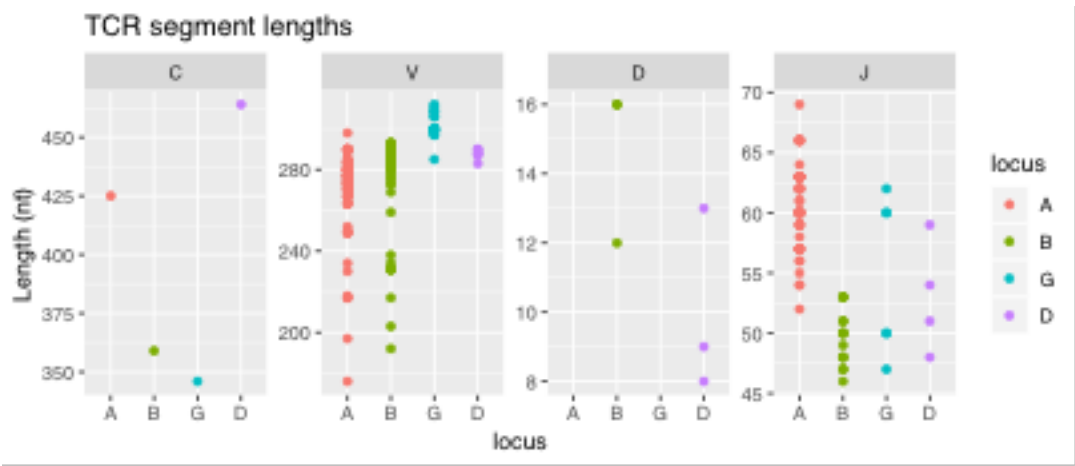


The curious case of the missing T cell receptor transcripts: part 3

Bonus post: TCR summary statistics

As long as I had all the human TCR segments in a tidy dataframe, I decided to display some basic properties. Here's the length and number of known segments by type and by locus.



Counts of known human TCR segments

	C	V	D	J
A	1	119	0	68
B	1	142	3	16
G	1	22	0	6
D	1	6	3	4

These counts are different from what my immunology textbook gives. (I'm reading "Primer to the Immune Response" by Tak Mak and Mary Saunders, and the TCR segment counts are in table 8-2, page 137.) Most of them are bigger here, and a few are much bigger. For instance, Mak and Saunders say there are only 42 V segments in the human TCRA locus, whereas the table above has 119.

Why? Because the TRACER recombinome lists a different segment for every allele, whereas the textbook may count two different allele as belonging to the same segment. I found this out because the names look like

TRAV1-1*02 , and the [IMGT IG/TR nomenclature page](#) says "Allele names comprise the IMGT gene name followed by an asterisk and a two-figure number." (That IMGT site is amazing. If you ever need an IGK allele table for the common brush-tailed possum, [look no further](#).)

Mak and Saunders claim there are 10 V segments in TCRD locus, whereas the recombinese only lists 6. This is the only category for which Mak and Saunders indicate more segments than TRACER does. I find this especially puzzling. Maybe the best available information has changed since 2008 when this copy of the textbook was printed.