Glossary of Terms Used in VDJbase

Allele and gene annotations in figures

Unk: In genotype graph, if a gene does not appear in a specific genotype sample, they are marked as unknown(Unk).
In haplotype inference, If the evidence for a gene is not strong enough (K is lower than a certain threshold (set to 1000 by default)), this allele is set to "Unk".

Del:Single or double chromosome deletion. Single chromosome deletion, deletion polymorphism, is inferred from RAbHIT haplotype function, where the K threshold for an Unknown assignment of a gene in a given chromosome is larger than a certain threshold (by default 1000). Double chromosome deletion is homozygousity for the deletion event, this is inferred using RAbHIT binomial test. In haplotype graphs both deletion events are present, however, in genotype graphs only double chromosome deletion is present.

NR: A non reliable gene (NR) is defined when the ratio of the multiple assignments with a gene in a haplotype is below the threshold. We use the RabHIT ratio threshold.

NRA: A non-reliable allele is a term for an ambiguous call. We collapse ambiguous allele using the RabHIT reliability scores. Each allele for which more than 60% of alignments are ambiguous calls are marked as unreliable (NRA), and later collapsed. For example, if IGHV3-23 allele 01 frequently appeared in ambiguous calls with allele 02 in more than 60% of the sequences, then the alignments are said to be 01_02 allele. The annotation is ordered numerically, for example IGHV270*01_10_11_13_15 will not give credence to *10.
NRA marked genes arealso designated '[*01]' in figures, and the full name is displayed at the bottom of the figure.

Certainty level of genotype and haplotype inference

Genotype Kdiff : The resulting genotypes include a measure of certainty of the genotype call for each gene (Kdiff). The larger the Kdiff, the greater the certainty of the allele chromosomal inference.

Haplotype Kdiff: The log of the Bayes factor (K) obtained from the haplotype Bayesian inference for each allele for a given gene. The larger the lK, the greater the certainty of the haplotype inference.