. 6: 1547–1562.e13.
- McVean, G., P. Awadalla, and P. Fearnhead. 2002. "A Coalescent-Based Method for Detecting and Estimating Recombination From Gene Sequences." *Genetics* 160, no. 3: 1231–1241.
- McVean, G. A. T., S. R. Myers, S. Hunt, P. Deloukas, D. R. Bentley, and P. Donnelly. 2004. "The Fine-Scale Structure of Recombination Rate Variation in the Human Genome." *Science* 304, no. 5670: 581–584.
- Miller, C. T., W. A. Freiwald, D. A. Leopold, J. F. Mitchell, A. C. Silva, and X. Wang. 2016. "Marmosets: A Neuroscientific Model of Human Social Behavior." *Neuron* 90, no. 2: 219–233.
- Myers, S., L. Bottolo, C. Freeman, G. McVean, and P. Donnelly. 2005. "A Fine-Scale Map of Recombination Rates and Hotspots Across the Human Genome." *Science* 310, no. 5746: 321–324.
- Myers, S., R. Bowden, A. Tumian, et al. 2010. "Drive Against Hotspot Motifs in Primates Implicates the PRDM9 Gene in Meiotic Recombination." *Science* 327, no. 5967: 876–879.
- Nurk, S., S. Koren, A. Rhie, et al. 2022. "The Complete Sequence of a Human Genome." *Science* 376, no. 6588: 44–53.
- Okano, H., K. Hikishima, A. Iriki, and E. Sasaki. 2012. "The Common Marmoset as a Novel Animal Model System for Biomedical and Neuroscience Research Applications." *Seminars in Fetal and Neonatal Medicine* 17, no. 6: 336–340.
- Parvanov, E. D., P. M. Petkov, and K. Paigen. 2010. "PRDM9 Controls Activation of Mammalian Recombination Hotspots." *Science* 327, no. 5967: 835.
- Peñalba, J. V., and J. B. W. Wolf. 2020. "From Molecules to Populations: Appreciating and Estimating Recombination Rate Variation." *Nature Reviews Genetics* 21, no. 8: 476–492.
- Pfeifer, S. P. 2017. "From Next-Generation Resequencing Reads to a High-Quality Variant Data Set." *Heredity* 118, no. 2: 111–124.
- Pfeifer, S. P. 2020a. "A Fine-Scale Genetic Map for Vervet Monkeys." *Molecular Biology and Evolution* 37, no. 7: 1855–1865.

- 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.
- Pfeifer, S. P. 2021. "Studying Mutation Rate Evolution in Primates—The Effects of Computational Pipelines and Parameter Choices." *GigaScience* 10, no. 10: giab069.
- Pfeifer, S. P., and J. D. Jensen. 2016. "The Impact of Linked Selection in Chimpanzees: A Comparative Study." *Genome Biology and Evolution* 8, no. 10: 3202–3208.
- Philippens, I. H. C. H. M., and J. A. M. Langermans. 2021. "Preclinical Marmoset Model for Targeting Chronic Inflammation as a Strategy to Prevent Alzheimer's Disease." *Vaccines* 9, no. 4: 388.
- Pratto, F., K. Brick, P. Khil, F. Smagulova, G. V. Petukhova, and R. D. Camerini-Otero. 2014. "Recombination Initiation Maps of Individual Human Genomes." *Science* 346, no. 6211: 1256442.
- Quinlan, A. R. 2014. "BEDTools: The Swiss-Army Tool for Genome Feature Analysis." *Current Protocols in Bioinformatics* 47, no. 11.12: 1–34.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Raney, B. J., G. P. Barber, A. Benet-Pagès, et al. 2024. "The UCSC Genome Browser Database: 2024 Update." *Nucleic Acids Research* 52, no. D1: D1082–D1088.
- Ritz, K. R., M. A. F. Noor, and N. D. Singh. 2017. "Variation in Recombination Rate: Adaptive or Not?." *Trends in Genetics* 33, no. 5: 364–374.
- Ross, C. N., J. A. French, and G. Ortí. 2007. "Germ-Line Chimerism and Paternal Care in Marmosets (*Callithrix kuhlii*)." In *Proceedings of the National Academy of Sciences of the United States of America* 104, no. 15, 6278–6282.
- Samuk, K., and M. A. F. Noor. 2022. "Gene Flow Biases Population Genetic Inference of Recombination Rate." *G3: Genes|Genomes|Genetics* 12, no. 11: jkac236.
- Schultz-Darken, N., K. M. Braun, and M. E. Emborg. 2016. "Neurobehavioral Development of Common Marmoset Monkeys." *Developmental Psychobiology* 58, no. 2: 141–158.
- Schwartz, J. J., D. J. Roach, J. H. Thomas, and J. Shendure. 2014. "Primate Evolution of the Recombination Regulator PRDM9." *Nature Communications* 5: 4370.
- Serre, D., R. Nadon, and T. J. Hudson. 2005. "Large-Scale Recombination Rate Patterns Are Conserved Among Human Populations." *Genome Research* 15, no. 11: 1547–1552.
- Soni, V., and J. D. Jensen. 2025. "Inferring Demographic and Selective Histories From Population Genomic Data Using a Two-Step Approach in Species With Coding-Sparse Genomes: An Application to Human Data." *G3* 15: jkaf019.
- Soni, V., S. P. Pfeifer, and J. D. Jensen. 2024. "The Effect of Mutation and Recombination Rate Heterogeneity on the Inference of Demography and the Distribution of Fitness Effects." *Genome Biology and Evolution* 16, no. 2: evae004.
- Soni, V., S. P. Pfeifer, and J. D. Jensen. 2025. "Recent Insights Into the Evolutionary Genomics of the Critically Endangered Aye-Aye (*Daubentonia madagascariensis*)." *American Journal of Primatology* 87, no. 12: e70105.
- Soni, V., C. J. Versoza, S. P. Pfeifer, and J. D. Jensen. 2025. "Investigating the Effects of Chimerism on the Inference of Selection: Quantifying Genomic Targets of Purifying, Positive, and Balancing Selection in Common Marmosets (*Callithrix jacchus*)." *Heredity* 134, no. 10–11: 645–657.
- Soni, V., C. J. Versoza, D. Shah, S. P. Pfeifer, and J. D. Jensen. 2026. "Inferring the Relative Contributions of Evolutionary Processes Shaping X Chromosome Dynamics in the Common Marmoset (*Callithrix jacchus*) in the Presence of Twinning and Hematopoietic Chimerism." *Genome Biology and Evolution* 18, no. 8: evag200.
- Soni, V., C. J. Versoza, J. W. Terbot, J. D. Jensen, and S. P. Pfeifer. 2025. "Inferring Fine-Scale Mutation and Recombination Rate Maps in Aye-Ayes (*Daubentonia madagascariensis*)." *Ecology and Evolution* 15, no. 11: e72314.
- Soni, V., C. J. Versoza, E. J. Vallender, J. D. Jensen, and S. P. Pfeifer. 2025. "Accounting for Chimerism in Demographic Inference: Reconstructing the History of Common Marmosets (*Callithrix jacchus*) From High-Quality, Whole-Genome, Population-Level Data." *Molecular Biology and Evolution* 42, no. 6: msaf119.
- Spatola, G. J., C. J. Versoza, V. Soni, E. J. Heenkenda, J. D. Jensen, and S. P. Pfeifer. 2026. "Comparing Fine-Scale Mutation and Recombination Landscapes in Rhesus Macaque (*Macaca mulatta*) Populations of Chinese and Indian Descent Inferred From Both Short- and Long-Read Sequencing Data." Preprint, BioRxiv.
- 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.
- Spencer, C. C. A., P. Deloukas, S. Hunt, et al. 2006. "The Influence of Recombination on Human Genetic Diversity." *PLoS Genetics* 2, no. 9: e148.
- 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.
- Stumpf, M. P. H., and G. A. T. McVean. 2003. "Estimating Recombination Rates From Population-Genetic Data." *Nature Reviews Genetics* 4, no. 12: 959–968.
- Sweeney, C. G., E. Curran, S. V. Westmoreland, K. G. Mansfield, and E. J. Vallender. 2012. "Quantitative Molecular Assessment of Chimerism Across Tissues in Marmosets and Tamarins." *BMC Genomics* 13: 98.
- Tardif, S. D., and C. E. Jaquish. 1997. "Number of Ovulations in the Marmoset Monkey (*Callithrix jacchus*): Relation to Body Weight, Age and Repeatability." *American Journal of Primatology* 42, no. 4: 323–329.
- Tardif, S. D., D. A. Smucny, D. H. Abbott, K. Mansfield, N. Schultz-Darken, and M. E. Yamamoto. 2003. "Reproduction in Captive Common Marmosets (*Callithrix jacchus*)." *Comparative Medicine* 53, no. 4: 364–368.
- Thomas, G. W. C., R. J. Wang, A. Puri, et al. 2018. "Reproductive Longevity Predicts Mutation Rates in Primates." *Current Biology* 28, no. 19: 3193–3197.e5.
- van der Auwera, G., and B. O'Connor. 2020. *Genomics in the Cloud*. 1st ed. O'Reilly Media.
- Versoza, C. J., K. L. Bales, J. D. Jensen, and S. P. Pfeifer. 2026. "Characterization of *De Novo* Germline Mutations Suggests a Strong Male Mutation Bias in Coppery Titi Monkeys (*Plecturocebus cupreus*)." *Molecular Biology and Evolution*.
- Versoza, C. J., E. E. Ehmke, J. D. Jensen, and S. P. Pfeifer. 2025. "Characterizing the Rates and Patterns of *De Novo* Germline Mutations in the Aye-Aye (*Daubentonia madagascariensis*)." *Molecular Biology and Evolution* 42, no. 3: msaf034.
- 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, no. 5: 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 Non-Crossover Events in Rhesus Macaques (*Macaca mulatta*)."
Genome Biology and Evolution 16, no. 1: 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.
- Ward, J. M., A. M. Buslov, and E. J. Vallender. 2014. "Twinning and Survivorship of Captive Common Marmosets (*Callithrix jacchus*) and Cotton-Top Tamarins (*Saguinus oedipus*)."
Journal of the American Association for Laboratory Animal Science: JAALAS 53, no. 1: 7–11.
- Williams, A. L., N. Patterson, J. Glessner, H. Hakonarson, and D. Reich. 2012. "Phasing of Many Thousands of Genotyped Samples."
American Journal of Human Genetics 91, no. 2: 238–251.
- Wislocki, G. B. 1939. "Observations on Twinning in Marmosets."
American Journal of Anatomy 64: 445–483.
- 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.
- 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.
- Yang, C., Y. Zhou, S. Marcus, et al. 2021. "Evolutionary and Biomedical Insights From a Marmoset Diploid Genome Assembly."
Nature 594, no. 7862: 227–233.
- Zoonomia Consortium. 2020. "Comparative Genomics Multitool for Scientific Discovery and Conservation."
Nature 587, no. 7833: 240–245.

Supporting Information

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

Supporting File: ajp70208-sup-0001-SUPP.pdf.