Agenda of IARC meeting 81, October 4th, 2021

In attendance: Corey Watson, Ayelet Peres, Gur Yaari, Andrew Collins, Martin Corcoran, William Lees, Mats Ohlin

1. **Approval of minutes of meeting 79**
   Approved

2. **Approval of minutes of meeting 80**
   Approved

3. **Reflection on learning of IARC's allele approval process**
   postponed

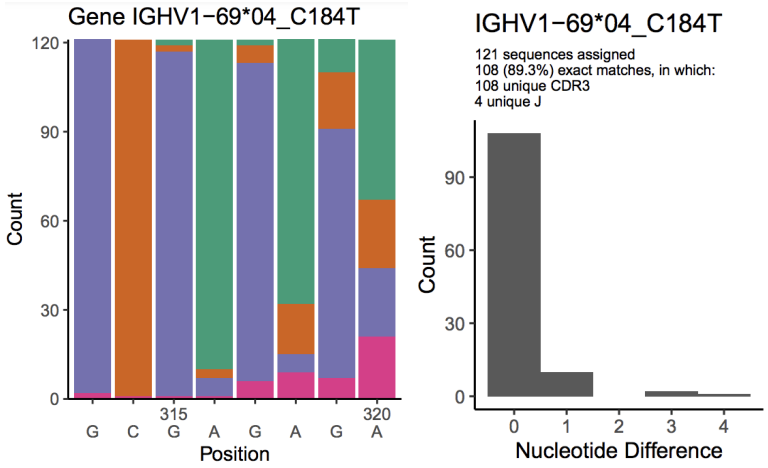
4. **Naming of alleles with no absolute, defined genomic location - using closest allele names as in the past or just OGRDB seq ID when reporting new alleles to IUIS**

   Discussion. It was determined that for human IGHV genes we will continue to report names involving the most similar human IGHV gene as in the past. Care will be taken to report uncertainty regarding the gene involved in cases of duplicated genes and highly similar genes. Other principles might be applied in species with less well characterized immunoglobulin loci.

5. **Assessment of inference IGHV1-69*04_C184T**
   IGHV1-69*04_C184T was inferred in one genotype in VDJbase P1 data sets (VDJbase P1_I50_S1). This inference has previously been pre-assessed at IARC meeting 60 (https://www.antibodysociety.org/wordpress/wp-content/uploads/2020/12/Meeting-60-1_9_20-minutes.pdf) in P1_I77_S1. This new allele represents a new variant that affects amino acid residue 62, implicated in some stereotypic immune responses. The genotype is also implied to carry IGHV1-69*02, IGHV1-69*04 and IGHV1-69*08, the latter of which are all mutated and likely represent rearrangements derived from IGHV1-69*02 (a short allele in the reference database that differs by one base from IGHV1-69*08) that also

incorporated the final base(s) of the IGHV gene into their sequence and thus were misassigned to IGHV1-69*08 by IgBLAST. The novel allele is the most expressed allele in the repertoire (40% allelic frequency; 2.4% of the total unmutated population). It is represented by 108 unmutated sequences and 108 unique CDR3s in the unmutated set. Haplotyping based on allelic diversity in IGHJ6 is not possible. IARC infers the sequence at Level 1 only up to and including base 319 in agreement with past practice. It is acknowledged that the allele most likely carries one additional base, typically A, at base position 320. A trailing "." indicates IARC's opinion that the sequence is likely to contain an additional 3' nucleotide for which there is insufficient evidence to make an affirmation.

OGRDBplots data on IGHV1-69*04_C184T:



The allele is given the name IGHV1-69*i03.

>IGHV1-69*i03 (IGHV1-69*04_C184T)
CAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGGGTCC
TCGGTGAAGGTCTCCTGCAAGGCTTCTGGAGGCACCTTCAGCAGCTATG
CTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCTTGAGTGGATGG
GAAGGATCATCCCTATCTTTGGTATAGCAAACTACGCACAGAAGTTCCAG
GGCAGAGTCACGATTACCGCGGACAAATCCACGAGCACAGCCTACATGG
AGCTGAGCAGCCTGAGATCTGAGGACACGGCCGTGTATTACTGTGCGAG
AG.

We recognize that this allele might be located at IGHV1-69 and/or IGHV1-69D and IARC gene naming does not reflect a position on this matter.

6. **Followup: The problem of identical sequences associated to different/duplicated genes in germline gene databases and the impact of**

multiple assignments in sequence analysis and VDJbase's representation of genotypes/haplotypes

Presentation and initial discussion of AP's visualization/description of data describing assessment of human IGHV genes that are affected by multiple assignments.

7. **Next meeting**
   To be decided.