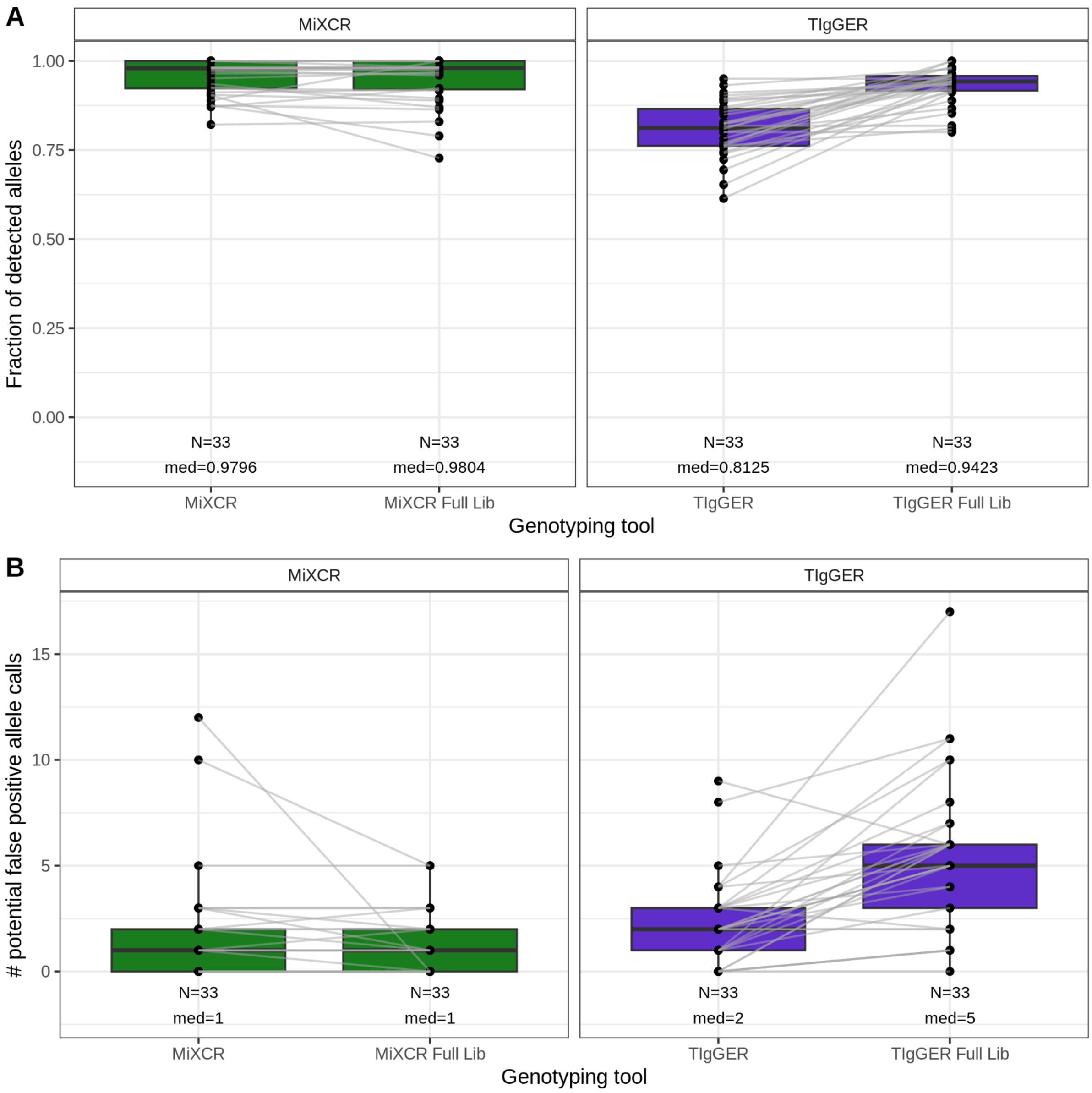
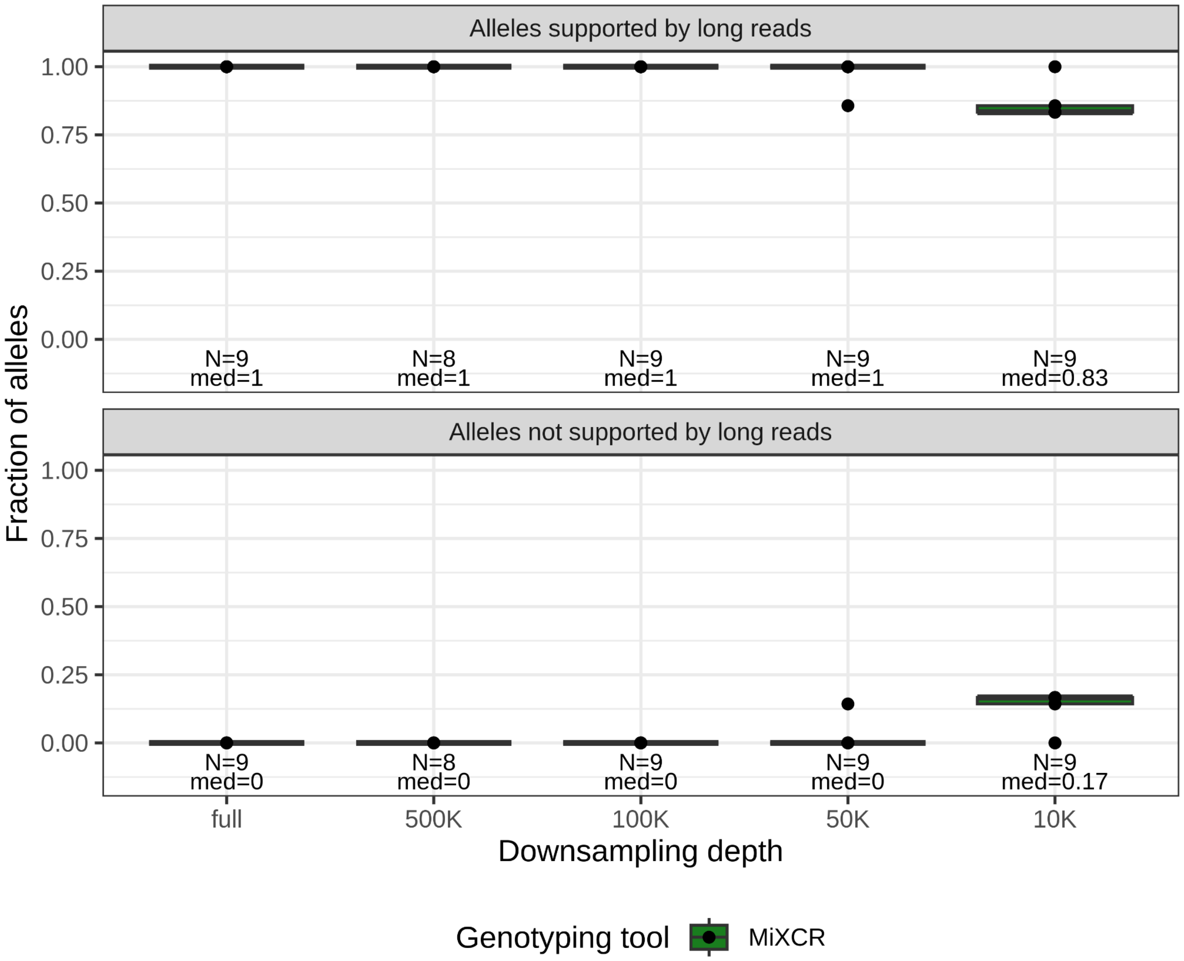
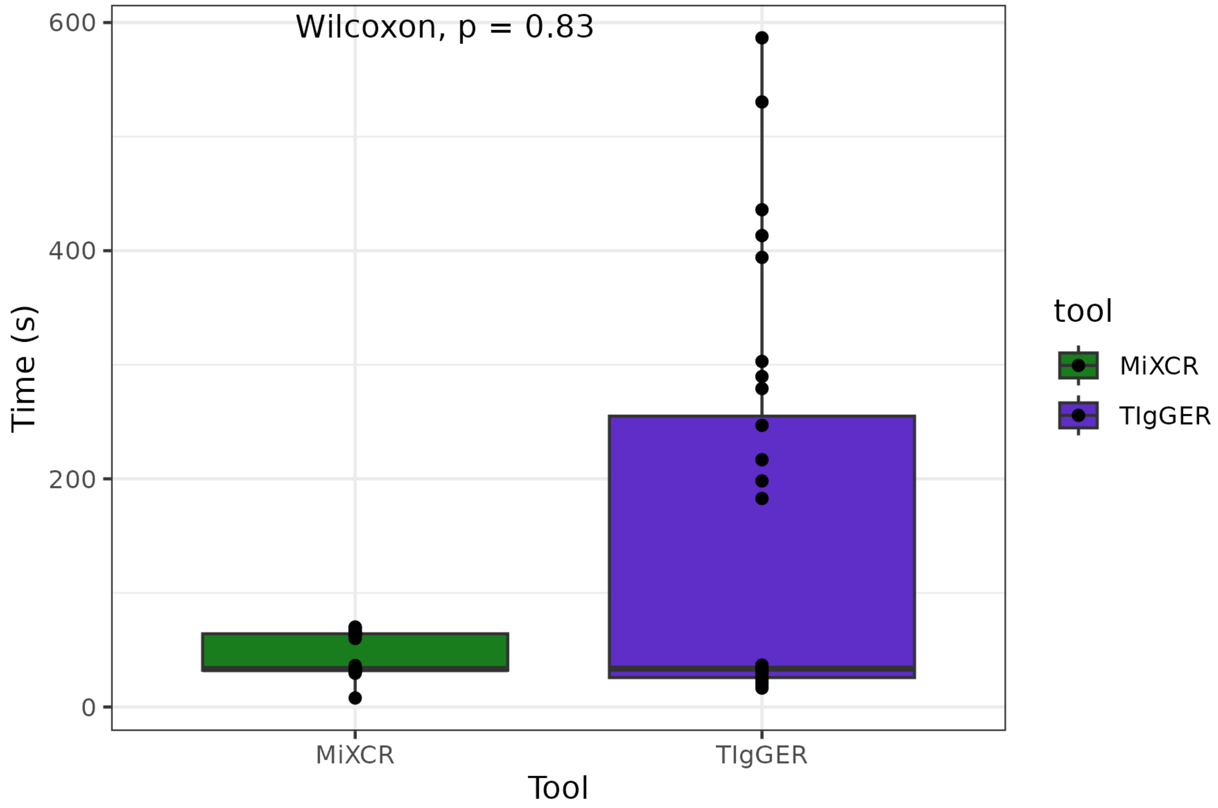


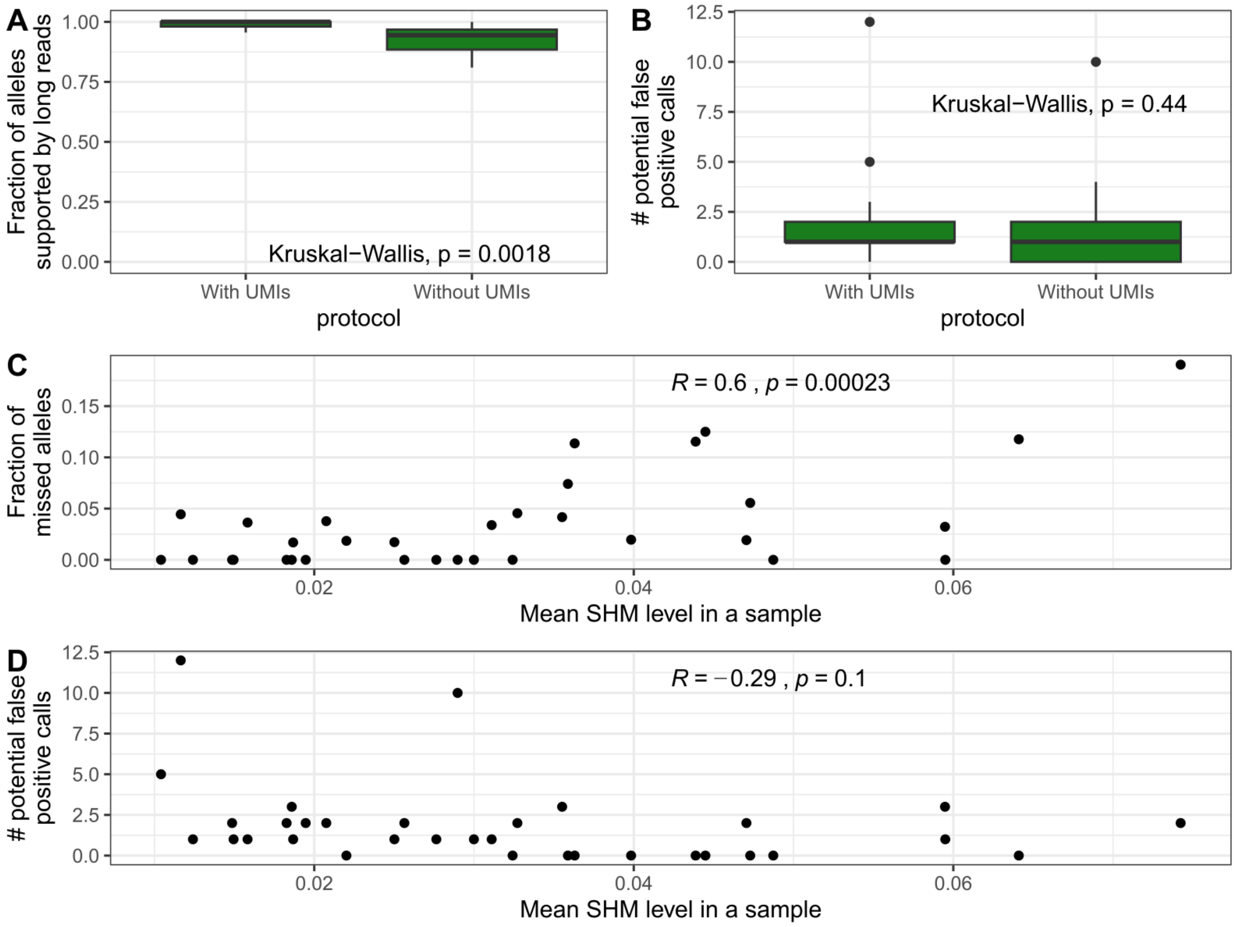
Supplemental Figure S1. Allele tool inference benchmarking using full reference libraries for each of the tools.



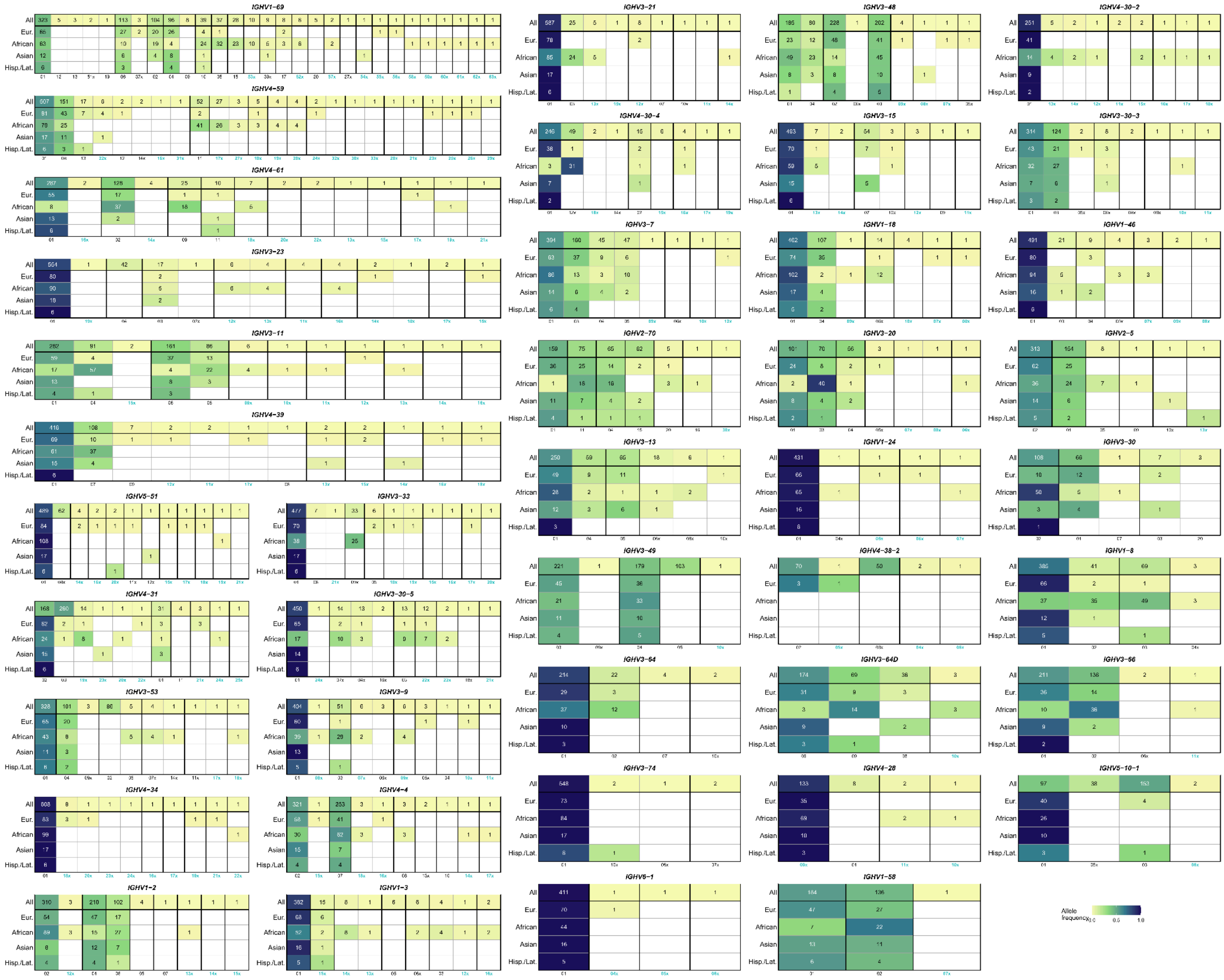
Supplemental Figure S2. Allele inference and genotyping of J gene allelic variants by MiXCR.



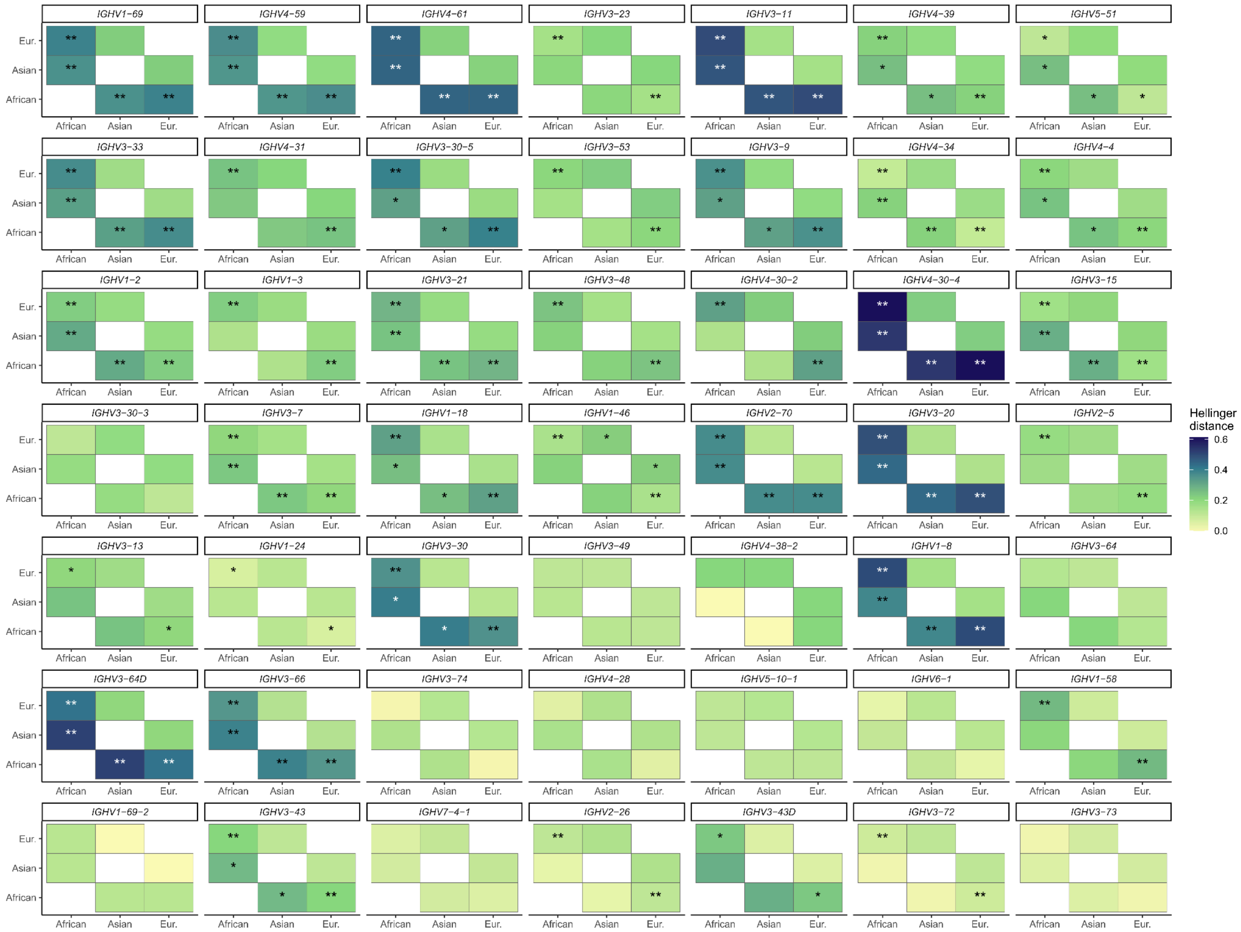
Supplemental Figure S3. Runtime comparison for allele inference and genotyping of V gene allelic variants with MiXCR and TIgGER.



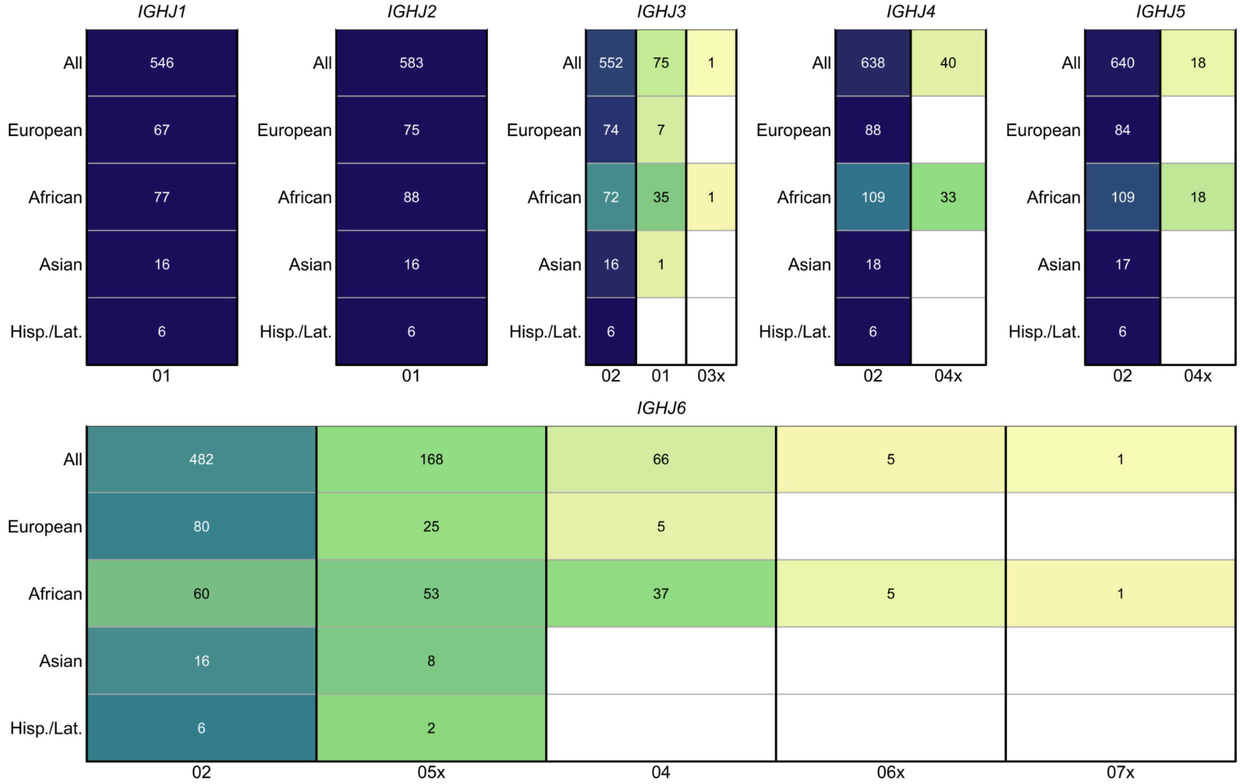
Supplemental Figure S4. Allele inference and genotyping of V gene allelic variants with MiXCR depending on SHM level and UMIs utilization. In a and b, the Kruskal-Wallis test is used for statistical comparison; in c and d, Pearson correlation coefficients (R) with corresponding p-values (p) are used.



Supplemental Figure S5. IGHV gene allele frequencies in major ethnic groups for expressed IGHV genes. Each column in heatmaps represents a particular allele; numbers for novel alleles, first reported in this study, are colored in green; “x” letter designates alleles inferred from AIRR-seq data (either in this study or previously, with the same sequences already present in OGRDB); bold lines separate groups of alleles with different amino acid sequences; groups of alleles with the same amino acid sequence are ordered by the aggregated frequencies of alleles; alleles within groups are order by allele frequency in the general population. Color represents the allele frequency within the ethnic group; numbers in cells represent the number of occurrences of the corresponding allele.



Supplemental Figure S6. Hellinger distance of allele frequency distributions between major ethnic groups. Comparison between ethnicities in each gene was performed using permutation test (1000 permutations, * = p ≤ 0.05, ** = p ≤ 0.01, non-significant not shown)



IGHJ1

	01
All	546
European	67
African	77
Asian	16
Hisp./Lat.	6

IGHJ2

	01
All	583
European	75
African	88
Asian	16
Hisp./Lat.	6

IGHJ3

	02	01	03x
All	552	75	1
European	74	7	
African	72	35	1
Asian	16	1	
Hisp./Lat.	6		

IGHJ4

	02	04x
All	638	40
European	88	
African	109	33
Asian	18	
Hisp./Lat.	6	

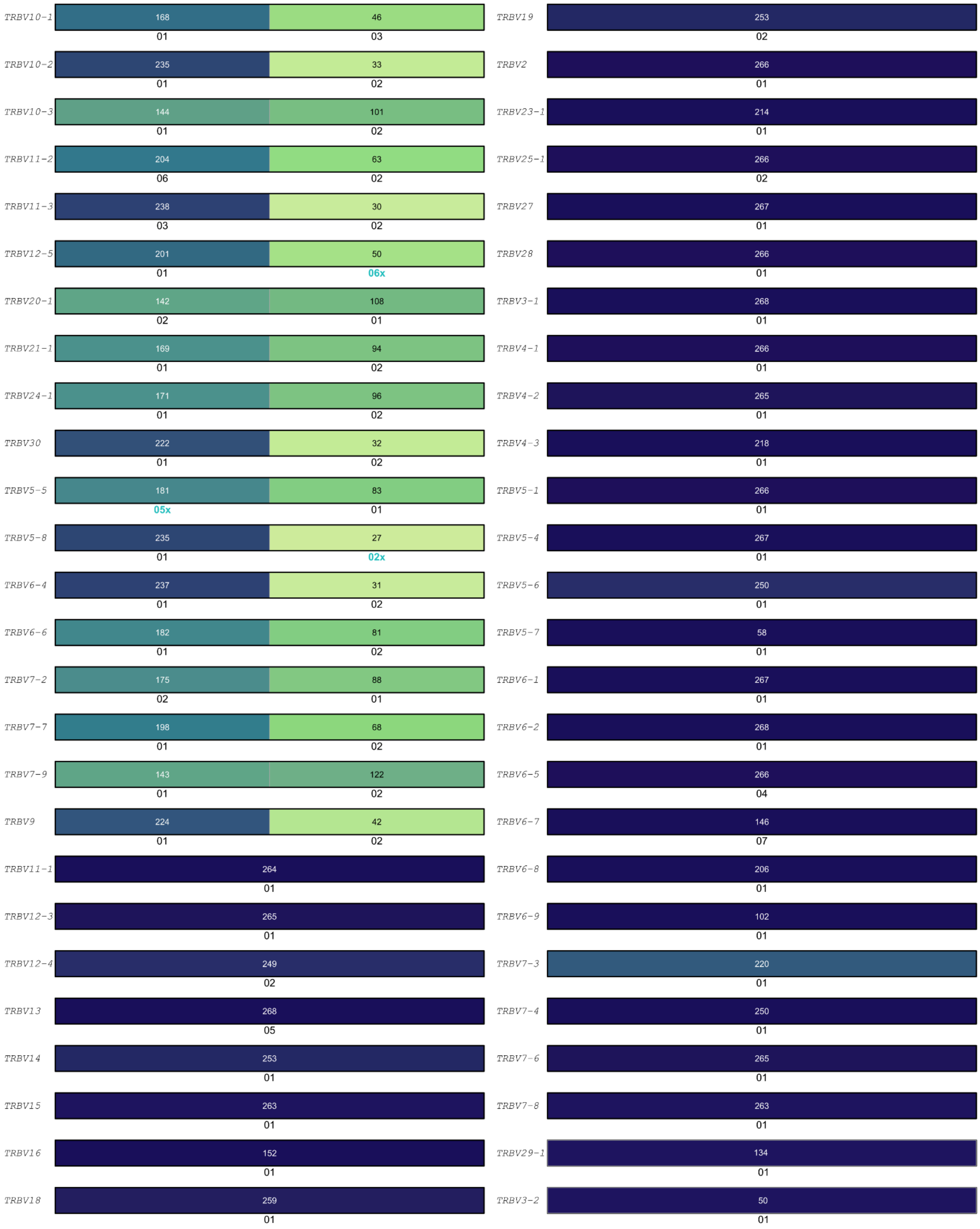
IGHJ5

	02	04x
All	640	18
European	84	
African	109	18
Asian	17	
Hisp./Lat.	6	

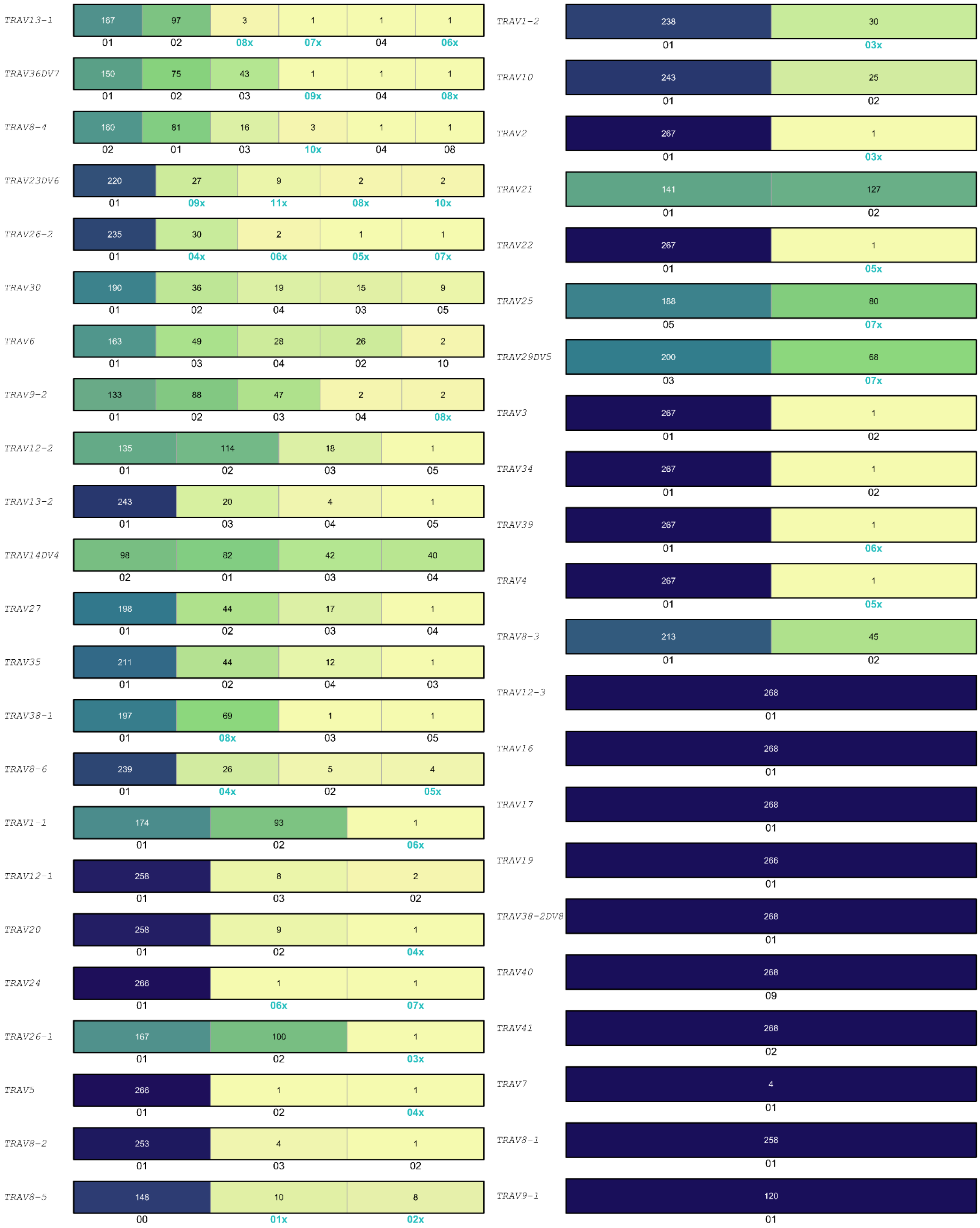
IGHJ6

	02	05x	04	06x	07x
All	482	168	66	5	1
European	80	25	5		
African	60	53	37	5	1
Asian	16	8			
Hisp./Lat.	6	2			

Supplemental Figure S7. *IGHJ* gene allele frequencies in major ethnic groups. Each column in heatmaps represents a particular allele, each row - major ethnic group. The letter 'x' designates alleles inferred from AIRR-seq data, with the sequences present in OGRDB. The alleles are ordered by allele frequency in the general population. Color represents the allele frequency within ethnic groups; numbers in cells represent the number of occurrences of the corresponding allele.



Supplemental Figure S8. *TRBV* gene allele frequencies. Each column in heatmaps represents a particular allele. Numbers for novel alleles are colored in green, and concatenated with letter ‘x’; alleles are ordered by allele frequency in the general population. Color represents the allele frequency within the general population; numbers in cells represent the number occurrences of the corresponding allele.



Supplemental Figure S9. *TRAV* gene allele frequencies. Each column in heatmaps represents a particular allele. Numbers for novel alleles are colored in green, and concatenated with letter ‘x’; alleles are ordered by allele frequency in the general population. Color represents the allele frequency within the general population; numbers in cells represent the number of occurrences of the corresponding allele.