

Agenda and minutes of IARC meeting 70, June 15th, 2021 at 10.00 UTC

Present: Ayelet Peres (AP), William Lees (WL), Andrew Collins (AC), Martin Corcoran (MC), Corey Watson (CW), Mats Ohlin (MO)

1. Approval of minutes of IARC meeting 69

Approved

2. VDJbase – current developments Naming conventions/nomenclature

AP describes poor relation to pmIG data, problems with short sequences in IMGT, duplicated and highly similar alleles/genes, use of gene clusters.

Discussion on

- how gene grouping will affect biological and computational interpretation.
- implications on our inability to assess genomic location unless genomic data is available.
- what the implications are of having reference sets with/without duplicated genes/alleles and which recommendations should be made regarding this matter.

It was noted that the IARC's main objective is to identify novel alleles made by inference. Naming process could be made based on group names rather than current, most similar allele's name.

We will communicate with IMGT on our recommendations.

3. Update on the process of GenBank/SRA_ENA submission of novel inferred alleles provisionally assessed by IARC

AP has submitted data to ENA after discussion with ENA helpdesk. Waiting for final GenBank records.

4. IARC manuscript 2021 – focus

Germline focus

Naming structure and effect on analysis output. AP describes examples from VDJbase. MC suggests that similar analysis can be performed with IgDiscover. AP, MC, and GY will do some early studies and present at next IARC meeting.

5. Database generated from 1000 Genomes data

Discussion on germline gene databases generated from short read genomic sequencing and on an approach to establish a public discussion on such data.

6. Next Meeting: June 29, 2021 at 10.00 UTC