

A

Human protein-coding sequence

Orthologous Chimp protein-coding sequence

Kmerize

Maybe encode to dayhoff
or hydrophobic-polar

Maybe encode to dayhoff
or hydrophobic-polar

Compute overlapping k-mers
Jaccard similarity

B

Orthologous pairs

Background: Shuffled pairs

C

Diagram of all-by-all k-merization comparison