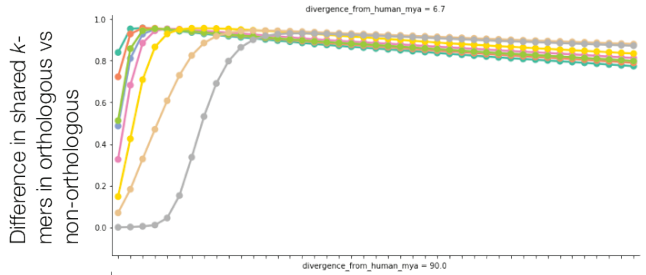


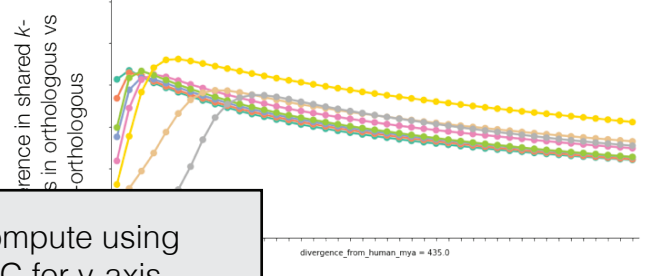
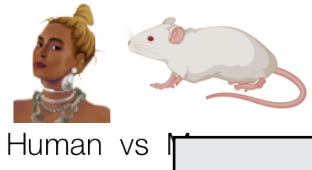
A

Divergence time
(millions of years ago, mya)

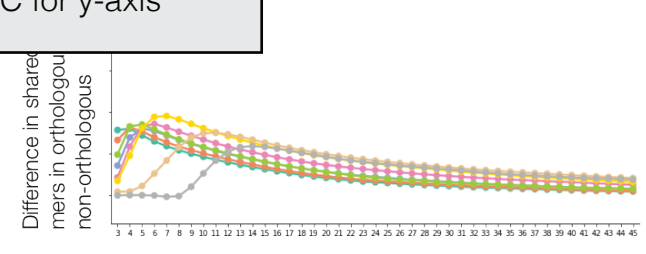
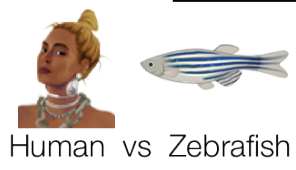
6.7 mya



90 mya



435 mya



Recompute using
AUC for y-axis

- encoding
- protein20
 - hsdm17
 - sdm12
 - aa9
 - botvinnik8
 - dayhoff6
 - gbmr4
 - hp2

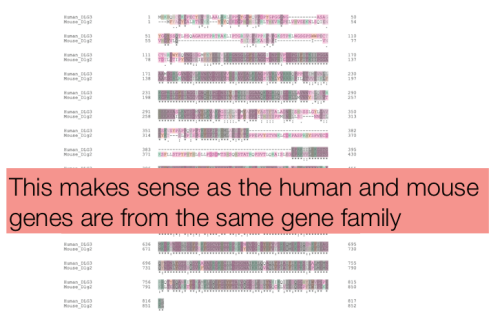
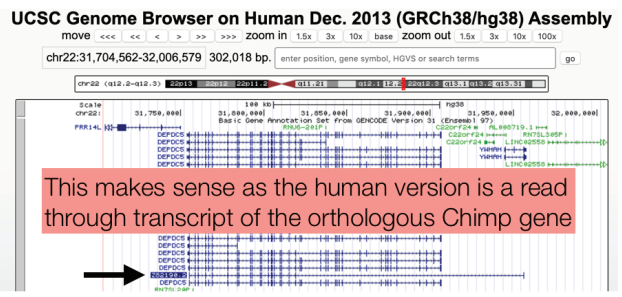
B



Human gene: Z82190.2: novel protein, DEPDC5-YWHAH readthrough
Chimp gene: DEPDC5
High jaccard (0.91) but not labeled as homologous by ENSEMBL COMPARA
Dayhoff encoding, peptide ksize=16



Human gene: DLG3
Mouse gene: Dlg2
Medium jaccard similarity (0.41)
Not homologous by ENSEMBL COMPARA
Dayhoff encoding, peptide ksize=8



C

Apply kmerslay compare-kmer-content to
OrthoDB/BUSCO - bigger datasets