Minutes of IARC meeting 121, May 8th, 2023

In attendance: Ayelet Peres, Gur Yaari, Andrew Collins, Martin Corcoran, William Lees, Corey Watson, Mats Ohlin, James Heather
Guest: Tahel Ronel

1. **Approval of minutes of meeting 120**
   Approved

2. **Next meeting**
   May 15th, 2023 at 10.00 UTC

3. **Assessment of inference TRBV7-7*01_C315T in P4_I9_S1 (S00036)**
   TRBV7-7*01_C315T has been inferred in seventeen genotypes in the VDJbase P4 data set, including in VDJbase P4_I9_S1, a haplotypable data set (based on heterozygocity in TRBJ1-6). The genotype is also implied to carry TRBV7-7*01. No other gene apart from IGHV7-6 (alleles of which also carry C315) in the IMGT database is highly similar to these alleles of TRBV7-7. The novel allele is the most expressed allele in the repertoire (58% allelic frequency; 0.16% of the total error-free population). It is represented by 37 error-free sequences and 33 unique CDR3s in the error-free set. Haplotyping based on allelic diversity in TRBJ1-6 demonstrates association of TRBV7-7*01_C315T with only one of the haplotypes (only few recorded cases; TRBV7-7*01 was not associated with any allele of TRBJ1-6).

   IARC affirms the sequence based on inference of expression data alone at Level 1 up to and including base 325. It is acknowledged that the allele most likely carries 1 additional base, typically C, at base position 326. Trailing "." indicates IARC's opinion that the sequence is likely to contain additional 3'-nucleotides for which there is insufficient evidence to make an affirmation.

```
>TRBV7-7*i01 (TRBV7-7*01_C315T)
GGTGCTGGAGTCTCCCAGTCTCCCAGGTACAAAGTCACAAAGAGGGGACAGGAT
GTAACTCTCAGGTGTGATCCAATTTCGAGTCATGCAACCCTTTATTGGTATCAA
CAGGCCCTGGGGCAGGGCCCAGAGTTTCTGACTTACTTCAATTATGAAGCTCAA
CCAGACAAATCAGGGCTGCCCAGTGATCGGTTCTCTGCAGAGAGGCCTGAGGGA
TCCATCTCCACTCTGACGATTCAGCGCACAGAGCAGCGGGACTCAGCCATGTAT
CGCTGTGCTAGCAGCTTAG.
```

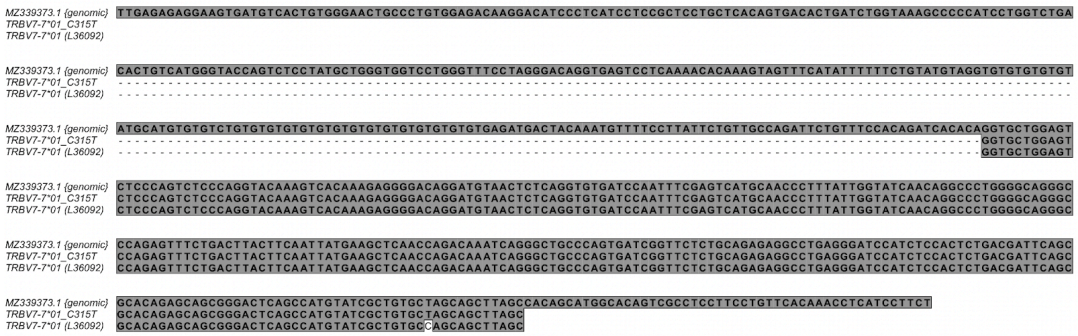
The allele has also been identified in a separate study as TRBV7-7*01_S0326 and Sanger validated (GenBank MZ339373) (Corcoran et al. (2023) Immunity 56, 635-652.E6 (DOI: 10.1016/j.immuni.2023.01.026)). It has been seen in a total of six Sanger sequenced genomic clones derived from two subjects.

>MZ339373
TTGAGAGAGGAAGTGATGTCACTGTGGGAACTGCCCTGTGGAGACAAGGACATC
CCTCATCCTCCGCTCCTGCTCACAGTGACACTGATCTGGTAAAGCCCCCATCCT
GGTCTGACACTGTC**ATGGGTACCAGTCTCCTATGCTGGGTGGTCCTGGGTTTCC**
TAGGGACAGGTGAGTCCTCAAAACACAAAGTAGTTTCATATTTTTTCTGTATGT
AGGTGTGTGTGTGTATGCATGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT
GTGTGAGATGACTACAAATGTTTTCCTTATTCTGTTGCCAGATTCTGTTTCCAC
AG**ATCACACAGGTGCTGGAGTCTCCCAGTCTCCCAGGTACAAAGTCACAAAGAG**
GGGACAGGATGTAACTCTCAGGTGTGATCCAATTTCGAGTCATGCAACCCTTTA
TTGGTATCAACAGGCCCTGGGGCAGGGCCCAGAGTTTCTGACTTACTTCAATTA
TGAAGCTCAACCAGACAAATCAGGGCTGCCCAGTGATCGGTTCTCTGCAGAGAG
GCCTGAGGGATCCATCTCCACTCTGACGATTCAGCGCACAGAGCAGCGGGACTC
AGCCATGTATCGCTGTGCTAGCAGCTTAGCCACAGCATGGCACAGTCGCCTCCT
TCCTGTTCACAAACCTCATCCTTCT
heptamer 23-mer nonamer

It has also been identified in a BAC clone with accession number AC229888

IARC affirms the sequence based on combined inference and genomic evidence at Level 1 up to and including base 326.

>TRBV7-7*i01 (TRBV7-7*01_C315T)
GGTGCTGGAGTCTCCCAGTCTCCCAGGTACAAAGTCACAAAGAGGGGACAGGAT
GTAACTCTCAGGTGTGATCCAATTTCGAGTCATGCAACCCTTTATTGGTATCAA
CAGGCCCTGGGGCAGGGCCCAGAGTTTCTGACTTACTTCAATTATGAAGCTCAA
CCAGACAAATCAGGGCTGCCCAGTGATCGGTTCTCTGCAGAGAGGCCTGAGGGA
TCCATCTCCACTCTGACGATTCAGCGCACAGAGCAGCGGGACTCAGCCATGTAT
CGCTGTGCTAGCAGCTTAGC



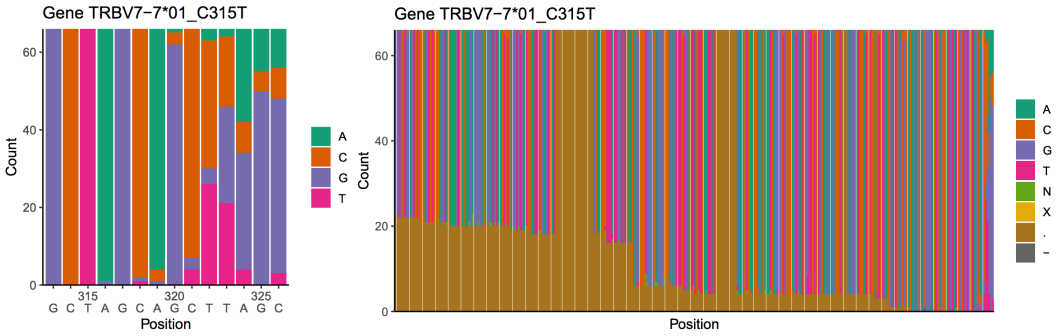
Result summary: TRBV7-7*01_C315T	No rearrangement found		
V-GENE and allele	Homsap TRBV7-7*01 F	score = 1380	identity = 100.00% (276/276 nt)
FR-IMGT lengths, CDR-IMGT lengths	[5.6.X]		

1. Alignment for V-GENE and allele identification

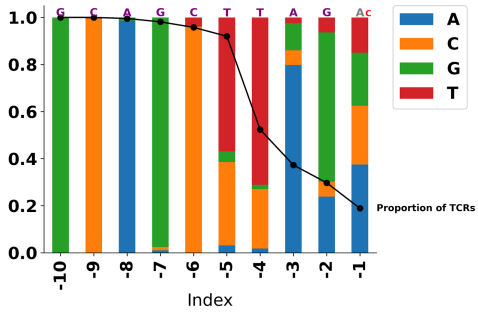
Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

		Score	Identity
L36092	Homsap TRBV7-7*01 F	1380	100.00% (276/276 nt)
X57607	Homsap TRBV7-7*02 (F)	1371	99.64% (275/276 nt)
L36092	Homsap TRBV7-6*01 F	1281	96.01% (265/276 nt)
X58806	Homsap TRBV7-6*02 (F)	1272	95.65% (264/276 nt)
M11953	Homsap TRBV7-8*01 F	1119	89.49% (247/276 nt)

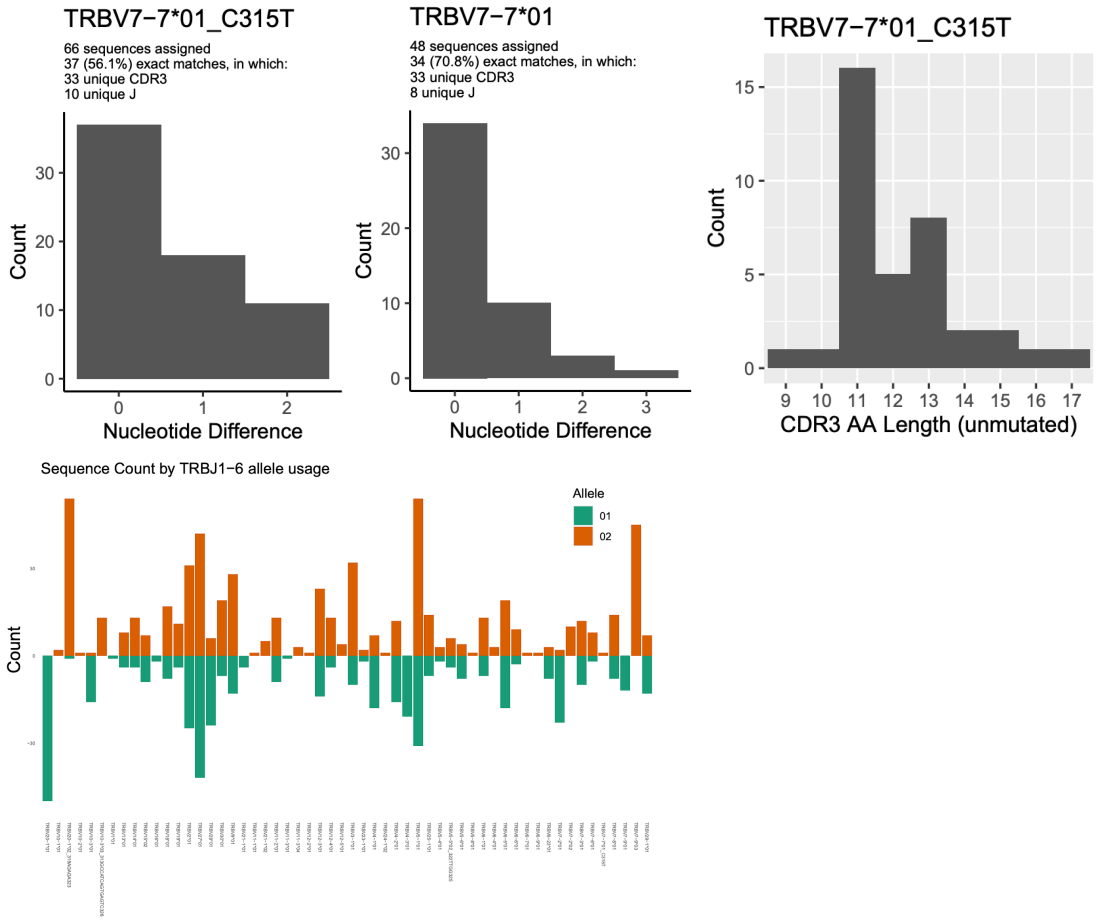
(Note: the difference between TRBV7-7*01_C315T and TRBV7-7*01 is outside of the range assessed by IMGT/V-QUEST)



Consensus plot of 3'-end:



All data:



4. **Germline reference set manuscript**
   AC presented analysis of data with a specific focus on the duplicated IGKV locus. Of note, information of evidence of different alleles have not been collected from the IMGT database.