

IARC Meeting 46: December 2nd 2019: minutes

The meeting commenced at 10:30 UTC. AC, MC, CS, MO and WL were in attendance.

1. The minutes of meeting 45 were considered, but required decisions regarding the nomenclature for Level 0 sequences before they could be finalised. It was agreed that the *i01-type nomenclature will be used for Level 0 sequences. The minutes of Meeting 45 will be edited accordingly, and considered again at Meeting 47.
2. The committee considered the views of CS regarding three sequences that were discussed at Meeting 45, and came to final decisions. Details regarding the frequency of these sequences within the S00010 dataset, and other relevant information, can be found in the minutes of the IARC Meeting 45.

IGHV4-61*01_S5549 (C136G A138G) was affirmed as the Level 1 sequence, IGHV4-61*i02.

IGHV1-69*01_S7220 (G163A) was affirmed as the Level 1 sequence, IGHV1-69*i02.

IGHV2-70*04_S4496 (C6T) was affirmed as the Level 0 sequence, IGHV2-70*i02.

3. The committee commenced the evaluation of additional inferences in S00010. These sequences were parts of the genotypes previously submitted as S00004 and S00006. However these particular sequences, although assessed during Meetings 7-12, had not been assigned as new inferences in these OGRDB submissions. The current submission was made after submission of sequences to Genbank and select sets to SRA, and has now been reported as part of a new IgDiscover analysis.
 - The committee considered IGHV4-4*08_S8301 (T165C,C237A) of Genotype B16, which was previously evaluated as IGHV4-4*08_S4567 at Meeting 12. It was affirmed as a Level 0 sequence at that time, because of haplotype analysis that at the time was considered inconclusive. The submitted sequence is as follows:

```
>IGHV4-4*08_S8301
CAGGTGCAGCTGCAGGAGTCGGGCCAGGACTGGTGAAGCCTTCG
GAGACCCTGTCCCTCACCTGCACTGTCTCTGGTGGCTCCATCAGTA
GTTACTACTGGAGCTGGATCCGGCAGCCCCAGGGAAGGGACTGG
AGTGGATTGGGTACATCTATACCAGTGGGAGCACCAACTACAACCC
CTCCCTCAAGAGTCGAGTCACCATATCAGTAGACACGTCCAAGAACC
AGTTCTCCCTGAAGCTGAGCTCTGTGACCGCCGCAGACACGGCCGT
GTATTACTGTGCGAGAGA
```

The sequence was seen in 1.81% of all unmutated rearrangements, with 12648 sequences including 5107 perfect matches to the inferred allele. There was abundant variation in the CDR3 regions of the aligned sequences. One other IGHV4-4 allele was present in the genotype, as well as two likely IGHV4-59 sequences and one IGHV4-61 sequence. Haplotype data is supportive of the inference, though the haplotype ratios likely reflect the presence of some artefacts within the dataset. The ratios for the two IGHV4-4 sequences were 75:25 and 81:19, suggesting that both sequences are carried on the IGHJ6*02-bearing chromosome. The sequence, up to and including nucleotide 318, was affirmed as a Level 1 sequence. Uncertainty regarding nucleotides 319 and 320 will be indicated in IARC and OGRDB publications by two dots at the end of the affirmed sequence. The committee believes it is most likely that the 319th nucleotide is 'C'.

The affirmed sequence is as follows:

```
>IGHV4-4*02
CAGGTGCAGCTGCAGGAGTCGGGCCCAGGACTGGTGAAGCCTTCG
GAGACCCTGTCCCTCACCTGCACTGTCTCTGGTGGCTCCATCAGTA
GTTACTACTGGAGCTGGATCCGGCAGCCCCCAGGGAAGGGACTGG
AGTGGATTGGGTACATCTATAACAGTGGGAGCACCAACTACAACCC
CTCCCTCAAGAGTCGAGTCACCATATCAGTAGACACGTCCAAGAACC
AGTTCTCCCTGAAGCTGAGCTCTGTGACCGCCGCAGACACGGCCGT
GTATTACTGTGCGAGA..
```

- The committee considered IGHV4-59*01_S8652 (T109C) of Genotype B16, which was previously evaluated as IGHV4-59*01_S5140 at Meeting 12. It was affirmed as a Level 0 sequence at that time. The submitted sequence is as follows:

```
>IGHV4-59*01_S8652
CAGGTGCAGCTGCAGGAGTCGGGCCCAGGACTGGTGAAGCCTTCG
GAGACCCTGTCCCTCACCTGCACTGTCTCTGGTGGCTCCATCAGTA
GTCACTACTGGAGCTGGATCCGGCAGCCCCCAGGGAAGGGACTGG
AGTGGATTGGGTATATCTATTACAGTGGGAGCACCAACTACAACCC
TCCCTCAAGAGTCGAGTCACCATATCAGTAGACACGTCCAAGAACCA
GTTCTCCCTGAAGCTGAGCTCTGTGACCGCTGCGGACACGGCCGTG
TATTACTGTGCGAGAGA
```

The sequence was seen in 3.03% of all unmutated rearrangements, with 18133 sequences including 8545 perfect matches to the inferred allele. There was abundant variation in the CDR3 regions of the aligned

sequences. One other IGHV4-59 allele was present in the genotype, as well as two IGHV4-4 sequences and one IGHV4-61 sequences. Haplotype data is supportive of the inference, though the haplotype ratios likely reflect the presence of some chimeric sequences within the dataset. The ratios for the two IGHV4-59 sequences were 77:23 for the inference and 19:81 for the *01 allele, suggesting that the sequences are carried on different chromosomes. The sequence, up to and including nucleotide 319, was affirmed as a Level 1 sequence, but was subsequently identified as the recently reported allele IGHV4-59*11. The affirmed sequence will be submitted to IMGT as IGHV4-59*i02, as a record of the rearrangability (and therefore likely functionality) of the sequence.

The affirmed sequence is as follows:

```
>IGHV4-59*i02
CAGGTGCAGCTGCAGGAGTCGGGCCCAGGACTGGTGAAGCCTTCG
GAGACCCTGTCCCTCACCTGCACTGTCTCTGGTGGCTCCATCAGTA
GTCACTACTGGAGCTGGATCCGGCAGCCCCAGGGAAGGGACTGG
AGTGGATTGGGTATATCTATTACAGTGGGAGCACCAACTACAACCCC
TCCCTCAAGAGTCGAGTCACCATATCAGTAGACACGTCCAAGAACCA
GTTCTCCCTGAAGCTGAGCTCTGTGACCGCTGCGGACACGGCCGTG
TATTACTGTGCGAGAG.
```

3. The next meeting (Meeting 47) will be held on December 9th at 10:30 UTC.

The meeting ended at 11:40 UTC.