IARC Meeting 60: September 1st 2020: minutes

The meeting commenced at 10:00 UTC. AC, MC, MO, CS and WL were in attendance. Corey Watson and Gur Yaari were also present.

1. The minutes of the previous two meetings were considered. Although members were generally happy with the minutes, some members felt they had not given them sufficient review. It was therefore agreed that the minutes would be accepted, but that if concerns were raised in the following week, the minutes would be modified and agreement reached via email .
2. It was agreed that the meeting would continue to consider VDJbase inferences, focusing first on inferences from P1 datasets that had been given a preliminary review via the excel spreadsheet googledoc that was set up following meeting 59. These sequences have not been formally submitted to the IARC, and this cannot happen until submissions have been made to Genbank. It was agreed that if these reviews led the committee to judge an inference to be true, post-submission consideration of the inference would be confined to the determination of the terminal nucleotides of the sequence. Decisions would then be ratified, and IMGT notified.
3. The meeting considered the inference of the variant IGHV1-69*06_g240a. The sequence was highlighted from the sample P1_I48_S1. The sequence was seen in 2.43% of all unmutated rearrangements, with 648 sequences including 598 perfect matches to the inferred allele. There was abundant variation in the CDR3 regions of the aligned sequences. IGHV1-69*01 and IGHV1-69*06 were also present in the genotype, at generally similar frequencies (*01: 6.7% of all unmutated sequences, 1755 sequences,1646 unmutated sequences; *06: 2.07% of all unmutated sequences, 564 sequences,509 unmutated sequences). The greater expression of the *01 allele is a familiar feature of datasets seen by members. CW noted that the presence of IGHV1-69-2*01 in the genotype stands as evidence of a likely gene duplication of IGHV1-69. Plots of the final 3' nucleotides are presently unavailable The sequence was affirmed as a Level 1 sequence, and the final 3' nucleotides will be considered at a later date, at which time the affirmed sequence will be noted in the IARC minutes.
4. The meeting considered the inference of the variant IGHV1-69*04_c184t. The sequence was highlighted in samples P1_I77_S1 and P4_I27_S1. The committee particularly focused on the evidence from the P1 sample, but noted the strong evidence in support of this inference from haplotype data of the P4 sample. In the P1 sample, the sequence was seen in 1.79% of all unmutated rearrangements, with 410 sequences including 351 perfect matches to the inferred allele. There was abundant variation in the CDR3 regions of the aligned sequences. IGHV1-69*02 and IGHV1-69*08 were also present in the genotype. IGHV1-69*02 was present at a similar frequency (1.52% of all unmutated sequences, 334 sequences, 297 unmutated sequences). The IGHV1-69*08 sequence was seen in 232 alignments, but only 3 of these were perfect matches. MO pointed out that the *08 allele is almost identical to the *02 allele, and they are both quite different to most other reported alleles. Although a

full-length version of the *02 allele has been published, the IMGT reference sequence is 2 nucleotides shorter than the *08 sequence. This results in frequency misalignments. MO and CW agreed that it is likely that the *08 ‘allele’ was reported in error. Plots of the final 3’ nucleotides of the IGHV1-69*04_c184t inference are presently unavailable The sequence was affirmed as a Level 1 sequence, and the final 3’ nucleotides will be considered at a later date, at which time the affirmed sequence will be noted in the IARC minutes.

5. The next meeting (Meeting 61) will be held on Tuesday September 22nd at 10:00 UTC.

The meeting ended at 11:05 UTC.