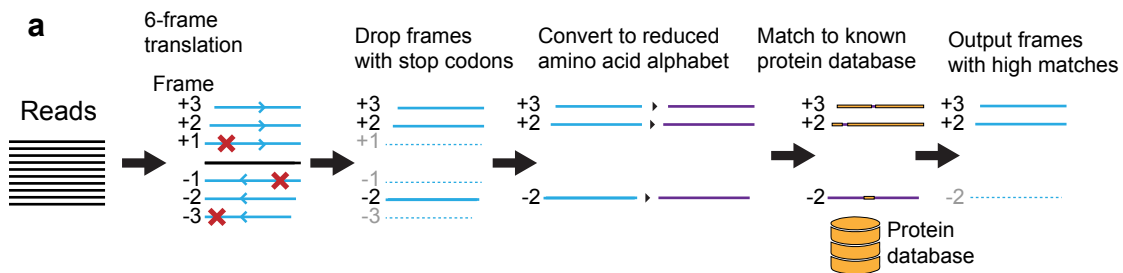


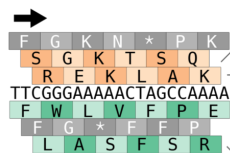
a

For each read, do six-frame translation
 Discard reading frames with stop codons

***k*-merize each "non-stop" reading frame**

Check against database of known human amino acid *k*-mers

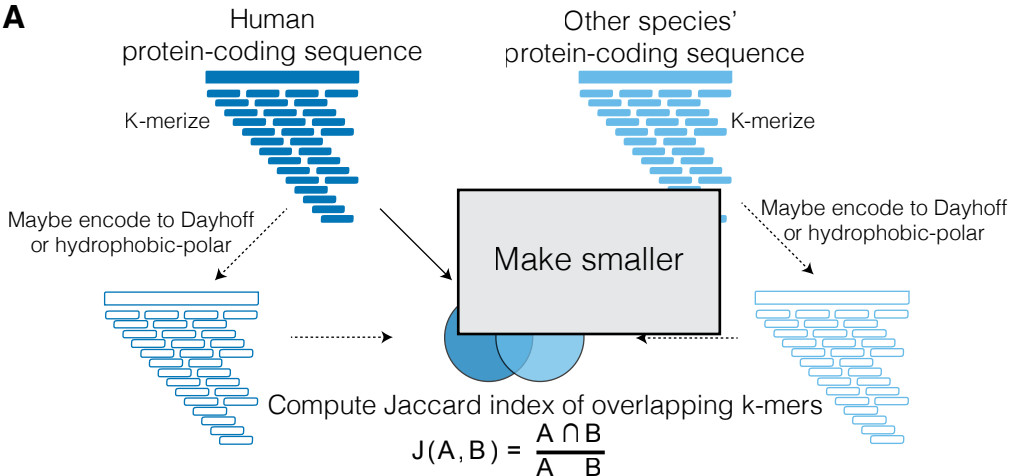
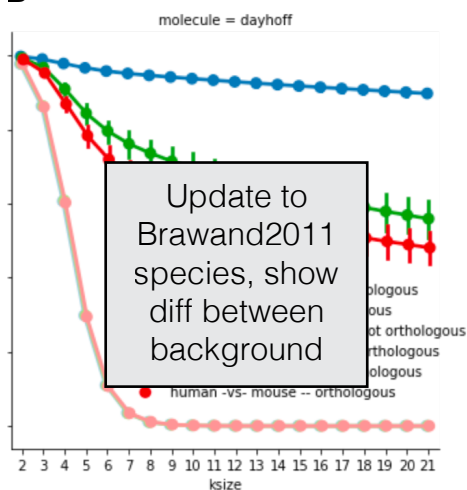
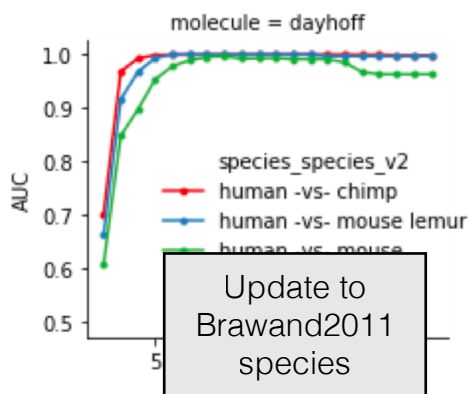
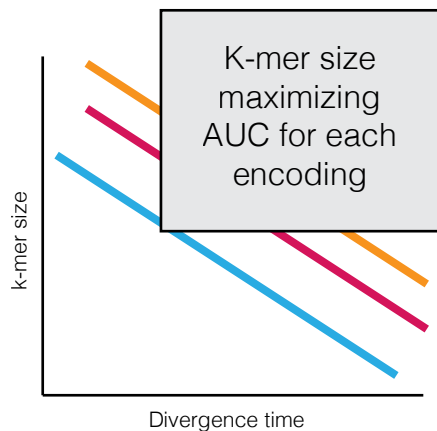
Infer reading frame from presence of known *k*-mers



ENSEMBL known peptide sequences for human stored as a Bloom filter

Mostly hits:
inferred reading frame(s)

Mostly hits:
inferred reading frame(s)

A**B****C****D****E**