

# **Near-complete Middle Eastern genomes refine autozygosity and enhance disease-causing and population-specific variant discovery**

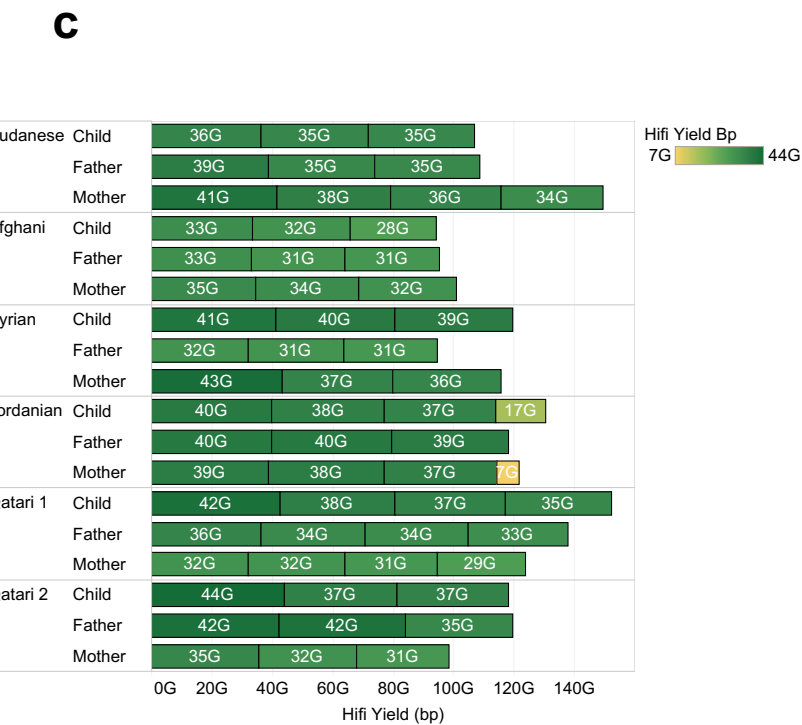
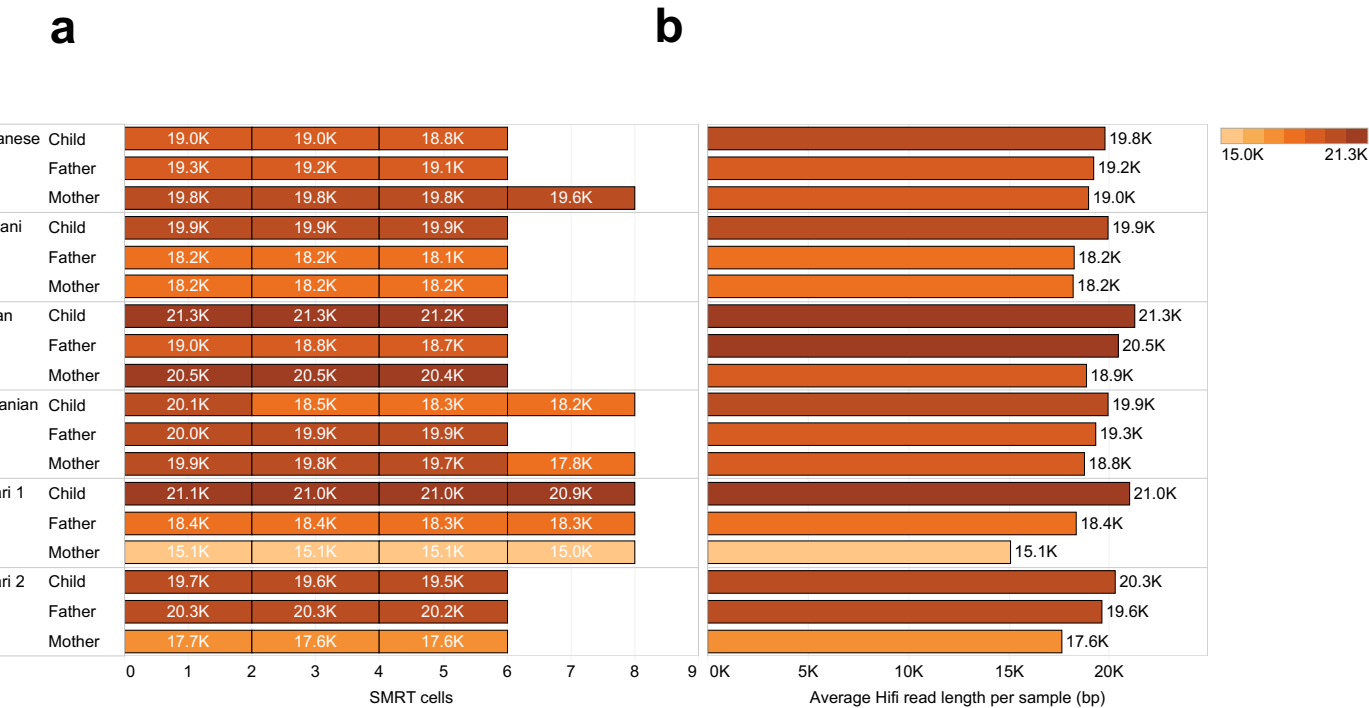
---

In the format provided by the  
authors and unedited

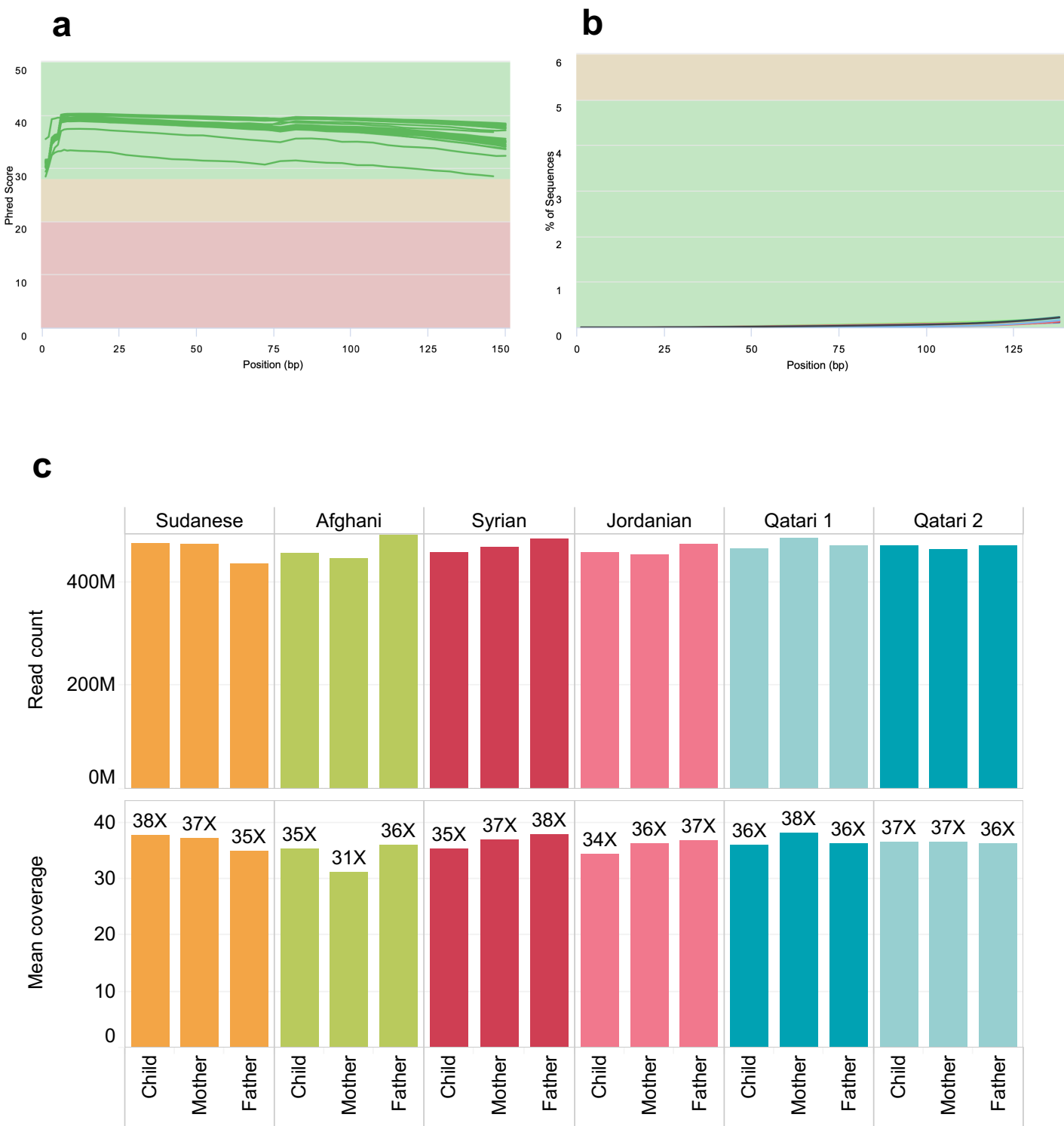
# **Contents**

## Supplementary Figures

**Supplementary Fig. 1: HiFi read data QC metrics. (a)** Number of SMRT cells used per sample. **(b)** Average HiFi read length per sample (bp) **(c)** Yield for each SMRT cell per sample

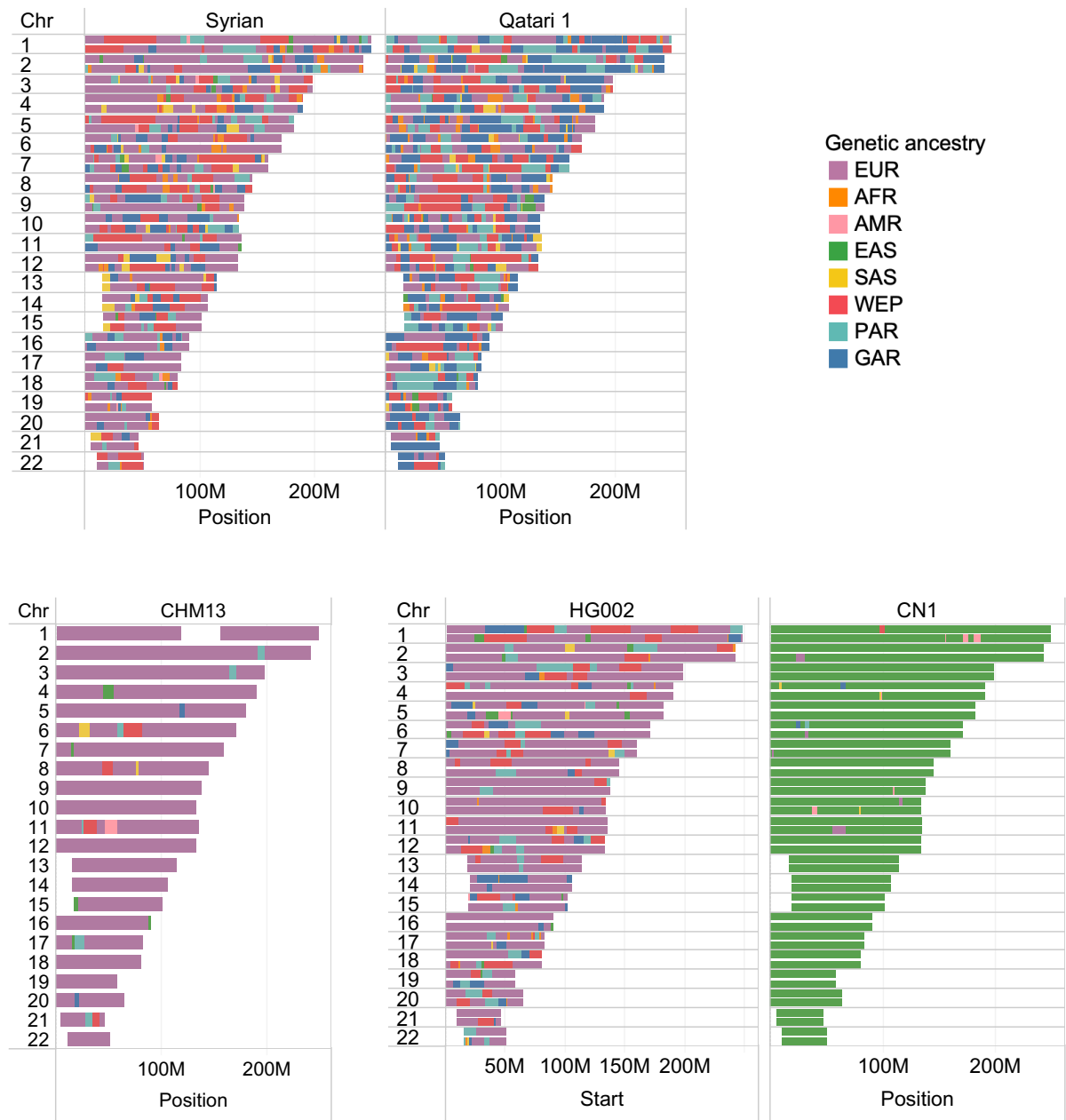


**Supplementary Fig. 2: Illumina read data QC metrics. (a)** Phred quality score distribution over the length of sequencing reads, indicating data quality per base position. **(b)** Distribution of the number of sequence tags over the read length, showing the consistency of tag presence across the length of the reads. **(c)** Read count and mean coverage per sample

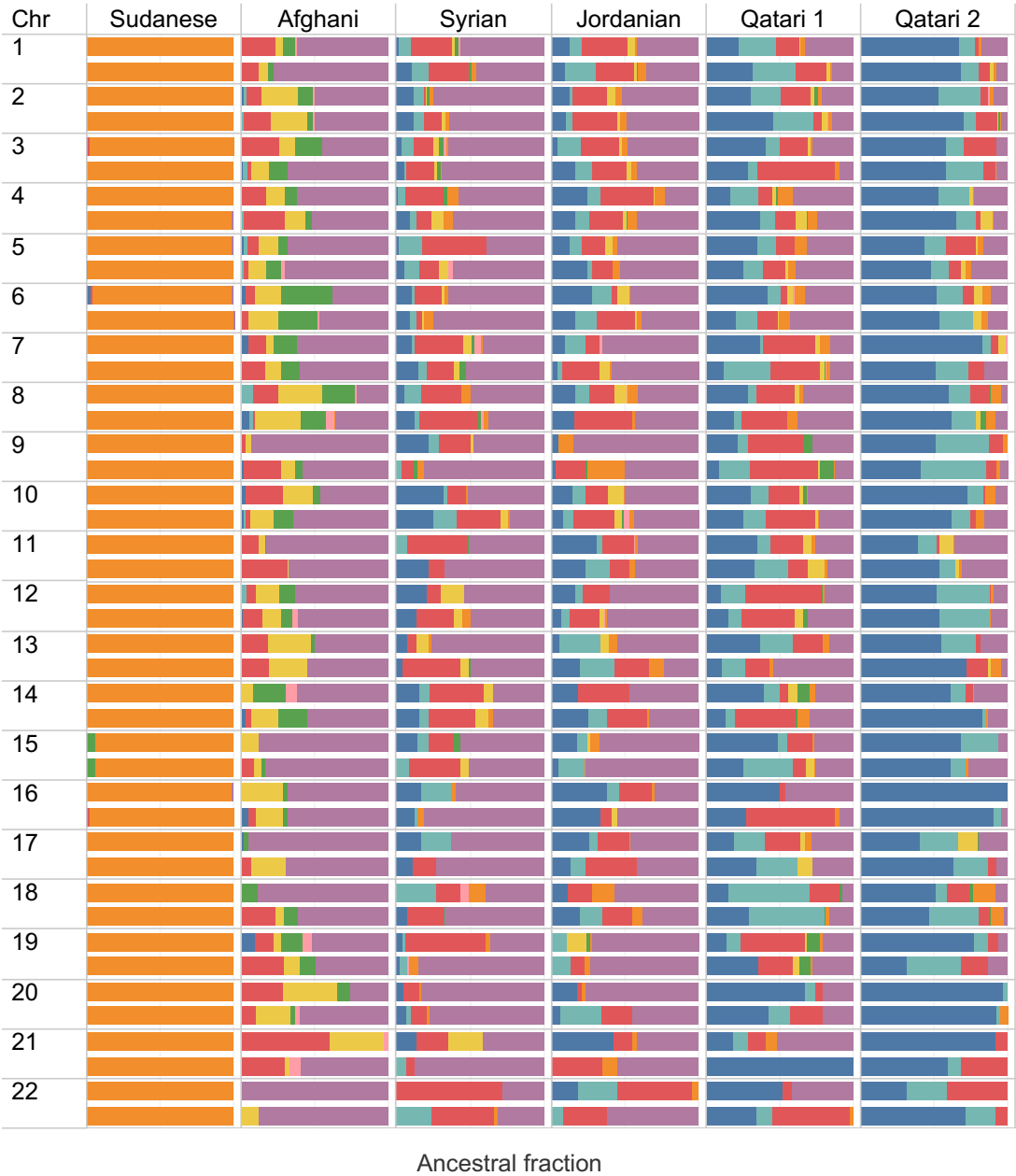


**Supplementary Fig. 3: Local and global genetic ancestry of child subjects and other publicly available T2T genomes. (a)** Local ancestry at chromosomal level for Syrian, Qatari 1, CHM13, HG002 and CN1. **(b)** Global ancestral assignments for all child study subjects

**a**

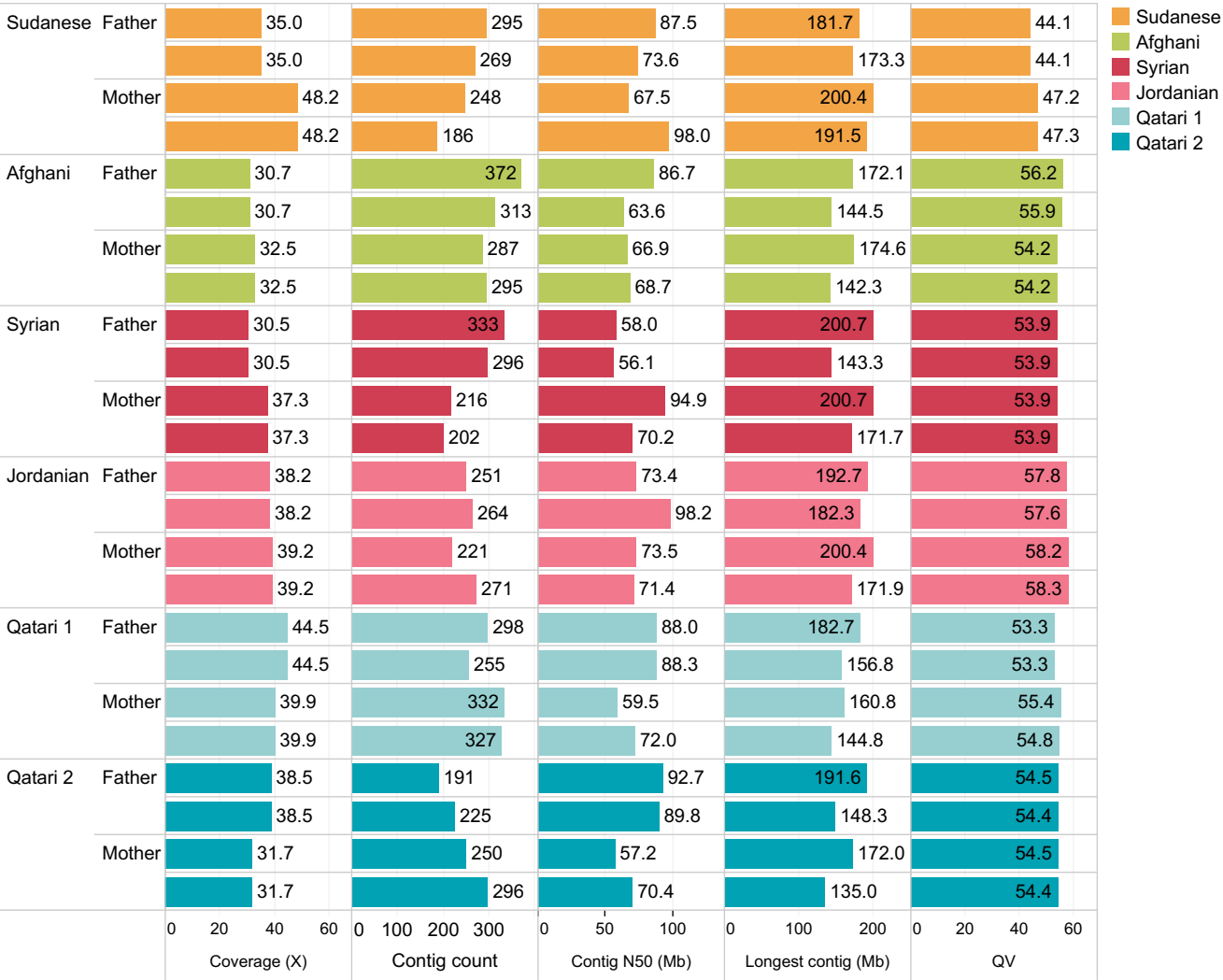


**b**

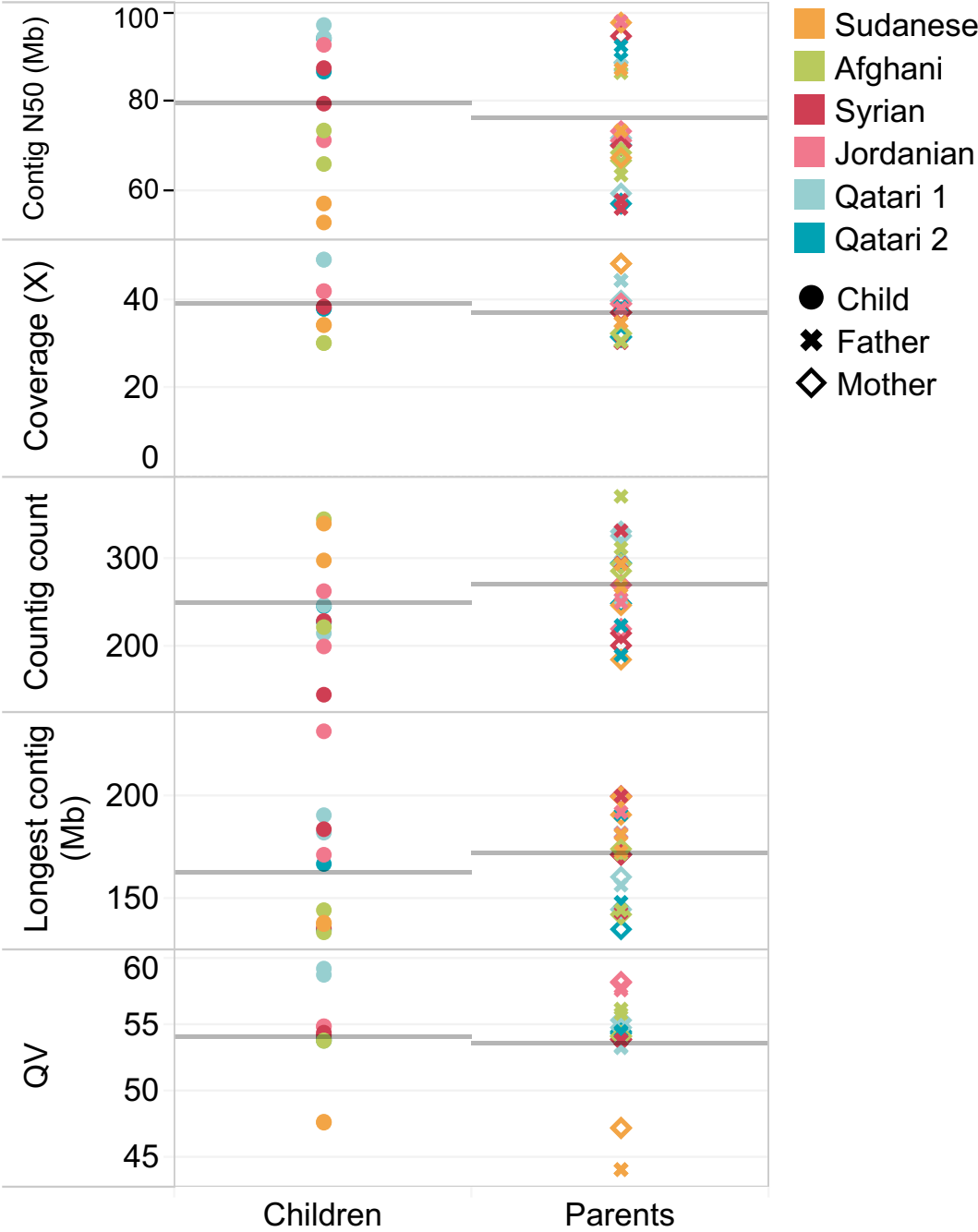


**Supplementary Fig. 4: Assembly results and QC. (a)** Assembly QC of parental subjects, including coverage depth, total contigs, contig N50, maximum contig length and QV. **(b)** Comparison of QC metrics of child and parental assemblies.

**a**

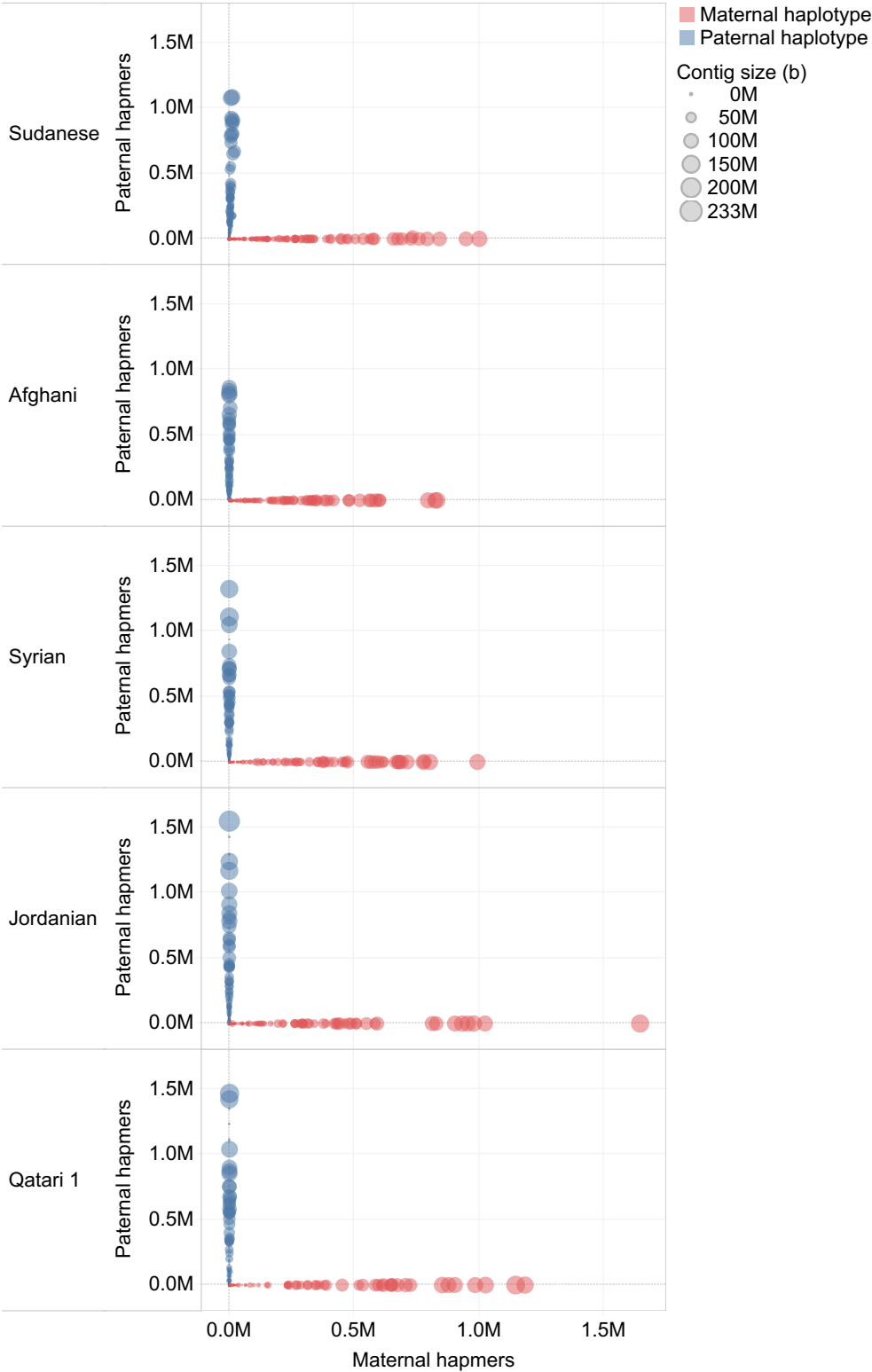


**b**

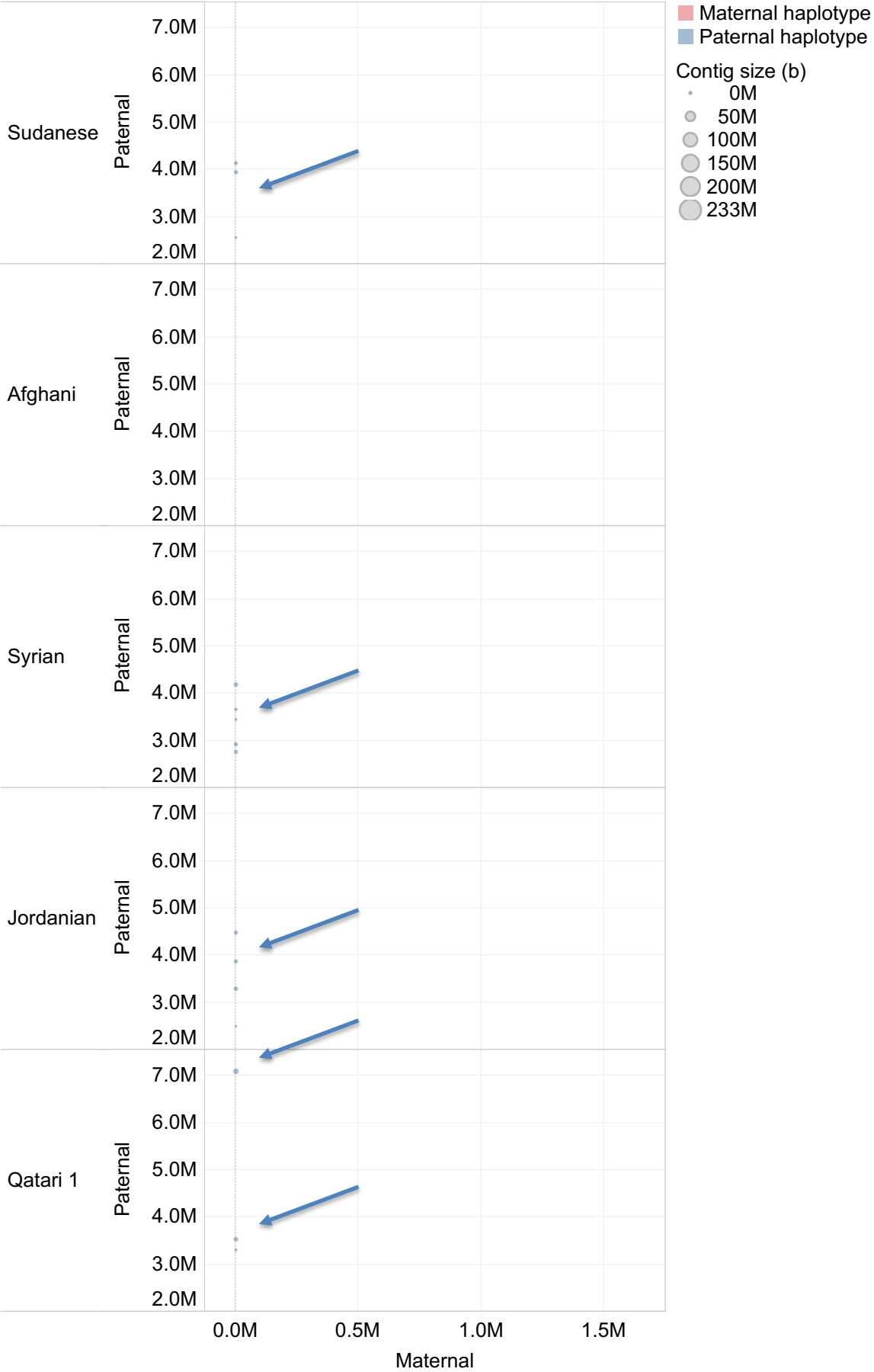


**Supplementary Fig. 5: Hapmer blob plots for other child assemblies (Phasing QC).** (a) Plots show separation of maternal (red) and paternal (blue) haplotypes, the blob size being proportional to contig size. (b) Complementary plot showing data for y axis > 2 M. Arrows point to the data points, for added clarity.

a

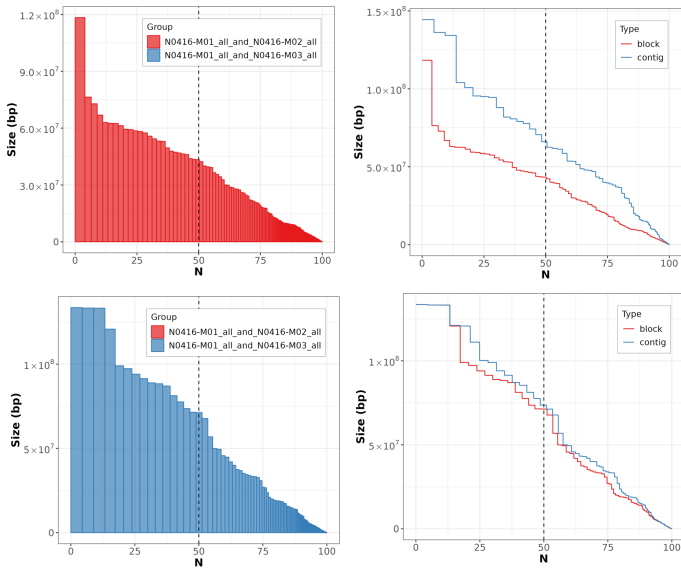


b

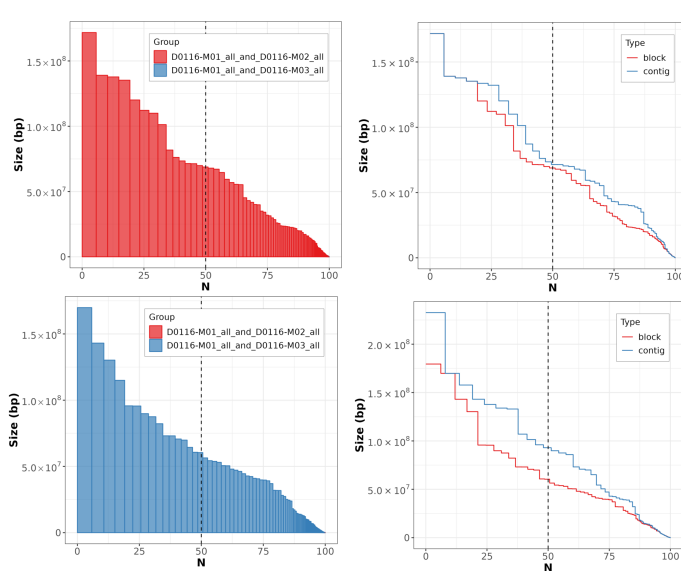


**Supplementary Fig. 6: Switch error plots for other child assemblies (Phasing QC).** Phase block NG plots of haplotype-resolved assembly sorted by size.

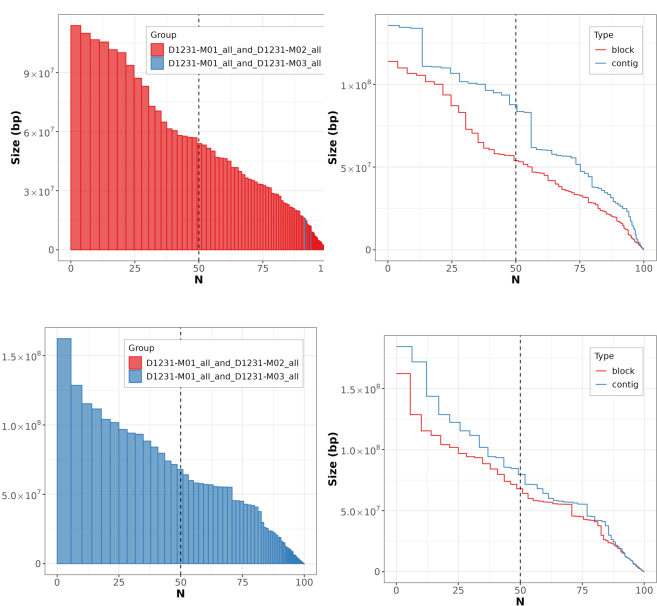
Afghani



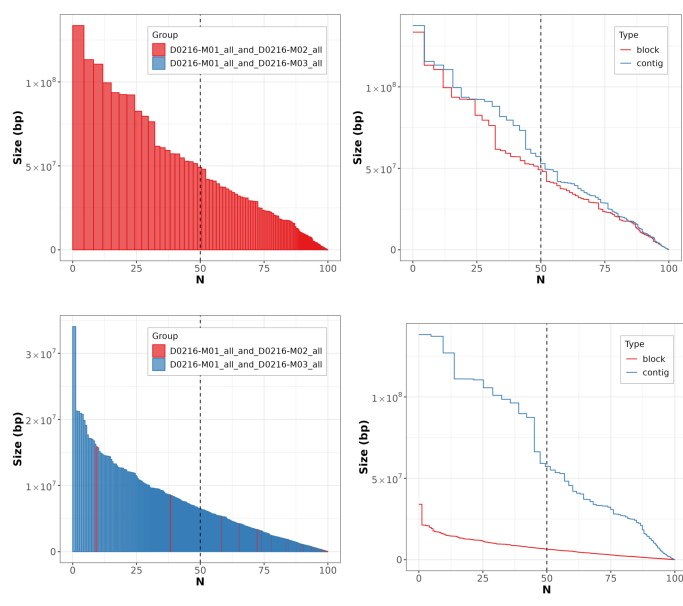
Jordanian



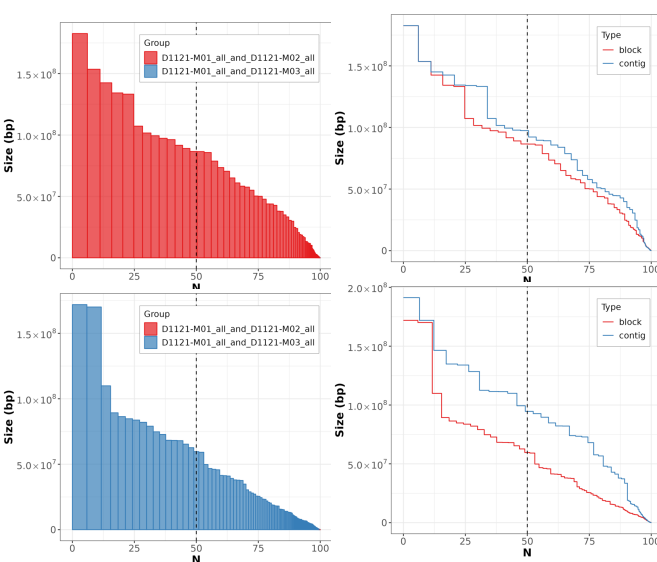
Syrian



Sudanese

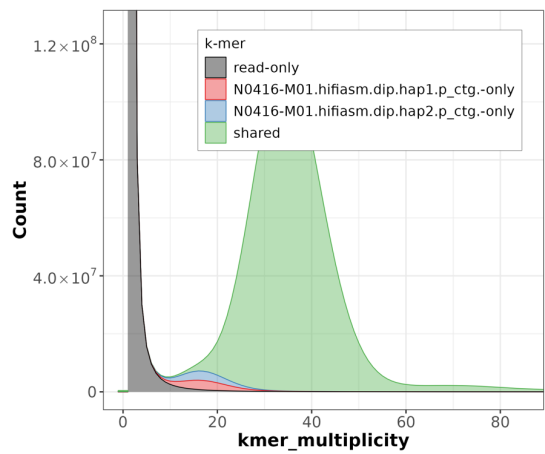


Qatari 1

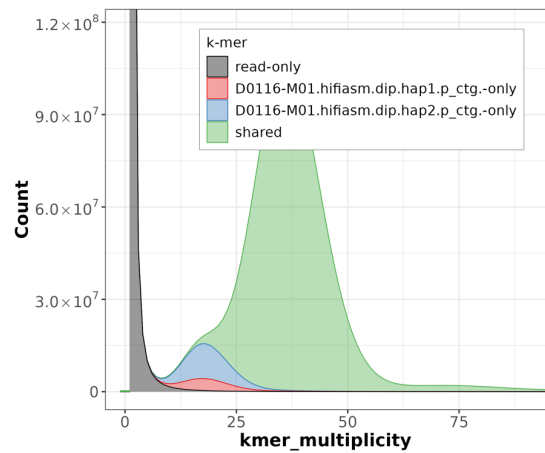


**Supplementary Fig.7.** Inherited hapmer plots for the child assemblies.  
k-mer multiplicity of the child read set colored by paternal (red) and maternal (blue) inheritance and those that are shared (green).

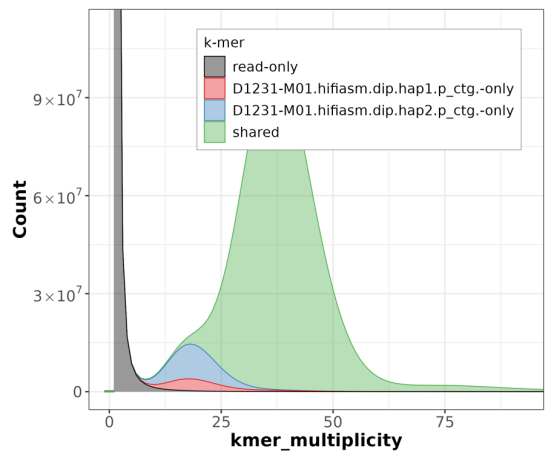
Afghani



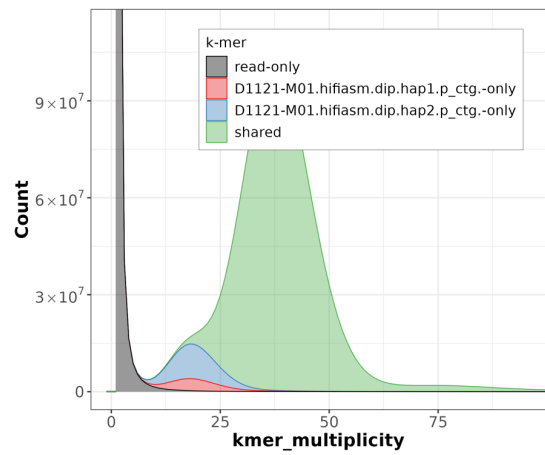
Jordanian



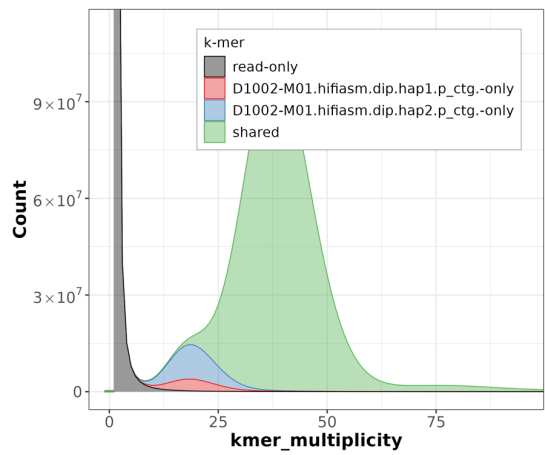
Syrian



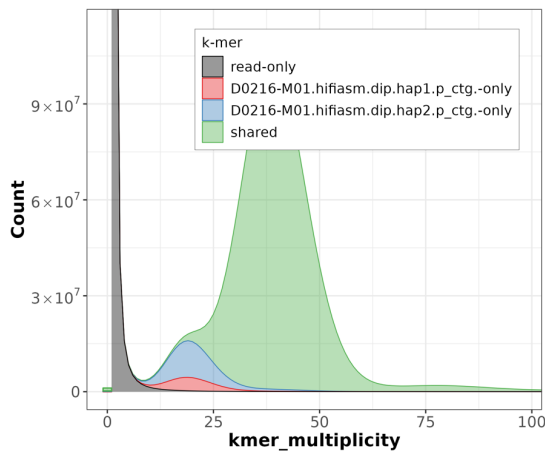
Qatari 1



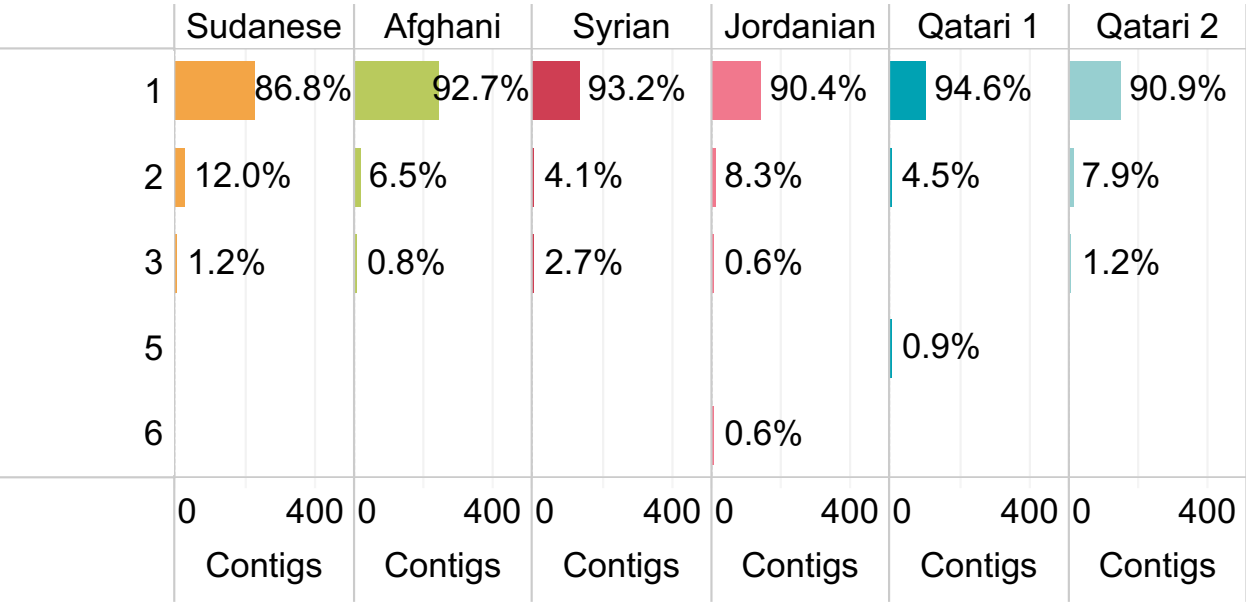
Qatari 2



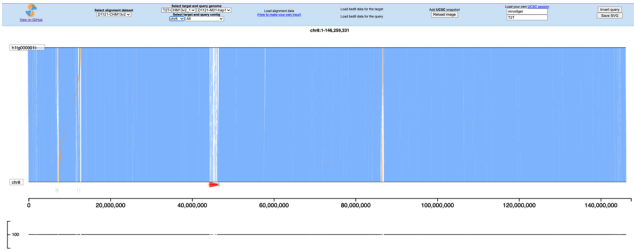
Sudanese



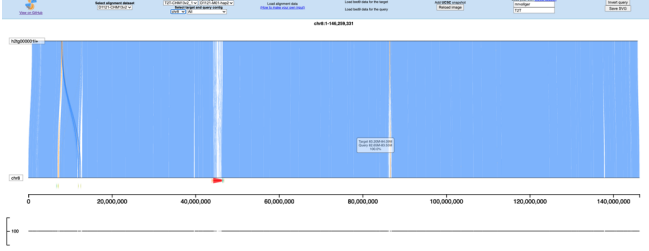
**Supplementary Fig. 8: Number of mapped locations for contigs with alignment coverage  $\geq 0.9$ .**



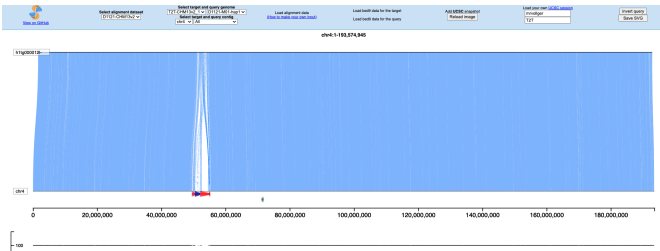
Supplementary Fig. 9: Other chromosome haplotypes spanned by a single contig in the child assemblies. Aligned to CHM13 and visualized on Saffire.



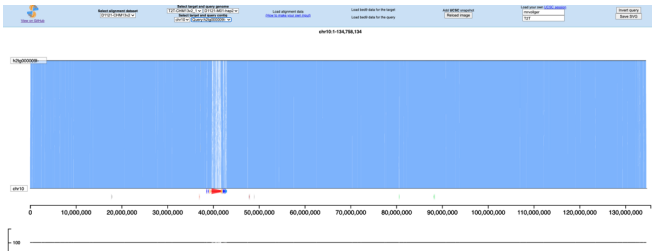
Qatari 1 hap1 :chr8



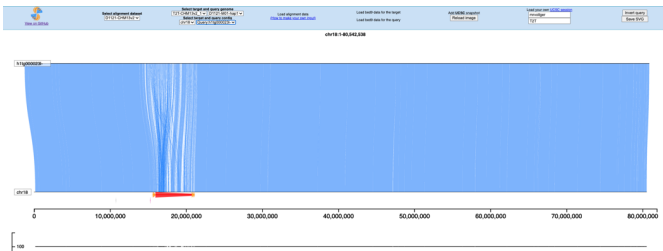
Qatari1 hap2 :chr8



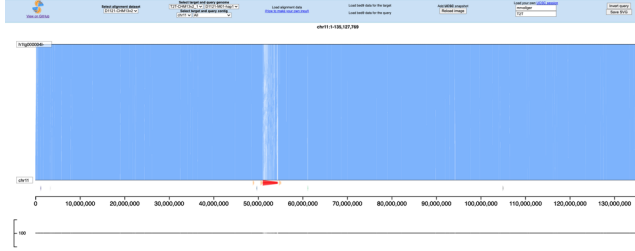
Qatari1 hap1 :chr4



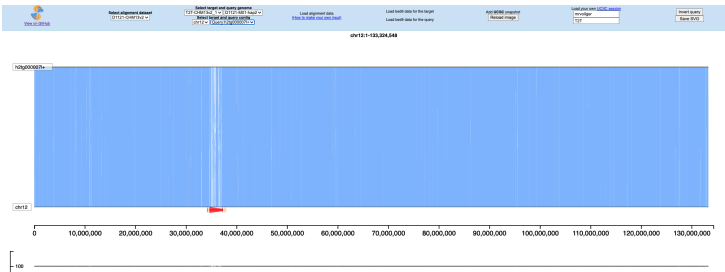
Qatari 1 hap2 :chr10



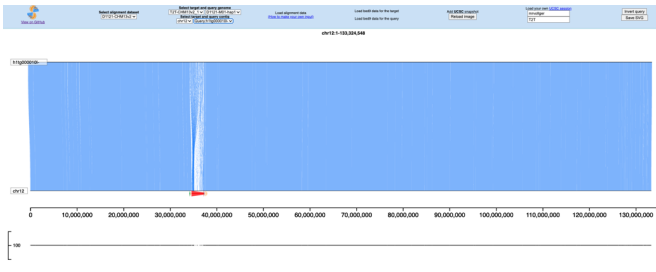
Qatari 1 hap1 :chr18



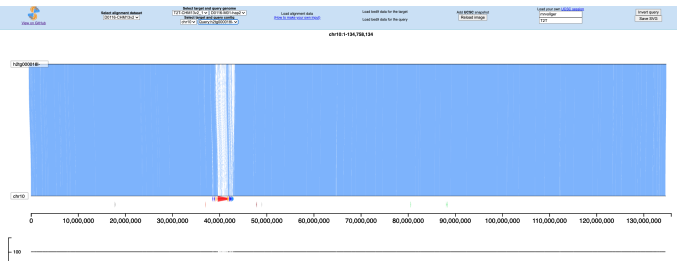
Qatari 1 hap1 :chr11



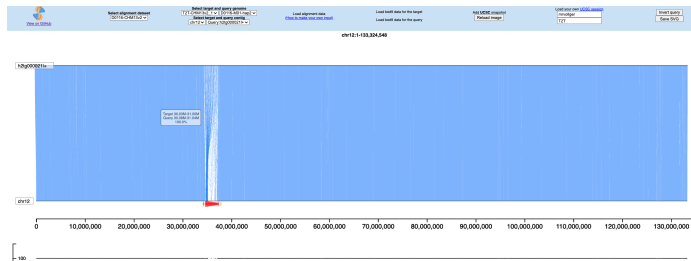
Qatari 1 hap2 :chr12



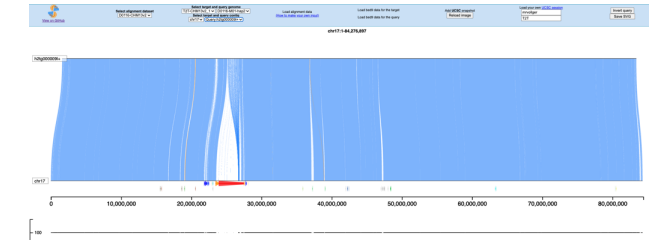
Qatari 1 hap1 :chr12



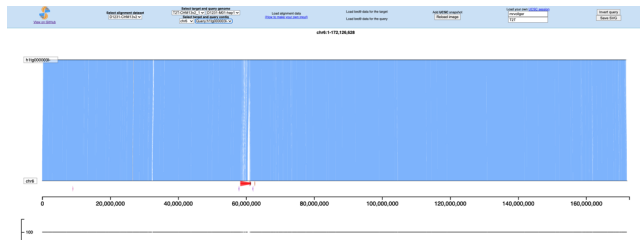
Jordanian hap2 :chr10



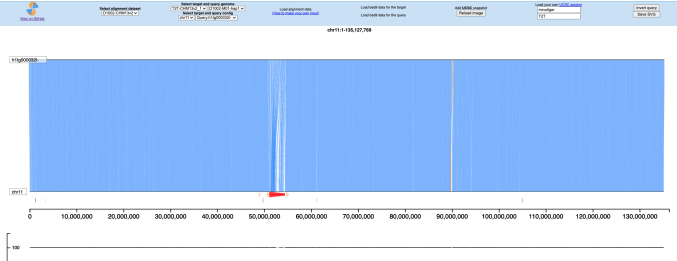
Jordanian hap 2 :chr12



Jordanian hap2 :chr17



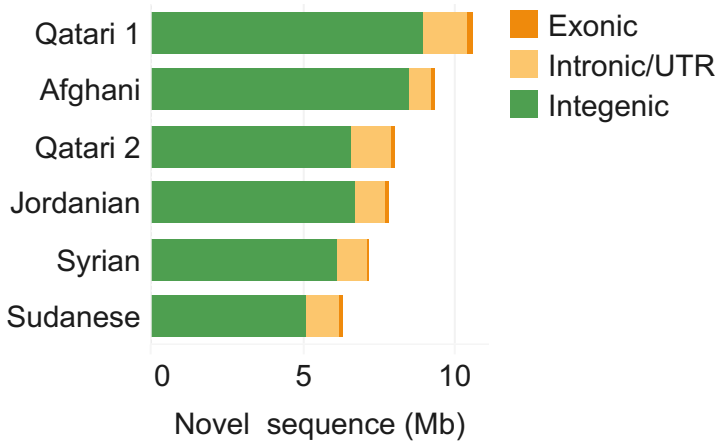
Syrian hap1 :chr6



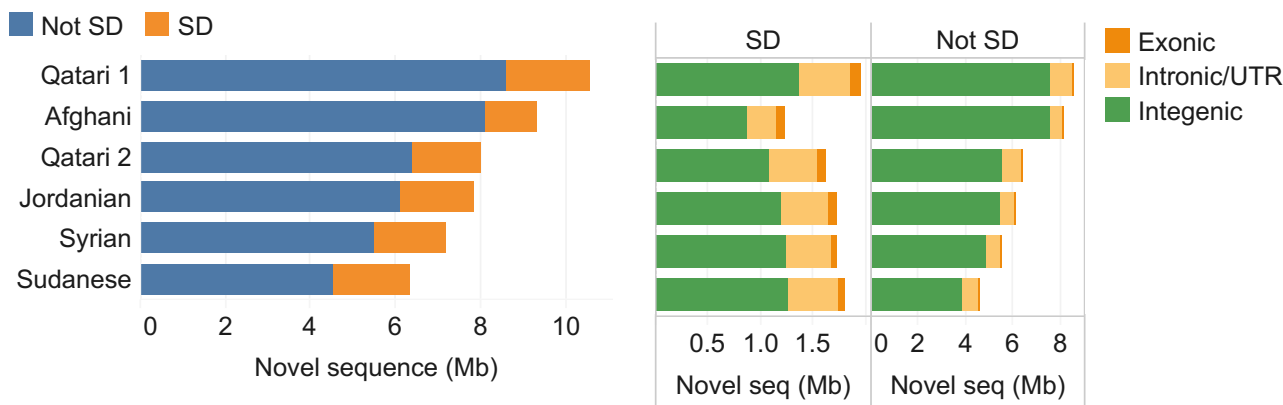
Qatar 2 hap1 :chr11

**Supplementary Fig. 10: Novel sequences occurrence in various genetic features, as a function of comparison to other T2T genomes and patterns across chromosomes. (a)** Total length of novel sequences in intergenic, intronic/UTR and exonic regions **(b)** Total length of novel sequences in segmental duplications (left) and stratified by occurrence in intergenic, intronic/UTR and exonic regions (right) **(c)** Novel sequences identified against CHM13 and HPRC upon incremental comparison with HG002 and CN1. **(d)** Length of novel sequences per chromosome. Dot size indicates the cumulative sequence length, while colors denote distinct assemblies. **(e)** Spread across the chromosomes highlighting the centromeric regions.

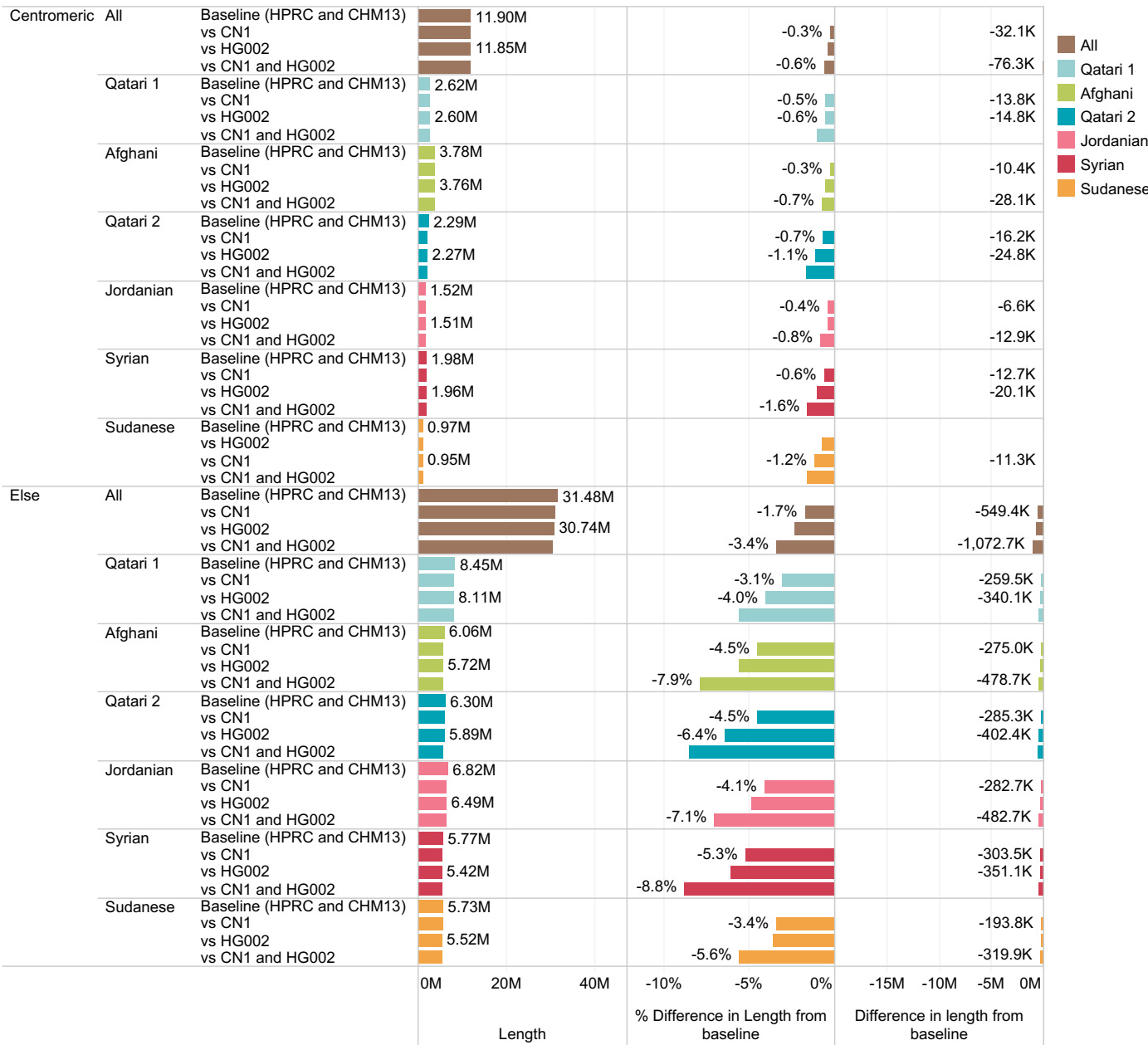
**a**



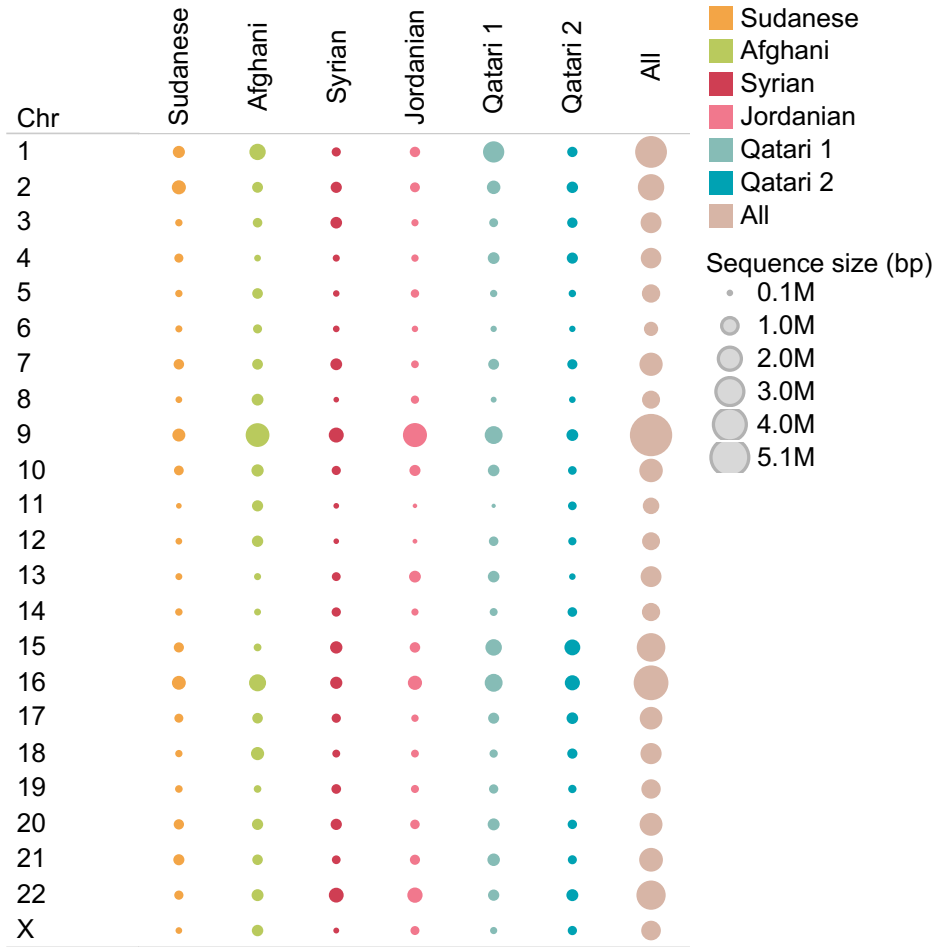
**b**



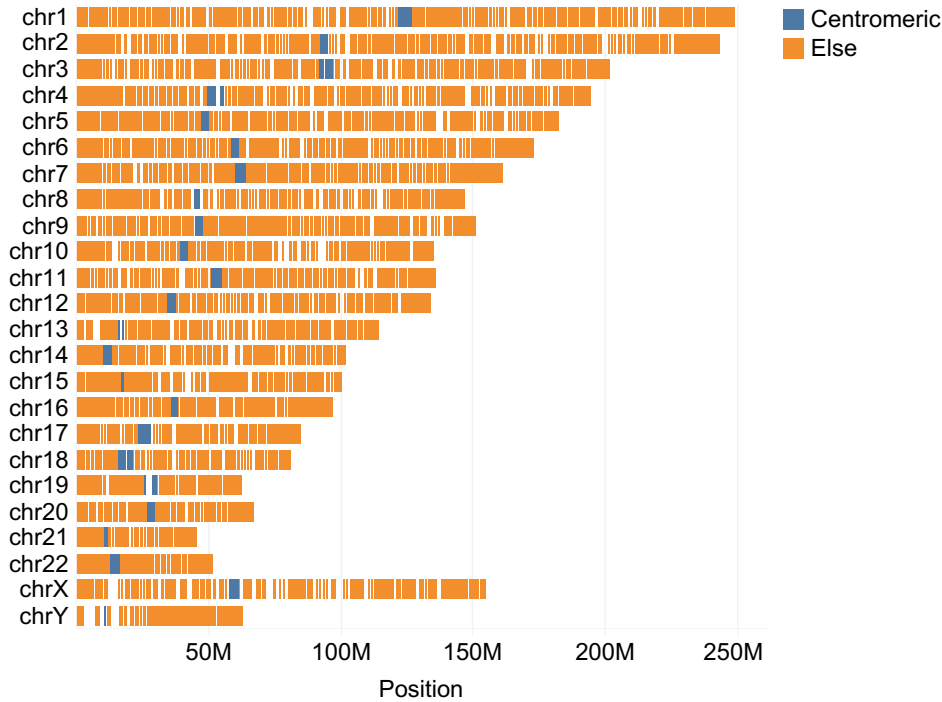
**C**



d

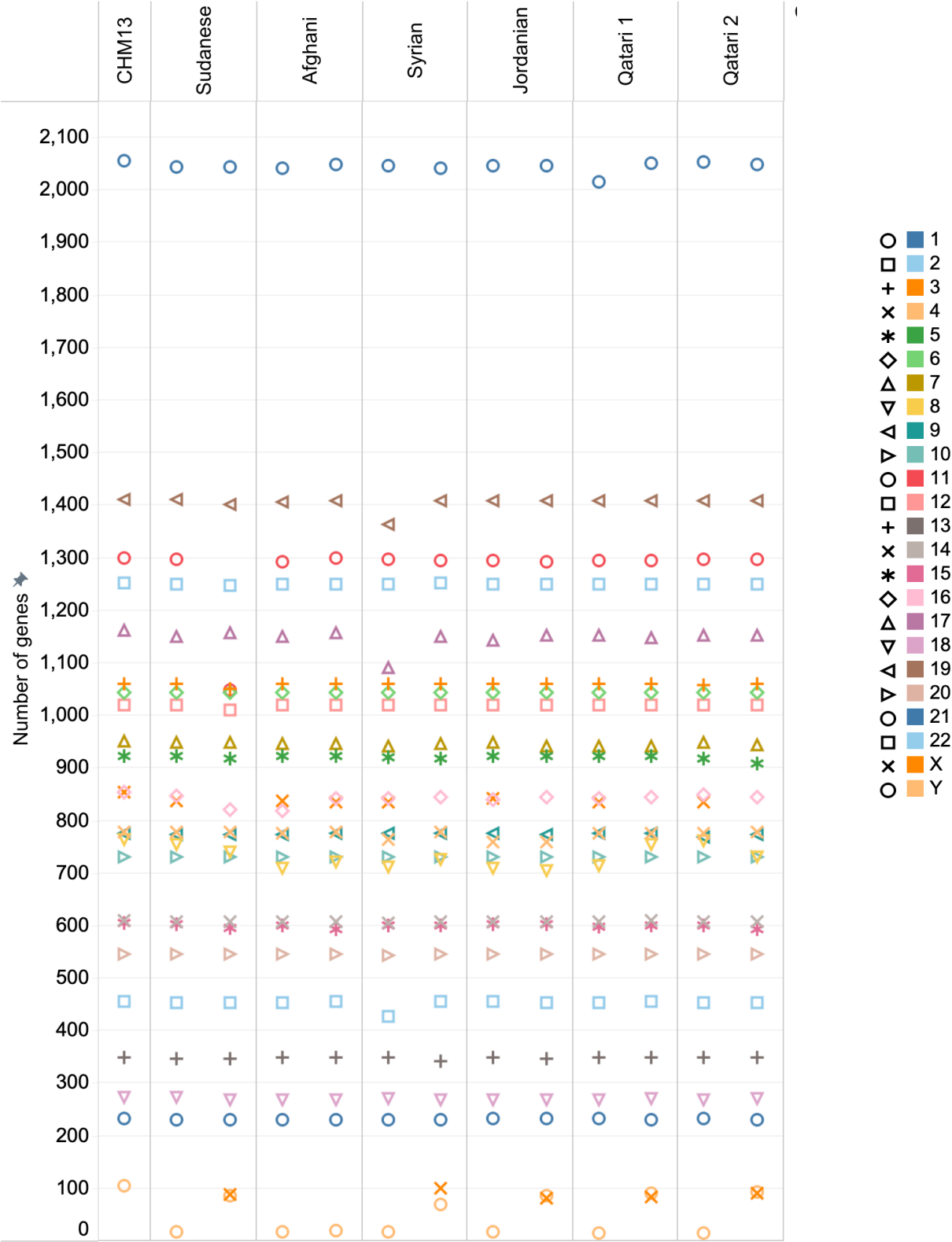


e

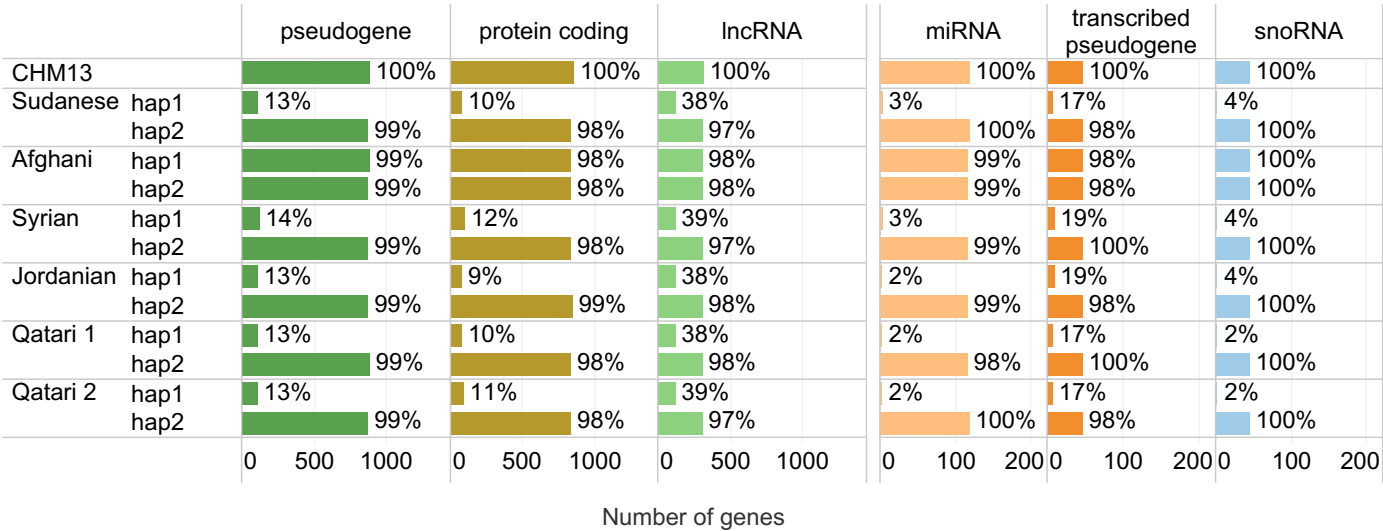


**Supplementary Fig. 11: Gene coverage of child assemblies relative to CHM13. (a)** Number of genes annotated per chromosome haplotype for the child assemblies and CHM13. **(b and c)** Coverage of annotated genes for the six largest gene categories in chromosome X and Y respectively.

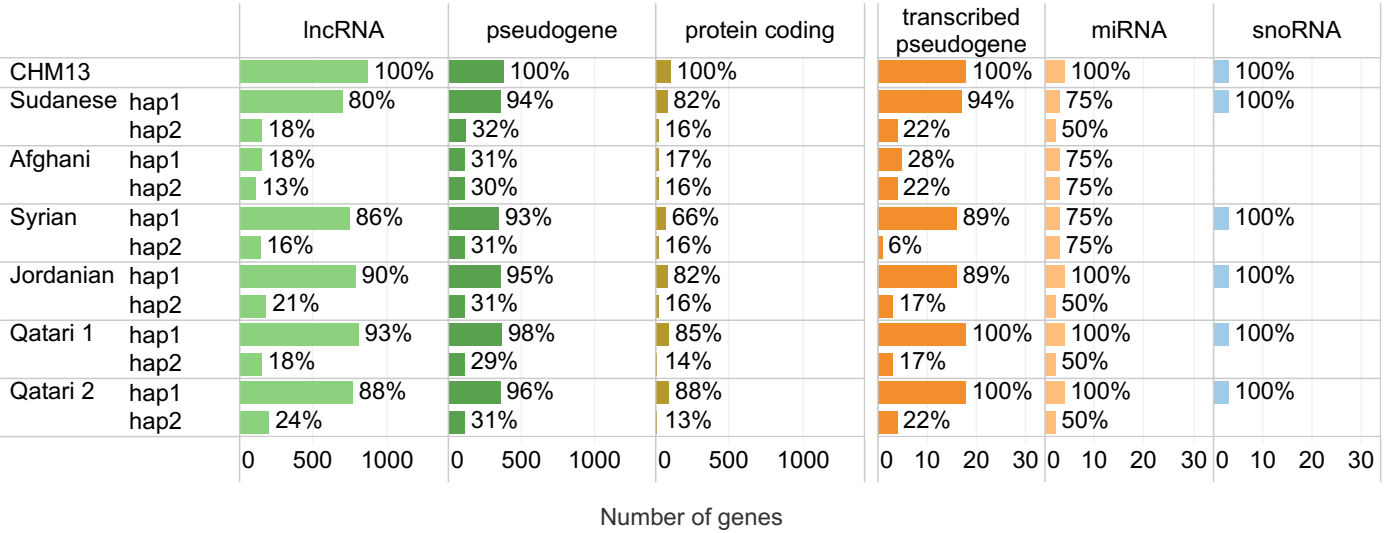
**a**



**b**



**c**



**Supplementary Fig. 12: *HLA* and *KIR* gene copies.** Shown are annotated *HLA* and *KIR* genes across ME child assemblies, benchmarked against 220 reference haplotypes. Each dot denotes the presence of gene copies, with colors indicating the number of copies detected.

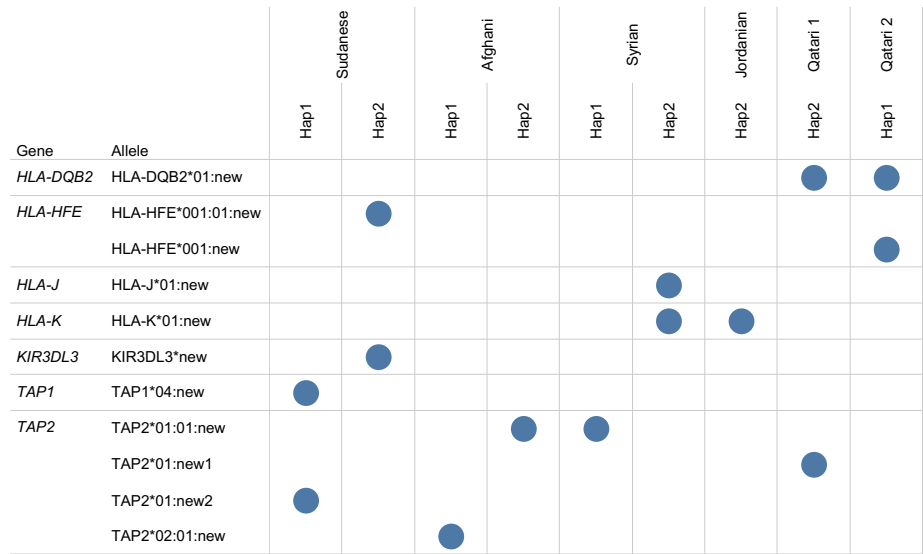


**Supplementary Fig. 13. Allele frequencies of *HLA-A* alleles in the child assemblies.** For each child subject, log allele frequency of each haplotype *HLA-A* allele is shown across various world populations. Alleles are categorized in each population by the CIWD catalogue v 3.0 as *C*, *I*, *WD* and *NA* denoting common, intermediate, well-documented and not applicable, respectively. The populations are defined by CIWD and denoted as follows: AFA (African/African American), API (Asian/Pacific Islands), EUR (European/European descent), ME (Middle East/North Coast of Africa), HIS (South or Central America/Hispanic/Latino), NAM (Native American populations) and UNK (unknown/not asked/multiple ancestries/other). Total is the overall population i.e., all groups combined.



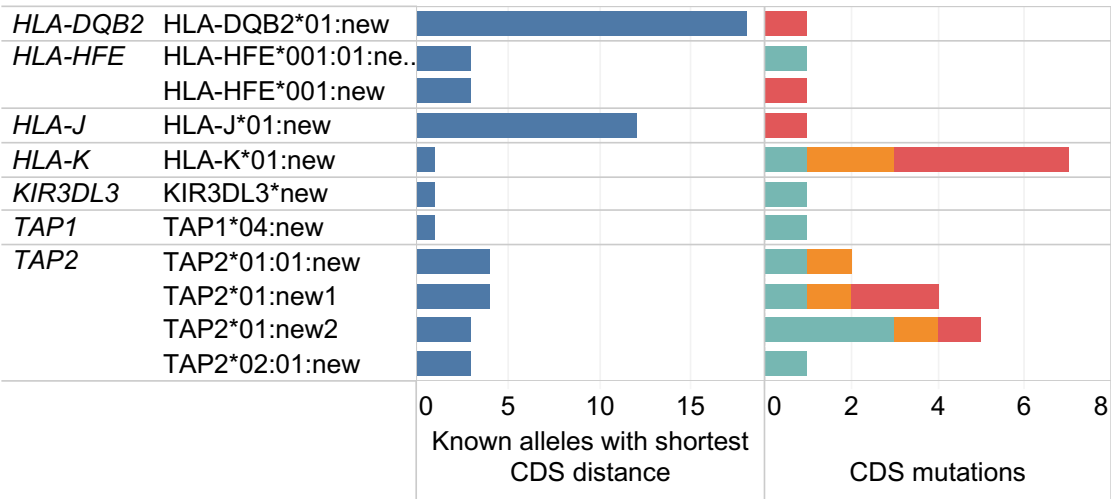
**Supplementary Fig. 14. Novel *HLA* and *KIR* alleles with mutations in the CDS in the child assemblies. (a)** Novel alleles are shown together with the corresponding genes and haplotypes for each subject **(b)** Per allele and subject, the number of known alleles with the shortest CDS distance and the number of CDS mutations that are either missense novel, missense known or synonymous. **(c)** Details of the novel alleles including the DNA alterations relative to known alleles with the closest CDS, as well as the population ancestry of the source samples in IPD for those related alleles.

**a**



**b**

■ Missense - novel  
■ Missense - known  
■ Synonymous



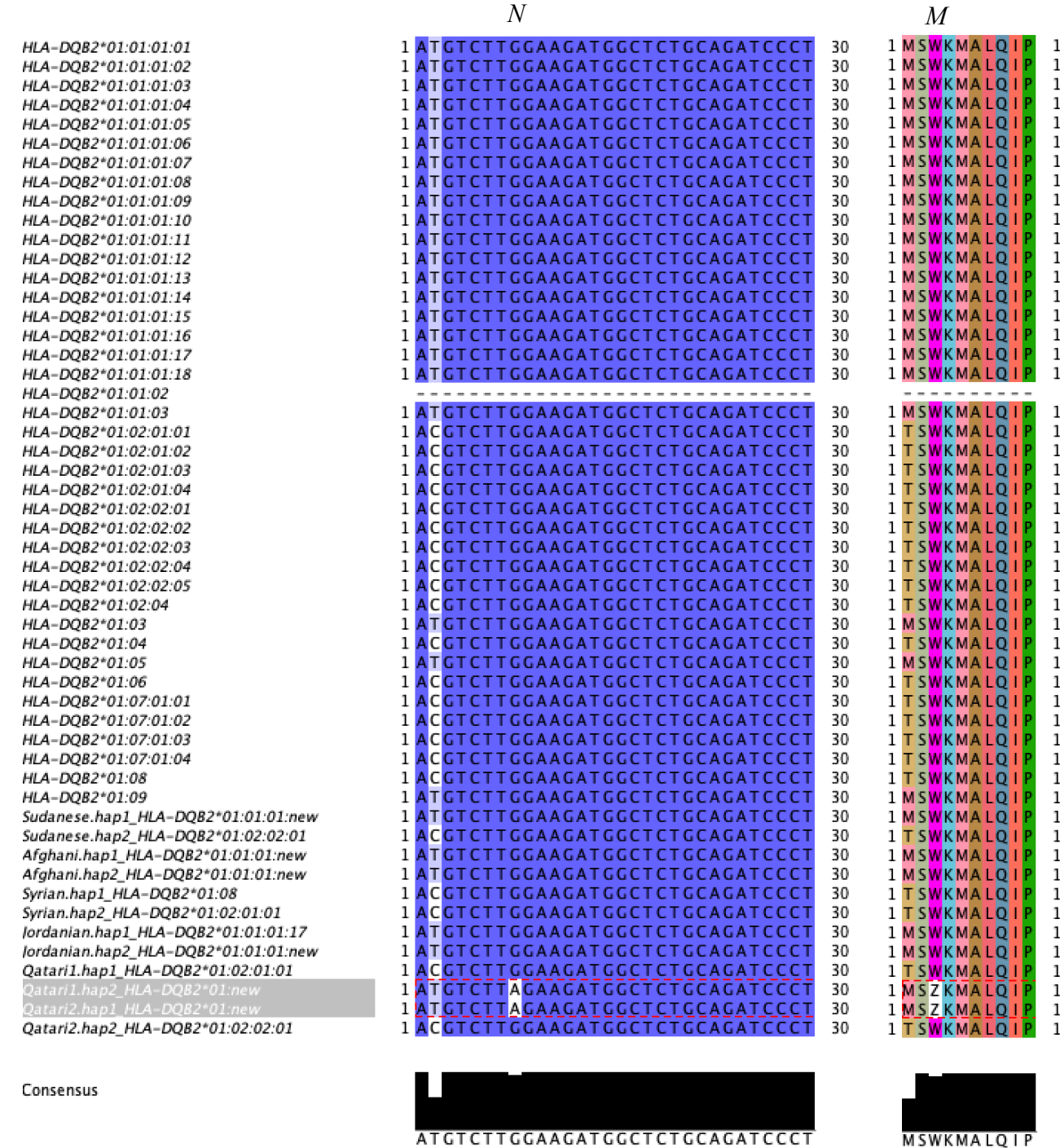
Ancestry of source samples in IPD for related alleles

■ Sub-Saharan African
 ■ East Asian  
■ European
 ■ Null  
■ Native American

| Allele             | CDS Distance | Mutation      | Novel DNA mutation | Alteration          | Category         | Related Allele       |                                                                            |
|--------------------|--------------|---------------|--------------------|---------------------|------------------|----------------------|----------------------------------------------------------------------------|
| HLA-DQB2*01:new    | 1            | :7_ag:799     | Yes                | W(TGG)<_(TAG)       | Missense - novel | HLA-DQB2*01:01:01:01 | <span style="color: purple;">●</span> <span style="color: green;">●</span> |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:02 | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:03 | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:04 | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:05 | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:06 | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:07 | <span style="color: green;">●</span>                                       |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:08 | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:09 | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:10 | <span style="color: orange;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:11 | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:12 | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:13 | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:14 | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:15 | <span style="color: green;">●</span>                                       |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:16 | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:17 | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | HLA-DQB2*01:01:01:18 | <span style="color: grey;">●</span>                                        |
| HLA-HFE*001:01:new | 1            | :1025_tc:21   | Yes                | Y(TAC)<Y(TAT)       | Synonymous       | HLA-HFE*001:01:01    | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-HFE*001:01:02    | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | HLA-HFE*001:01:03    | <span style="color: grey;">●</span>                                        |
| HLA-HFE*001:new    | 1            | :828_ag:218   | Yes                | E(GAG)<K(AAG)       | Missense - novel | HLA-HFE*001:01:01    | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-HFE*001:01:02    | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | HLA-HFE*001:01:03    | <span style="color: grey;">●</span>                                        |
| HLA-J*01:new       | 1            | :931_ct:163   | Yes                | I(ATT)<T(ACT)       | Missense - novel | HLA-J*01:01:01:01    | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-J*01:01:01:02    | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-J*01:01:01:03    | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | HLA-J*01:01:01:04    | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-J*01:01:01:05    | <span style="color: purple;">●</span> <span style="color: pink;">●</span>  |
|                    |              |               |                    |                     |                  | HLA-J*01:01:01:11    | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-J*01:01:01:14    | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-J*01:01:01:15    | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-J*01:01:01:17    | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-J*01:01:01:27    | <span style="color: green;">●</span>                                       |
|                    |              |               |                    |                     |                  | HLA-J*01:01:01:28    | <span style="color: purple;">●</span>                                      |
|                    |              |               |                    |                     |                  | HLA-J*01:01:01:30    | <span style="color: grey;">●</span>                                        |
| HLA-K*01:new       | 7            | :7_ct:123     | No                 | L(TTA)<S(TCA)       | Missense - known | HLA-K*01:02          | <span style="color: purple;">●</span>                                      |
|                    |              | ag:8          | No                 | N(AAC)<K(AAG)       | Missense - known | HLA-K*01:02          | <span style="color: purple;">●</span>                                      |
|                    |              | gc:238        | No                 | N(AAC)<K(AAG)       | Missense - novel | HLA-K*01:02          | <span style="color: purple;">●</span>                                      |
|                    |              | gc:238-cg:465 | Yes                | R(CGC)<CCGGC(CCGGC) | Missense - novel | HLA-K*01:02          | <span style="color: purple;">●</span>                                      |
|                    |              | gc:246        | Yes                | P(CCC)<P(CCG)       | Synonymous       | HLA-K*01:02          | <span style="color: purple;">●</span>                                      |
|                    |              | tc:1          | No                 | R(CGG)<_(TGA)       | Missense - novel | HLA-K*01:02          | <span style="color: purple;">●</span>                                      |
|                    |              | tc:1, ag:8    | Yes                | R(CGG)<_(TGA)       | Missense - novel | HLA-K*01:02          | <span style="color: purple;">●</span>                                      |
| KIR3DL3*new        | 1            | :167_ct:1065  | No                 | N(AAT)<N(AAC)       | Synonymous       | KIR3DL3*035          | <span style="color: grey;">●</span>                                        |
| TAP1*04:new        | 1            | :158_ga:2088  | Yes                | P(CCA)<P(CCG)       | Synonymous       | TAP1*04:01           | <span style="color: grey;">●</span>                                        |
| TAP2*01:01:new     | 1            | :1157_tg:903  | No                 | G(GGG)<G(GGT)       | Synonymous       | TAP2*01:01:03:01     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:02     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:03     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:02           | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:01     | <span style="color: grey;">●</span>                                        |
| TAP2*01:new1       | 2            | :1692_ga:368  | No                 | T(ACT)<A(GCT)       | Missense - known | TAP2*01:02           | <span style="color: grey;">●</span>                                        |
|                    |              | :1157_tg:258  | No                 | G(GGG)<G(GGT)       | Synonymous       | TAP2*01:01:03:01     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:02     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:03     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:02           | <span style="color: grey;">●</span>                                        |
| TAP2*01:new2       | 5            | :1416_ag:275  | Yes                | V(GTC)<I(ATC)       | Missense - novel | TAP2*01:01:03:01     | <span style="color: grey;">●</span>                                        |
|                    |              | ag:644        | Yes                | V(GTC)<I(ATC)       | Missense - novel | TAP2*01:01:03:02     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:03     | <span style="color: grey;">●</span>                                        |
|                    |              | ga:368        | No                 | T(ACT)<A(GCT)       | Missense - known | TAP2*01:02           | <span style="color: grey;">●</span>                                        |
|                    |              | :657_ac:461   | Yes                | R(CGA)<R(AGA)       | Synonymous       | TAP2*01:01:03:01     | <span style="color: grey;">●</span>                                        |
| TAP2*02:01:new     | 1            | :1157_tg:954  | No                 | G(GGG)<G(GGT)       | Synonymous       | TAP2*01:01:03:02     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:03     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:03     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:01     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:02     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:03     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:01     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:02     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:03     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:01     | <span style="color: grey;">●</span>                                        |
|                    |              |               |                    |                     |                  | TAP2*01:01:03:02     | <span style="color: grey;">●</span>                                        |

**Supplementary Fig. 15. Alignments of the Coding DNA sequences (CDS) and corresponding amino acid sequences for the novel alleles in the child subjects against reference sequences from IPD.** For each locus, CDS alignment is colored by position-level percentage identity and the amino acid alignment is colored by residue type. Consensus plot is indicated at the bottom. The novel allele names are highlighted in grey, and their sequence are contoured with a discontinued red line. At the top of the alignments, DNA mutations are annotated with *N* (Novel) or *n* (known) while amino acid mutations are labeled as *M* (Missense) or *s* (synonymous).

### HLA-DQB2



HLA-HFE

|                                  | N   |             | M   |                   |
|----------------------------------|-----|-------------|-----|-------------------|
| HLA-HFE*001:01:01                | 823 | GGGGAAGAGC  | 832 | 277 EQRYTCQVE 285 |
| HLA-HFE*001:01:02                | 823 | GGGGAAGAGC  | 832 | 277 EQRYTCQVE 285 |
| HLA-HFE*001:01:03                | 823 | GGGGAAGAGC  | 832 | 277 EQRYTCQVE 285 |
| HLA-HFE*002                      | 747 | GGGGAAGAGC  | 756 | 251 EQRYTYQVE 259 |
| HLA-HFE*003                      | 747 | GGGGAAGAGC  | 756 | 251 EQRYTCQVE 259 |
| HLA-HFE*004                      | 747 | GGGGAAGAGC  | 756 | 251 EQRYTCQVE 259 |
| Sudanese.hap1_HLA-HFE*001:01:new | 823 | GGGGAAGAGC  | 832 | 277 EQRYTCQVE 285 |
| Sudanese.hap2_HLA-HFE*001:01:new | 823 | GGGGAAGAGC  | 832 | 277 EQRYTCQVE 285 |
| Afghani.hap1_HLA-HFE*001:01:new  | 823 | GGGGAAGAGC  | 832 | 277 EQRYTCQVE 285 |
| Afghani.hap2_HLA-HFE*001:01:01   | 823 | GGGGAAGAGC  | 832 | 277 EQRYTCQVE 285 |
| Syrian.hap1_HLA-HFE*001:01:new   | 823 | GGGGAAGAGC  | 832 | 277 EQRYTCQVE 285 |
| Syrian.hap2_HLA-HFE*001:01:new   | 823 | GGGGAAGAGC  | 832 | 277 EQRYTCQVE 285 |
| Jordanian.hap1_HLA-HFE*001:01:01 | 823 | GGGGAAGAGC  | 832 | 277 EQRYTCQVE 285 |
| Jordanian.hap2_HLA-HFE*001:01:01 | 823 | GGGGAAGAGC  | 832 | 277 EQRYTCQVE 285 |
| Qatari1.hap1_HLA-HFE*001:01:new  | 823 | GGGGAAGAGC  | 832 | 277 EQRYTCQVE 285 |
| Qatari1.hap2_HLA-HFE*001:01:new  | 823 | GGGGAAGAGC  | 832 | 277 EQRYTCQVE 285 |
| Qatari2.hap1_HLA-HFE*001:new     | 823 | GGGGAAAGAGC | 832 | 277 KQRYTCQVE 285 |
| Qatari2.hap2_HLA-HFE*001:01:new  | 823 | GGGGAAGAGC  | 832 | 277 EQRYTCQVE 285 |

Consensus

GGGGAAGAGC

EQRYTCQVE

|                                  | N     |             | S     |                   |
|----------------------------------|-------|-------------|-------|-------------------|
| HLA-HFE*001:01:01                | 1023  | CTACGTCTTA  | 1032  | 340 GHYVLAERE 348 |
| HLA-HFE*001:01:02                | 1023  | CTACGTCTTA  | 1032  | 340 GHYVLAERE 348 |
| HLA-HFE*001:01:03                | 1023  | CTACGTCTTA  | 1032  | 340 GHYVLAERE 348 |
| HLA-HFE*002                      | ----- |             | ----- |                   |
| HLA-HFE*003                      | ----- |             | ----- |                   |
| HLA-HFE*004                      | ----- |             | ----- |                   |
| Sudanese.hap1_HLA-HFE*001:01:new | 1023  | CTACGTCTTA  | 1032  | 340 GHYVLAERE 348 |
| Sudanese.hap2_HLA-HFE*001:01:new | 1023  | CTATGTCCTTA | 1032  | 340 GHYVLAERE 348 |
| Afghani.hap1_HLA-HFE*001:01:new  | 1023  | CTACGTCTTA  | 1032  | 340 GHYVLAERE 348 |
| Afghani.hap2_HLA-HFE*001:01:01   | 1023  | CTACGTCTTA  | 1032  | 340 GHYVLAERE 348 |
| Syrian.hap1_HLA-HFE*001:01:new   | 1023  | CTACGTCTTA  | 1032  | 340 GHYVLAERE 348 |
| Syrian.hap2_HLA-HFE*001:01:new   | 1023  | CTACGTCTTA  | 1032  | 340 GHYVLAERE 348 |
| Jordanian.hap1_HLA-HFE*001:01:01 | 1023  | CTACGTCTTA  | 1032  | 340 GHYVLAERE 348 |
| Jordanian.hap2_HLA-HFE*001:01:01 | 1023  | CTACGTCTTA  | 1032  | 340 GHYVLAERE 348 |
| Qatari1.hap1_HLA-HFE*001:01:new  | 1023  | CTACGTCTTA  | 1032  | 340 GHYVLAERE 348 |
| Qatari1.hap2_HLA-HFE*001:01:new  | 1023  | CTACGTCTTA  | 1032  | 340 GHYVLAERE 348 |
| Qatari2.hap1_HLA-HFE*001:new     | 1023  | CTACGTCTTA  | 1032  | 340 GHYVLAERE 348 |
| Qatari2.hap2_HLA-HFE*001:01:new  | 1023  | CTACGTCTTA  | 1032  | 340 GHYVLAERE 348 |

Consensus

CTACGTCTTA

GHYVLAERE

# HLA-J

|                                  | N                  | M                   |
|----------------------------------|--------------------|---------------------|
| HLA-J*01:01:01:01                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:02                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:03                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:04                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:05                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:06                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:07                | 929 TATCATTGCT 938 | 310 VSLLAWF --- 316 |
| HLA-J*01:01:01:08                | 928 TATCATTGCT 937 | 308 CGYHC-W --- 313 |
| HLA-J*01:01:01:09                | 920 TATCATTGCT 929 | 307 VSLLAWF --- 313 |
| HLA-J*01:01:01:10                | 921 TATCATTGCT 930 | 306 VGIIAGLVLL 315  |
| HLA-J*01:01:01:11                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:12                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:13                | 920 TATCATTGCT 929 | 307 VSLLAWF --- 313 |
| HLA-J*01:01:01:14                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:15                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:16                | 928 TATCATTGCT 937 | 308 CGYHC-W --- 313 |
| HLA-J*01:01:01:17                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:18                | 928 TATCATTGCT 937 | 308 CGYHC-W --- 313 |
| HLA-J*01:01:01:19                | 929 TATCATTGCT 938 | 310 VSLLAWF --- 316 |
| HLA-J*01:01:01:20                | 928 TATCATTGCT 937 | 308 CGYHC-W --- 313 |
| HLA-J*01:01:01:21                | 928 TATCATTGCT 937 | 308 CGYHC-W --- 313 |
| HLA-J*01:01:01:22                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:23                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:24                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:25                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:26                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:27                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:28                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:29                | 920 TATCATTGCT 929 | 307 VSLLAWF --- 313 |
| HLA-J*01:01:01:30                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:31                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*01:01:01:32                | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| HLA-J*02:01                      | 926 TATCATTGCT 935 | 309 VSLLAWF --- 315 |
| Sudanese.hap1_HLA-J*01:01:01:15  | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| Sudanese.hap2_HLA-J*01:01:01:01  | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| Afghani.hap1_HLA-J*01:01:01:02   | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| Afghani.hap2_HLA-J*01:01:01:02   | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| Syrian.hap1_HLA-J*01:01:01:04    | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| Syrian.hap2_HLA-J*01:new         | 927 TATCACTGCT 936 | 308 VGIIAGLVLL 317  |
| Jordanian.hap1_HLA-J*01:01:01:05 | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| Jordanian.hap2_HLA-J*01:01:01:01 | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| Qatari1.hap1_HLA-J*01:01:01:04   | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| Qatari1.hap2_HLA-J*01:01:01:17   | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| Qatari2.hap1_HLA-J*01:01:01:01   | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |
| Qatari2.hap2_HLA-J*01:01:01:17   | 927 TATCATTGCT 936 | 308 VGIIAGLVLL 317  |

Consensus

TATCATTGCT

VGIIAGLVLL

*HLA-K*

|                               | <i>N</i>                                      | <i>S</i>                                      |
|-------------------------------|-----------------------------------------------|-----------------------------------------------|
| HLA-K*01:01:01:01             | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| HLA-K*01:01:01:02             | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| HLA-K*01:01:01:03             | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| HLA-K*01:01:01:04             | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| HLA-K*01:02                   | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| HLA-K*01:03                   | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| Sudanese.hap2_HLA-K*01:03     | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| Syrian.hap1_HLA-K*01:01:01:01 | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| Syrian.hap2_HLA-K*01:new      | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| Jordanian.hap1_HLA-K*01:02    | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| Jordanian.hap2_HLA-K*01:new   | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| Qatari1.hap1_HLA-K*01:new     | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| Qatari1.hap2_HLA-K*01:02      | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| Qatari2.hap1_HLA-K*01:01:new  | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| Qatari2.hap2_HLA-K*01:02      | 115 C G G G T C G C G G G G A G C C C C G 133 | 41 A G S P G T S Q W A T W T T R S S C G S 60 |
| Consensus                     | C G G G T C G C G G G G A G C C C C G         | A G S P G T S Q W A T W T T R S S C G S       |

HLA-K\*01:01:01:01 361 TAATGTATGGCTGCGACTTGGGG 383 109 TTRARPGLTPSRZCMAATWG 128

HLA-K\*01:01:01:02 361 TAATGTATGGCTGCGACTTGGGG 383 109 TTRARPGLTPSRZCMAATWG 128

HLA-K\*01:01:01:03 361 TAATGTATGGCTGCGACTTGGGG 383 109 TTRARPGLTPSRZCMAATWG 128

HLA-K\*01:01:01:04 361 TAATGTATGGCTGCGACTTGGGG 383 109 TTRARPGLTPSRZCMAATWG 128

HLA-K\*01:02 361 TAATGTATGGCTGCGACTCGGGG 383 109 TTRARPGLTPSRZCMAATRG 128

HLA-K\*01:03 361 TAATGTATGGCTGCGACTTGGGG 383 109 TTRARPGLTPSRZCMAATWG 128

Sudanese.hap2\_HLA-K\*01:03 361 TAATGTATGGCTGCGACTTGGGG 383 109 TTRARPGLTPSRZCMAATWG 128

Syrian.hap1\_HLA-K\*01:01:01:01 361 TAATGTATGGCTGCGACTTGGGG 383 109 TTRARPGLTPSRZCMAATWG 128

Syrian.hap2\_HLA-K\*01:new 361 TAATGTATGGCTGCGACTTGAGG 383 109 TTRARPGLTPSRZCMAATZG 128

Jordanian.hap1\_HLA-K\*01:02 361 TAATGTATGGCTGCGACTCGGGG 383 109 TTRARPGLTPSRZCMAATRG 128

Jordanian.hap2\_HLA-K\*01:new 361 TAATGTATGGCTGCGACTTGAGG 383 109 TTRARPGLTPSRZCMAATZG 128

Qatar1.hap1\_HLA-K\*01:new 362 TAATGTATGGCTGCGACTTGGGG 384 109 TTRARPGLTPSRZCMAATRG 128

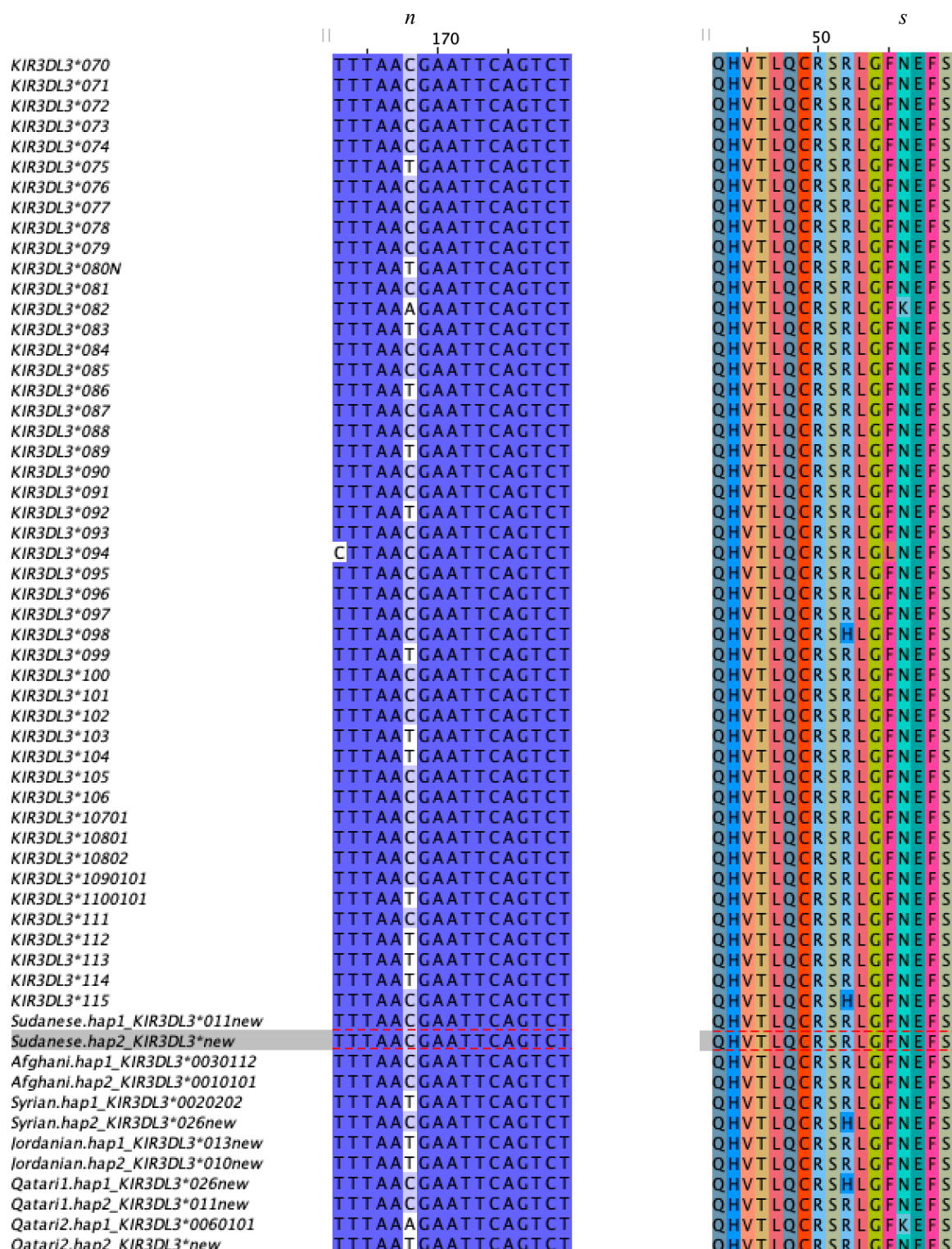
Qatar1.hap2\_HLA-K\*01:02 361 TAATGTATGGCTGCGACTCGGGG 383 109 TTRARPGLTPSRZCMAATWG 128

Qatar2.hap1\_HLA-K\*01:01:new 361 TAATGTATGGCTGCGACTTGGGG 383 109 TTRARPGLTPSRZCMAATRG 128

Qatar2.hap2\_HLA-K\*01:02 361 TAATGTATGGCTGCGACTCGGGG 383 109 TTRARPGLTPSRZCMAATRG 128

Consensus TAATGTATGGCTGCGACTTGGGG TTRARPGLTPSRZCMAATWG

*KIR3DL3*



## Consensus

TTTAACGAATTCAGTCT

**S**

OHVTLOCRSRLGFNEFS

TAP1

|                                 | N   |                                                                                   | S  |                                                                                   |
|---------------------------------|-----|-----------------------------------------------------------------------------------|----|-----------------------------------------------------------------------------------|
| TAP1*01:01:01:01                | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| TAP1*01:01:01:02                | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| TAP1*01:01:01:03                | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| TAP1*01:01:01:04                | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| TAP1*01:01:01:05                | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| TAP1*01:02N                     | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| TAP1*02:01:01                   | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| TAP1*02:01:02                   | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| TAP1*03:01                      | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| TAP1*04:01                      | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| TAP1*05:01                      | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| TAP1*06:01                      | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| Sudanese.hap1_TAP1*04.new       | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| Sudanese.hap2_TAP1*01:01:01.new | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| Afghani.hap1_TAP1*01:01:01:05   | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| Afghani.hap2_TAP1*01:01:01:05   | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| Syrian.hap1_TAP1*01:01:01:05    | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| Syrian.hap2_TAP1*02:01.new      | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| Jordanian.hap1_TAP1*01:01:01:01 | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| Jordanian.hap2_TAP1*01:01:01:01 | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| Qatari1.hap1_TAP1*02:01.new     | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| Qatari1.hap2_TAP1*01:01:01:02   | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| Qatari2.hap1_TAP1*01:01:01:01   | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| Qatari2.hap2_TAP1*01:01:01:01   | 151 | GGGCTGCCAC 160                                                                    | 51 | ALP L L R V W A V 60                                                              |
| Consensus                       |     |  |    |  |

TAP2

|                                  | <i>n</i>             | <i>s</i>           |
|----------------------------------|----------------------|--------------------|
| TAP2*01:01:01                    | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| TAP2*01:01:02                    | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| TAP2*01:01:03:01                 | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| TAP2*01:01:03:02                 | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| TAP2*01:01:03:03                 | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| TAP2*01:02                       | 1151 ACTTGGGTGT 1160 | 381 RVLHLGVQML 390 |
| TAP2*01:03                       | 961 ACTTGGGTGT 970   | 317 RVLHLGVQML 326 |
| TAP2*01:04                       | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| TAP2*02:01:01                    | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| TAP2*02:01:02:01                 | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| TAP2*02:01:02:02                 | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| TAP2*02:01:02:03                 | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| Sudanese.hap1_TAP2*01:new        | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| Sudanese.hap2_TAP2*02:01:02:new  | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| Afghani.hap1_TAP2*02:01:new      | 1151 ACTTGGGTGT 1160 | 381 RVLHLGVQML 390 |
| Afghani.hap2_TAP2*01:01:new      | 1151 ACTTGGGTGT 1160 | 381 RVLHLGVQML 390 |
| Syrian.hap1_TAP2*01:01:new       | 1151 ACTTGGGTGT 1160 | 381 RVLHLGVQML 390 |
| Syrian.hap2_TAP2*01:01:03:new    | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| Jordanian.hap1_TAP2*01:01:03:new | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| Jordanian.hap2_TAP2*02:01:new    | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| Qatari1.hap1_TAP2*01:01:03:new   | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| Qatari1.hap2_TAP2*01:new         | 1151 ACTTGGGTGT 1160 | 381 RVLHLGVQML 390 |
| Qatari2.hap1_TAP2*01:01:03:new   | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |
| Qatari2.hap2_TAP2*02:01:02:new   | 1151 ACTTGGGGGT 1160 | 381 RVLHLGVQML 390 |

Consensus

ACTTGGGGGT

RVLHLGVQML

|                                  | <i>N</i>             | <i>M</i>           |
|----------------------------------|----------------------|--------------------|
| TAP2*01:01:01                    | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| TAP2*01:01:02                    | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| TAP2*01:01:03:01                 | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| TAP2*01:01:03:02                 | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| TAP2*01:01:03:03                 | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| TAP2*01:02                       | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| TAP2*01:03                       | 1221 CAAGACGTCT 1230 | 407 QDVSFAYPNR 416 |
| TAP2*01:04                       | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| TAP2*02:01:01                    | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| TAP2*02:01:02:01                 | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| TAP2*02:01:02:02                 | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| TAP2*02:01:02:03                 | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| Sudanese.hap1_TAP2*01:new        | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| Sudanese.hap2_TAP2*02:01:02:new  | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| Afghani.hap1_TAP2*02:01:new      | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| Afghani.hap2_TAP2*01:01:new      | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| Syrian.hap1_TAP2*01:01:new       | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| Syrian.hap2_TAP2*01:01:03:new    | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| Jordanian.hap1_TAP2*01:01:03:new | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| Jordanian.hap2_TAP2*02:01:new    | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| Qatari1.hap1_TAP2*01:01:03:new   | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| Qatari1.hap2_TAP2*01:new         | 1411 CAAGACATCT 1420 | 471 QDVSFAYPNR 480 |
| Qatari2.hap1_TAP2*01:01:03:new   | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |
| Qatari2.hap2_TAP2*02:01:02:new   | 1411 CAAGACGTCT 1420 | 471 QDVSFAYPNR 480 |

Consensus

CAAGACGTCT

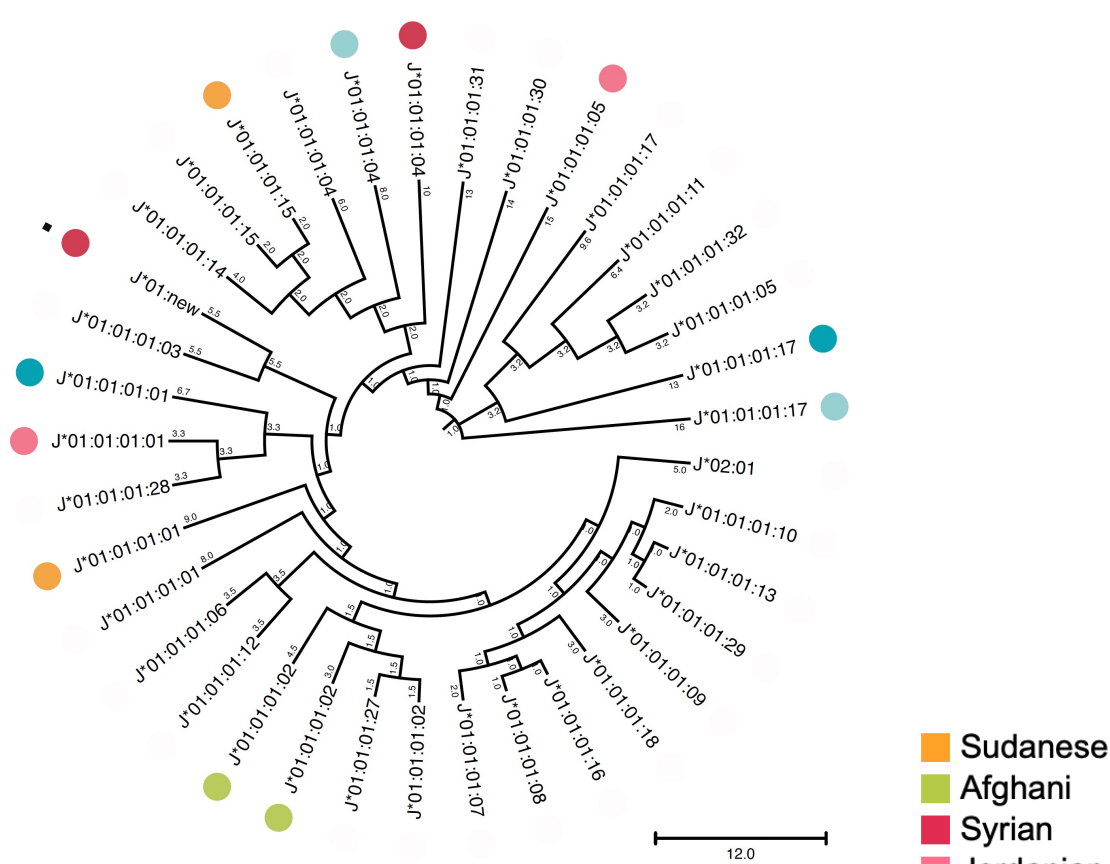
QDVSFAYPNR

|                                  | N                      | M                  |
|----------------------------------|------------------------|--------------------|
| TAP2*01:01:01                    | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| TAP2*01:01:02                    | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| TAP2*01:01:03:01                 | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| TAP2*01:01:03:02                 | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| TAP2*01:01:03:03                 | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| TAP2*01:02                       | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| TAP2*01:03                       | 921 CTGGAACGCGCC 932   | 307 LERALYLLVR 316 |
| TAP2*01:04                       | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| TAP2*02:01:01                    | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| TAP2*02:01:02:01                 | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| TAP2*02:01:02:02                 | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| TAP2*02:01:02:03                 | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| Sudanese.hap1_TAP2*01:new        | 1111 CTGGAACGCGCC 1122 | 371 LERTLYLLVR 380 |
| Sudanese.hap2_TAP2*02:01:02:new  | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| Afghani.hap1_TAP2*02:01:new      | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| Afghani.hap2_TAP2*01:01:new      | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| Syrian.hap1_TAP2*01:01:new       | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| Syrian.hap2_TAP2*01:01:03:new    | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| Jordanian.hap1_TAP2*01:01:03:new | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| Jordanian.hap2_TAP2*02:01:new    | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| Qatari1.hap1_TAP2*01:01:03:new   | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| Qatari1.hap2_TAP2*01:new         | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| Qatari2.hap1_TAP2*01:01:03:new   | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| Qatari2.hap2_TAP2*02:01:02:new   | 1111 CTGGAACGCGCC 1122 | 371 LERALYLLVR 380 |
| Consensus                        | CTGGAACGCGCC           | LERALYLLVR         |

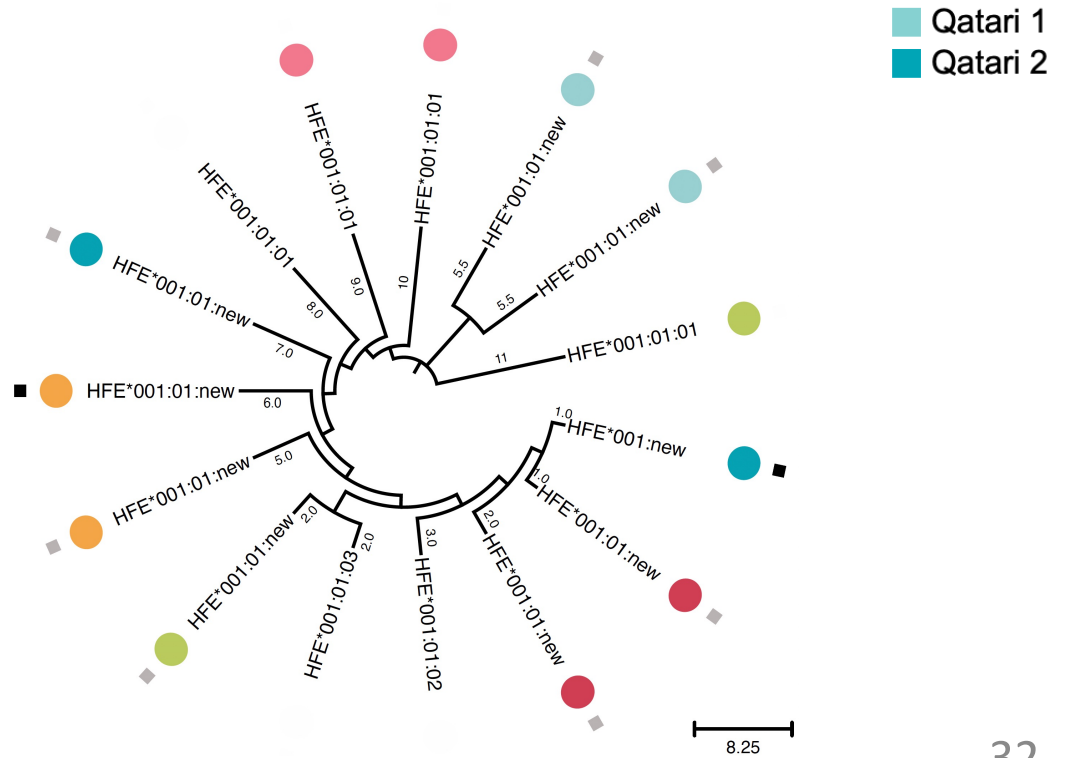
|                                  | N    | NN                                               | M   |                          |
|----------------------------------|------|--------------------------------------------------|-----|--------------------------|
| TAP2*01:01:01                    | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| TAP2*01:01:02                    | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| TAP2*01:01:03:01                 | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| TAP2*01:01:03:02                 | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| TAP2*01:01:03:03                 | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| TAP2*01:02                       | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| TAP2*01:03                       | 1173 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1214  | 387 | PNLPSPGTLAPTTLQGVVK 405  |
| TAP2*01:04                       | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| TAP2*02:01:01                    | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| TAP2*02:01:02:01                 | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| TAP2*02:01:02:02                 | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| TAP2*02:01:02:03                 | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| Sudanese.hap1_TAP2*01:new        | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGAATTGTG 1404 | 451 | PNLPSPGTLAPTTLQGLIVK 469 |
| Sudanese.hap2_TAP2*02:01:02:new  | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| Afghani.hap1_TAP2*02:01:new      | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| Afghani.hap2_TAP2*01:01:new      | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| Syrian.hap1_TAP2*01:01:new       | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| Syrian.hap2_TAP2*01:01:03:new    | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| Jordanian.hap1_TAP2*01:01:03:new | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| Jordanian.hap2_TAP2*02:01:new    | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| Qatari1.hap1_TAP2*01:01:03:new   | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| Qatari1.hap2_TAP2*01:new         | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| Qatari2.hap1_TAP2*01:01:03:new   | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| Qatari2.hap2_TAP2*02:01:02:new   | 1363 | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG 1404  | 451 | PNLPSPGTLAPTTLQGVVK 469  |
| Consensus                        |      | TCACCTGGCACGCTTGCCCCACCACCTCTGCAGGGGGTTGTG       |     | PNLPSPGTLAPTTLQGVVK      |

**Supplementary Fig. 16. Phylogenetic trees of *HLA* and *KIR* genes with novel alleles having mutation in the CDS region.** Colored circles indicate various child samples. Novel alleles are labeled showing those with mutations in the CDS region (black squares) and those with mutations in other part of the sequence (grey squares).

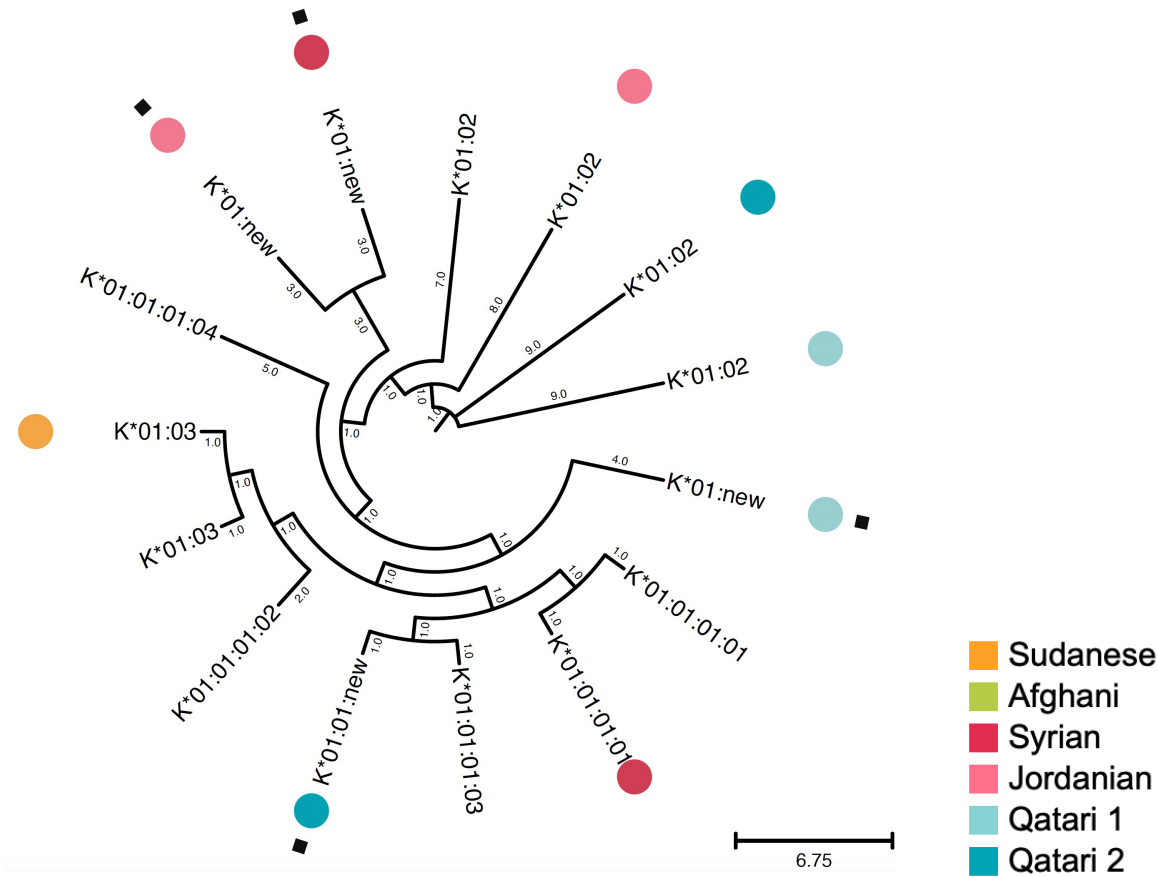
HLA-J



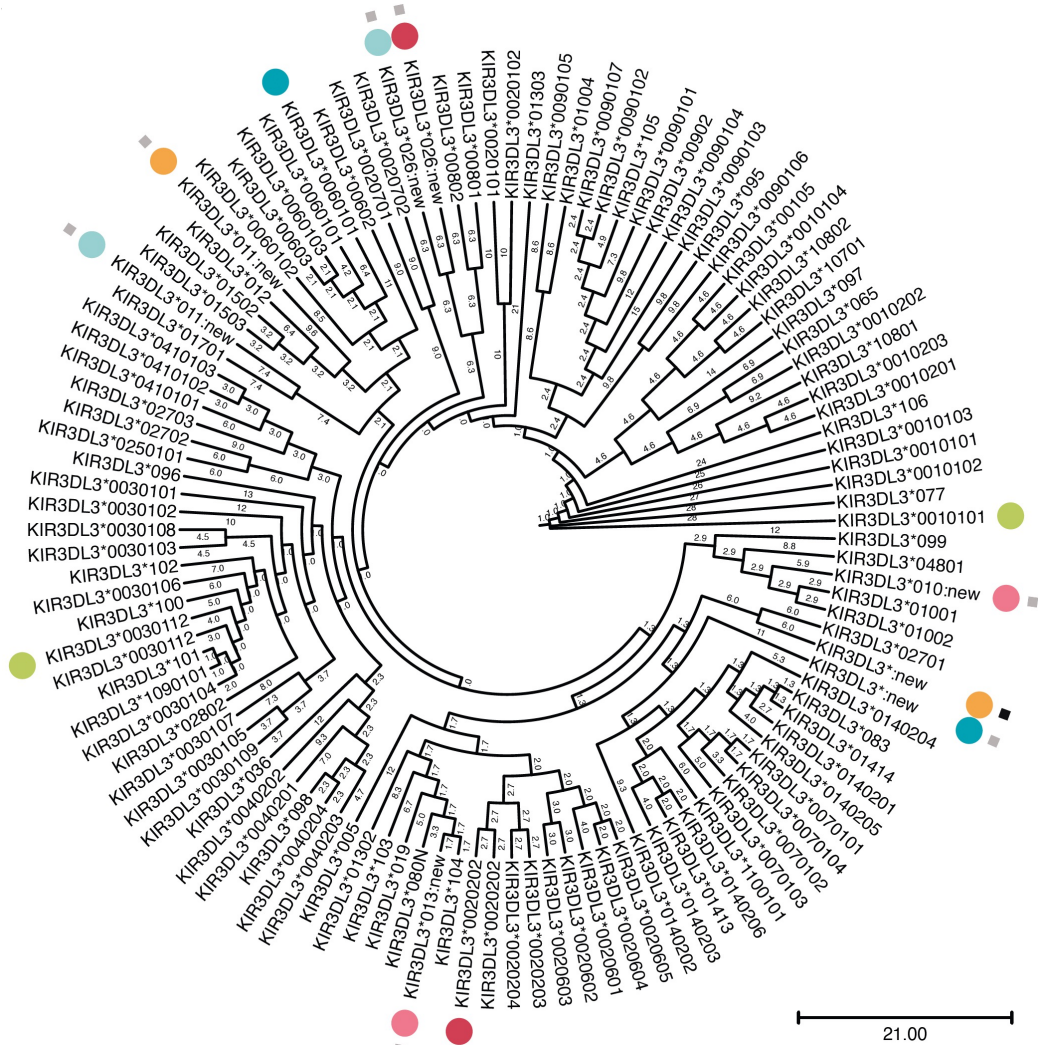
HLA-HFE



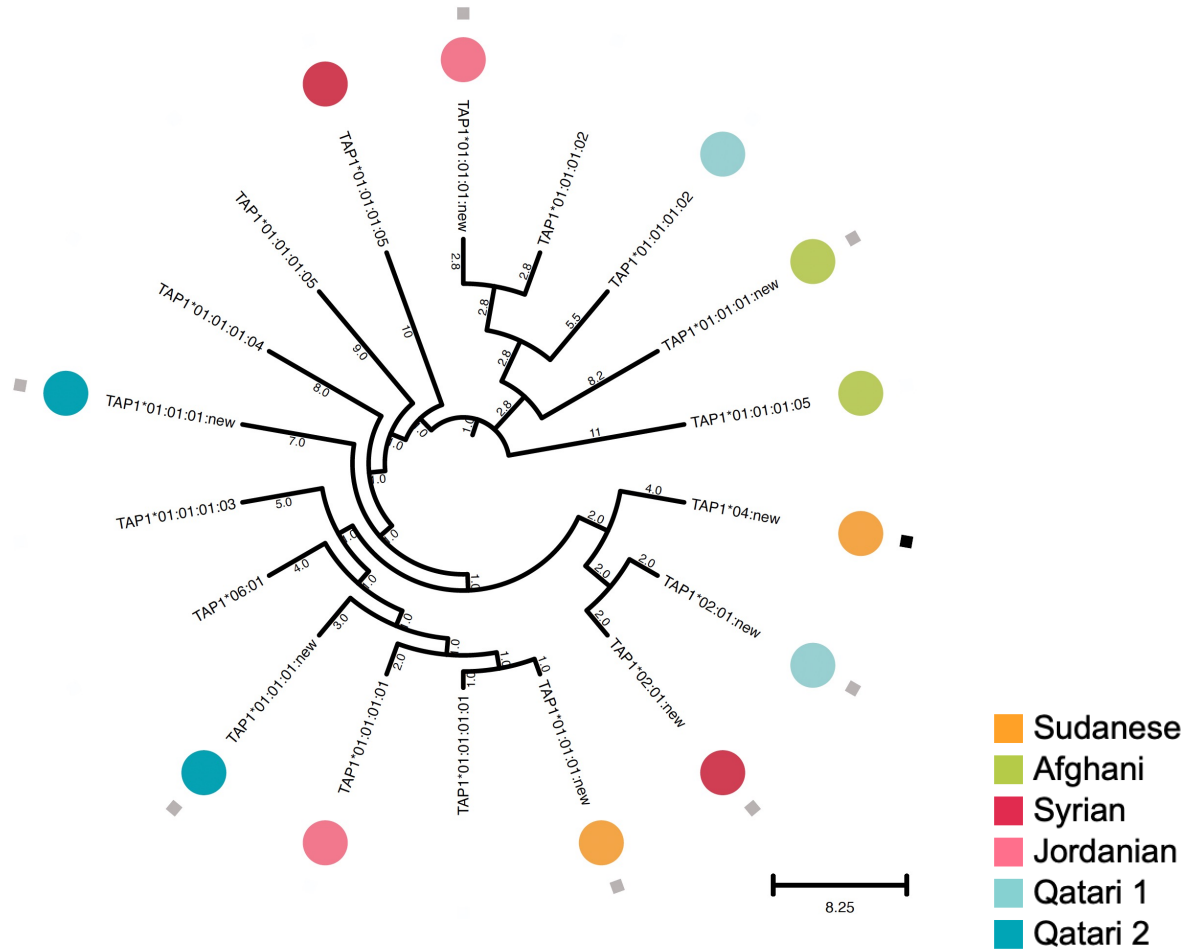
HLA-K



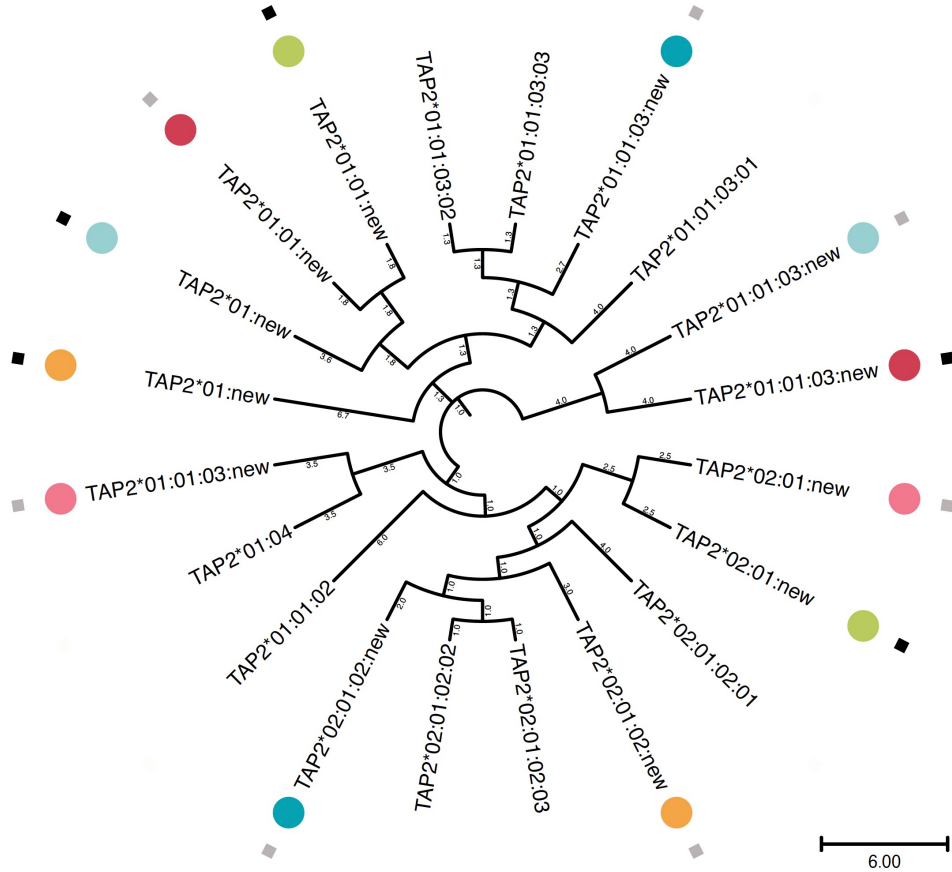
KIR3DL3



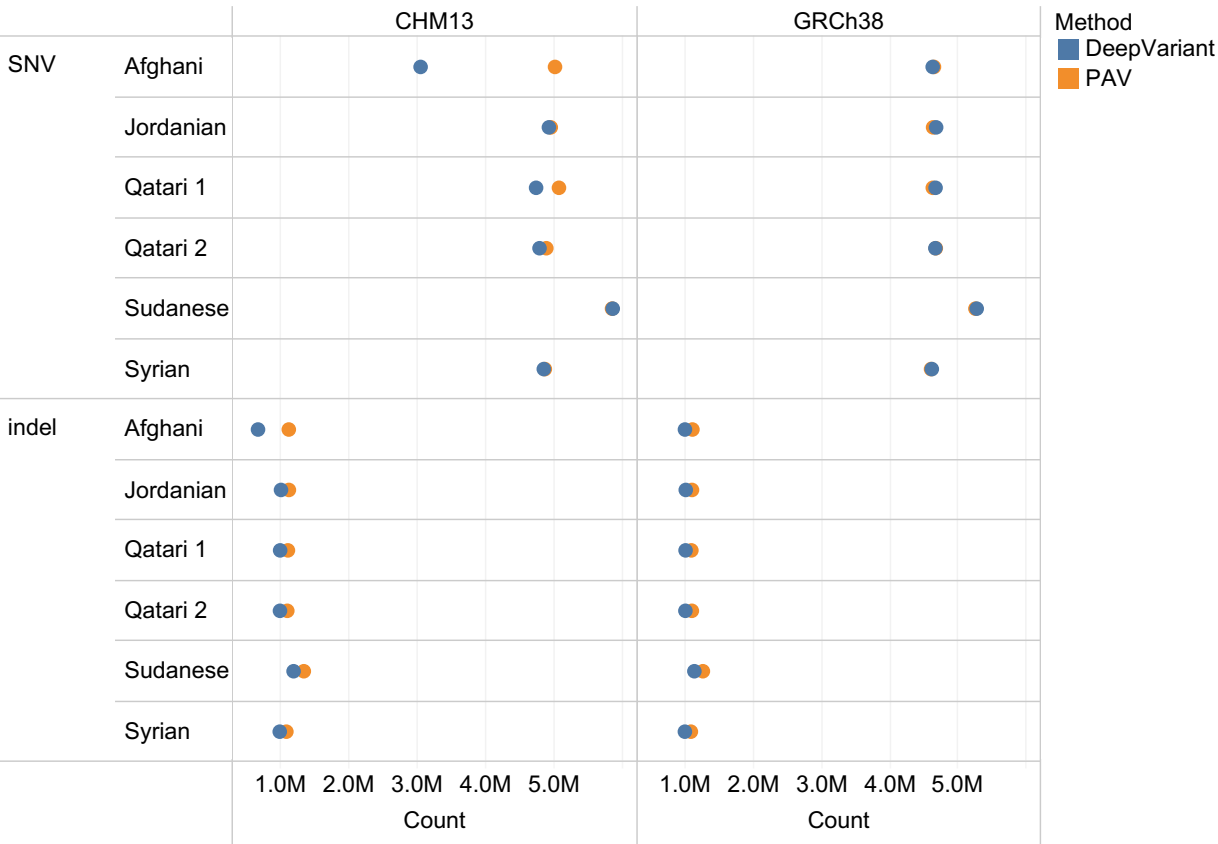
TAP1



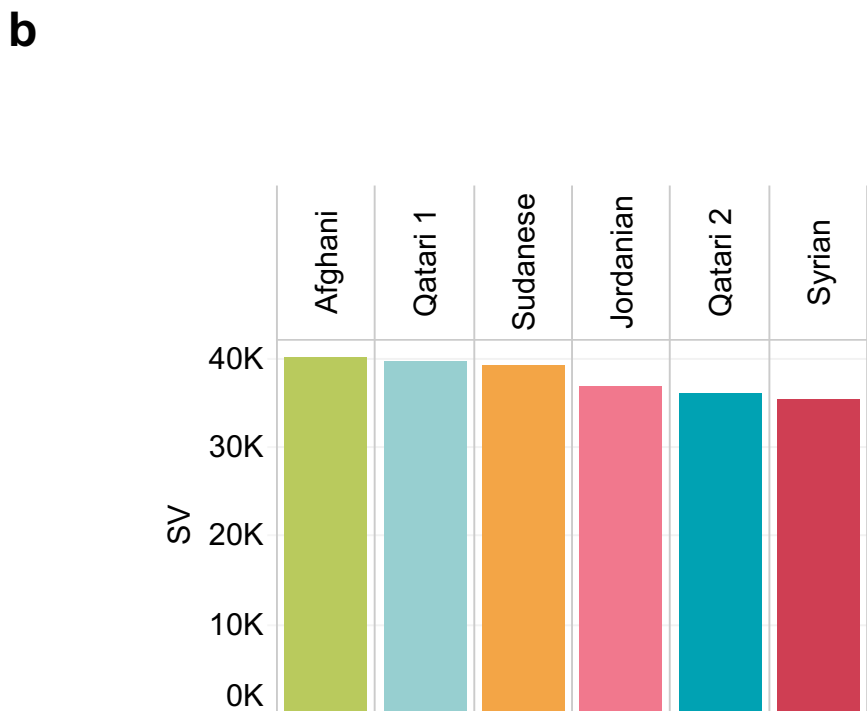
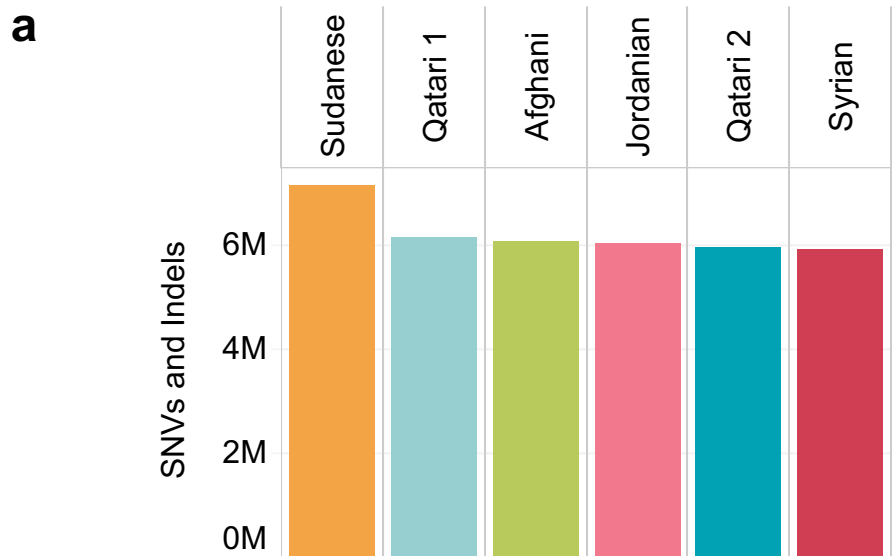
TAP2



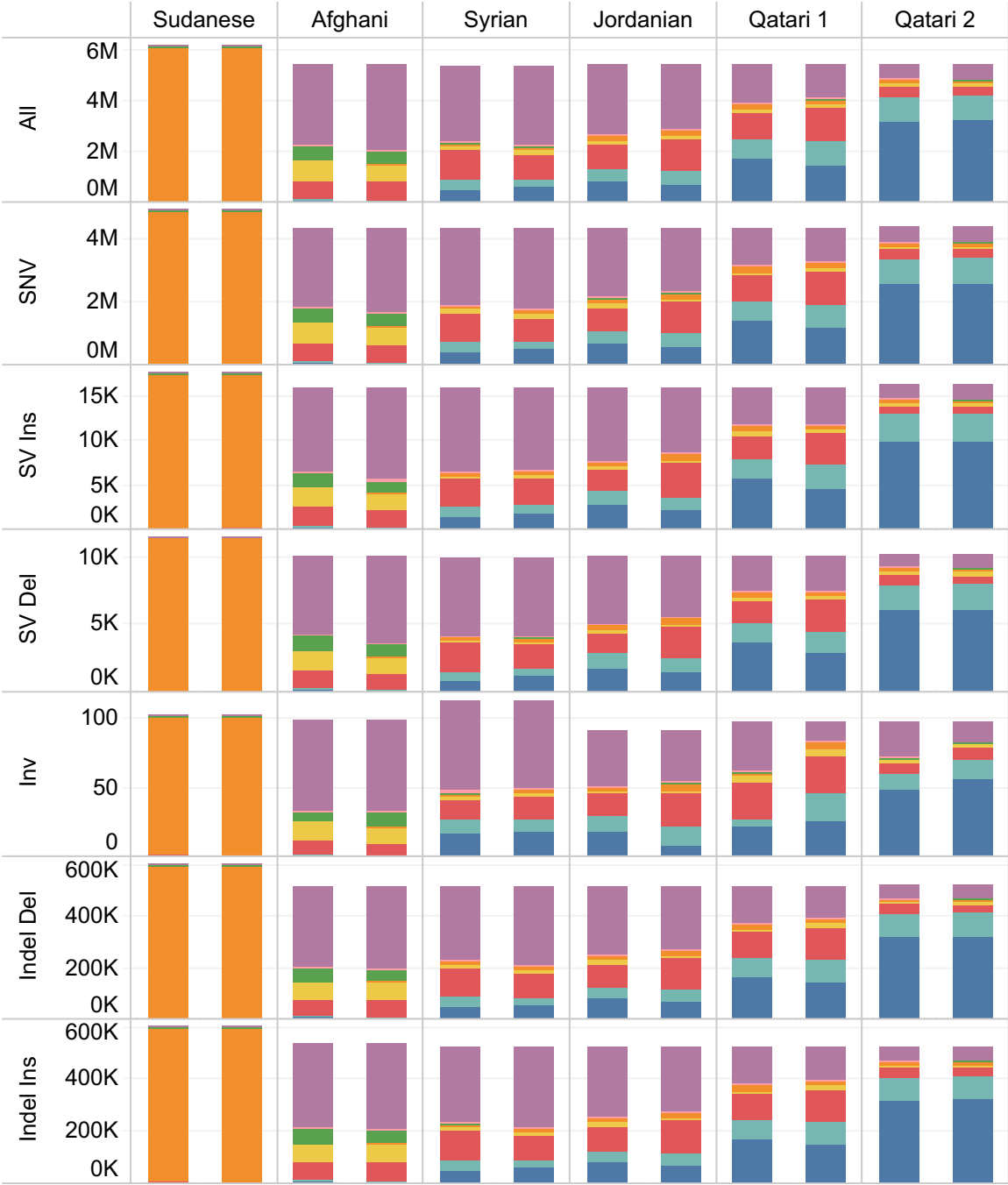
**Supplementary Fig. 17: Variant calls using read-based and assembly-based methods.** The plot shows a comparison of variant counts identified by DeepVariant (read-based method) and PAV (assembly-based method) against CHM13 and GRCh38.



**Supplementary Fig. 18: Variant calls against CHM13 for child assemblies. (a) SNV and indel calls. (b) SV calls.**



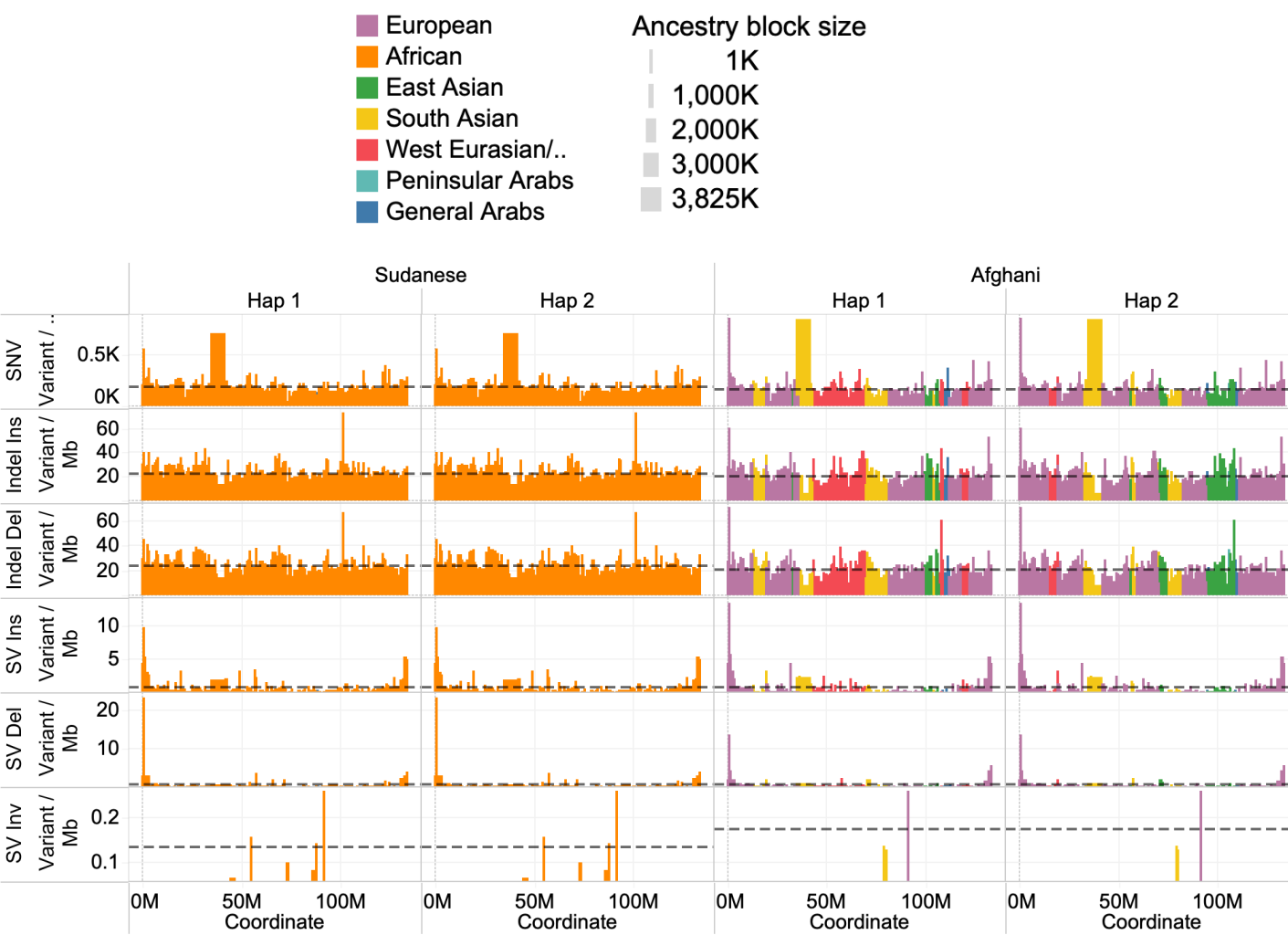
Supplementary Fig. 19: Variant calls binned per local ancestry.

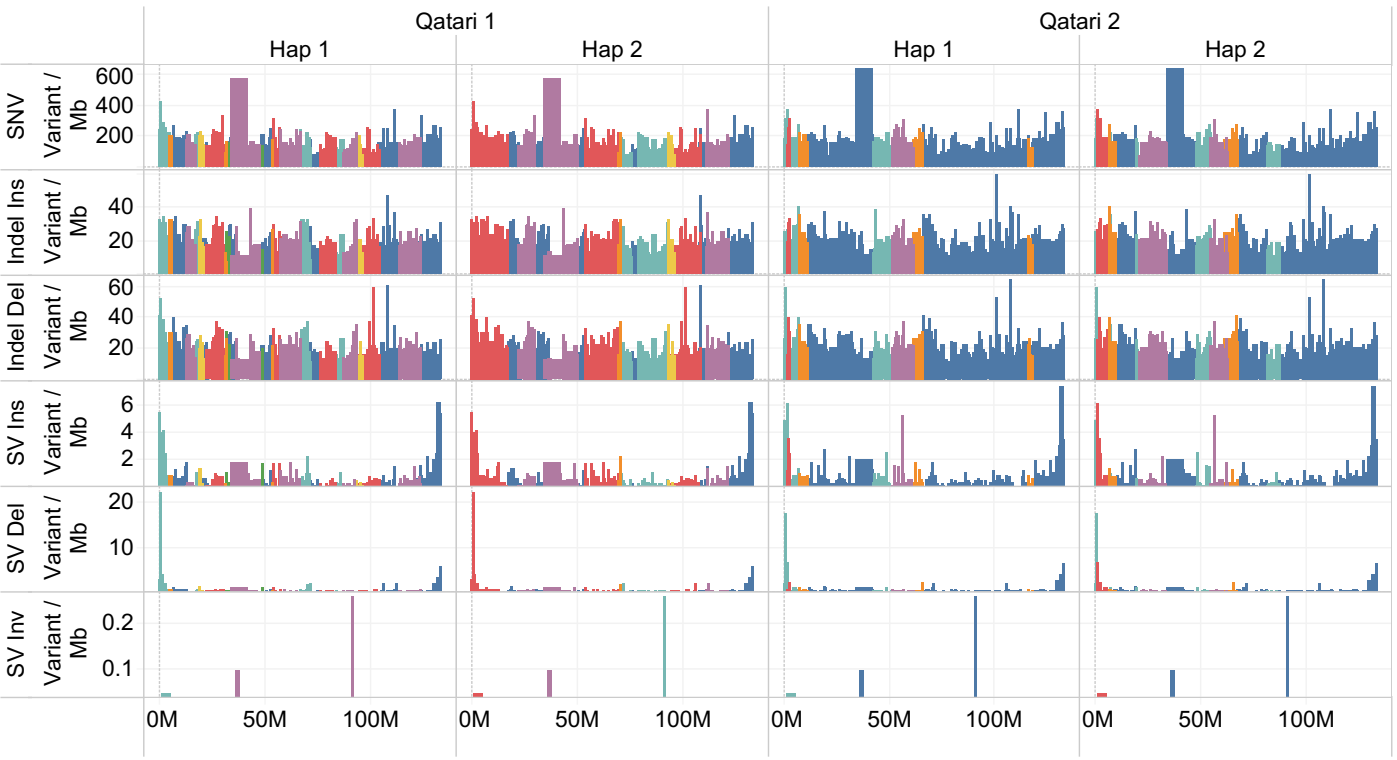
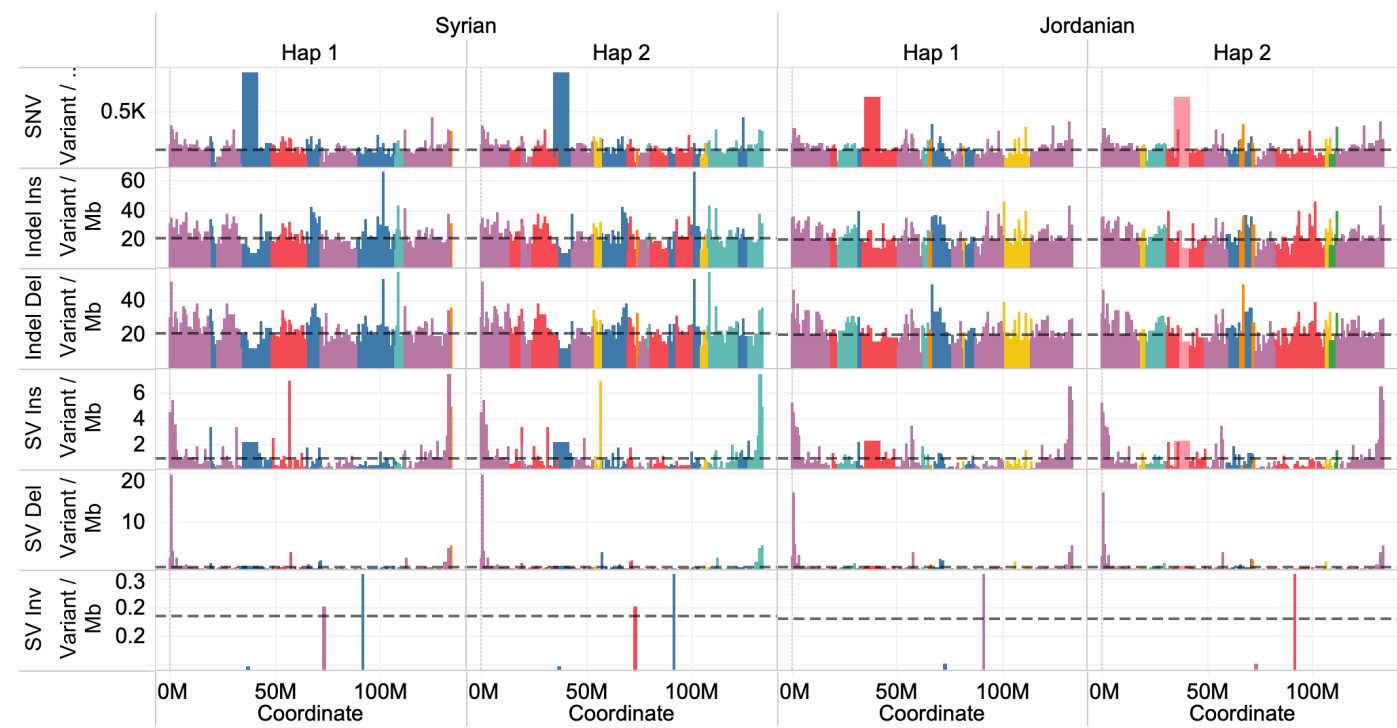


- European
- African
- South American
- South Asian
- East Asian
- West Eurasian/Persian
- Peninsular Arab
- General Arab

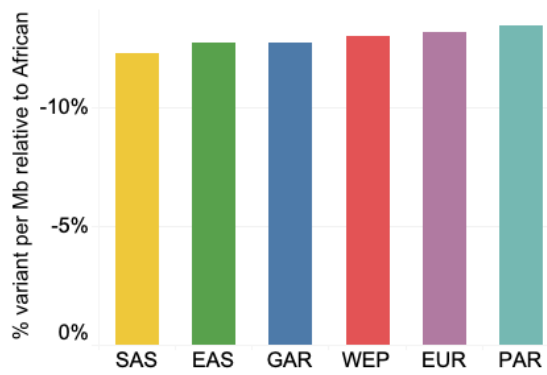
**Supplementary Fig. 20: Genetic variation as a function of local ancestry (a)** SNV, Indel and SV insertions and deletions per 1Mb of local ancestry segment across chromosome 10. Bar plot showing number of variant calls per 1 Mbp on chr10 across haplotypes. The color represents different ancestries, and the size of each bar corresponds to the length of the ancestry stretch. The horizontal dotted line for each SV type indicates the average count of variant per Mb. **(b)** Bar plot showing median counts of variants per MB relative to African segments in the same subjects aggregated by family, for various ancestries.

**a**

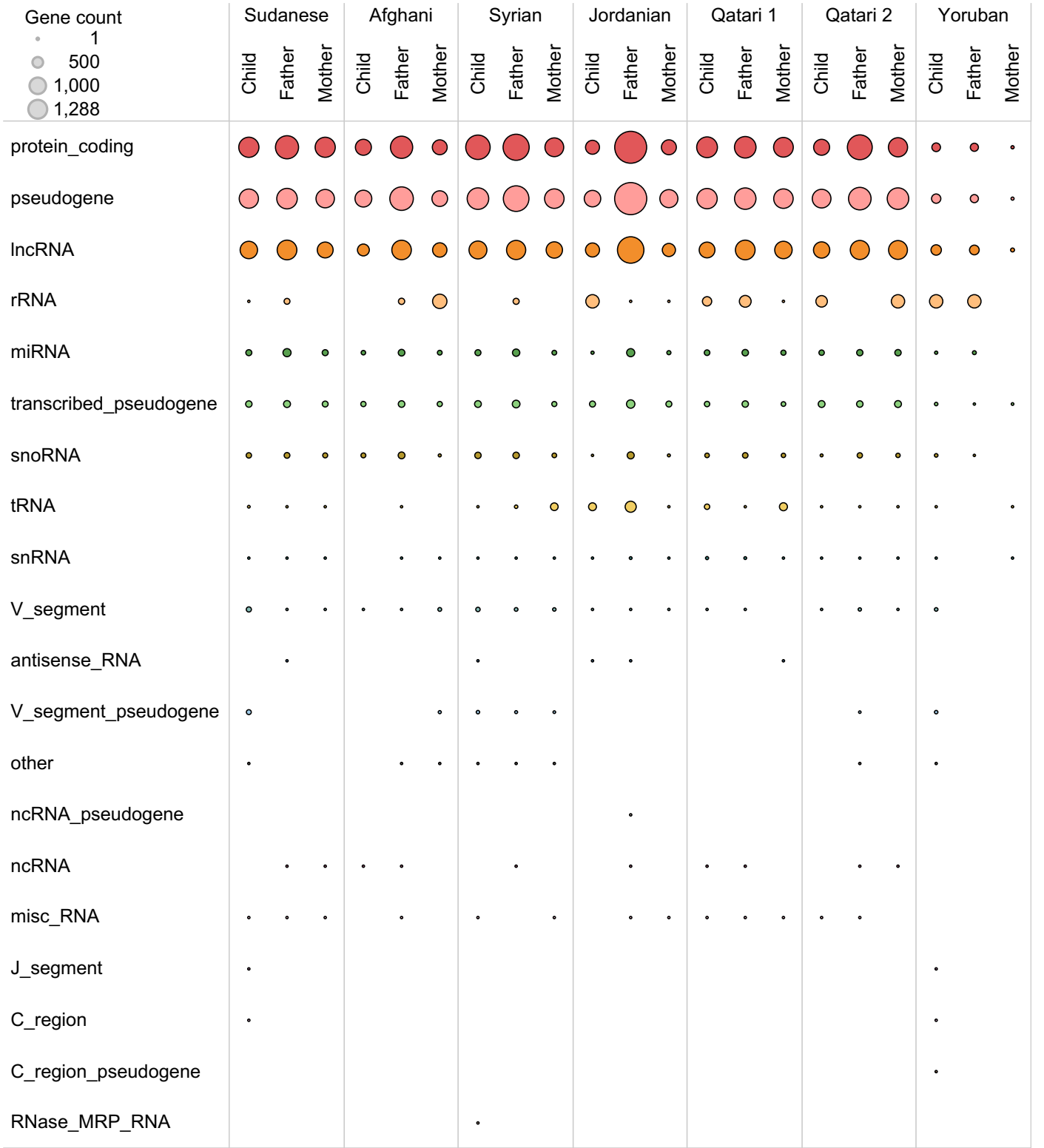




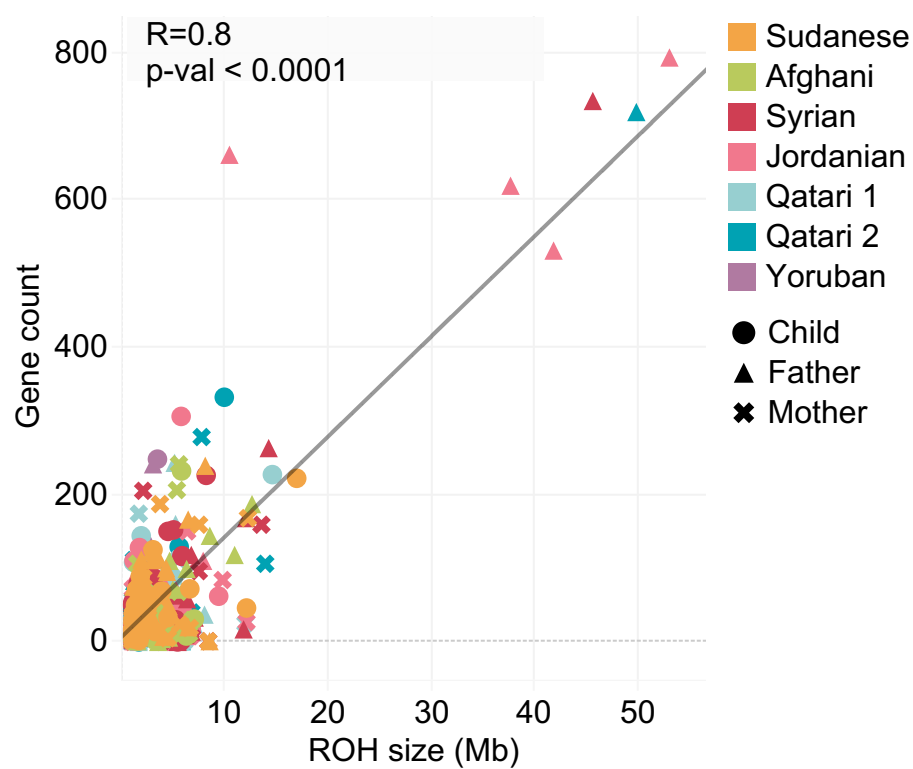
**b**



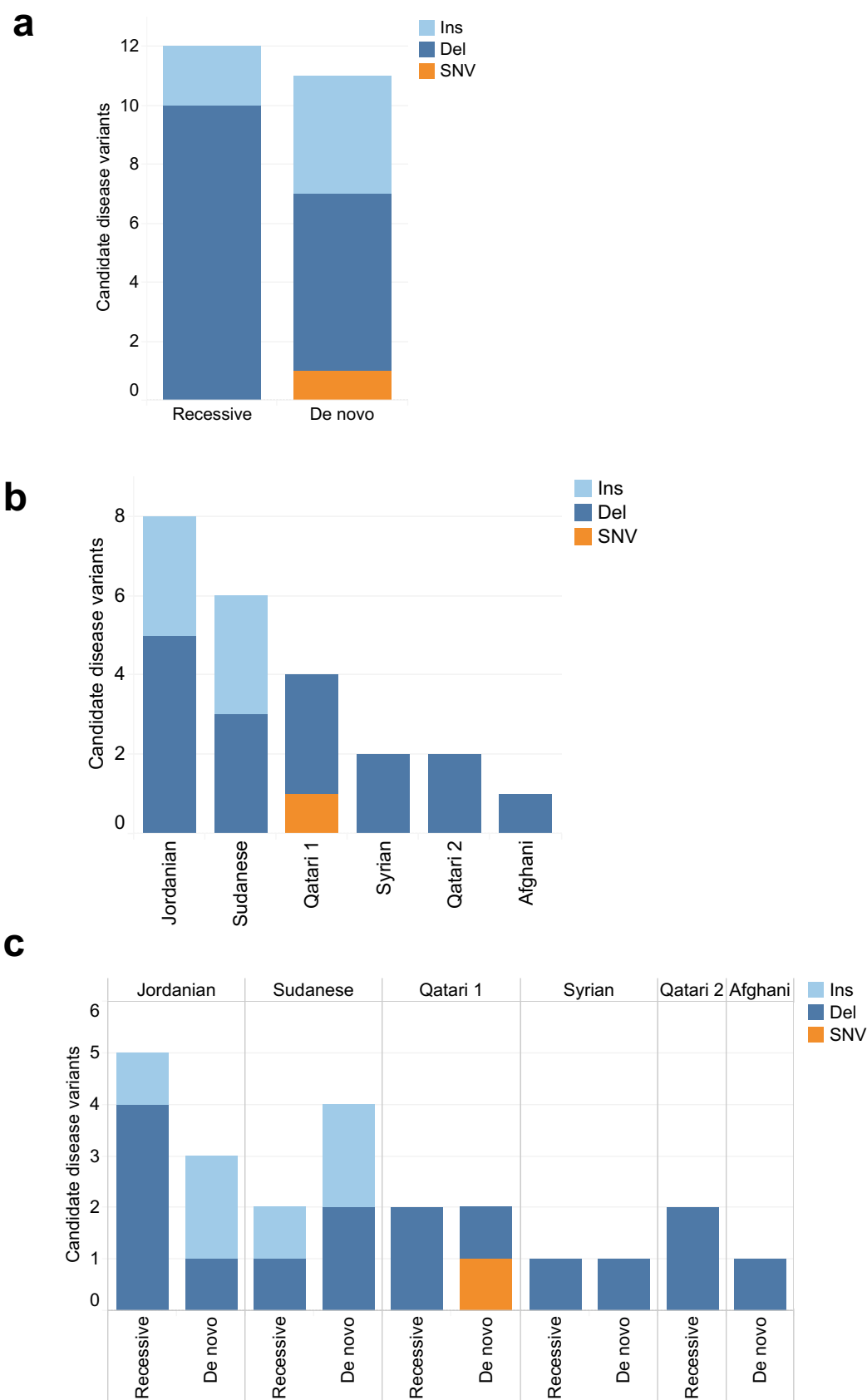
**Supplementary Fig. S21: Runs of Homozygosity (ROH) per gene type.**  
 Total count of genes per gene type in the ROH segments.



**Supplementary Fig. 22: Correlation between ROH size and count of encompassed genes.**

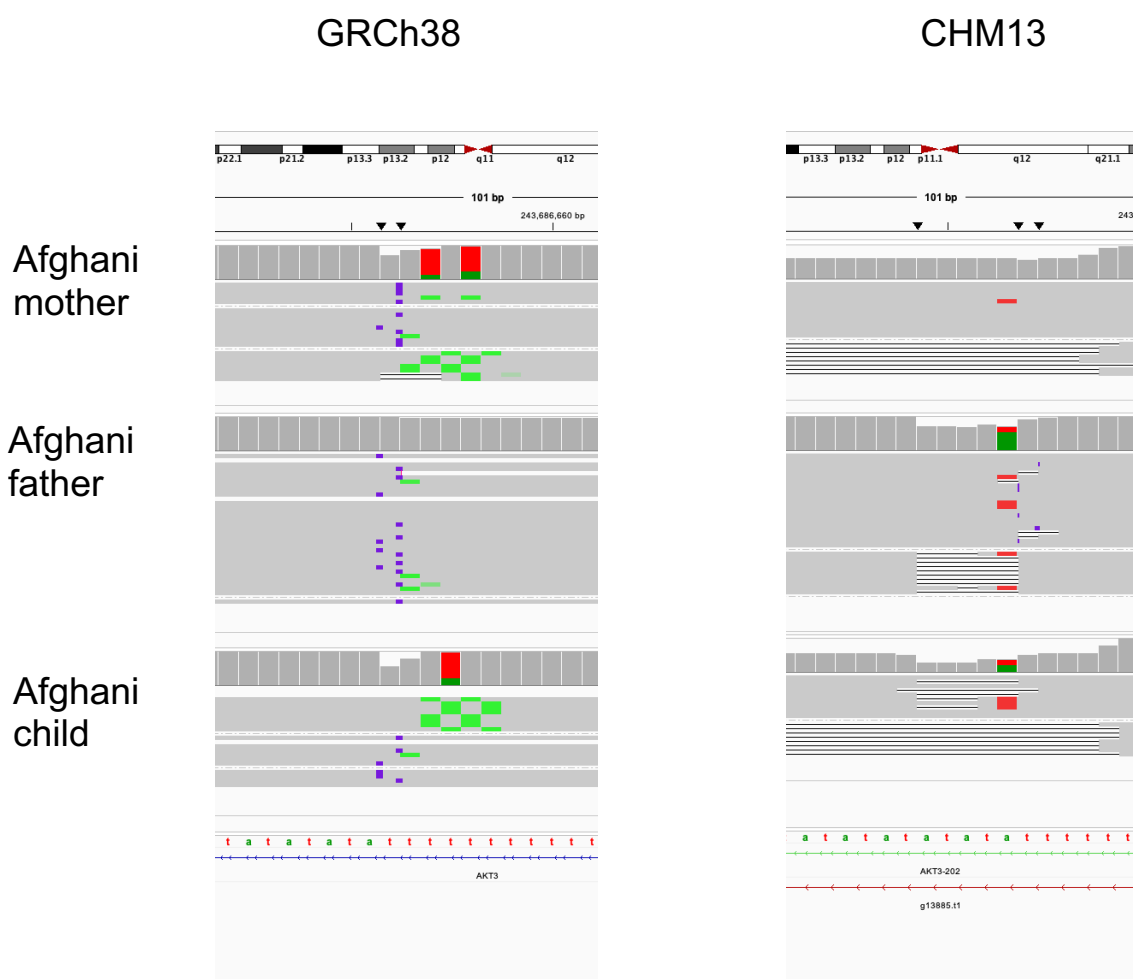


**Supplementary Fig. 23: Candidate disease variants short listed from family trio analysis. (a)** Count of variants per type and mode of inheritance **(b)** Count of variants per subject and variant type **(c)** Count of variants per subject, variant type and mode of inheritance.



**Supplementary Fig. 24: Prioritized variants from classical read-based approach found to be false positive *de novo*.** Panels show IGV view of alignments of long reads against either GRCh38 or CHM13, for the mother, father and child for a given trio for *AKT3*, *TRPM3* and *KMD4B*. For *KMD4B*, alignments are also shown against Jordanian father and HG002. Bar plots indicate coverage. Purple denotes insertion, while other colors denote different bases as per IGV default color scheme.

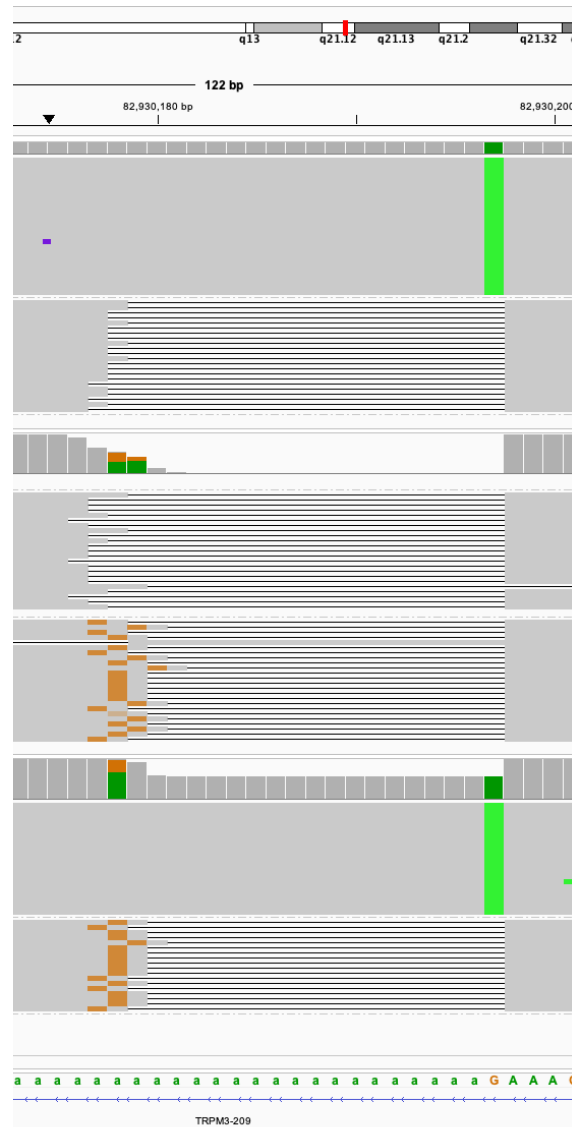
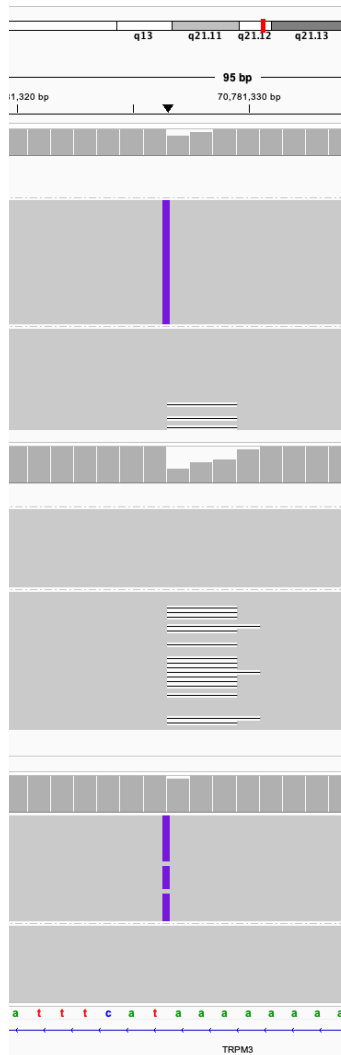
*AKT3*



*TRPM3*

GRCh38

CHM13



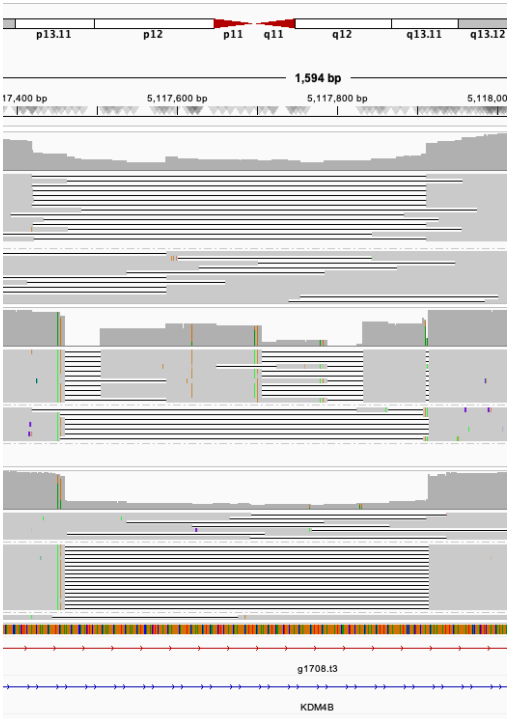
GRCh38

CHM13

Syrian mother

Syrian father

Syrian child



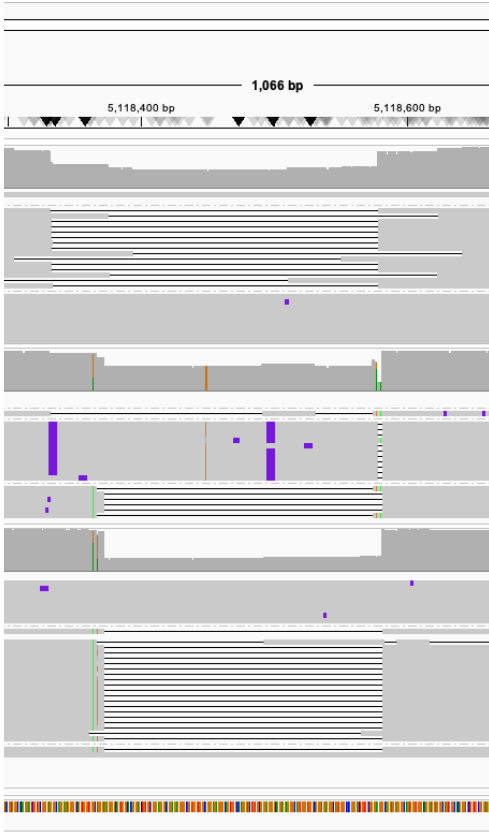
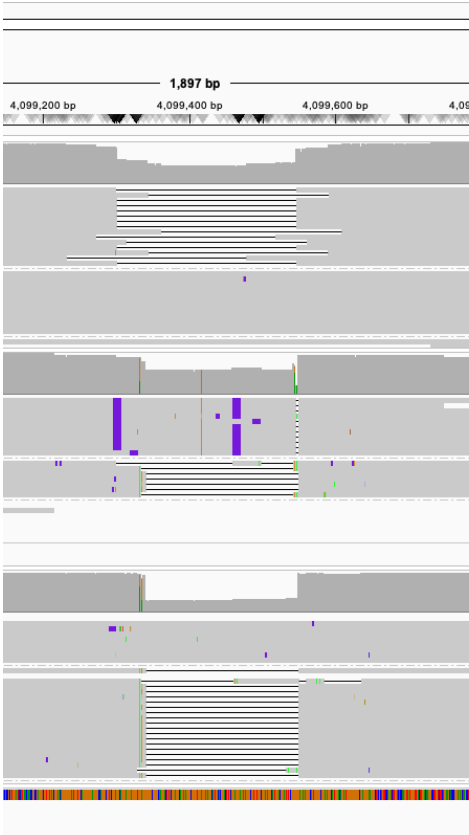
Jordanian father

HG002

Syrian mother

Syrian father

Syrian child



**Supplementary Fig. 25: Count of singletons for ME query samples called against MER and other genome references including GRCh38.**  
Difference in count of singletons in the replacement chromosomes relative to MER.

