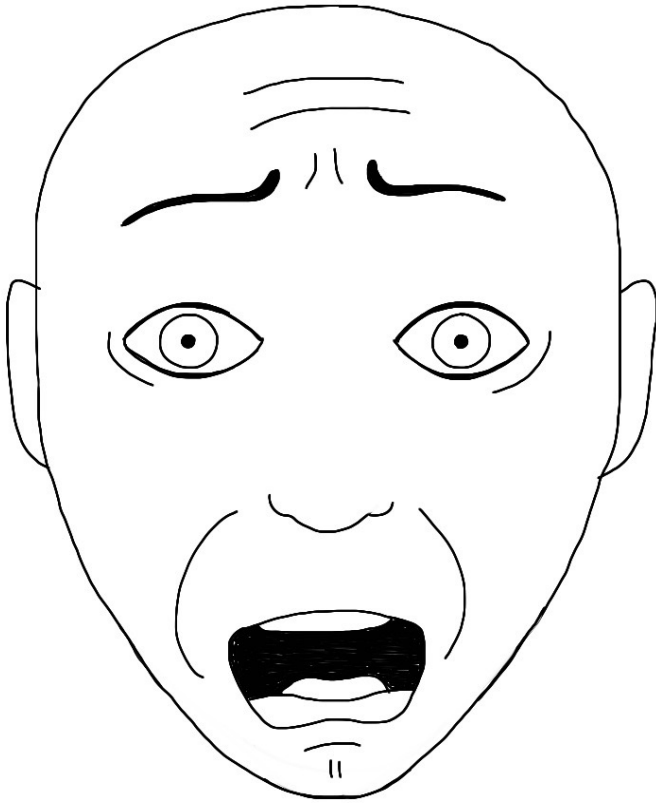


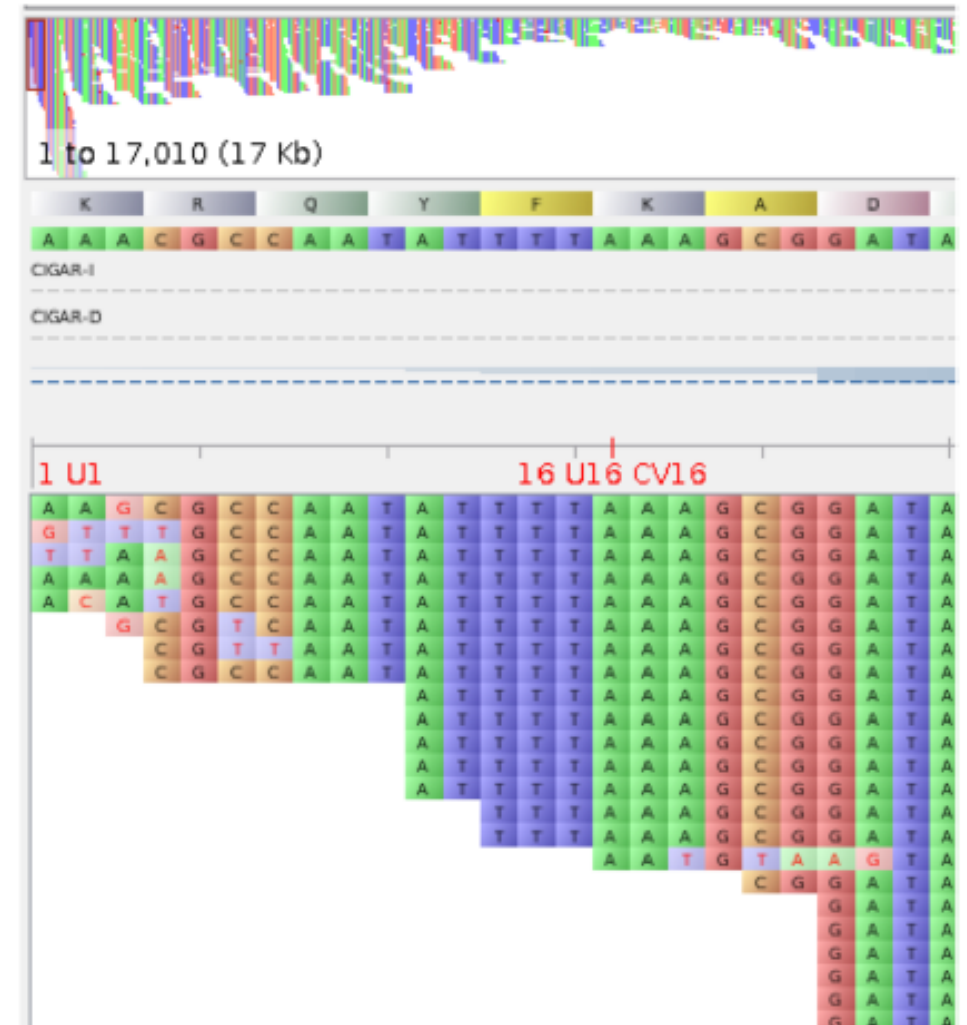
Day 5 recap



Command line: friend or foe?

Using a pipeline to process data

- An existing a pipeline
- Range of data files and formats
- Quality steps, alignment, filtering, variant calling



Alignments

- Important to have an accurate alignment. If it is not and you haven't checked:
 - Wrong tree
 - Wrong inference
 - Wrong result
 - WRONG

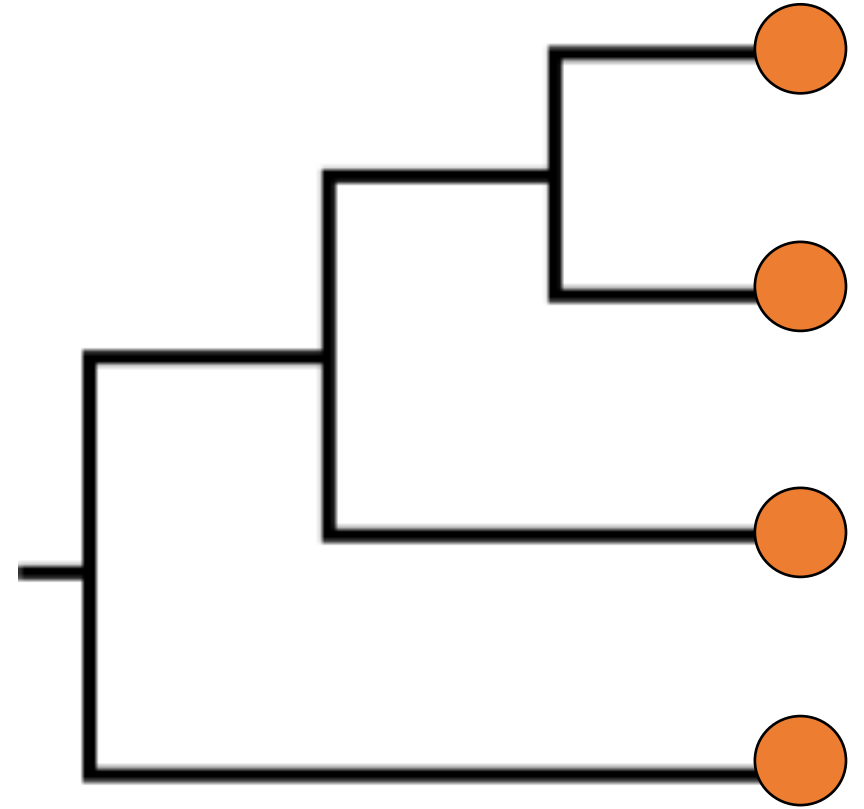


ATGCCTGGAAAAATTACTT
ATGCCTGGAAA--TTACTT
ATGCCTGGAAAAATTGCTT
ATGCCTGGAAAAATTGCTT

The diagram illustrates a sequence alignment between four DNA sequences. A vertical yellow bar highlights the column containing the 11th character of each sequence. In the first sequence, it is 'A'; in the second, it is a gap ('-'); in the third and fourth, it is 'G'. A red box highlights the 10th column, showing a mismatch: 'A' in the first sequence, '-' in the second, and 'A' in the third and fourth sequences.

Phylogenetics

- Key terminology and interpretation



Today's plan

830-1030	Phylogenetics: building and annotating trees
1030-1200	Molecular clocks explainer & practical
1200-1220	Clades and lineages explainer
1220-1230	RABV-GLUE demo
LUNCH	
130-230	MADDOG practical
230-430	Discussion groups
PHOTOS?	