

Primate SDs and CNVs in the Human genome



Figure S 1. Overview of the location of CNVRs. CNVRs detected with the oligonucleotide array on a human karyotype ideogram. Genomic positions of SDs of human, chimpanzee, orangutan and macaque are also shown.

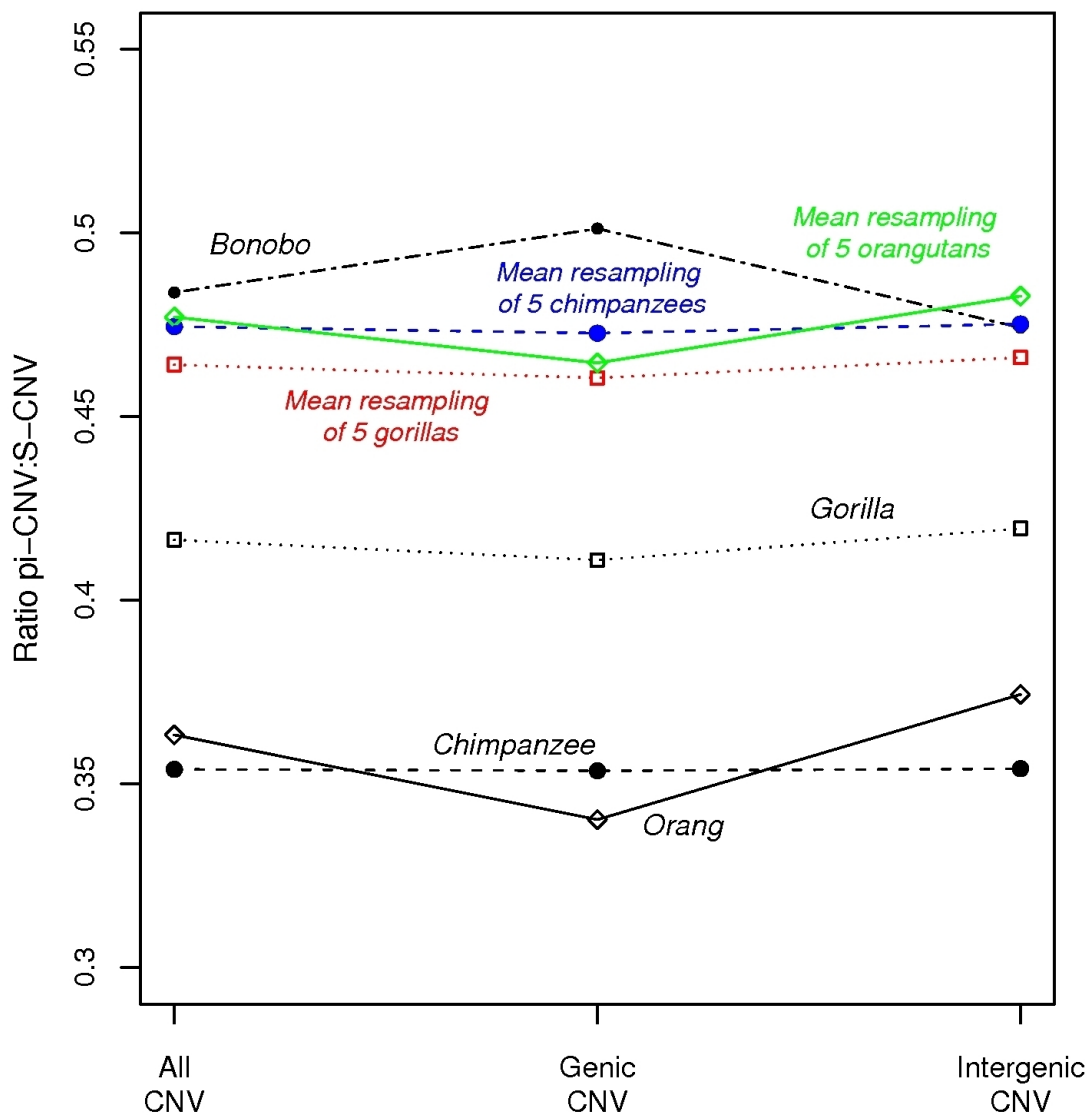


Figure S2. Effect of resampling individuals on the the ratio of CNV over SCNV. The ration is plotted for the four species and for each category of CNVR (all CNVR, genic CNVR and non-genic CNVR). Red, green and blue lines show the mean values of these same variables for the 6 combinations of 5 gorillas, 5 orangutans and 5 chimpanzee, respectively.

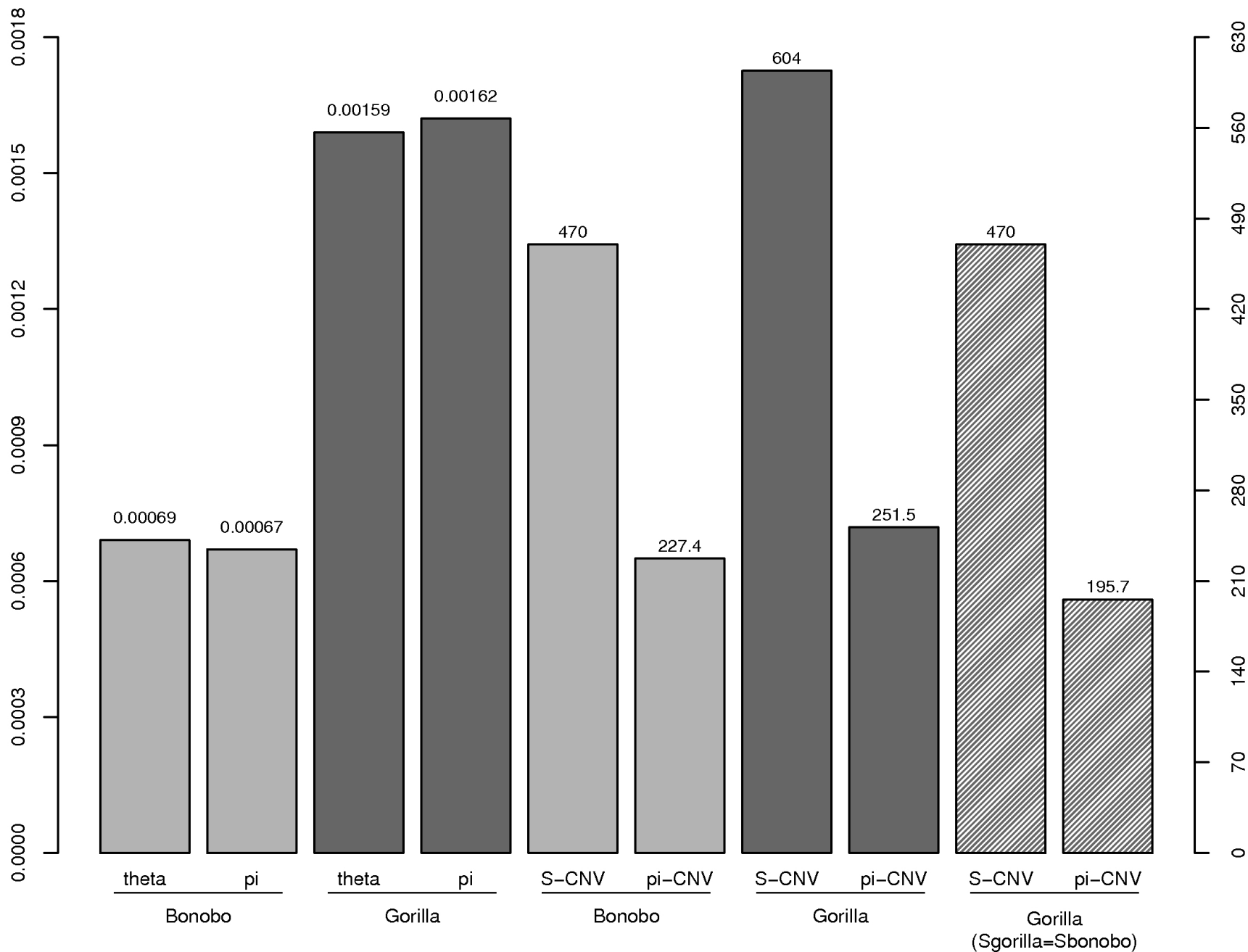


Figure S3. Nucleotide and CNV diversity patterns in bonobo and gorilla. The left y axis gives the scale of the nucleotide diversity values and the right y axis gives the values of CNV diversity data. The dashed bars show the results of the resampling procedure on gorillas, in which S-CNV gorilla is set N= 470 (the value of S-CNV bonobo). The pi-CNV value is the mean pi-CNV over 100,000 random resampling of N gorilla CNVR.

All CNV

Genic CNV

Intergenic CNV

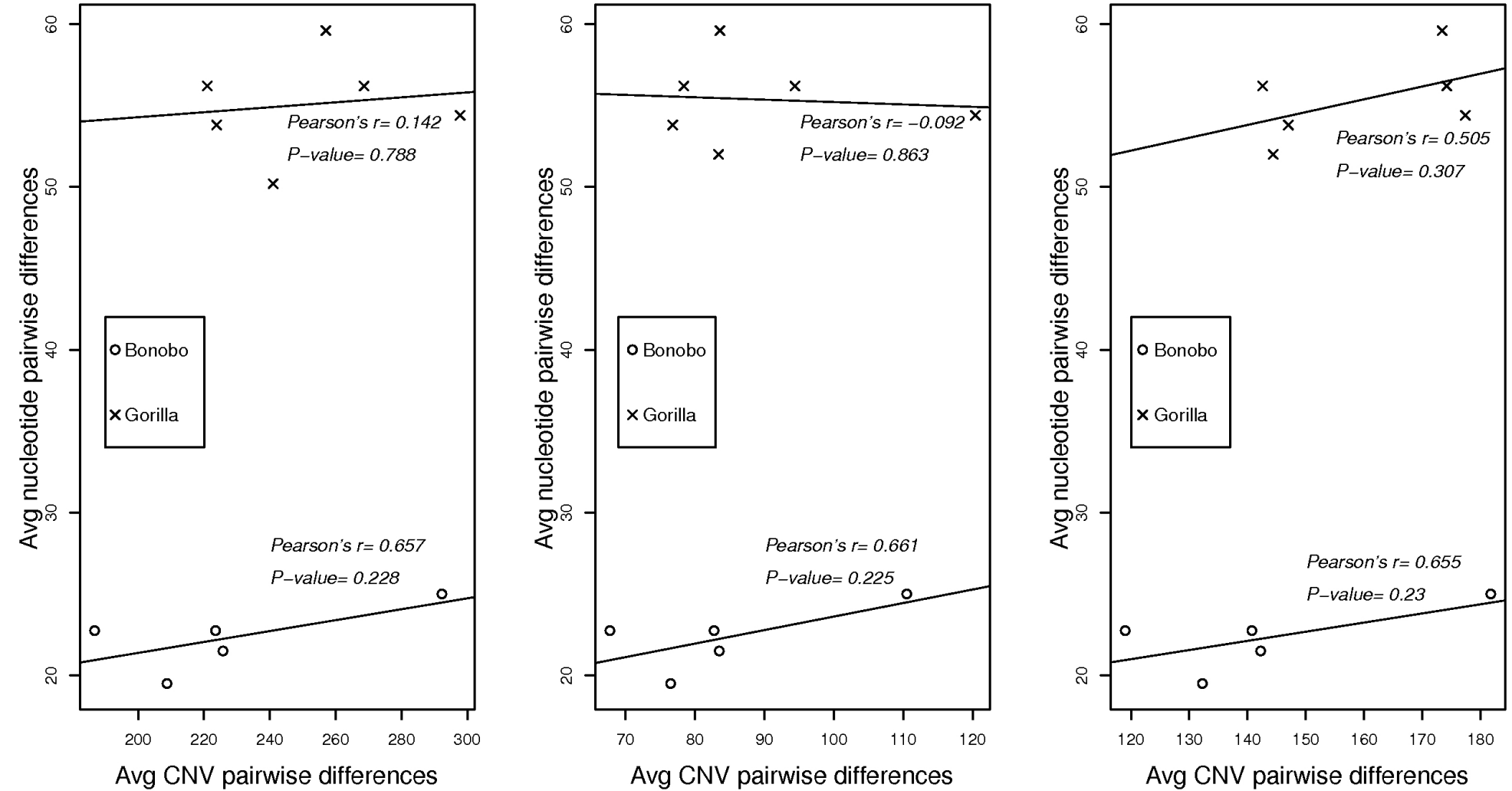


Figure S4. Relationship between nucleotide and CNV diversity. Each dot represents the mean pairwise difference in CNVRs (presence/absence) and nucleotide changes of each individual against the rest of the same species. Data are shown for all CNVR together, genic CNVR alone, and intergenic CNVR alone.

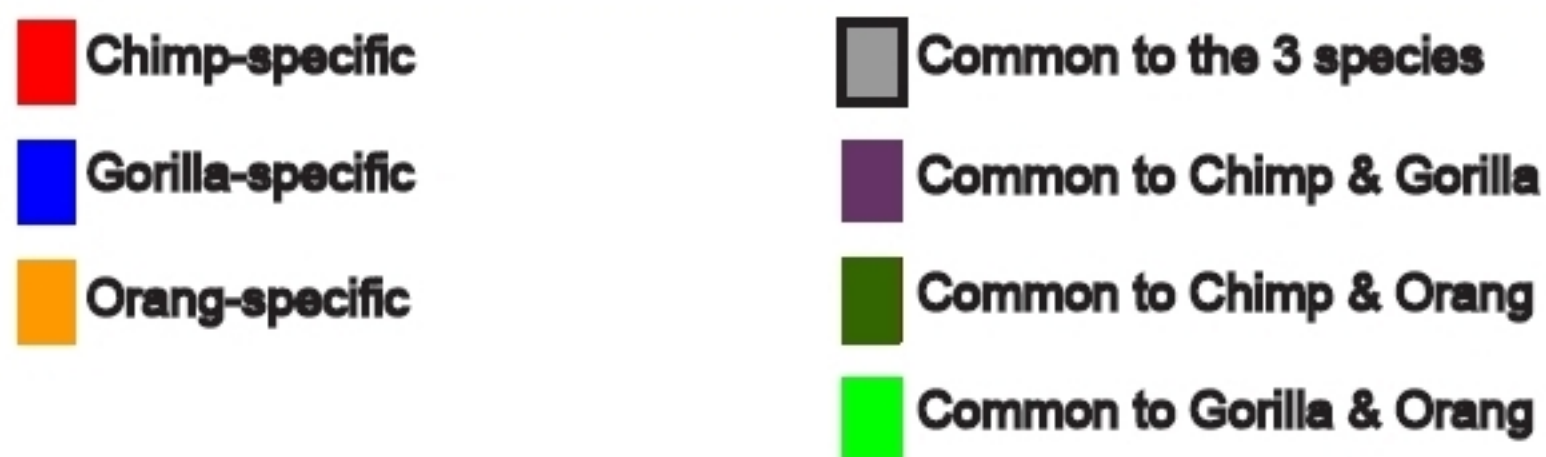
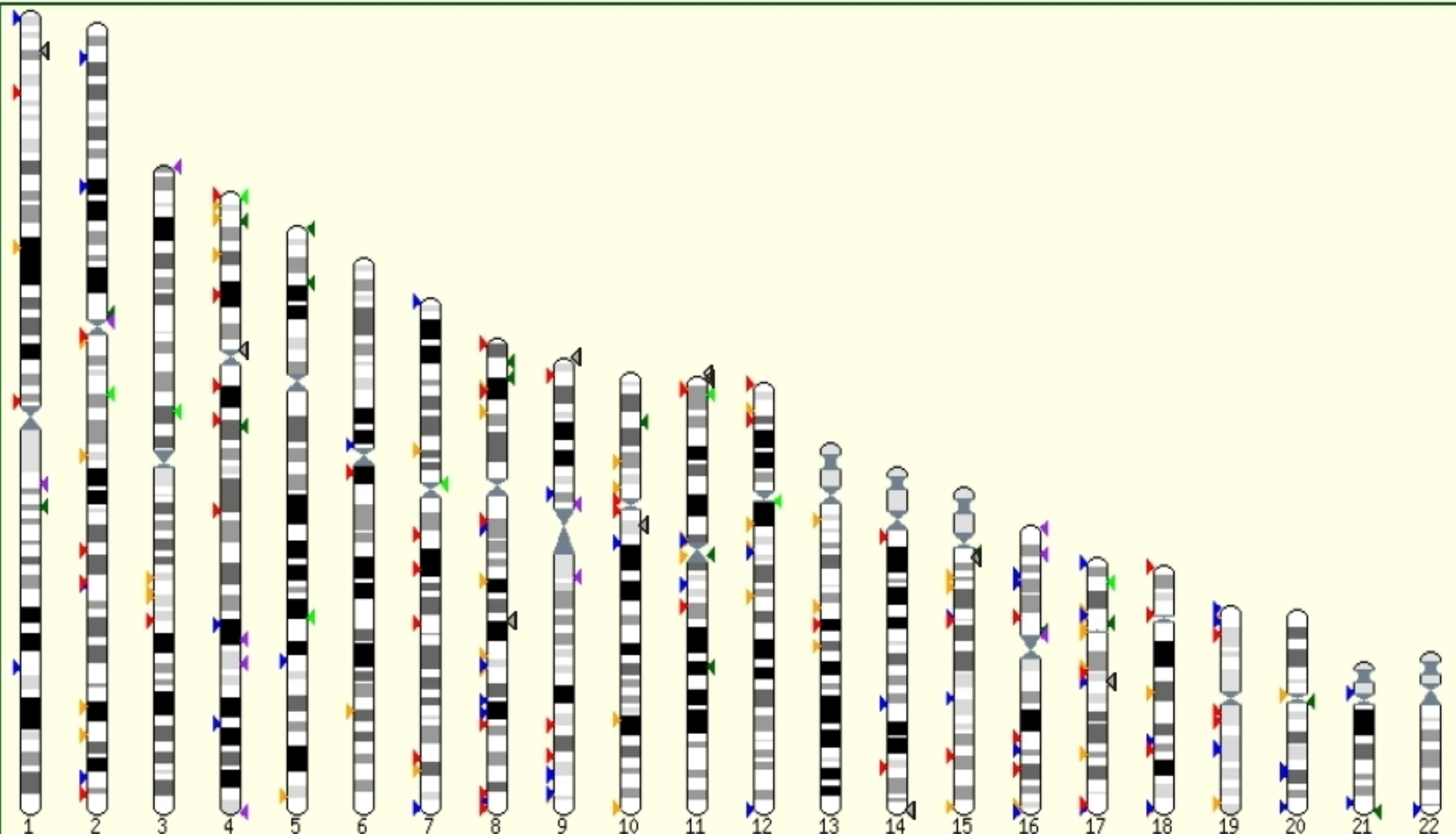


Figure S5. Location of BAC-detected CNVRs. CNVRs detected with the BAC array on a human karyotype ideogram.

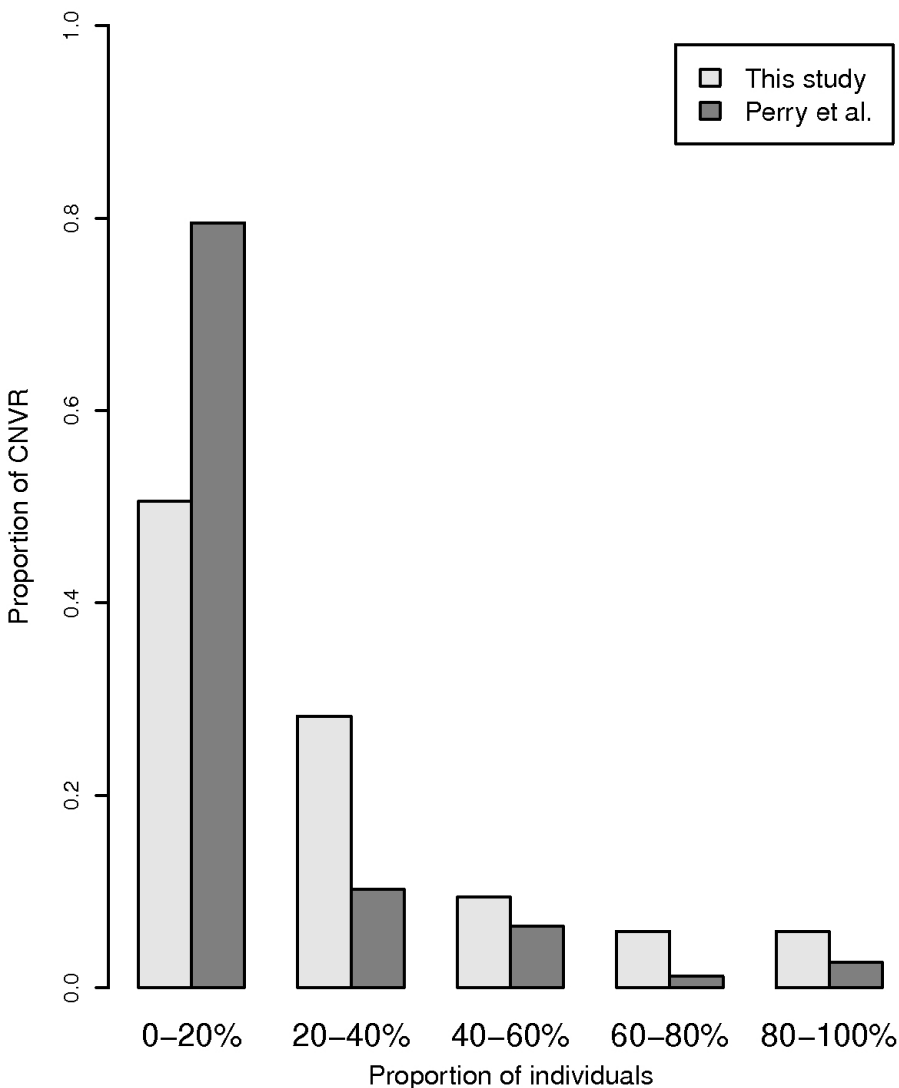


Figure S6. Chimpanzee frequency distribution comparison. Distribution of the proportion of individuals per CNVR for chimpanzees hybridized on the BAC array in the present study (light gray) and chimpanzees studied by Perry et al. (2008) (dark gray).

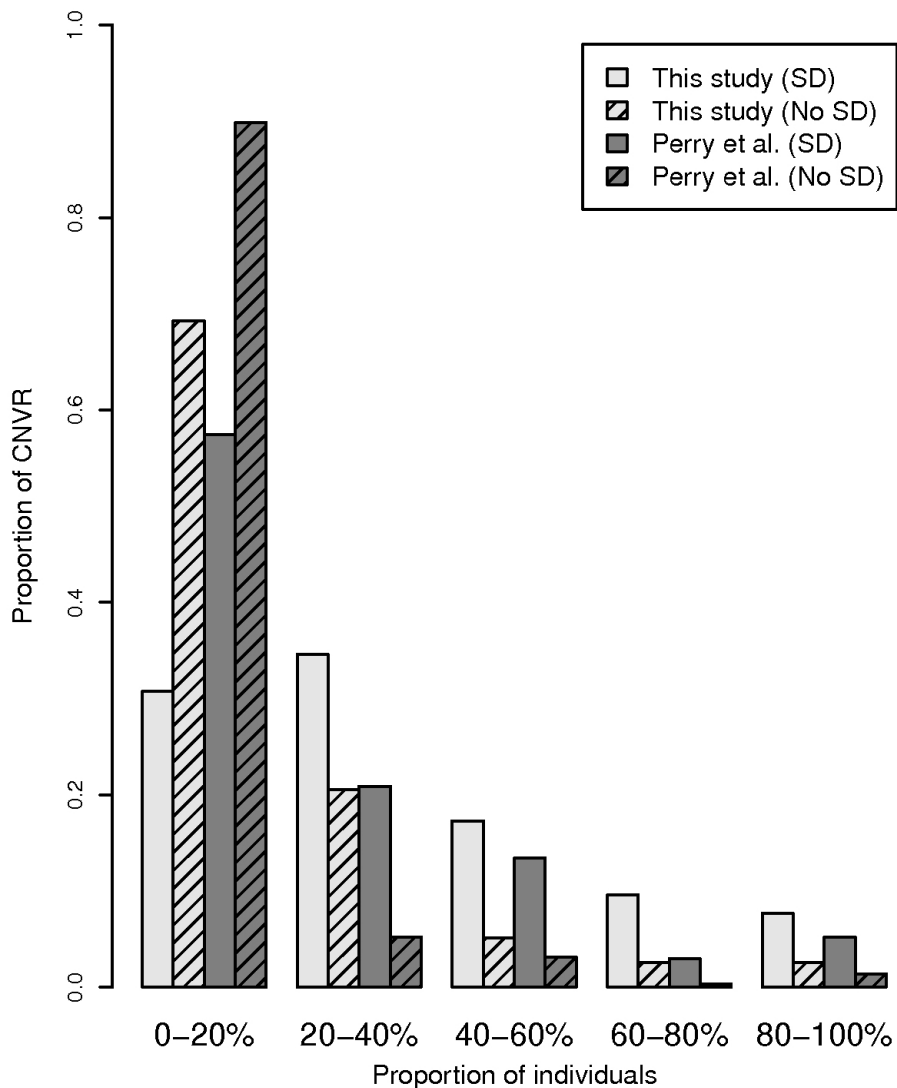


Figure S7. Chimpanzee frequency distribution according their overlap with SDs. The comparison between the chimpanzees hybridized on the BAC array in the present study (light gray) and the chimpanzees studied by Perry et al. (2008) (dark gray) is shown.

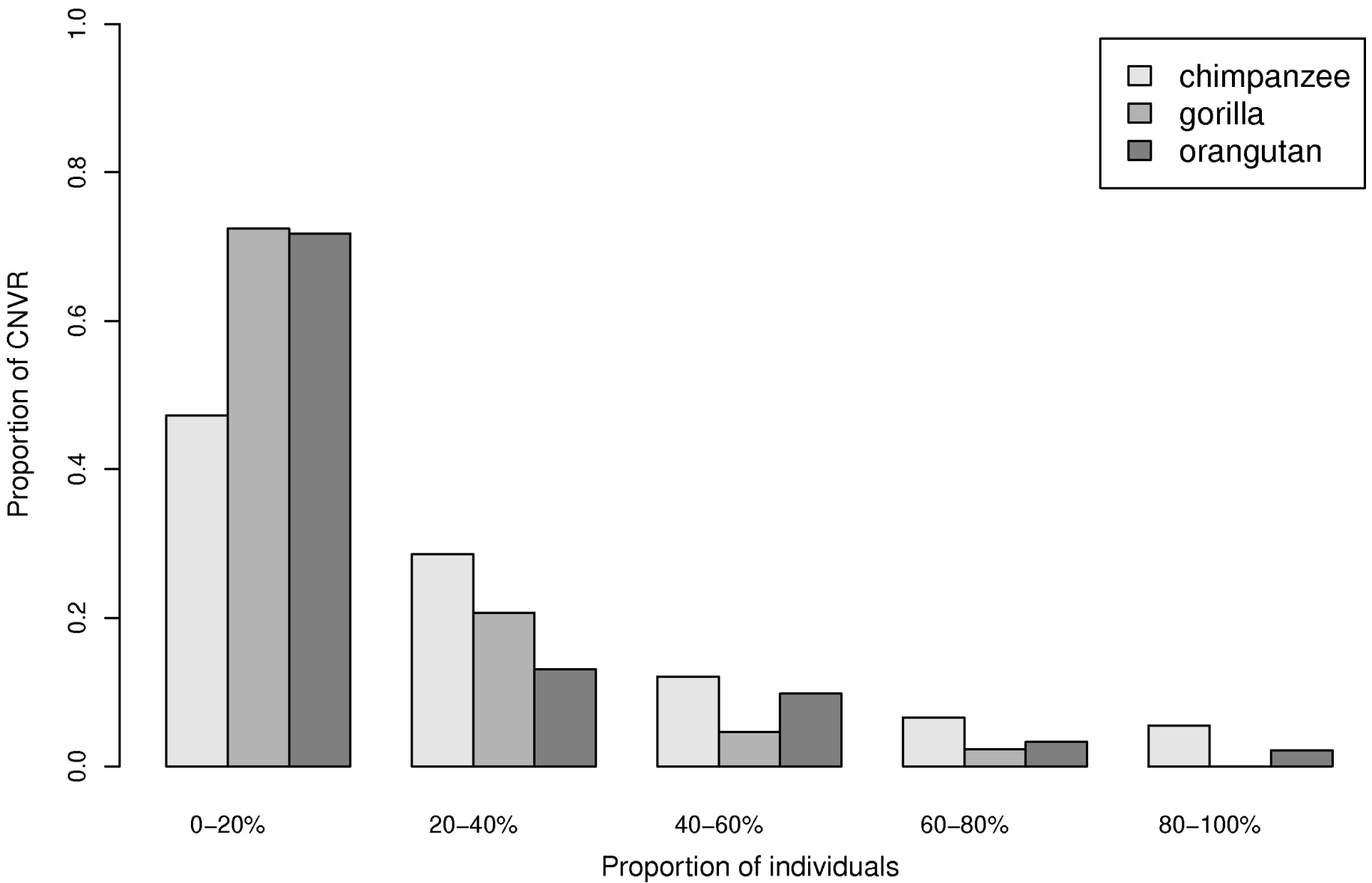


Figure S8. Frequency distribution of BAC-array CNVRs. Proportion of CNVR detected with the BAC array that are supported by a given proportion of the individuals of the sample.

Distribution of call length

a.

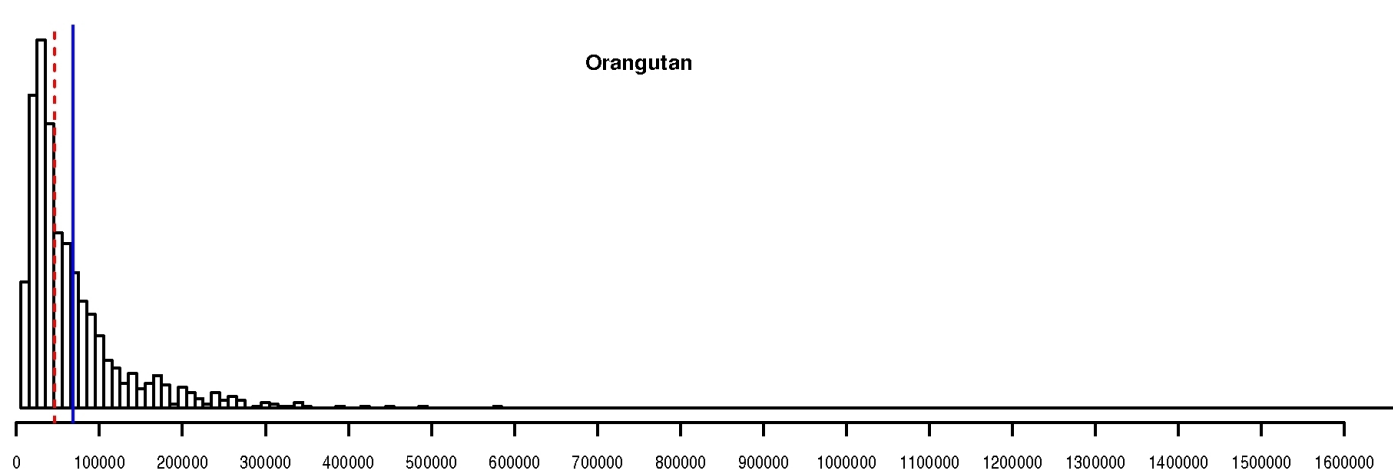
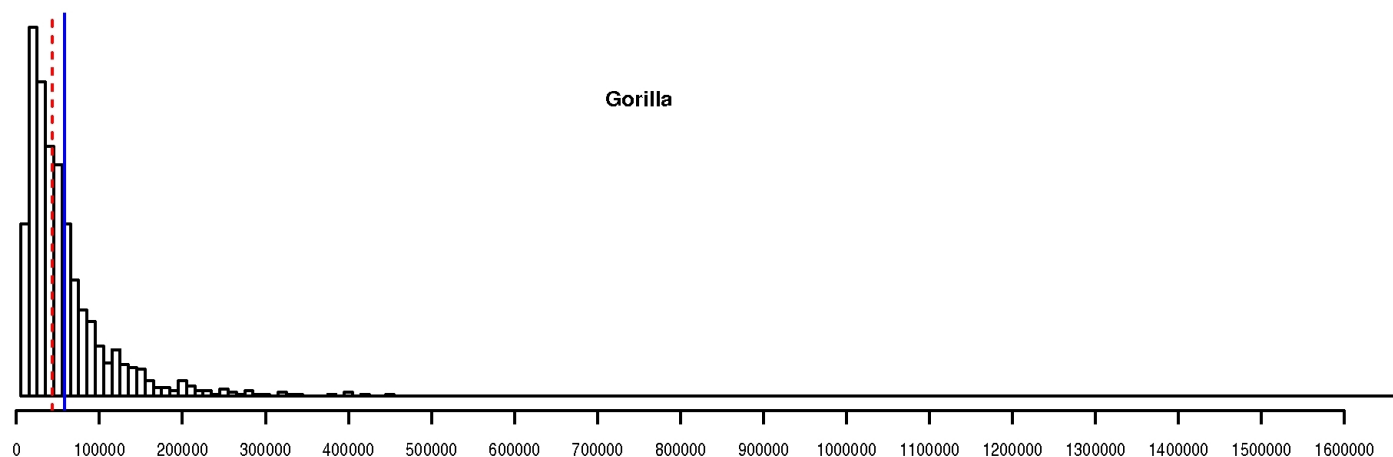
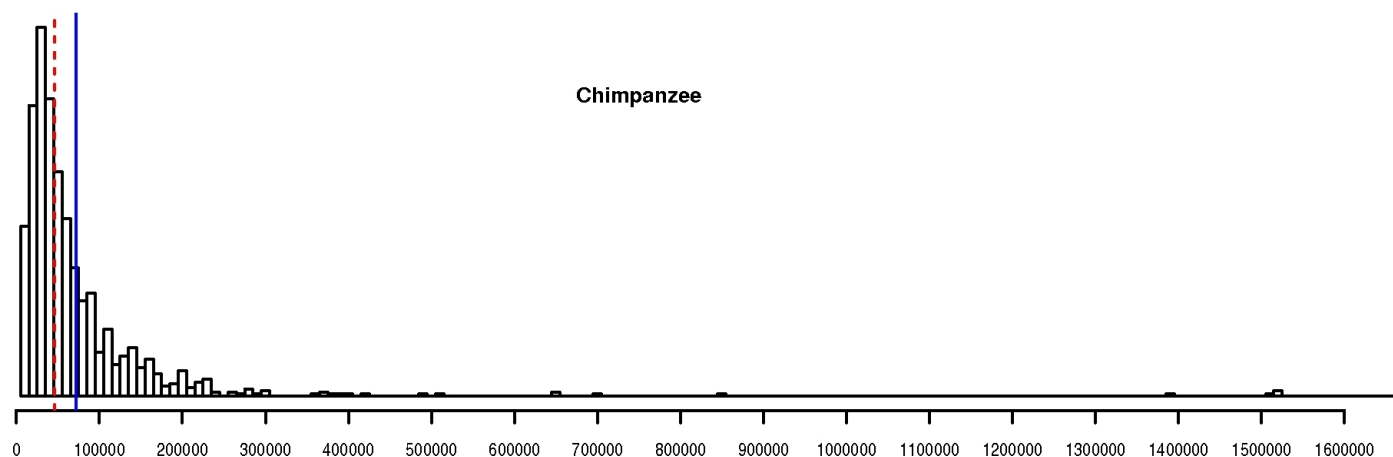
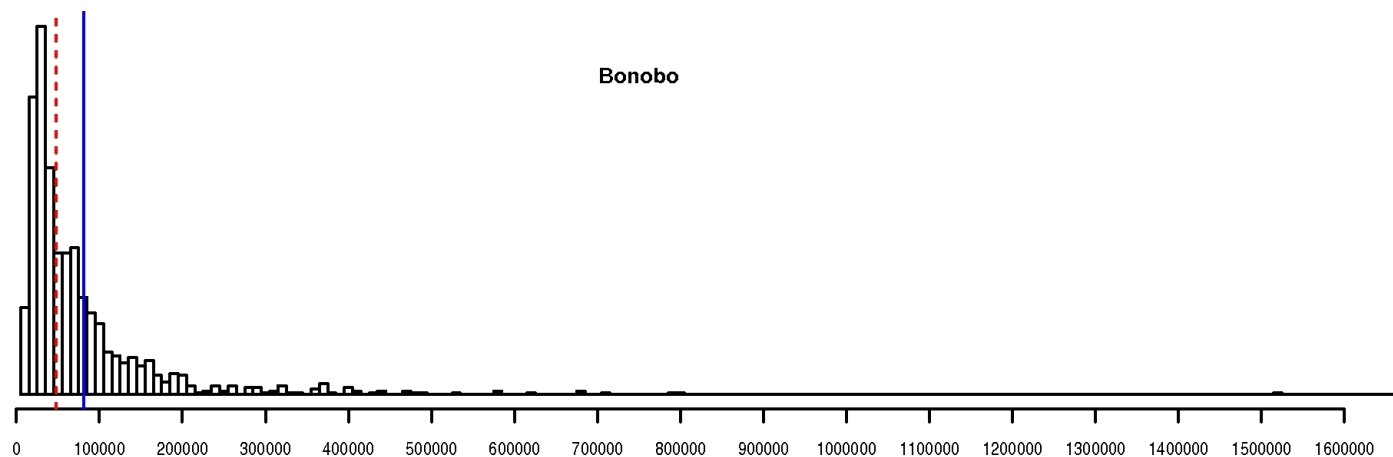


Figure S9. Distribution of lengths of CNVs and CNVRs. Solid lines (blue) indicate mean values of the distribution, dashed lines (red) indicate median values. (a) Full distribution of call length.

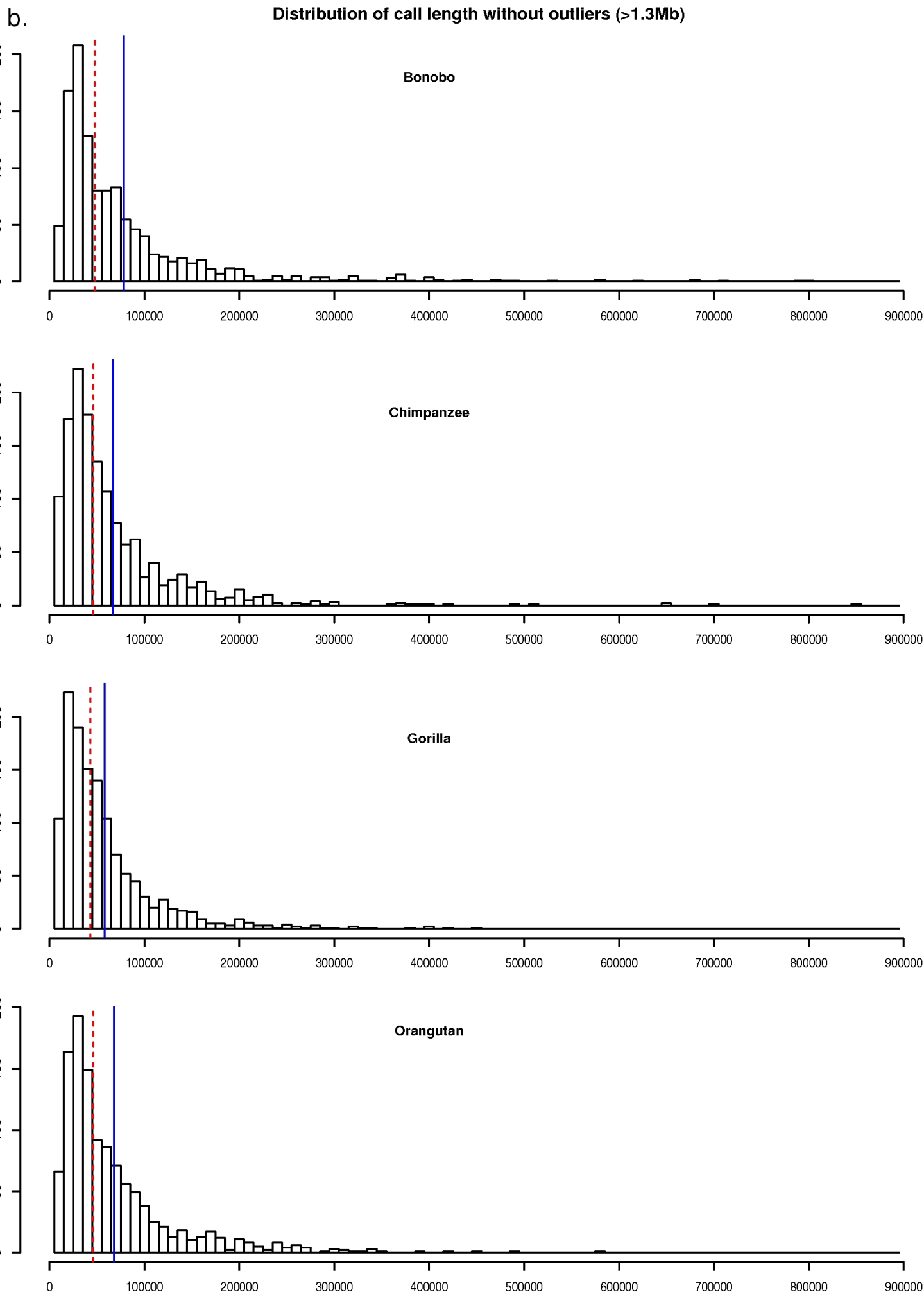


Figure S9. Distribution of lengths of CNVs and CNVRs. Solid lines (blue) indicate mean values of the distribution, dashed lines (red) indicate median values. (b) Distribution of call length without outliers (CNV > 1.3 Mb).

C.

Distribution of region length

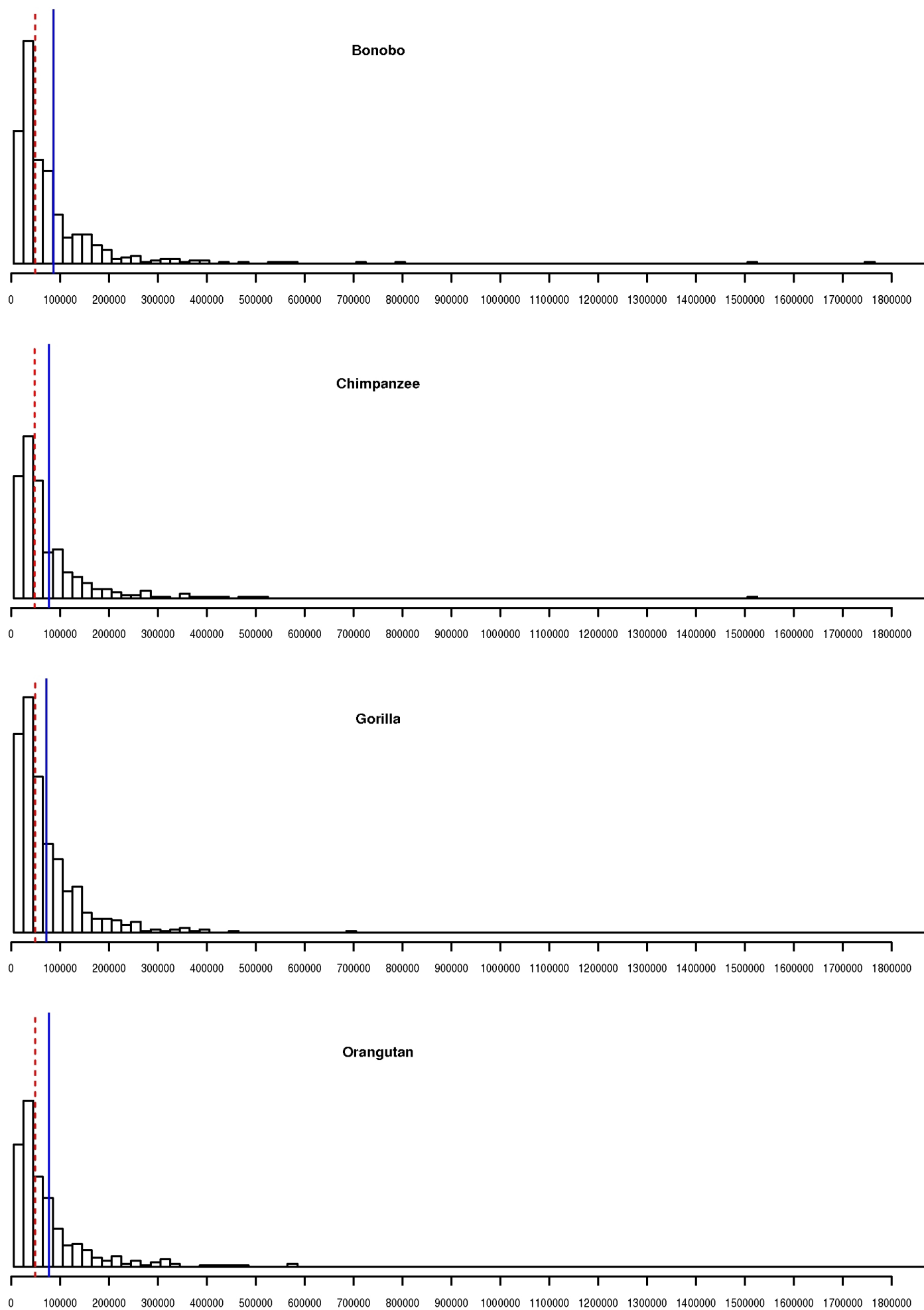


Figure S9. Distribution of lengths of CNVs and CNVRs. Solid lines (blue) indicate mean values of the distribution, dashed lines (red) indicate median values. (c) Full distribution of CNVR length.

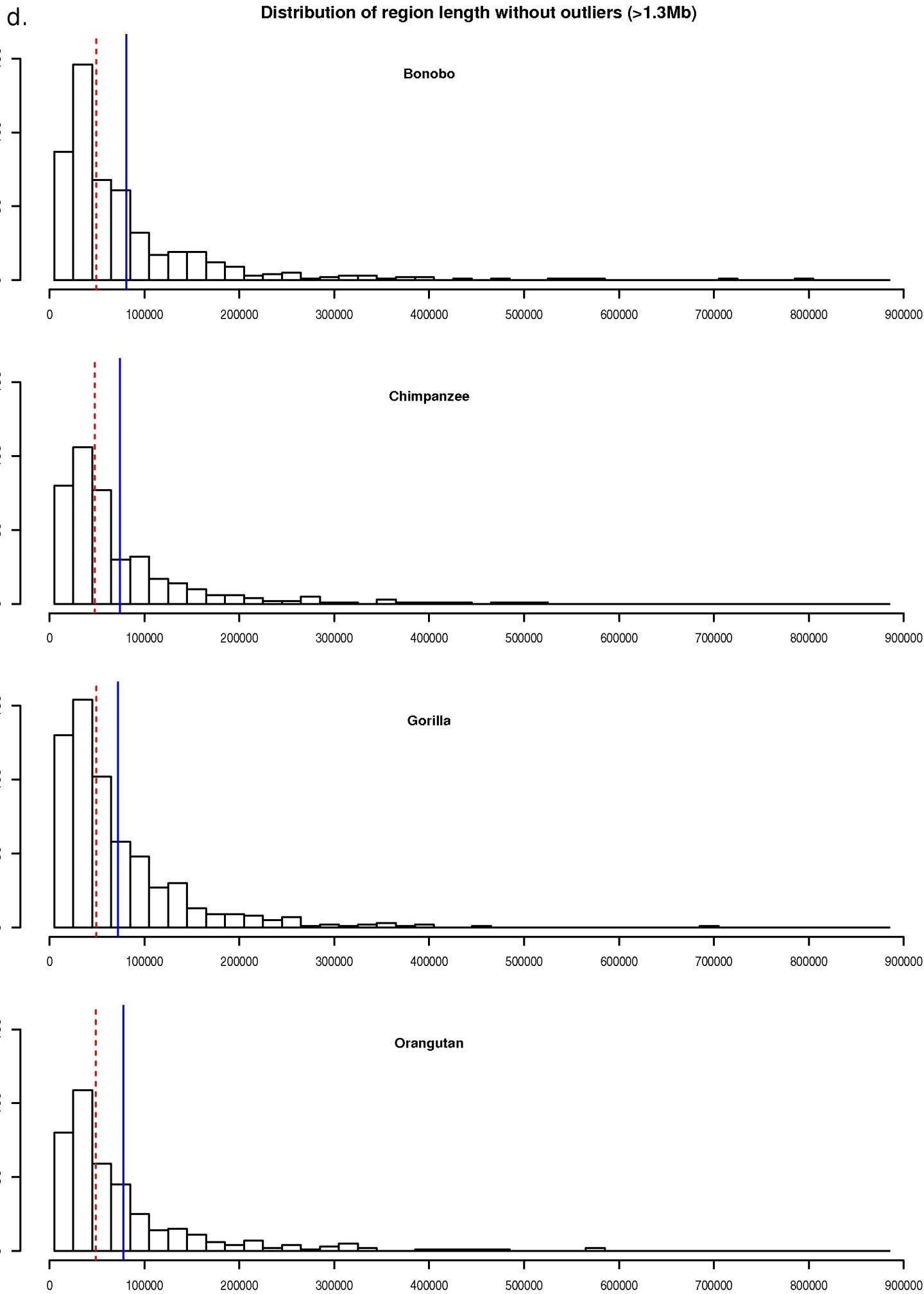


Figure S9. Distribution of lengths of CNVs and CNVRs. Solid lines (blue) indicate mean values of the distribution, dashed lines (red) indicate median values. (d) Distribution of lengths of CNVRs built after removing outlier calls.

