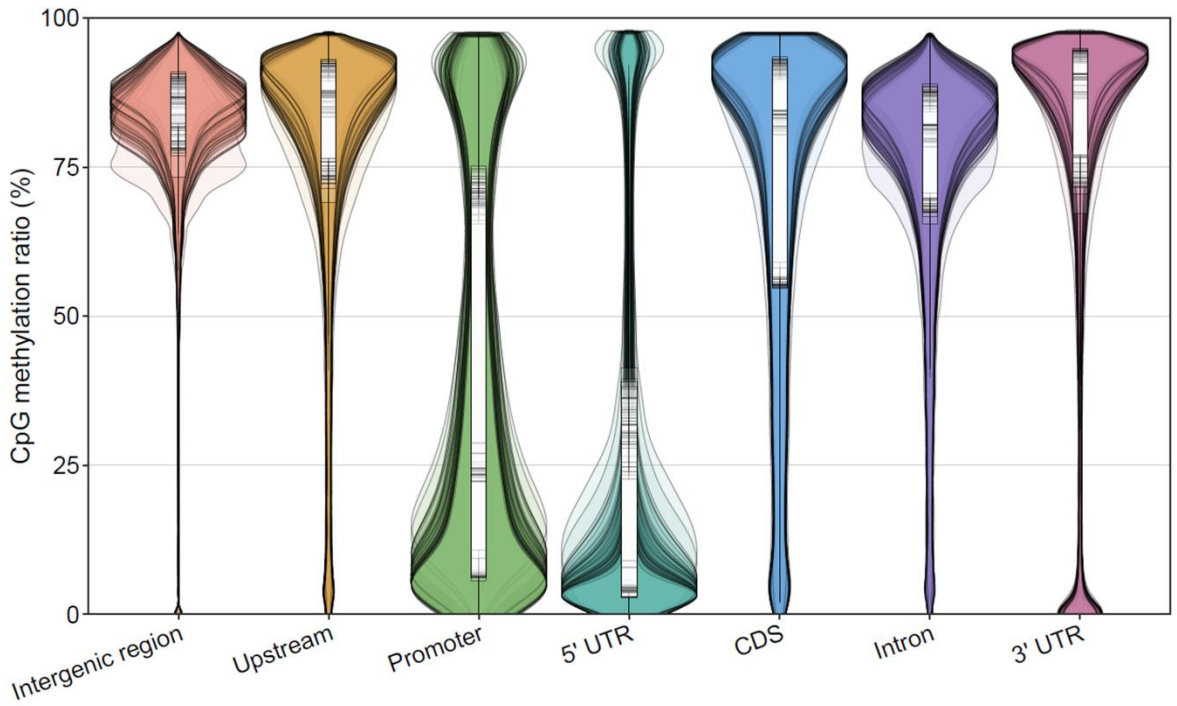
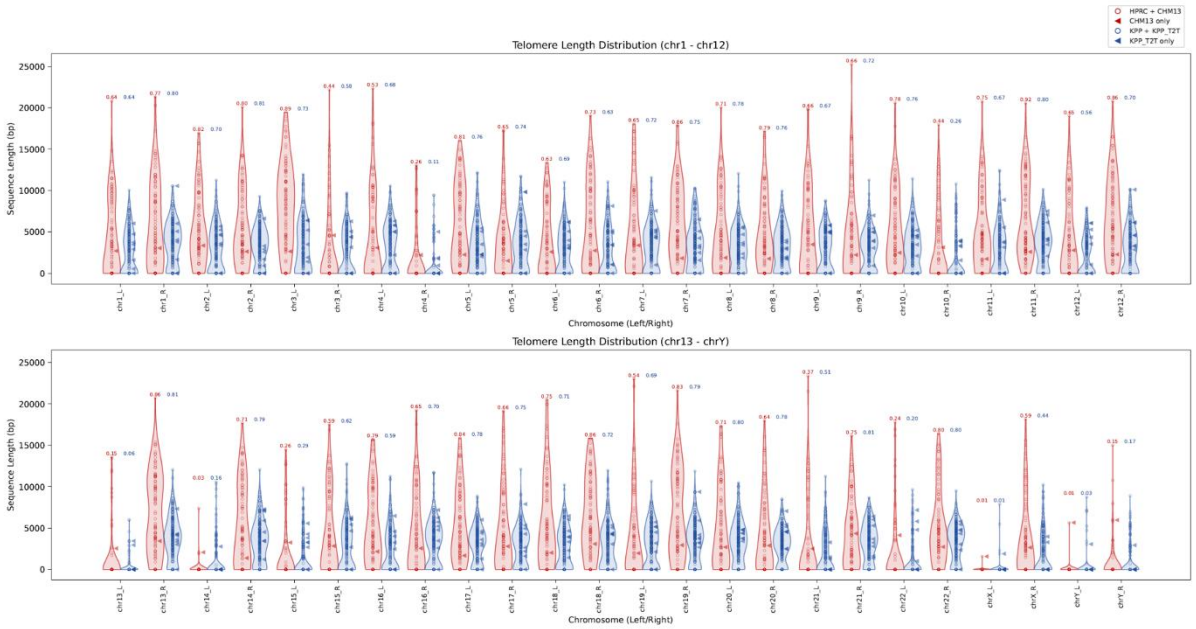
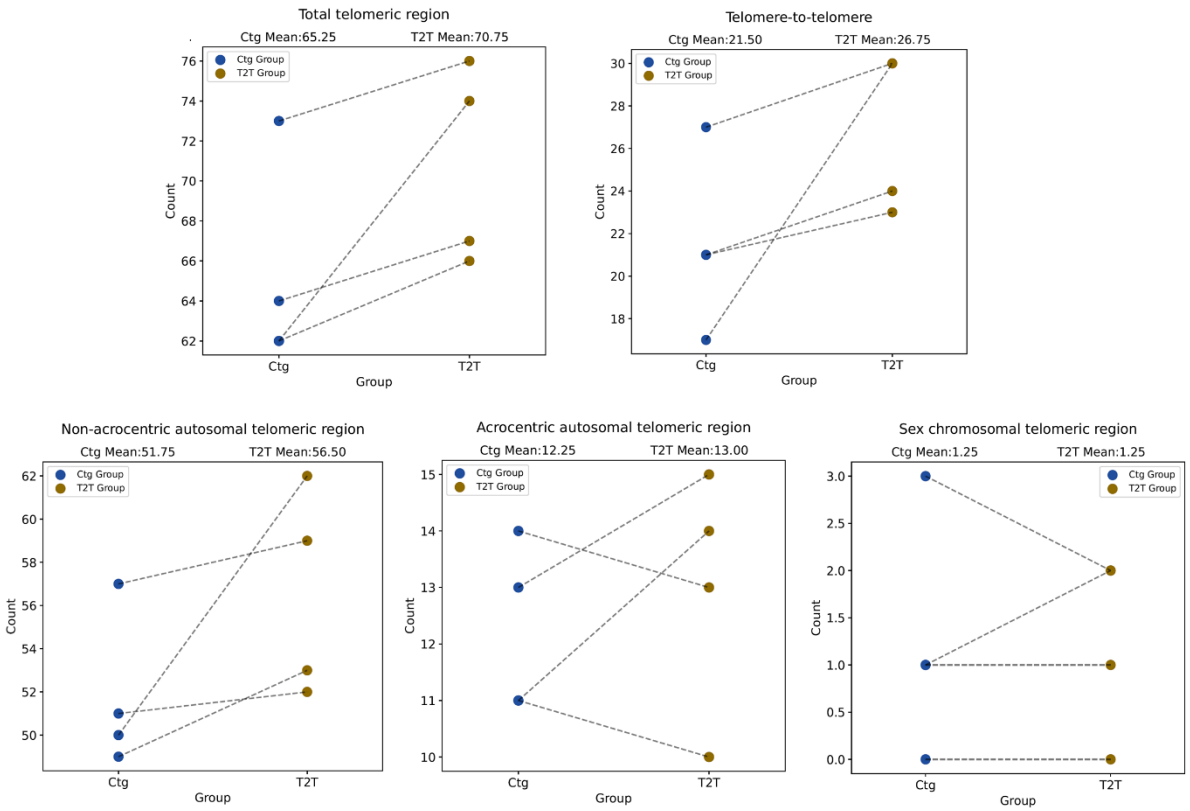


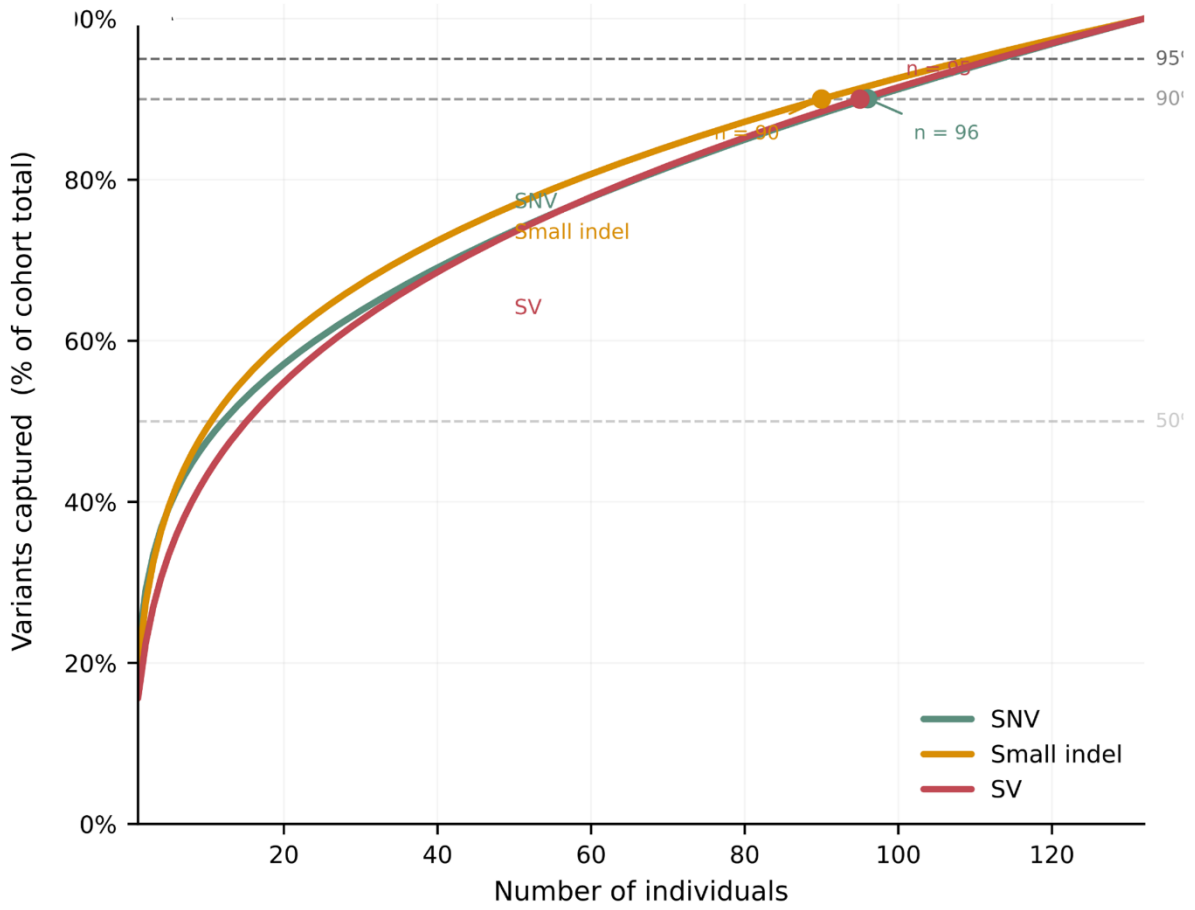
Supplementary Figure 1. Methylation profiles of intergenic and genic regions in the KPP samples. CpG methylation ratios for each genomic feature were calculated using methylation tag information contained in PacBio unmapped BAM files in combination with pb-CpG-tools. The analyzed regions were classified into upstream regions (1–5 kb from the transcription start site), promoter regions (0–1 kb from the transcription start site), 5′ untranslated regions (UTRs), coding sequences (CDS), intronic regions, 3′ UTRs, and intergenic regions. Methylation profiles from a total of 128 individuals were displayed as overlapping violin plots, with horizontal lines indicating the first quartile, median, and third quartile in the corresponding box plots.



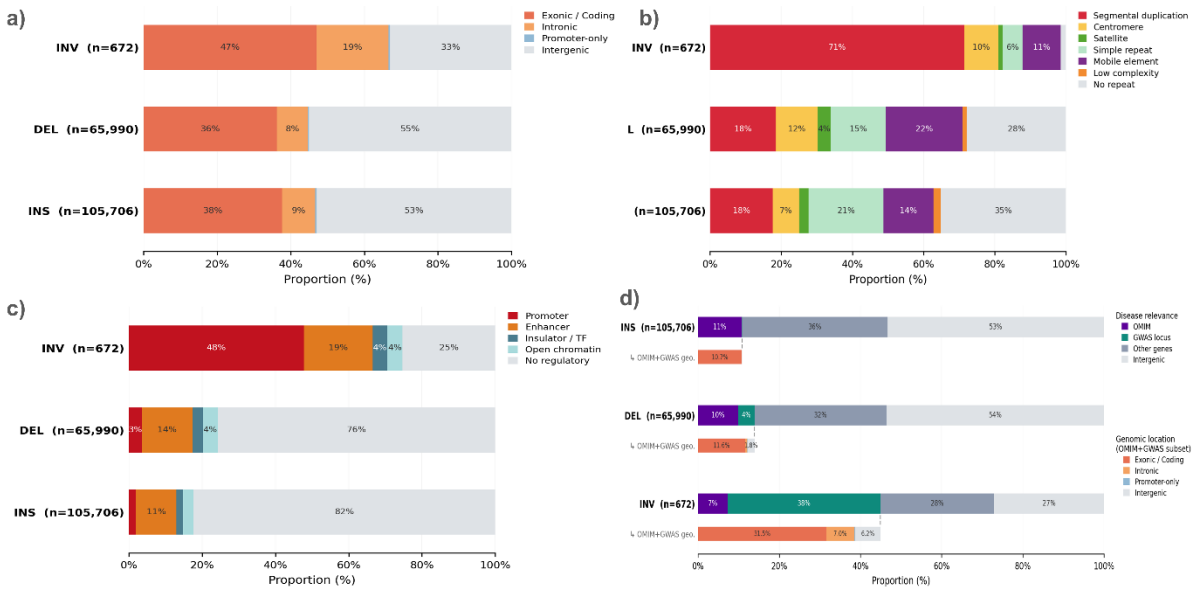
Supplementary Figure 2. Length distribution of telomeric repeats found in HPRC and KPP genome assemblies. Violin plots describe the distribution of telomeric repeat lengths within each group, with dots representing data points for telomeric regions of haplotype genome assemblies. Red and blue arrowheads highlight T2T-CHM13 and near-T2T-level KPP samples, respectively. The numbers above the violin plots indicate the frequency of telomeric repeat detection at chromosomal ends for each group.



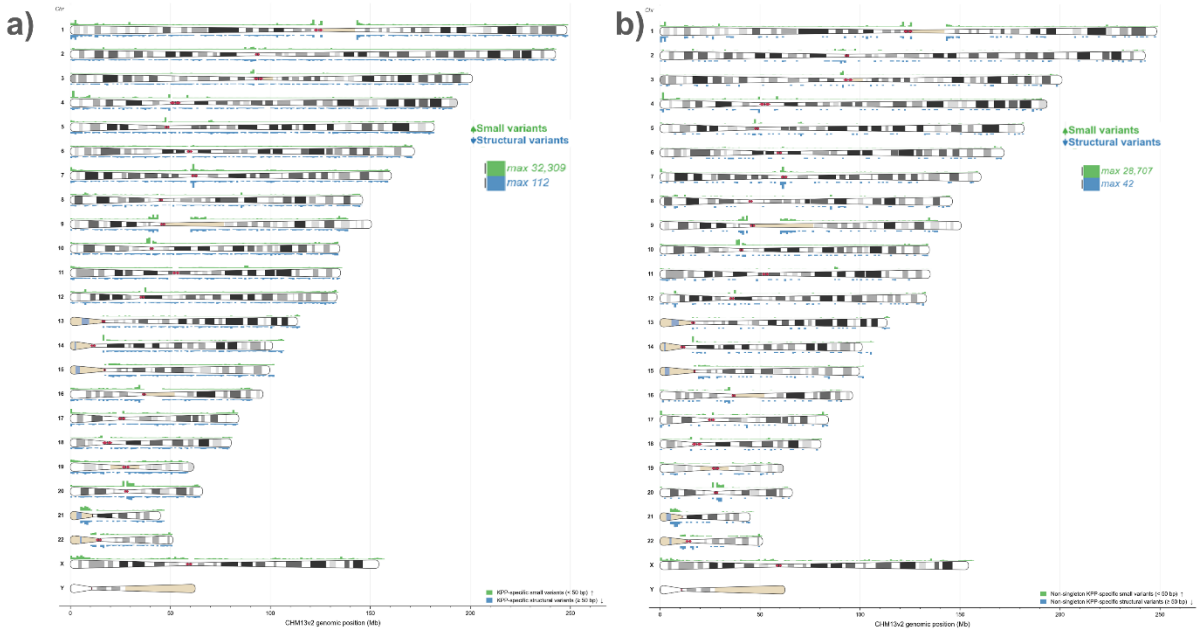
Supplementary Figure 3. Comparison of telomeric region and T2T chromosome counts in near-T2T-level KPP samples (KPPD129, KPPD130, KPPD131, and KPPD132). Each dot represents an individual, and dashed lines connect values for the same individual.



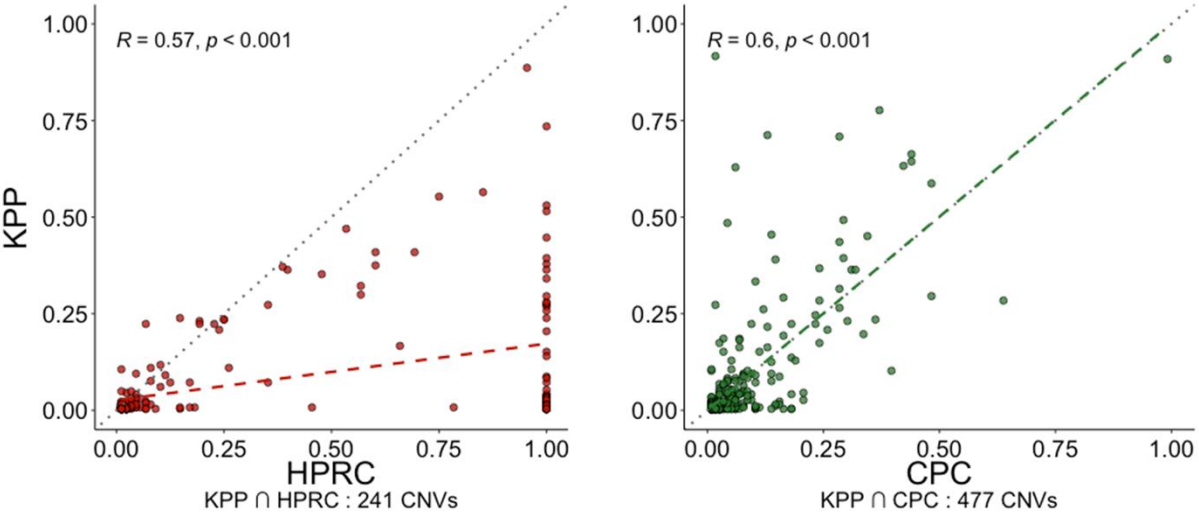
Supplementary Figure 4. Saturation portrait of cumulative variant discovery across the Korean Pangenome Pilot cohort. Curves show the fraction of the cohort-total variant set captured as individuals were sequentially added to the Korean Pangenome Pilot cohort. SNVs, small indels, and structural variants (SVs) are shown separately to illustrate class-specific saturation behavior across 132 individuals.



Supplementary Figure 5. Annotation composition of PAV variants by structural-variant class. a) Genomic location composition across inversion (INV), deletion (DEL), and insertion (INS) variants. Stacked bars show the proportion of variants assigned to exonic or coding sequence, intronic sequence, promoter-only sequence, or intergenic sequence. **b)** Repeat-region composition across the same variant classes. Stacked bars show the proportion of variants assigned to segmental duplication, centromeric sequence, satellite sequence, simple repeat sequence, mobile elements, low-complexity sequence, or no repeat annotation. **c)** Regulatory annotation overlap across the same variant classes. Stacked bars show the proportion of variants assigned to promoter, enhancer, insulator or transcription-factor-associated elements, open chromatin, or no regulatory annotation. **d)** Disease-relevance composition across the same variant classes. Upper bars show the proportion of variants assigned to OMIM genes, GWAS loci, other genes, or intergenic sequence, and lower bars show the genomic-location composition of the OMIM plus GWAS subset, scaled to its proportion within each variant class.

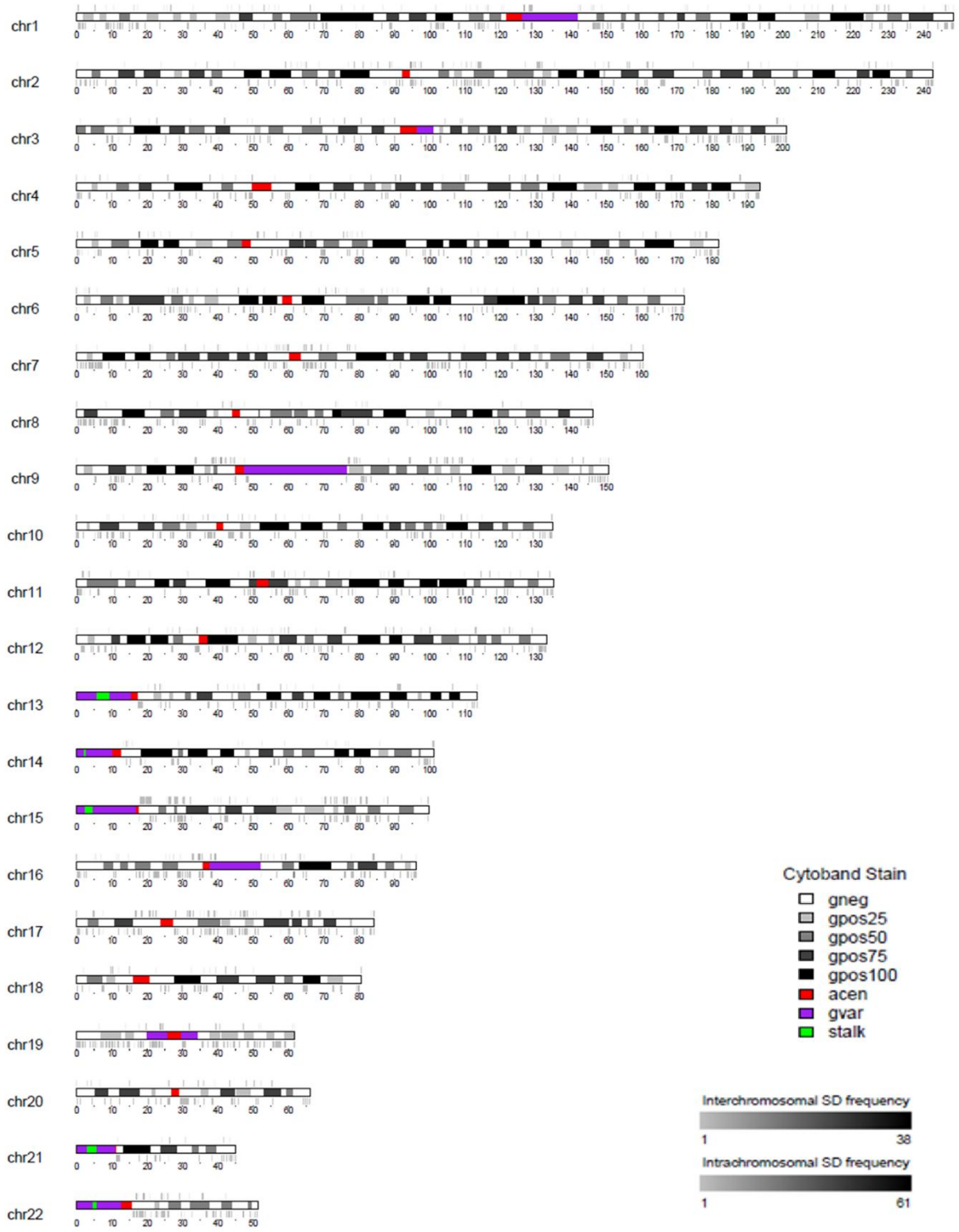


Supplementary Figure 6. Chromosomal distribution of assembly-based KPP-specific variants. **a)** Chromosomal distribution of assembly-based KPP-specific variants across the T2T-CHM13 (v2.0) genome. Green histograms indicate small variants and blue histograms indicate structural variants, plotted above and below each chromosome ideogram, respectively. **b)** Chromosomal distribution of the non-singleton subset of assembly-based KPP-specific variants across the same genome. Green histograms indicate non-singleton small variants and blue histograms indicate non-singleton structural variants, plotted above and below each chromosome ideogram, respectively.

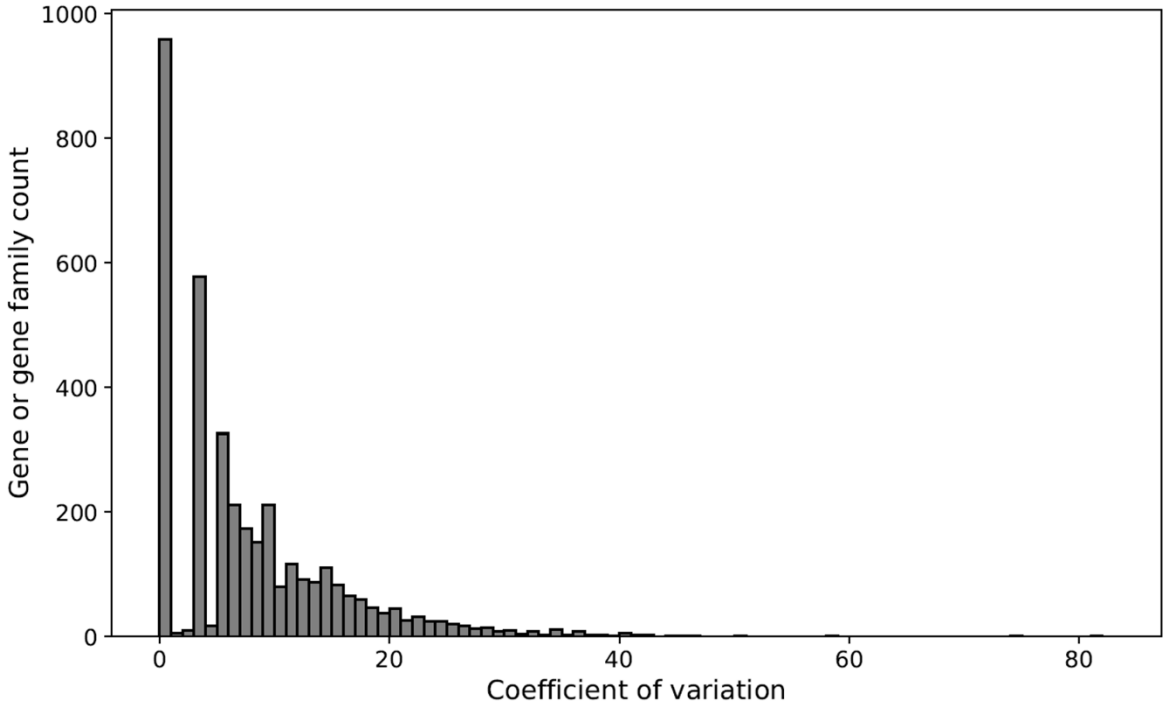


Supplementary Figure 7. CNV frequencies were compared between KPP and two external pangenome resources, HPRC and CPC. For KPP, frequency was calculated as the proportion of 264 haplotype assemblies carrying each CNV. HPRC and CPC frequencies were based on

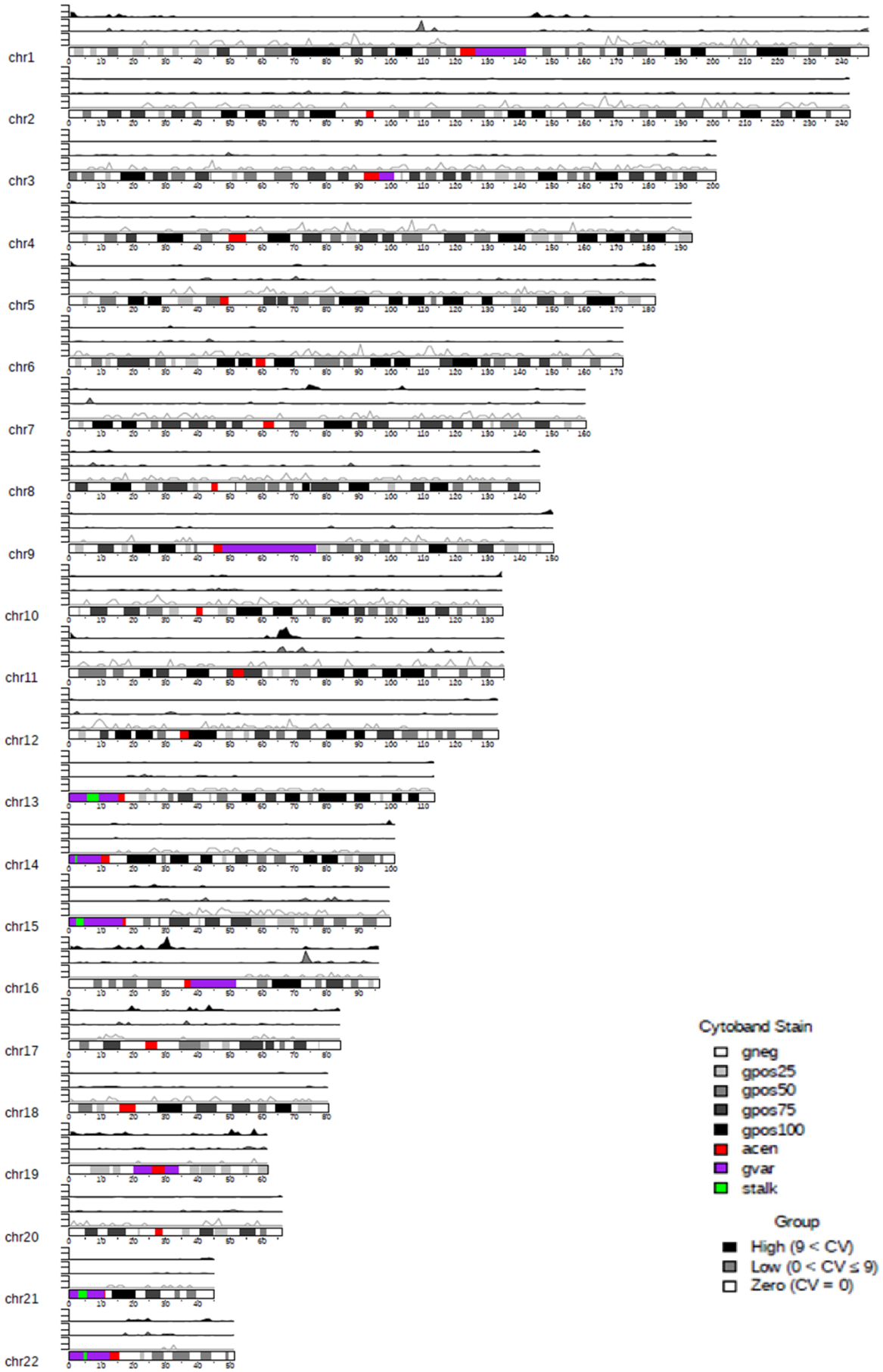
the reported CNV frequency values from the corresponding supplementary tables of each study. Spearman correlation coefficients and corresponding p-values are shown in each panel.



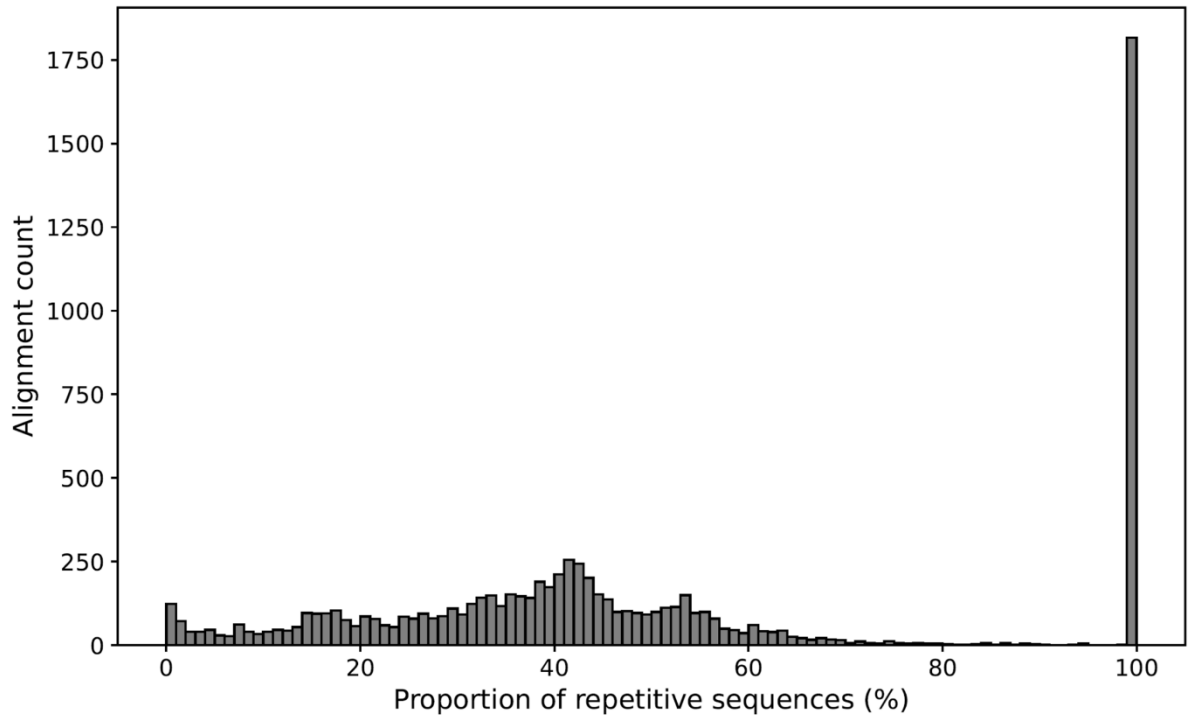
Supplementary Figure 8. Frequency distribution of KPP-specific SDs from haplotype genome assemblies. Data shown here correspond to the KPP-specific segmental duplications from Figure 3a.



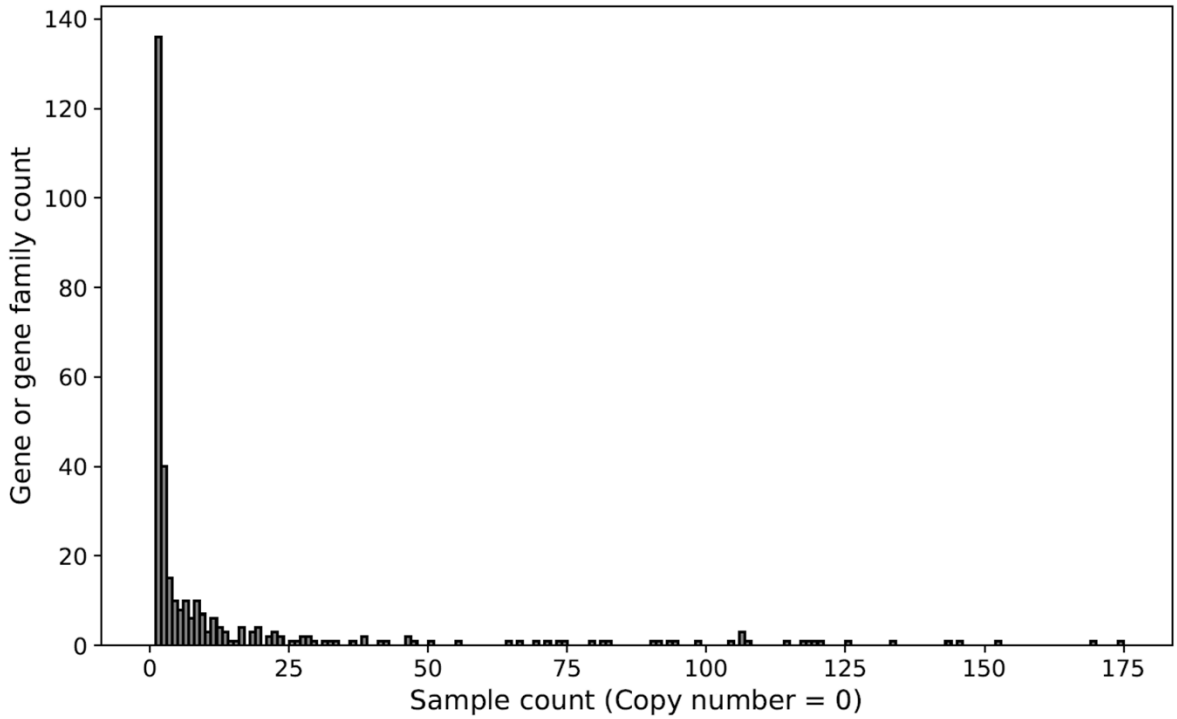
Supplementary Figure 9. Histogram of copy number coefficient of variation (CV) for genes or gene families. T2T-CHM13 was excluded from CV calculations because its copy number was counted from only a single haplotype genome assembly.


chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr19
chr20
chr21
chr22
Cytoband Stain
gneg
gpos25
gpos50
gpos75
gpos100
acen
gvar
stalk
Group
High (9 < CV)
Low (0 < CV ≤ 9)
Zero (CV = 0)

Supplementary Figure 10. Chromosomal distribution of genes or gene families categorized by coefficient of variations (CVs) of copy numbers. Each horizontal box represents a chromosome, with genomic positions labeled below in Mb. The three density plots above the chromosome ideograms illustrate the distribution of gene or gene family counts per 1 Mb non-overlapping window of high, low, and zero CV groups from top to bottom. The y-axis ranges from zero to maximum density values for each group.T2T- CHM13 was omitted from the CV analysis.



Supplementary Figure 11. Repetitive sequence proportion in genic alignment regions. Alignments from T2T-CHM13 passing coverage and identity filters are presented here.



Supplementary Figure 12. Histogram of genes or gene families with zero copy numbers in at least one individual. Genes or gene families with copy number = 0 in ≥16 samples excluded as artifacts.