Instructions for the production of bed files to annotate the Immunoglobulin locus.

Download or copy and paste gene tables from imgt source to Microsoft excel.

A picture containing table

Description automatically generated

Figure: Example of imgt gene tables data. This table contains the coordinates of the locus genes which can be extracted to form bed files.

1. Use Ctrl + Shift to highlight “IMGT subgroup”, “Accession numbers” and “Positions in the sequence” columns. Paste onto a new excel sheet to produce data tables with genes and their corresponding coordinates.
2. Open file in command line and select rows with accession numbers corresponding to the reference used using *grep*.
3. Create two tab separated columns for the gene coordinates by replacing the “..” character with a tab, using *sed*.
4. Select the first allele of each “IMGT subgroup” category.