

Agenda of IARC meeting 90, Dec 20th, 2021

In attendance: Ayelet Peres, Gur Yaari, Andrew Collins, Martin Corcoran, William Lees, Mats Ohlin, James Heather (guest)

1. Approval of minutes of meeting 88

Approved

2. Approval of minutes of meeting 89

Approved

3. AP's web interface for germline gene repertoire assessment of AIRR-seq data (Functional groups reference book) and integration of information with output of OGRDB

New features including e.g. thresholds that have been set for all alleles in “Functional groups reference book”.

Summary of all alleles have been added to “Functional groups reference book”

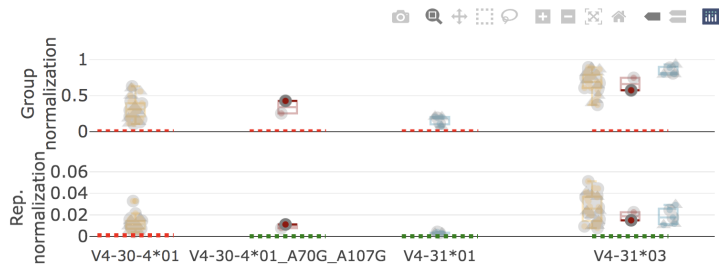
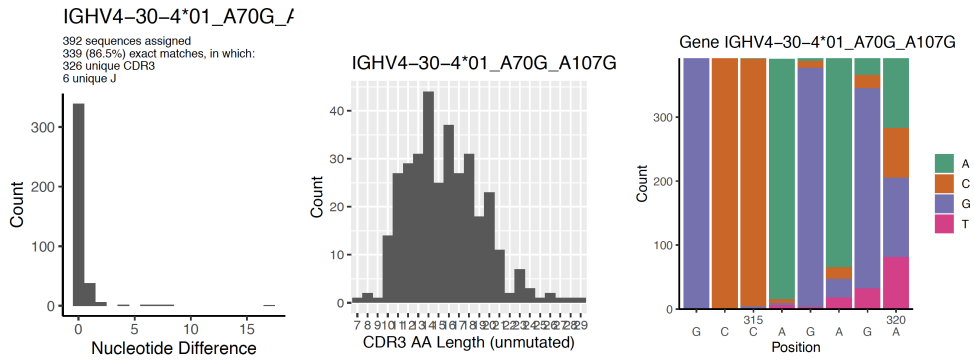
Discussion on how IGHD genes could be used for haplotyping, including deletion polymorphisms.

Report on upcoming information in VDJbase describing the status of novel alleles.

4. Additional novel alleles in VDJbase, study P1

VDJbase P1 study carries a number of other novel alleles that have not been affirmed in the past. Two of these are

- a. IGHV4-30-4*01_a70g_a107g (tentatively assigned to level 1 when pre-assessed at meeting 62) (VDJbase: P1_I41, P1_I49)
Identified in two haplotypable data sets. The quality of data is best for P1_I41 with many reads and a high frequency of perfect matches to the inferred allele. 43% frequency of assignments made to alleles of G35 of this subject in “Functional Groups Reference Book” (date: Dec 20th, 2022) (the other allele of G35 in this genotype is IGHV4-31*03; no allele of IGHV4-30-4 in this genotype is assigned to G34, but one allele of IGHV4-30-2). Both IGHV4-31*03 and IGHV4-30-4*01_a70g_a107g associate only to IGHJ6*03 (IGHV4-30-2*01 associate to both haplotypes).



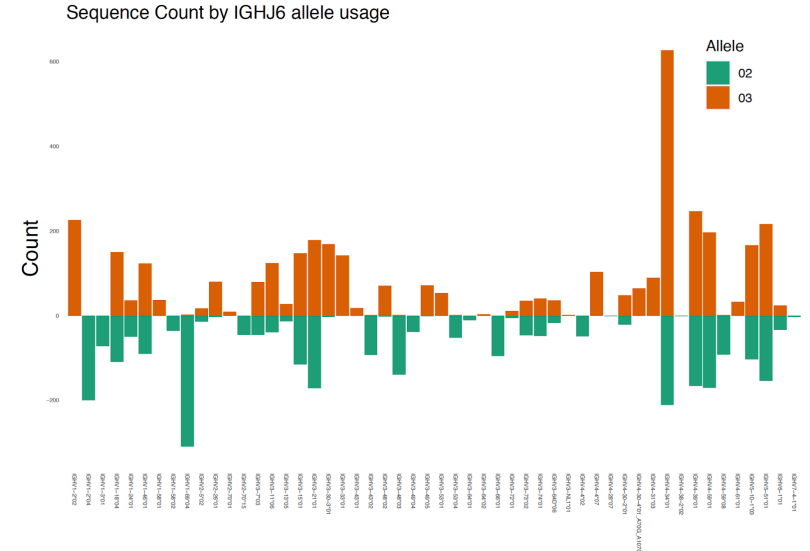
```

IGHV4-30-2*01 CAGCTGCAGCTGCAGGAGTCCGGCTCAGGACTGGTGAAGCCTTCACAGACCCTGTCCTCACCTGCCGTGTCTCT
IGHV4-30-4*01_a70g_a107g CAGGTGCAGCTGCAGGAGTCCGGCCAGGACTGGTGAAGCCTTCACAGACCCTGTCCTCACCTGCCGTGTCTCT
IGHV4-31*03 CAGGTGCAGCTGCAGGAGTCCGGCCAGGACTGGTGAAGCCTTCACAGACCCTGTCCTCACCTGCCGTGTCTCT

IGHV4-30-2*01 GGTGGCTCCATCAGCAGTGGTGGTTACTCTGGAGCTGGATCCGCAGCCACAGGGAAGGGCCTGGAGTGGATT
IGHV4-30-4*01_a70g_a107g GGTGGCTCCATCAGCAGTGGTGGTTACTACTGGAGTGGATCCGCCAGCCCCAGGGAAGGGCCTGGAGTGGATT
IGHV4-31*03 GGTGGCTCCATCAGCAGTGGTGGTTACTACTGGAGCTGGATCCGCCAGCCCCAGGGAAGGGCCTGGAGTGGATT

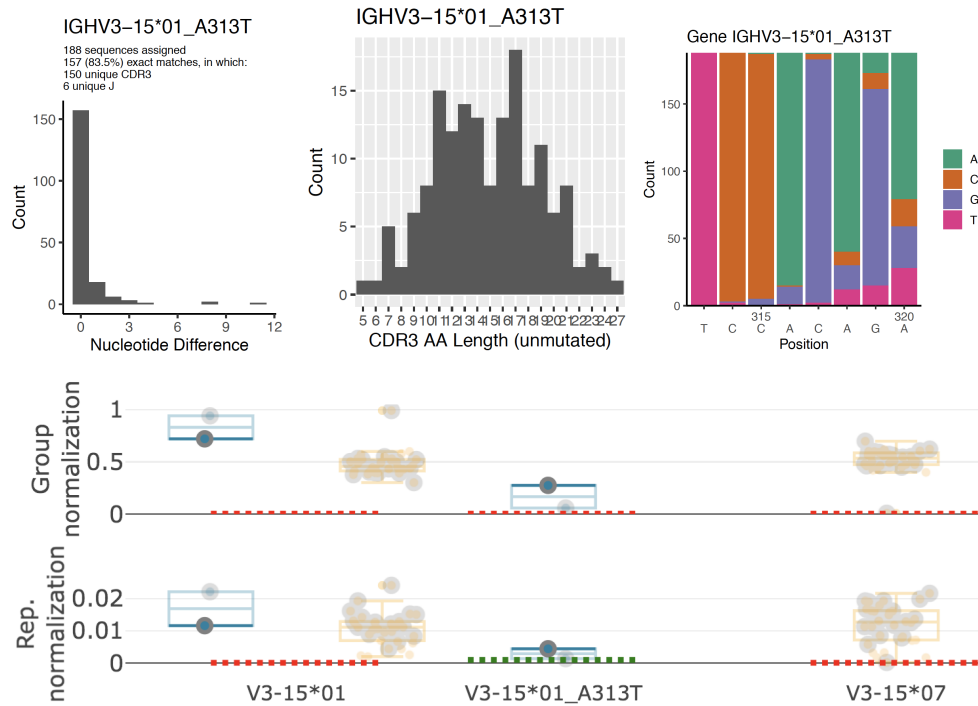
IGHV4-30-2*01 GGGTACATCTATCTAGTGGGAGCACCTACTACAACCCGTCCTCAAGAGTCGAGTCAACCATATCAGTAGACACG
IGHV4-30-4*01_a70g_a107g GGGTACATCTATACAGTGGGAGCACCTACTACAACCCGTCCTCAAGAGTCGAGTACCATATCAGTAGACACG
IGHV4-31*03 GGGTACATCTATACAGTGGGAGCACCTACTACAACCCGTCCTCAAGAGTCGAGTACCATATCAGTAGACACG

IGHV4-30-2*01 TCCAAGAACCAGTTCTCCCTGAAGCTGAGCTCTGTGACTGCCGGGACACGGCCGTGTATTACTGTGCCAGAGA
IGHV4-30-4*01_a70g_a107g TCCAAGAACCAGTTCTCCCTGAAGCTGAGCTCTGTGACTGCCGGGACACGGCCGTGTATTACTGTGCCAGAGA
IGHV4-31*03 TCCAAGAACCAGTTCTCCCTGAAGCTGAGCTCTGTGACTGCCGGGACACGGCCGTGTATTACTGTGCCAGAGA
  
```



b. IGHV3-15*01_a313t (or rather IGHV3-15*01_a313t_c317g) (VDJbase: P1_164)
 Allele was not found in a haplotypable data set. Good quality data in P1_164, many unmutated sequences, high diversity in CDR3. 28% frequency of assignments made to IGHV3-15 in “Functional Groups Reference Book” (date: Dec 20th, 2022) (the other allele of this genotype is IGHV3-15*01). OGRDB stats is convincing up to and

including base 319 but defines base 317 as G rather than C, i.e. the inference is IGHV3-15*01 A313T C317G. This finding sparked a discussion of how IgBLAST identifies allelic variants at the end of sequences. AP will conduct further studies.



5. Next meeting

January 10th, 2022 at 11.00 UTC