

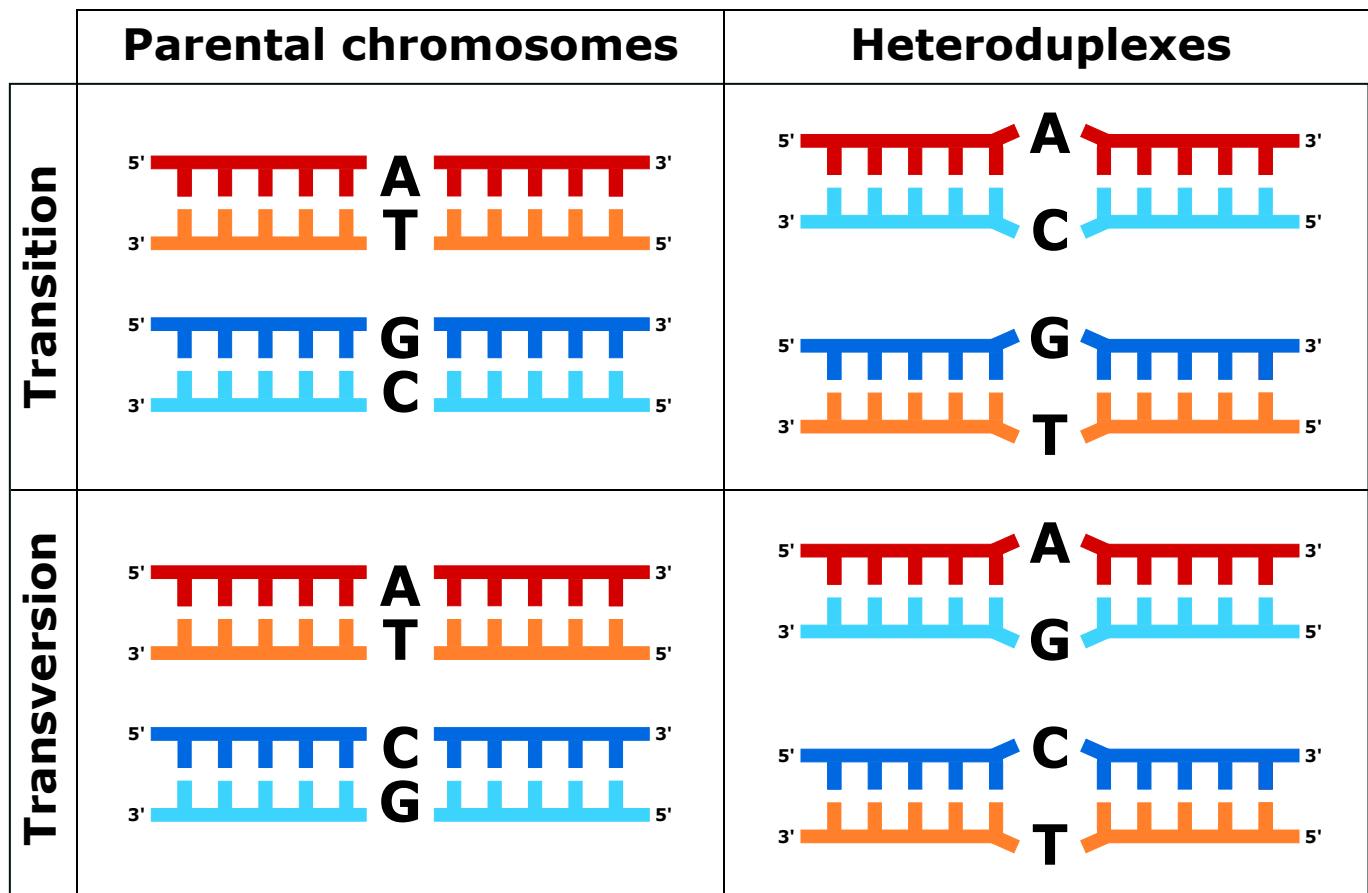
Supplemental Material

Supplementary table S1: Number of autosomal regions and sites that passed filtering criteria and are used in regression analyses of Tables 1 and 2.

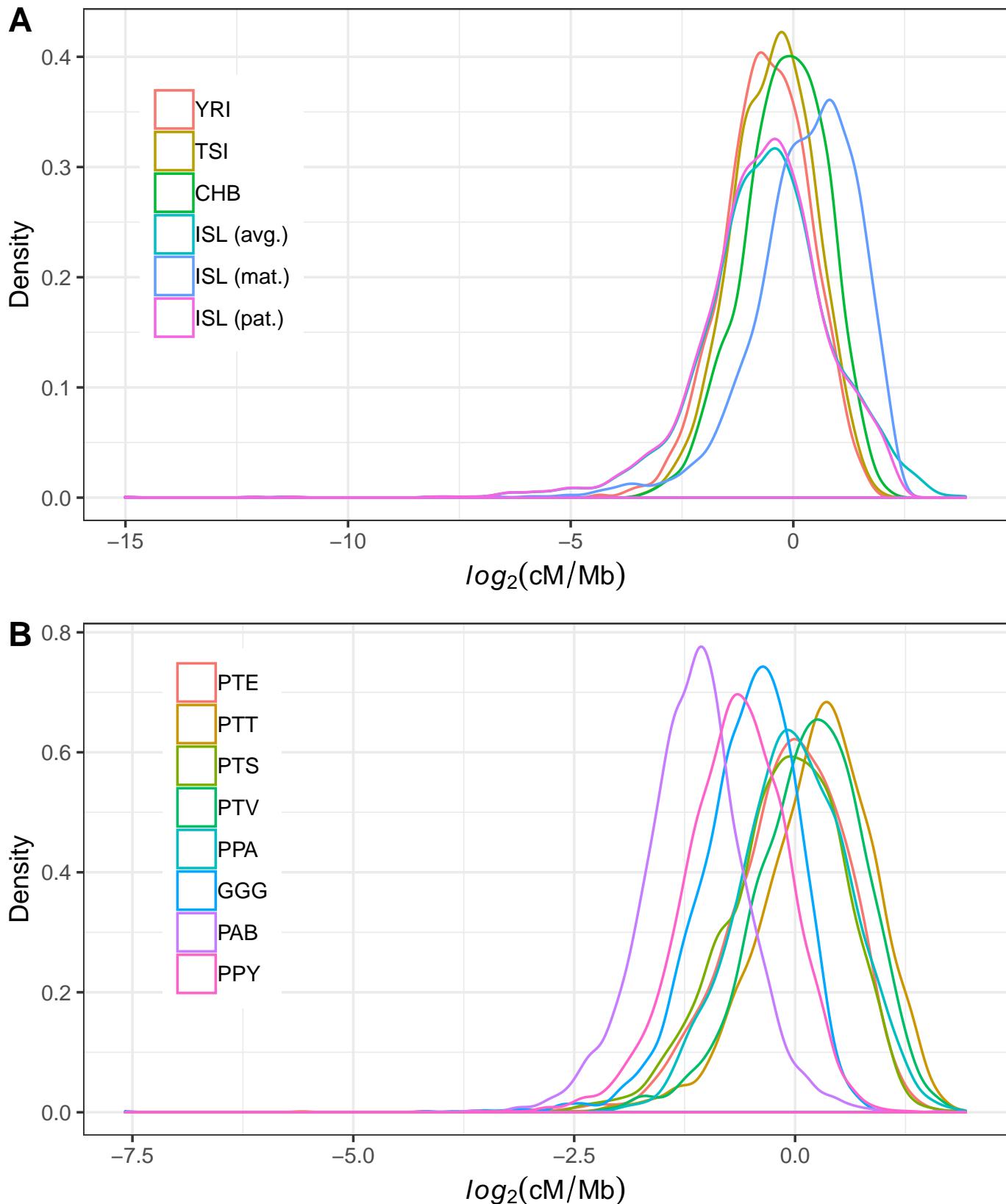
Population	#Regions	Sites			
		#All	#TS (CpG-)	#TS (CpG+)	#TV (CpG-)
YRI	2,666	7,934,810	4,215,989	2,204,730	1,264,783
TSI	2,534	4,629,559	2,463,756	1,292,461	727,313
CHB	2,408	3,890,522	2,041,254	1,103,720	620,772
ISL	2,529	4,635,328	2,505,744	1,256,398	725,925

Supplementary table S2: Count and mean/median GC frequency of bi-allelic transitions and transversion, and tri-allelic sites segregating at non-CpG and CpG sites in the African YRI population.

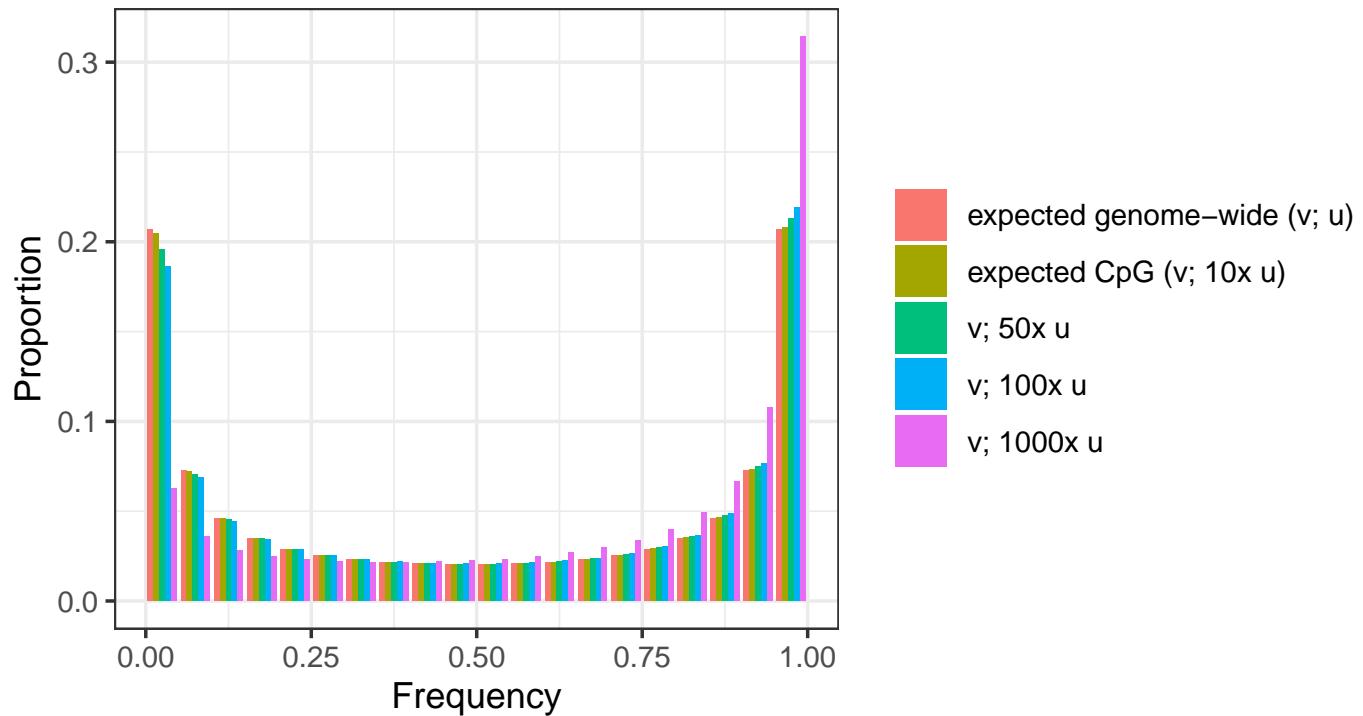
	#CpG-	#CpG+	mean/median GC (CpG-)	mean/median GC (CpG+)
TS	4,231,805	2,212,779	54.90%/67.59%	60.01%/79.17%
TV	1,269,834	250,334	58.21%/75.93%	40.27%/21.76%
Tri-allelic	11,264	18,168	62.16%/80.56%	72.96%/89.81%



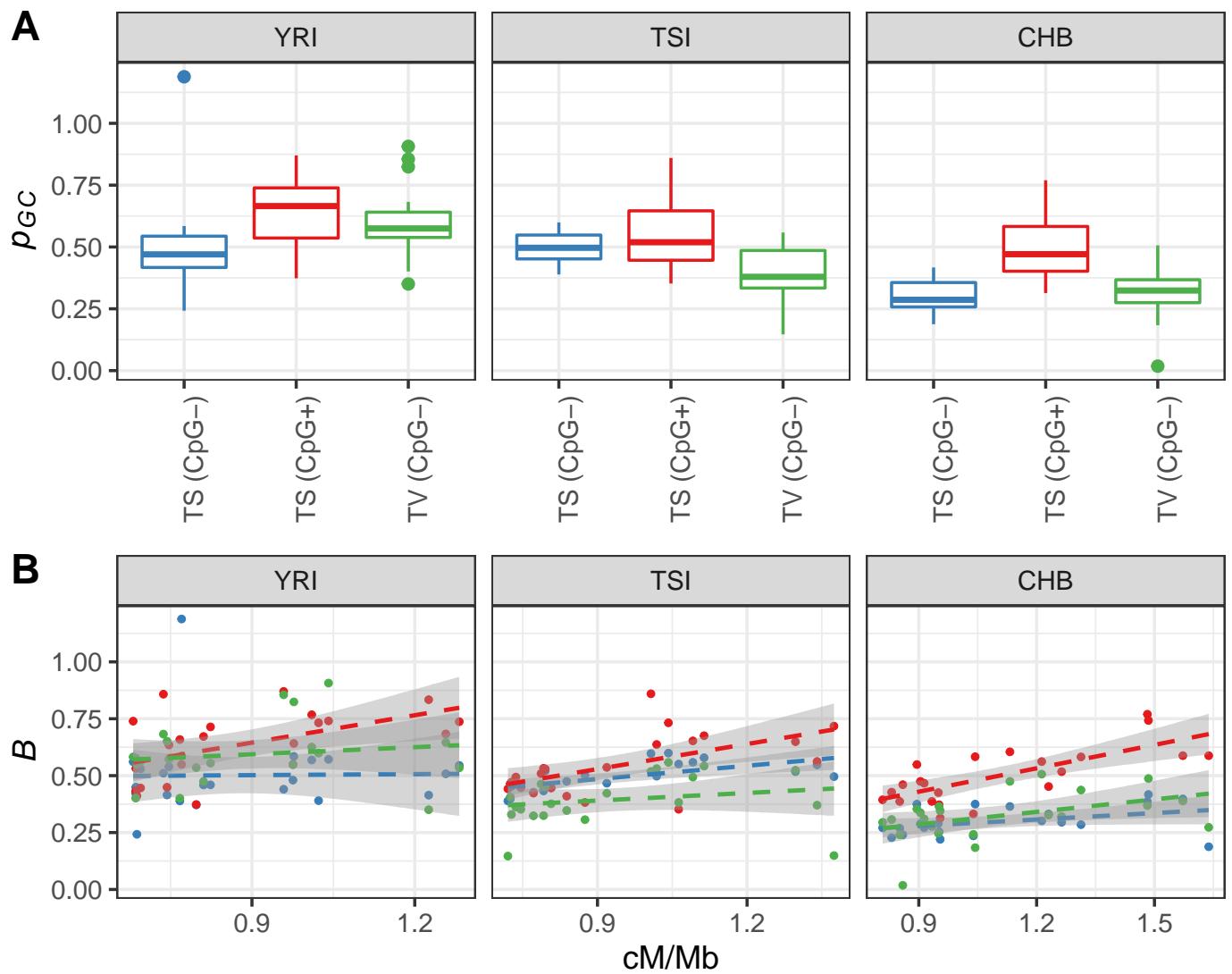
Supplementary figure S1: Two types of GC-changing mutations and the corresponding heteroduplex mismatches formed during recombination.



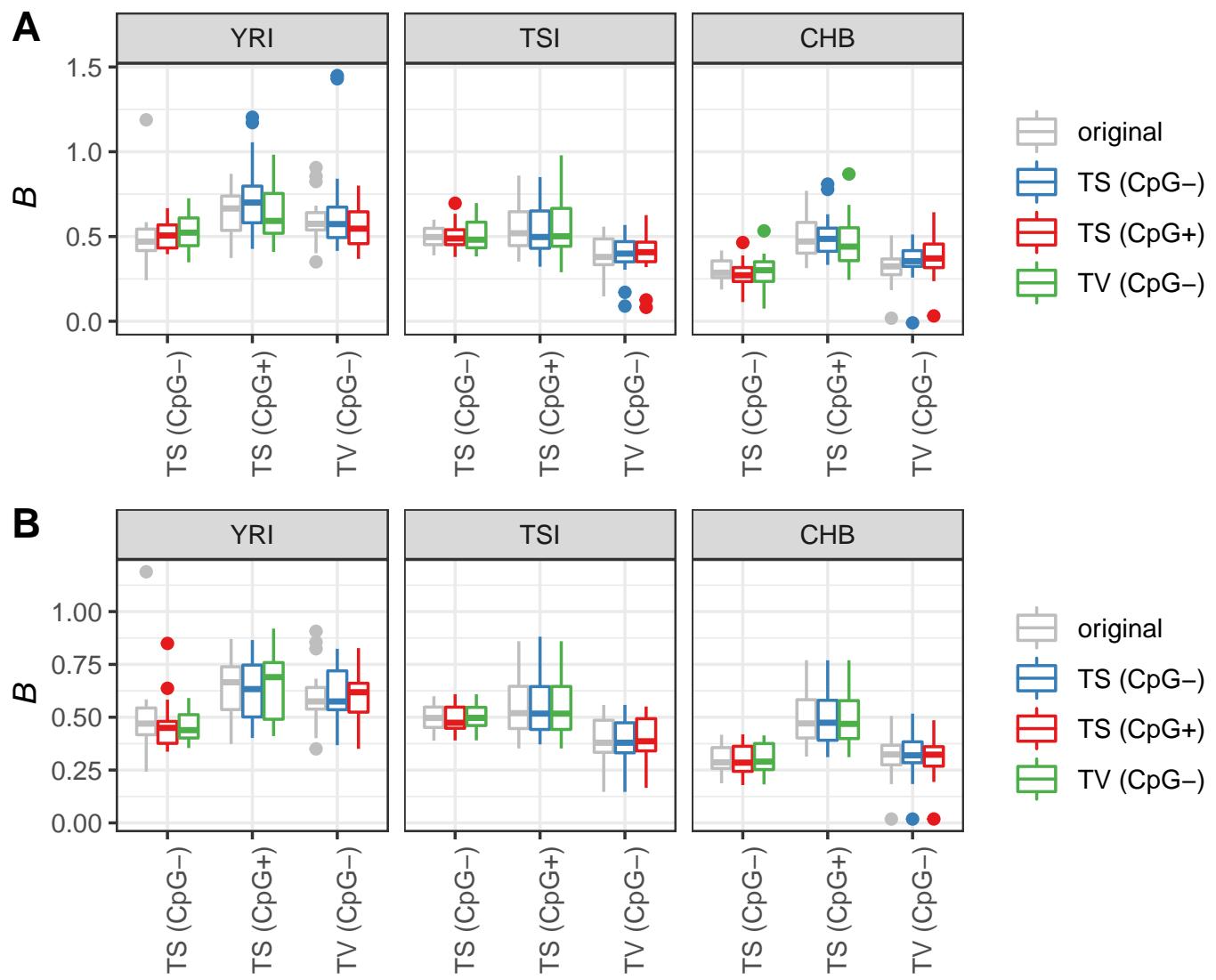
Supplementary figure S2: **A.** Distributions of average recombination rates in 1 Mb genomic regions for the Yoruba (YRI), Toscana (TSI), Han Chinese (CHB) and Iceland (ISL) populations. For the ISL, three different maps are available: sex-averaged (avg.), maternal (mat.) and paternal (pat.) map. **B.** Distributions of average recombination rates in 1 Mb genomic regions for different great ape populations (PTE = *P. troglodytes ellioti*; PTS = *P. troglodytes schweinfurthii*; PTT = *P. troglodytes troglodytes*; PTV = *P. troglodytes verus*; PPA = *P. paniscus*; GGG = *G. gorilla gorilla*; PAB = *P. abelii*; PPY = *P. pygmaeus*).



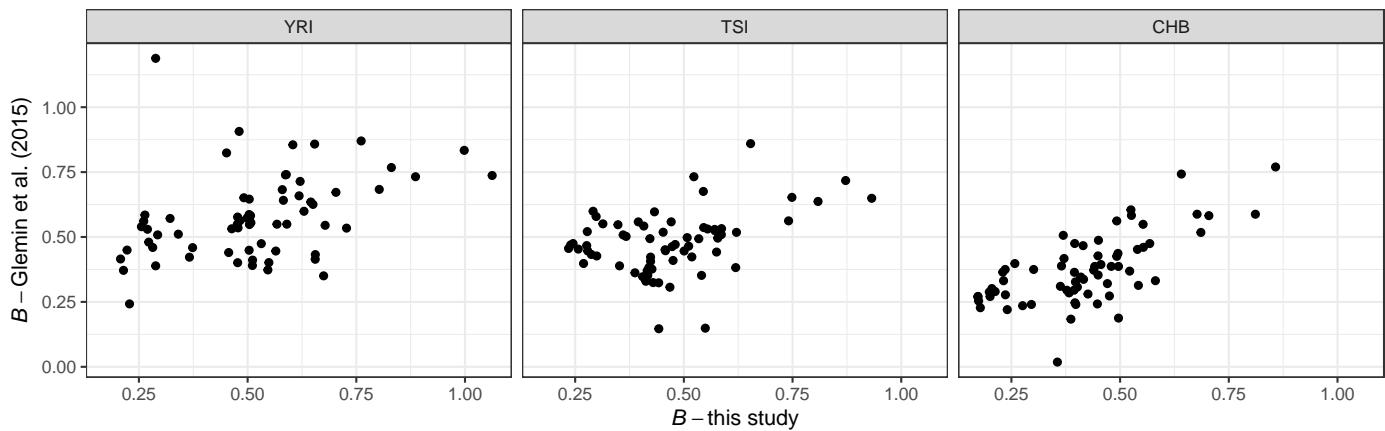
Supplementary figure S3: Expected distribution of allelic frequencies given a mutation-drift model and different mutation rates and biases. The mutation-drift expectation was derived by ? as $Cx^{4Nu-1}(1-x)^{4Nv-1}$, where N is the effective size of the population, u and v are the mutation rates from allele 0 to allele 1 and in reverse, respectively, x is the segregating frequency of allele 0, and C is the normalization constant. We plot this distribution assuming x is the GC frequency of a segregating variant (while $1-x$ is the segregating AT frequency), $N=10,000$ (equivalent to the human effective size) and u and v correspond to GC \rightarrow AT and AT \rightarrow GC mutation rates, respectively. As the genome-wide mutation rate for humans is 1.29×10^{-8} per basepair per generation (?), while mutations bias is approximately $\frac{1}{3} : \frac{2}{3}$ in favour of AT nucleotides, the genome-wide rates u and v are $\frac{2}{3}1.29 \times 10^{-8}$ and $\frac{1}{3}1.29 \times 10^{-8}$, respectively. Given these values, the expected distribution of allelic frequencies is plotted in red. On the other hand, the GC \rightarrow AT mutation rate at CpG sites is $10\times$ the genome average (*i.e.*, $10\times u$; plotted in light green). To demonstrate the conditions for which the frequency distribution becomes skewed under the mutation-drift equilibrium model, we also plot distributions where the GC \rightarrow AT mutation rate is equal to $50\times u$ (dark green), $100\times u$ (blue) and $1000\times u$ (pink).



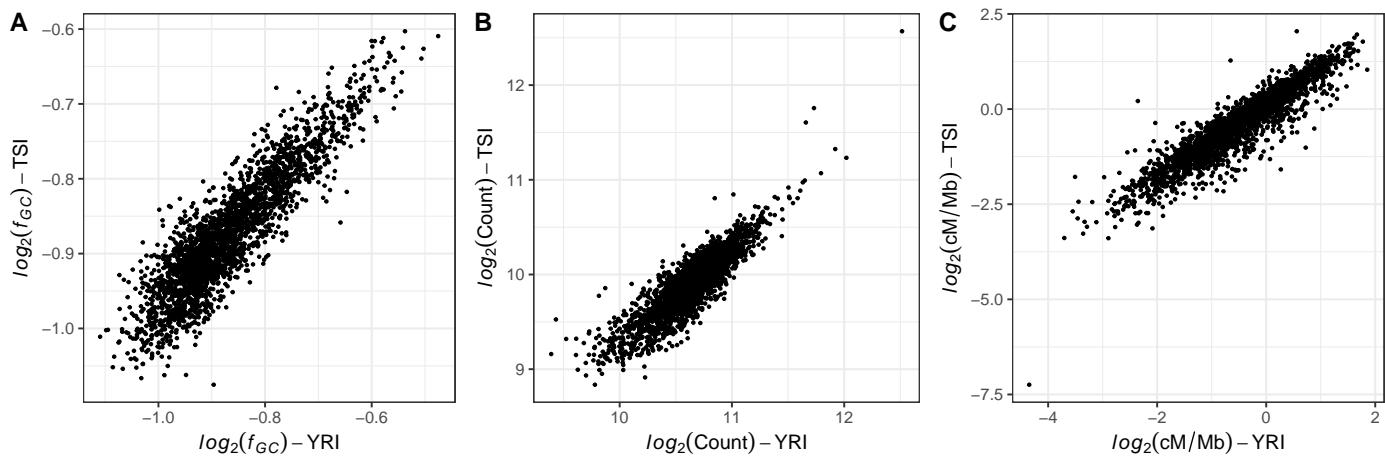
Supplementary figure S4: **A.** Distribution of the per-chromosome B estimates obtained using the ? method for non-CpG transitions (TS; CpG-), CpG transitions (TS; CpG+) and GC-changing transversions (TV; CpG-). **B.** Relationship between the non-CpG transitions (blue), CpG transitions (red) or GC-changing transversions (green) B parameter and the sex-averaged recombination rate across autosomes. Each point is an autosomal chromosome and the three panels in each figure correspond to the Yoruba (YRI), Toscana (TSI) and Han Chinese (CHB). AFSs used in the analysis are provided in Supplementary data S1.



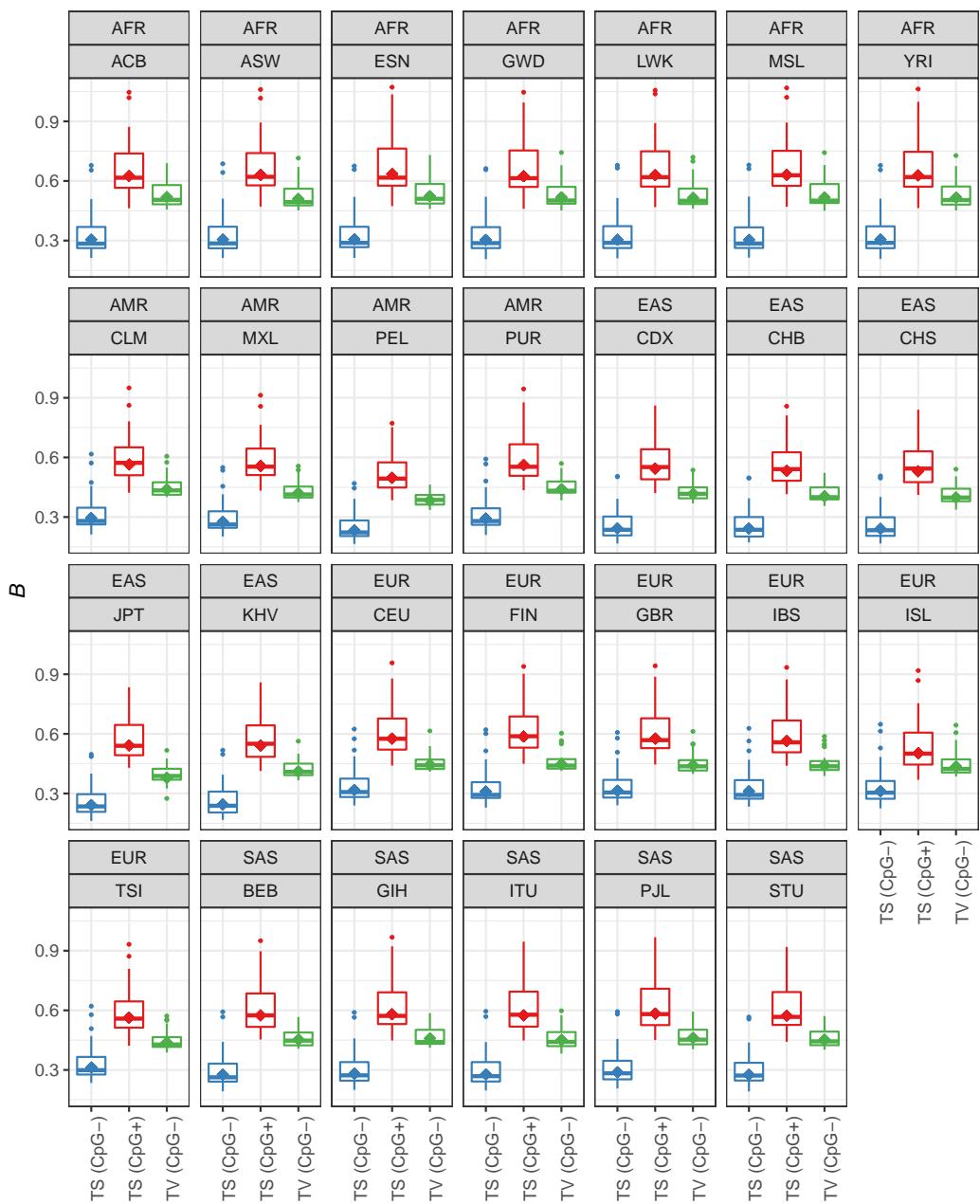
Supplementary figure S5: A. Distribution of the per-chromosome B estimates obtained using the $\hat{?}$ method for non-CpG transitions (TS; CpG $-$), CpG transitions (TS; CpG $+$) and GC-changing transversions (TV; CpG $-$) after resampling the AT \rightarrow GC and GC \rightarrow AT AFSs of the focal mutation type (x-axis) to match the sizes of the AFSs of a different mutation type (indicated by colors). **B.** Distribution of the per-chromosome B estimates for non-CpG transitions (TS; CpG $-$), CpG transitions (TS; CpG $+$) and GC-changing transversions (TV; CpG $-$) after exchanging the p_{GC} parameter of the focal mutation type (x-axis) with the p_{GC} parameter of a different mutation type (indicated by colors). The gray boxplots correspond to original B distributions from Supplementary figure S4A. Each point is an autosomal chromosome and the three panels in each figure correspond to the Yoruba (YRI), Toscana (TSI) and Han Chinese (CHB) populations.



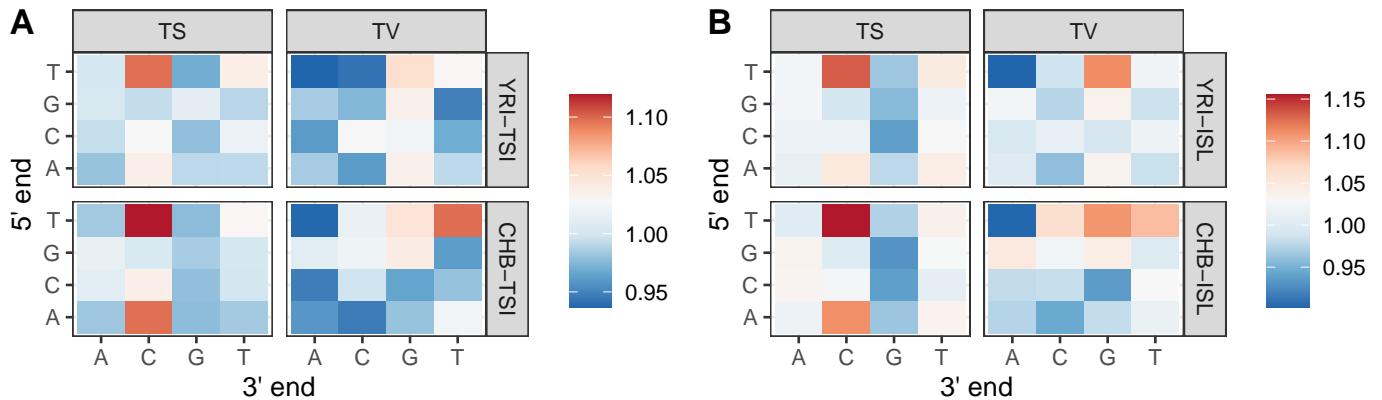
Supplementary figure S6: Correlation between per-chromosome B values for the YRI, TSI and CHB populations, estimated using the ? method and the method presented in this study. Each point corresponds to an autosomal B estimate for one of the three mutation types.



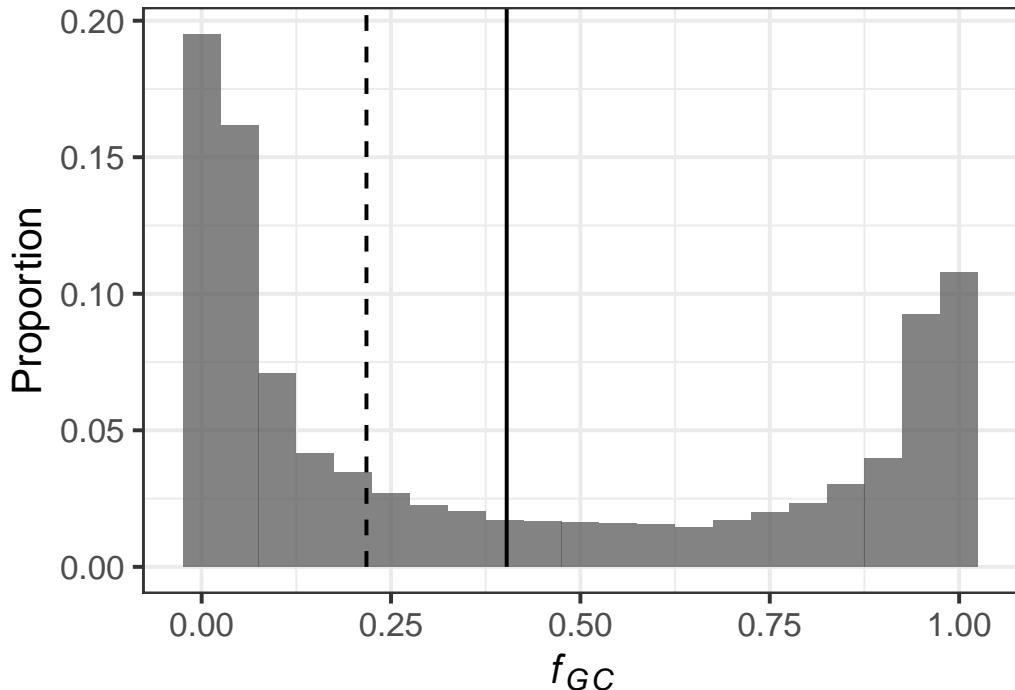
Supplementary figure S7: A. Correlation between the average GC frequency (f_{GC}) of non-CpG transitions in 1 Mb autosomal windows of the African YRI and European TSI populations. B. Correlation between the count of non-CpG transitions in 1 Mb autosomal windows of the African YRI and European TSI populations. C. Correlation between the average recombination rate in 1 Mb autosomal windows of the African YRI and European TSI populations.



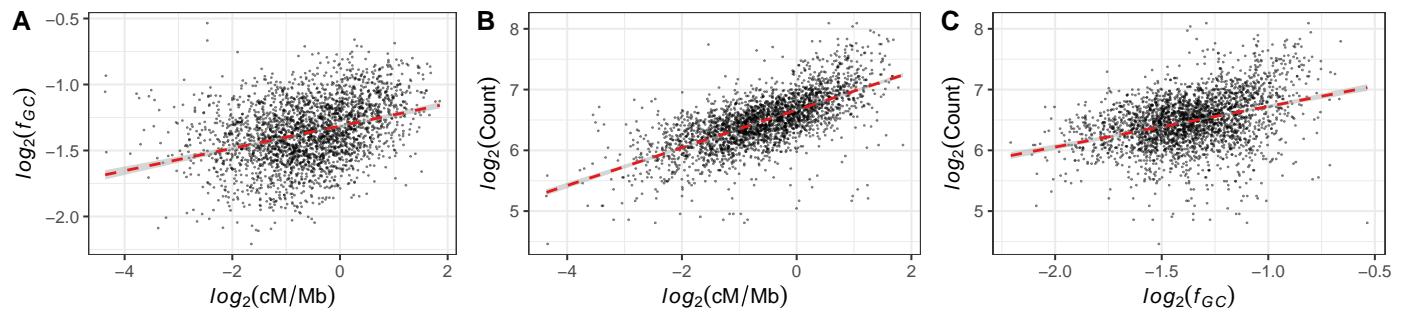
Supplementary figure S8: Distribution of the B parameter for non-CpG transitions (TS; CpG $-$), CpG transitions (TS; CpG $+$) and transversions (TV; CpG $-$) across autosomes for 26 human populations of the 1000 Genomes dataset and the Iceland population (ISL). In addition to specific population identifiers, we also indicate geographic superpopulation identifiers (AFR = African; SAS = South Asian; EAS = East Asian; EUR = European; AMR = Admixed American). AFSs used in the analysis are provided in Supplementary data S1.



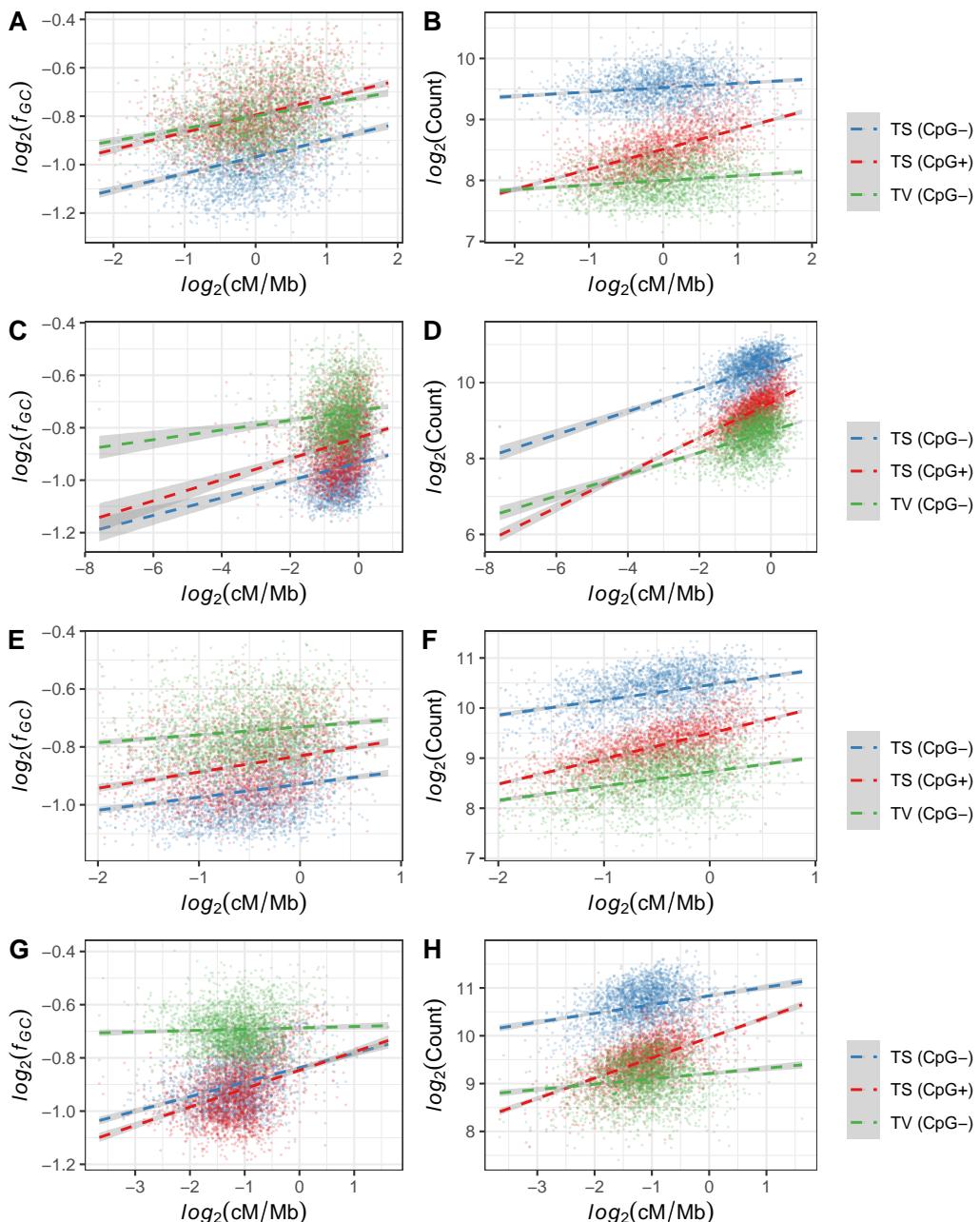
Supplementary figure S9: The ratios of proportions of the 32 different mutation types between the non-European populations and European **A.** TSI and **B.** ISL populations. Each block represents the value of a statistic across different mutation types (transitions; TS, and transversions; TV) and population pairs (e.g., YRI-TSI for the proportion ratios of the Yoruba and Toscana populations, and CHB-TSI for the proportion ratios of the Han Chinese and Toscana populations). Values are calculated separately based on their flanking context, where the x-axis and y-axis represent the 3' and 5' nucleotides, respectively. Complementary nucleotide contexts are grouped together. AFSs used in the analysis are provided in Supplementary data S2.



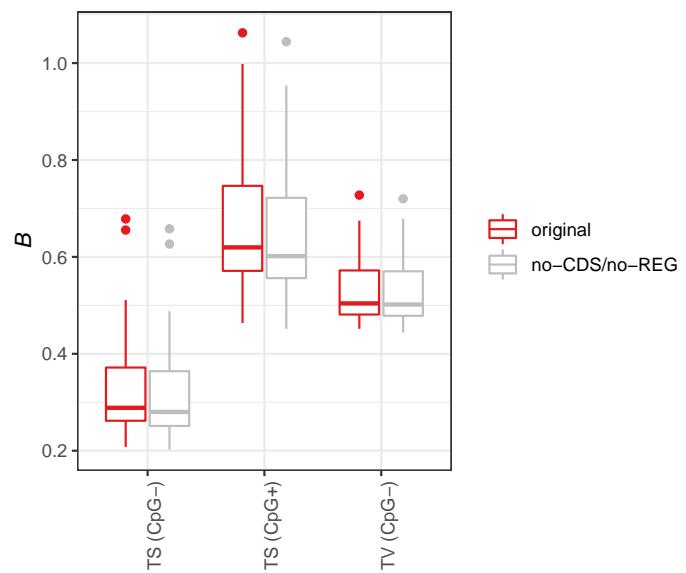
Supplementary figure S10: The distribution of GC frequencies (f_{GC}) for CpG transversions of the YRI population. The solid line represent the median, and the dashed line represent the mean of the distribution.



Supplementary figure S11: **A.** Relationship between the average GC frequency (f_{GC}) at segregating CpG transversion sites and average recombination rate in 1 Mb autosomal windows of the African Yoruba population. **B.** Relationship between the count of CpG transversions and recombination rate in 1 Mb autosomal windows of the African Yoruba population. **C.** Relationship between the count of CpG transversion and average GC frequency (f_{GC}) in 1 Mb autosomal windows of the African Yoruba population.



Supplementary figure S12: Relationship between the average GC frequency (f_{GC}) at segregating sites and average recombination rate in 1 Mb autosomal for non-CpG transitions (TS; CpG-), CpG transitions (TS; CpG+) and GC-changing transversions (TV; CpG-) of A. the *P. paniscus* population, C. the *G. gorilla gorilla* population, E. the *G. gorilla gorilla* population with $\text{cM}/\text{mb} > 0.25$, and G. the *P. abelii* population. Relationship between the count of GC-changing segregating sites and average recombination rate in 1 Mb autosomal for non-CpG transitions (TS; CpG-), CpG transitions (TS; CpG+) and GC-changing transversions (TV; CpG-) of B. the *P. paniscus* population, D. the *G. gorilla gorilla* population, F. the *G. gorilla gorilla* population with $\text{cM}/\text{mb} > 0.25$, and H. the *P. abelii* population.



Supplementary figure S13: Comparison between distributions of the per-chromosome B estimates for non-CpG transitions (TS; CpG $-$), CpG transitions (TS; CpG $+$) and GC-changing transversions (TV; CpG $-$), for the YRI population. The gray distributions correspond to the original distributions in Figure 3A, while the red distributions are estimated after exclusion of sites in coding and regulatory regions.