Multiple Alignments and Phylogenies

Mark Voorhies

3/29/2012

Comprehending our BLAST results

- We have a bunch of sequences that look similar to our query

Comprehending our BLAST results

- We have a bunch of sequences that look similar to our query
- We infer that they are homologous to each other

Comprehending our BLAST results

- We have a bunch of sequences that look similar to our query
- We infer that they are homologous to each other
- What does that mean, anyway?

Nomenclature

Homologs heritable elements with a common evolutionary origin.

Nomenclature

Homologs heritable elements with a common evolutionary origin.

Orthologs homologs arising from speciation.

Paralogs homologs arising from duplication and divergence within a single genome.

Nomenclature

Homologs heritable elements with a common evolutionary origin.

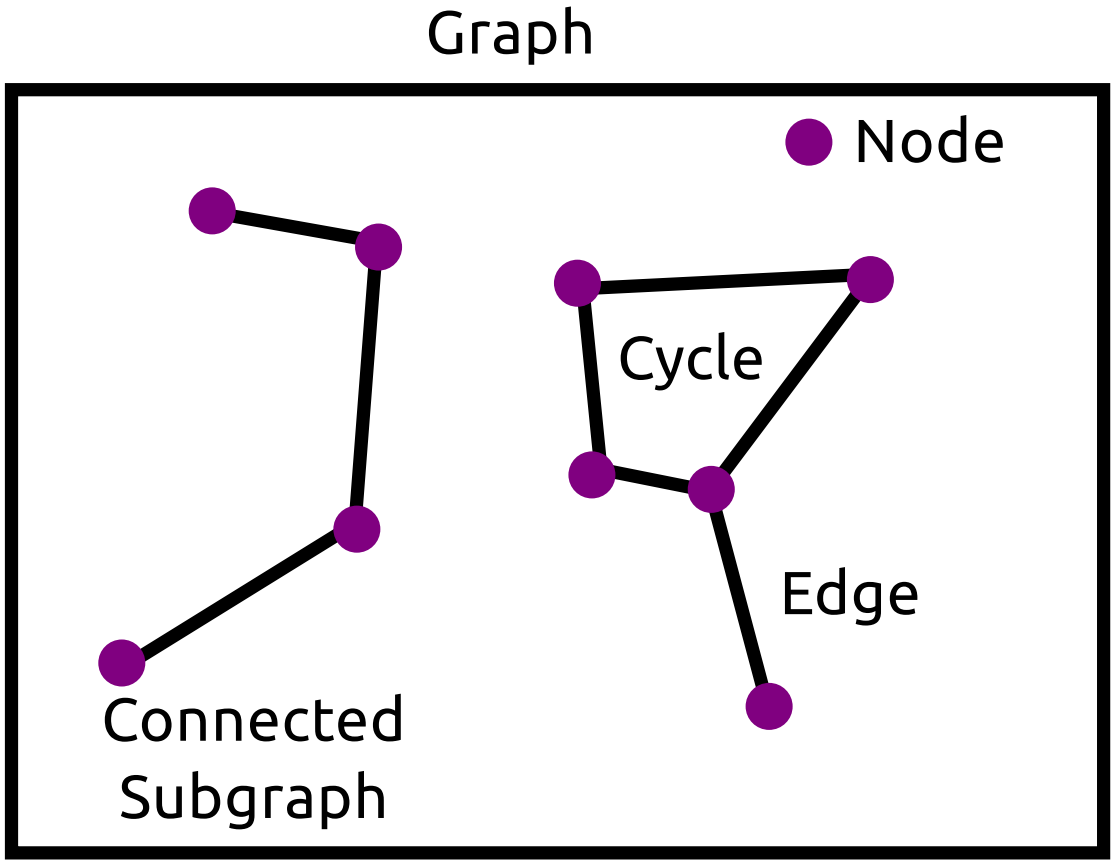
Orthologs homologs arising from speciation.

Paralogs homologs arising from duplication and divergence within a single genome.

Xenologs homologs arising from horizontal transfer.

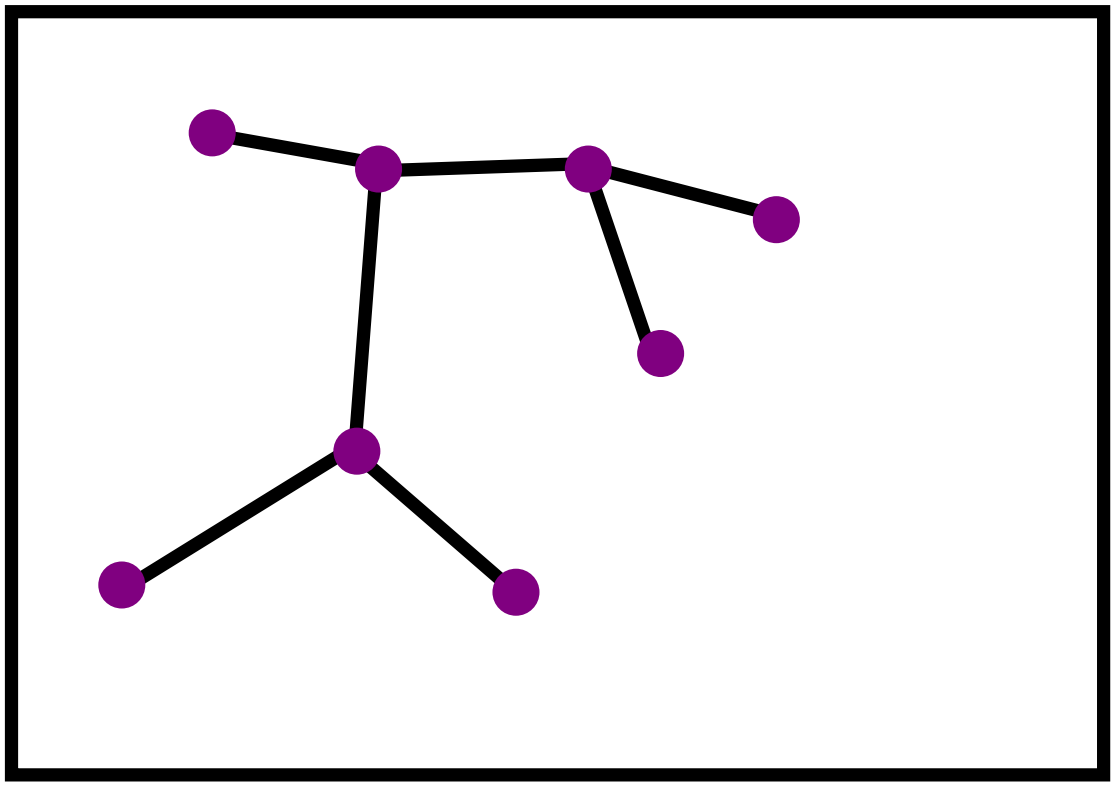
Onologs homologs arising from whole genome duplication.

Nomenclature

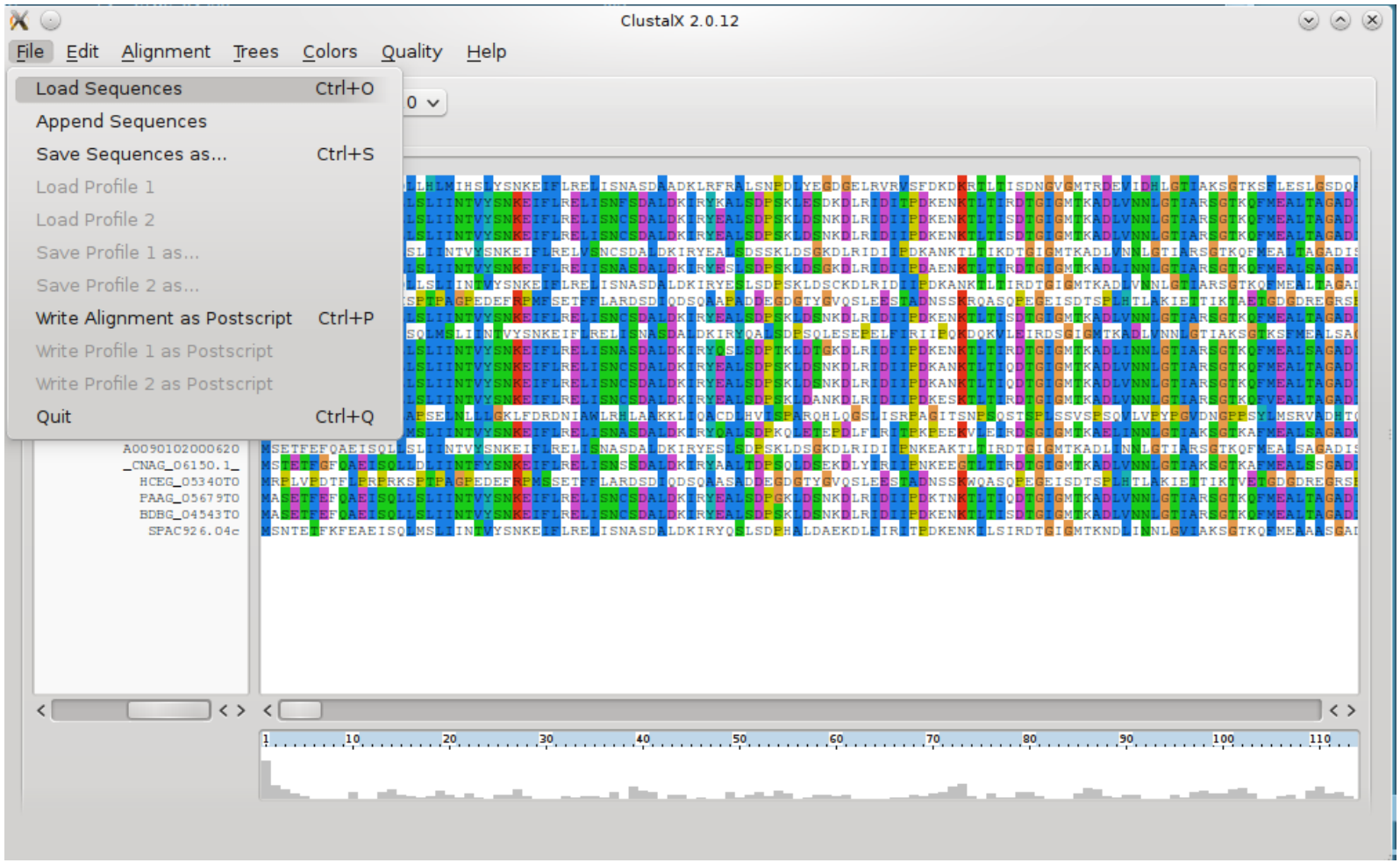


Nomenclature

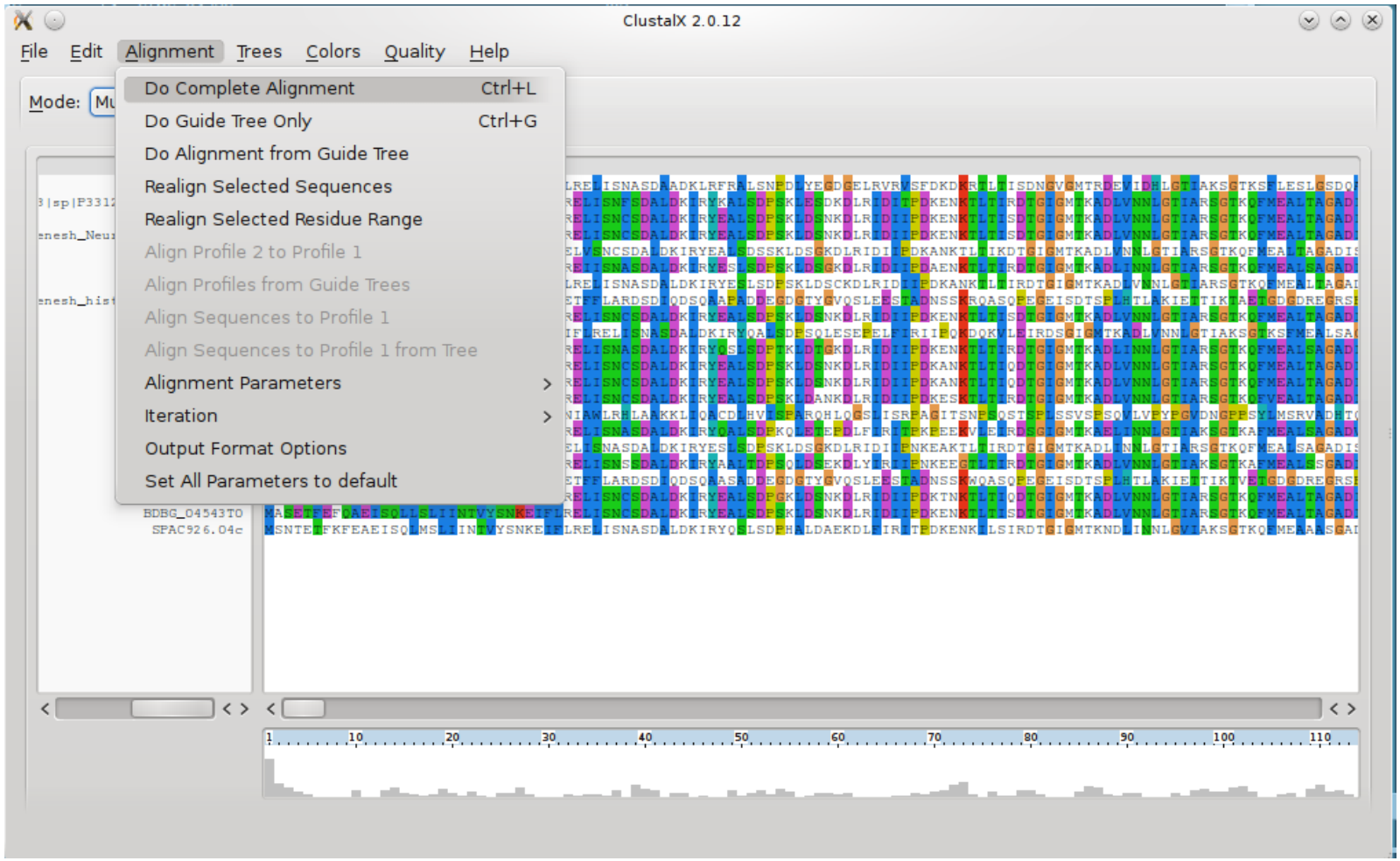
Tree = Connected Graph with no Cycles



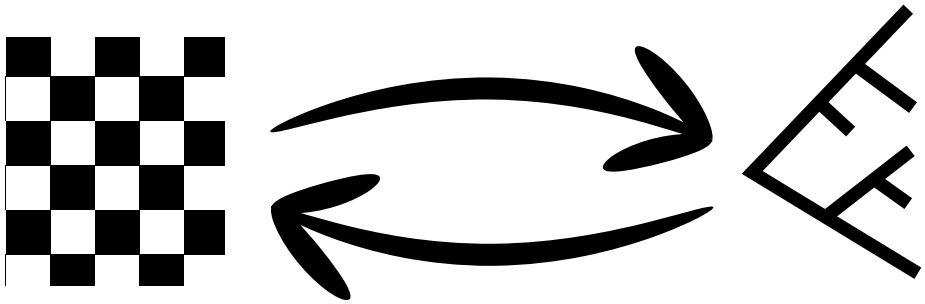
Generating a multiple alignment in CLUSTALX



Generating a multiple alignment in CLUSTALX



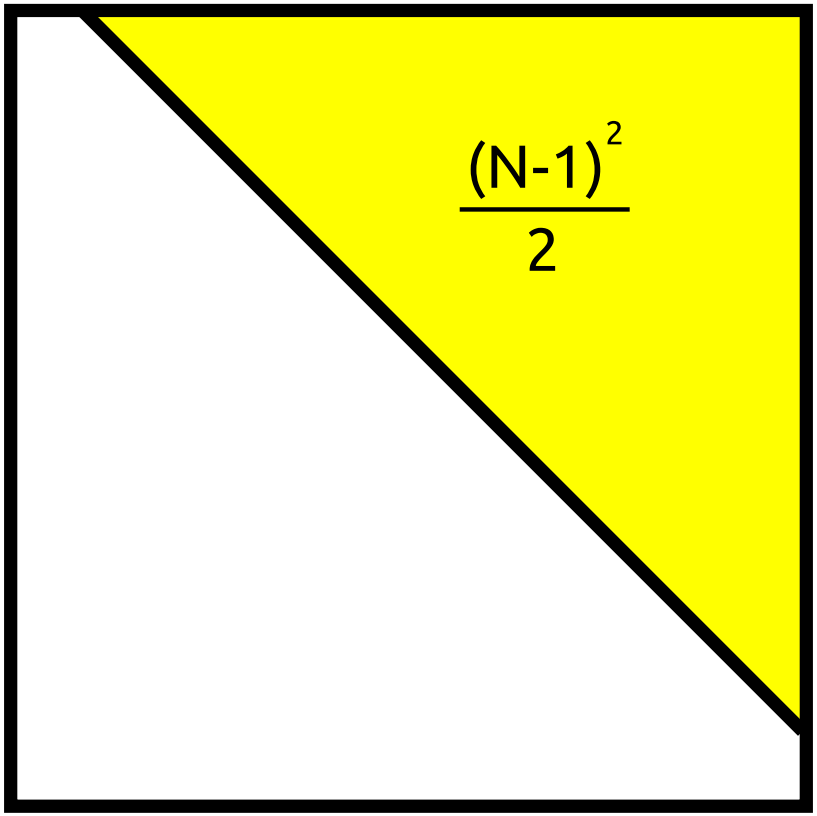
Evolution implies a self-consistent model



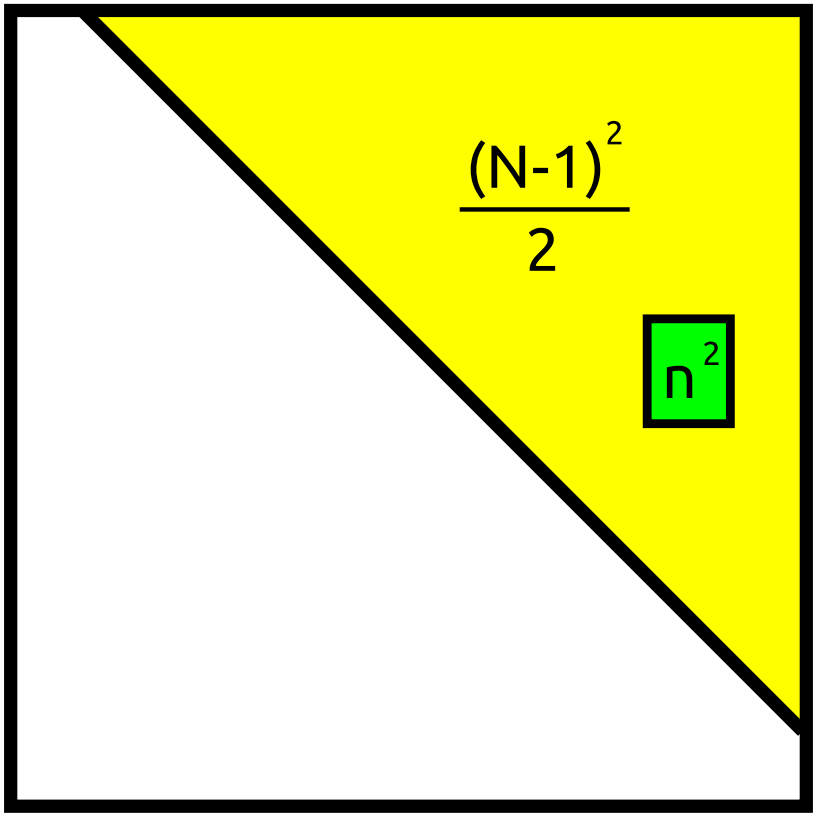
Distances
(Pairwise relationships)

Topology
(Evolutionary history)

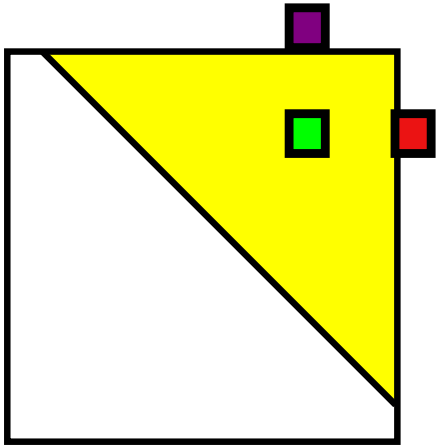
Measure all pairwise distances by dynamic programming

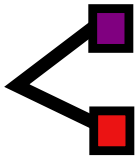


Measure all pairwise distances by dynamic programming

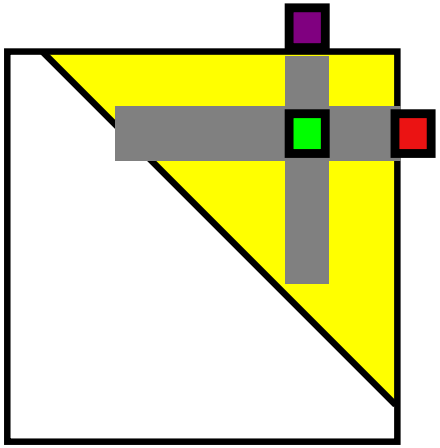


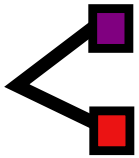
Generate a guide tree by UPGMA



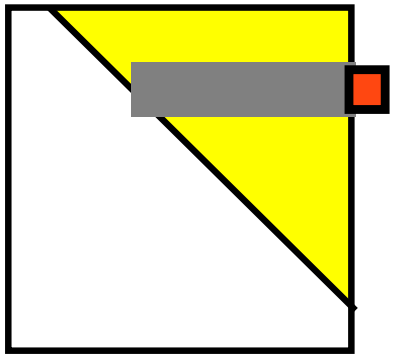


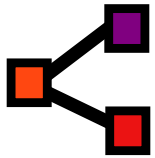
Generate a guide tree by UPGMA



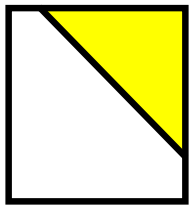


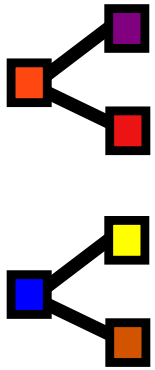
Generate a guide tree by UPGMA



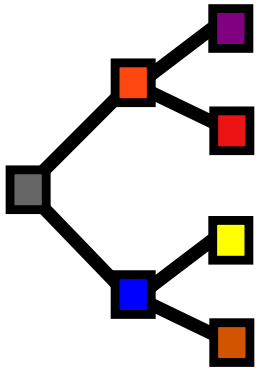


Generate a guide tree by UPGMA

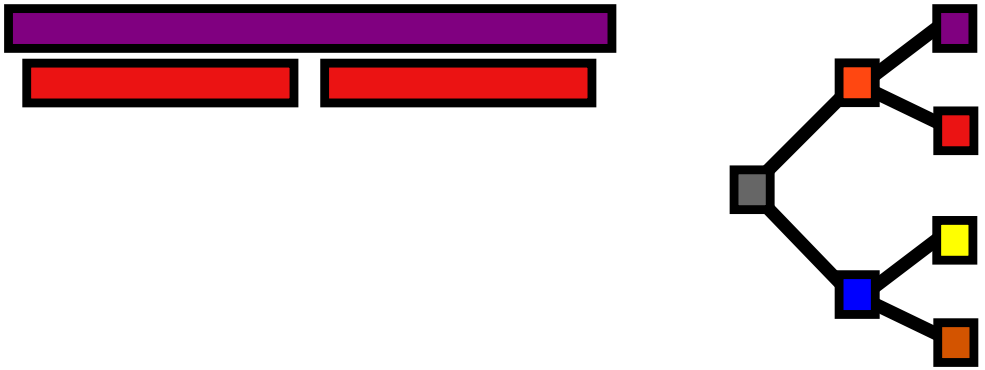




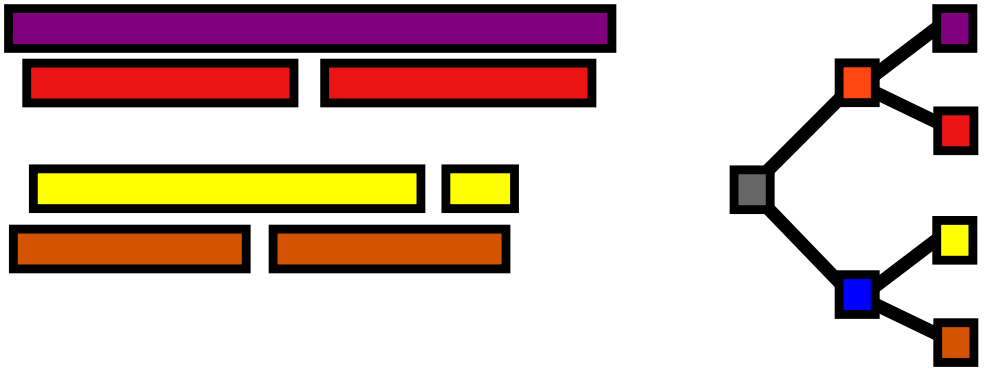
Generate a guide tree by UPGMA



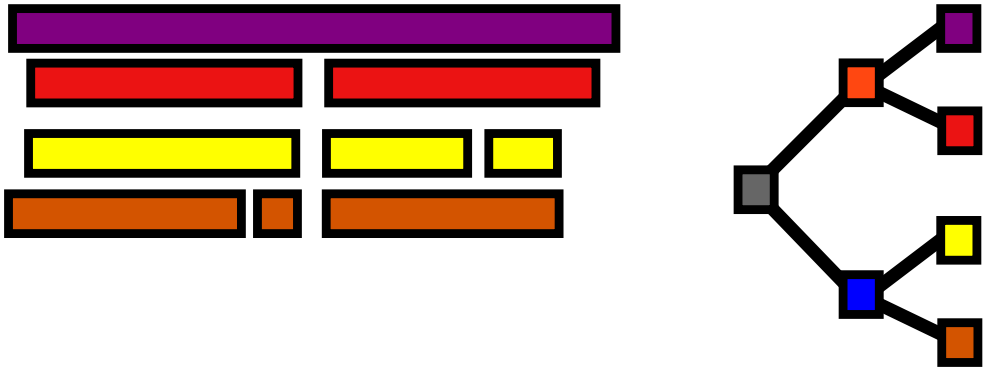
Progressive alignment following the guide tree



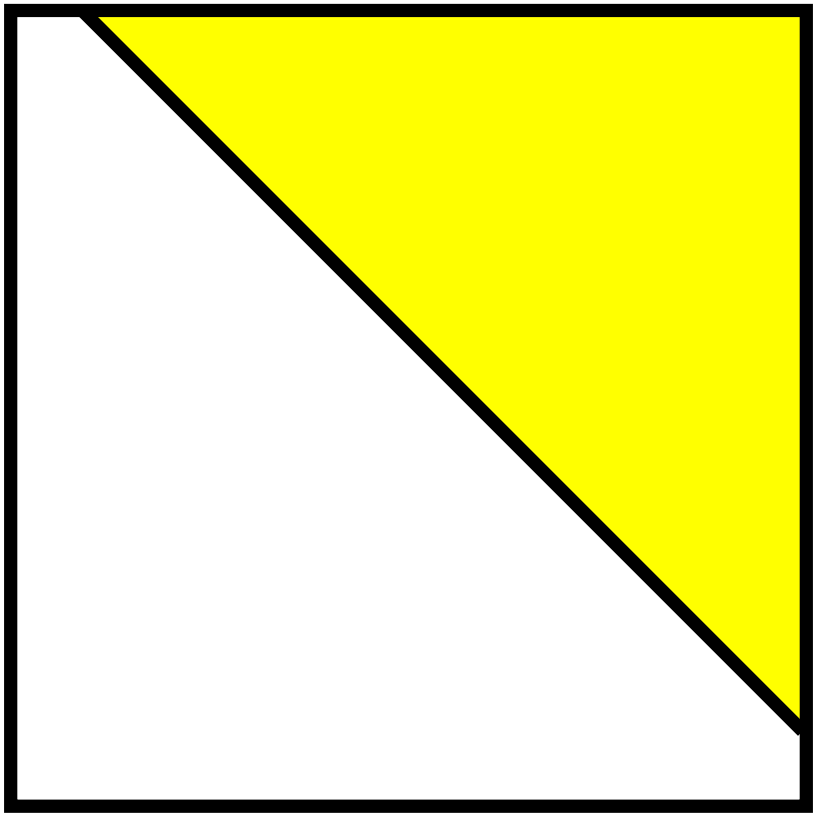
Progressive alignment following the guide tree



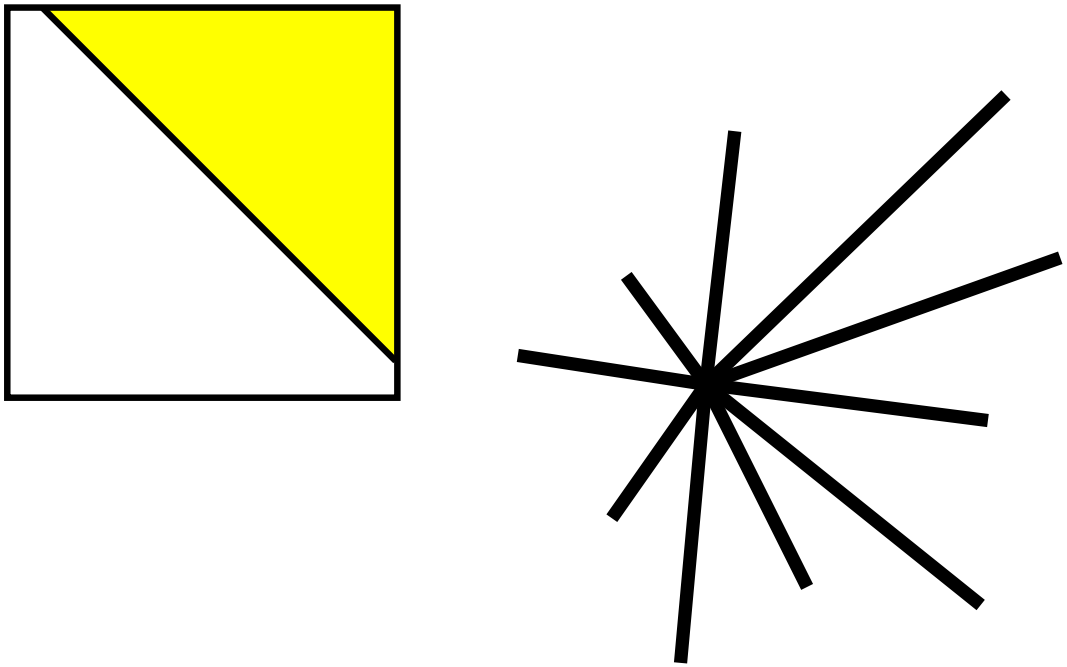
Progressive alignment following the guide tree



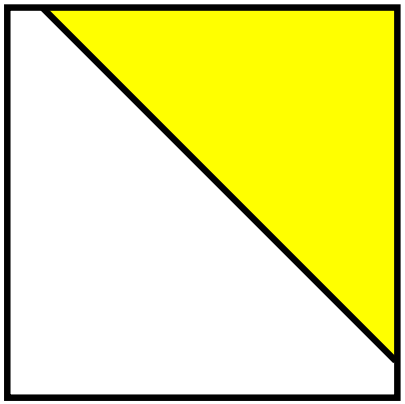
Measure distances directly from the alignment

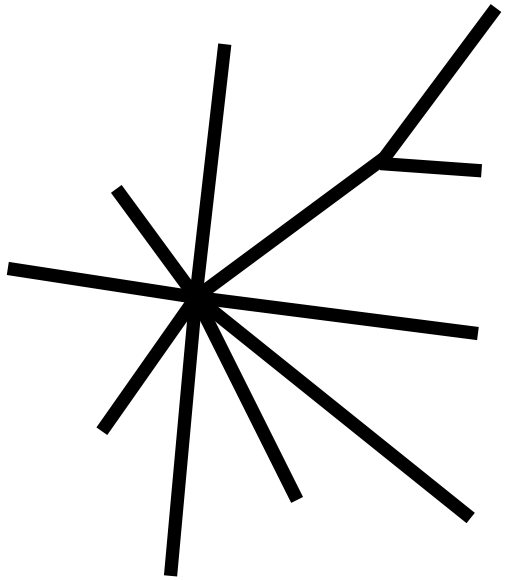


Generate neighbor-joining tree from new distances

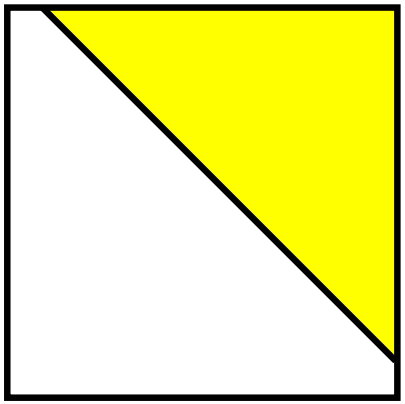


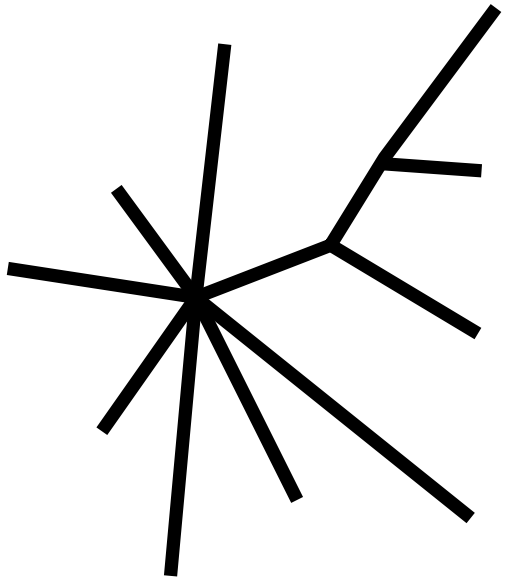
Generate neighbor-joining tree from new distances



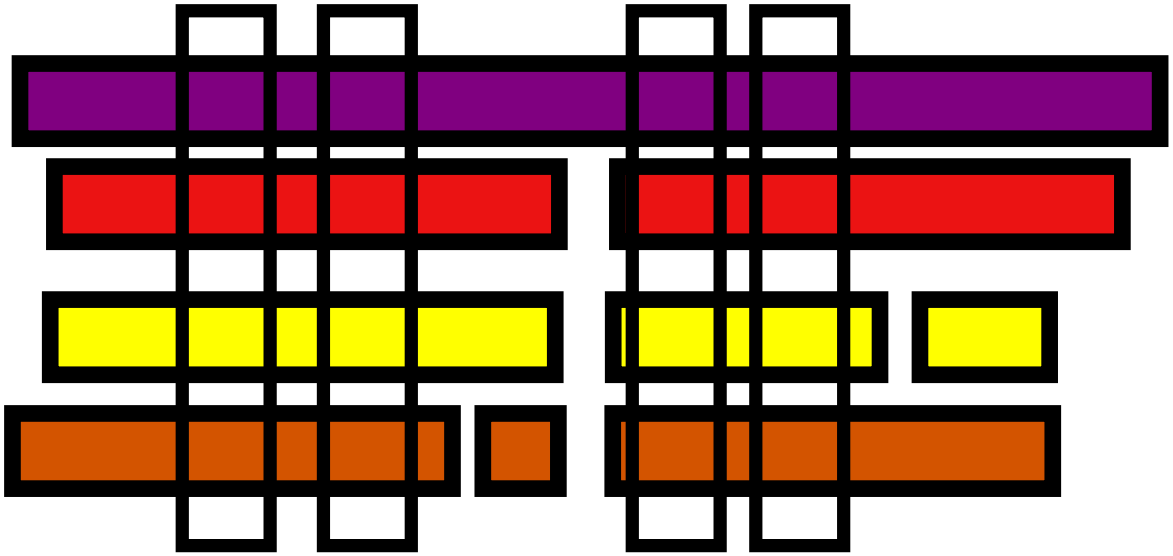


Generate neighbor-joining tree from new distances

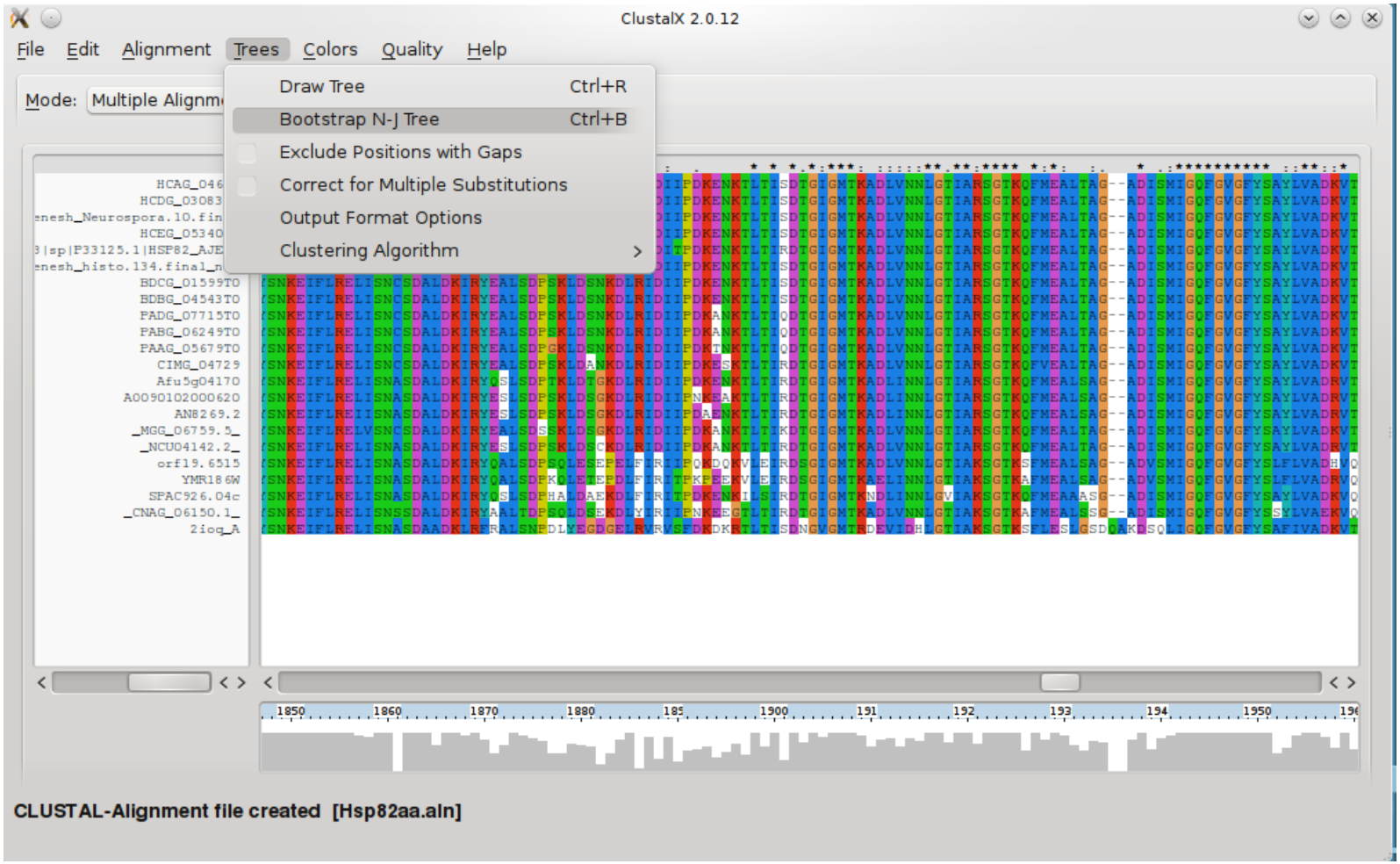




Generate bootstrap values from subsets of the alignment



Generating a neighbor joining tree in CLUSTALX

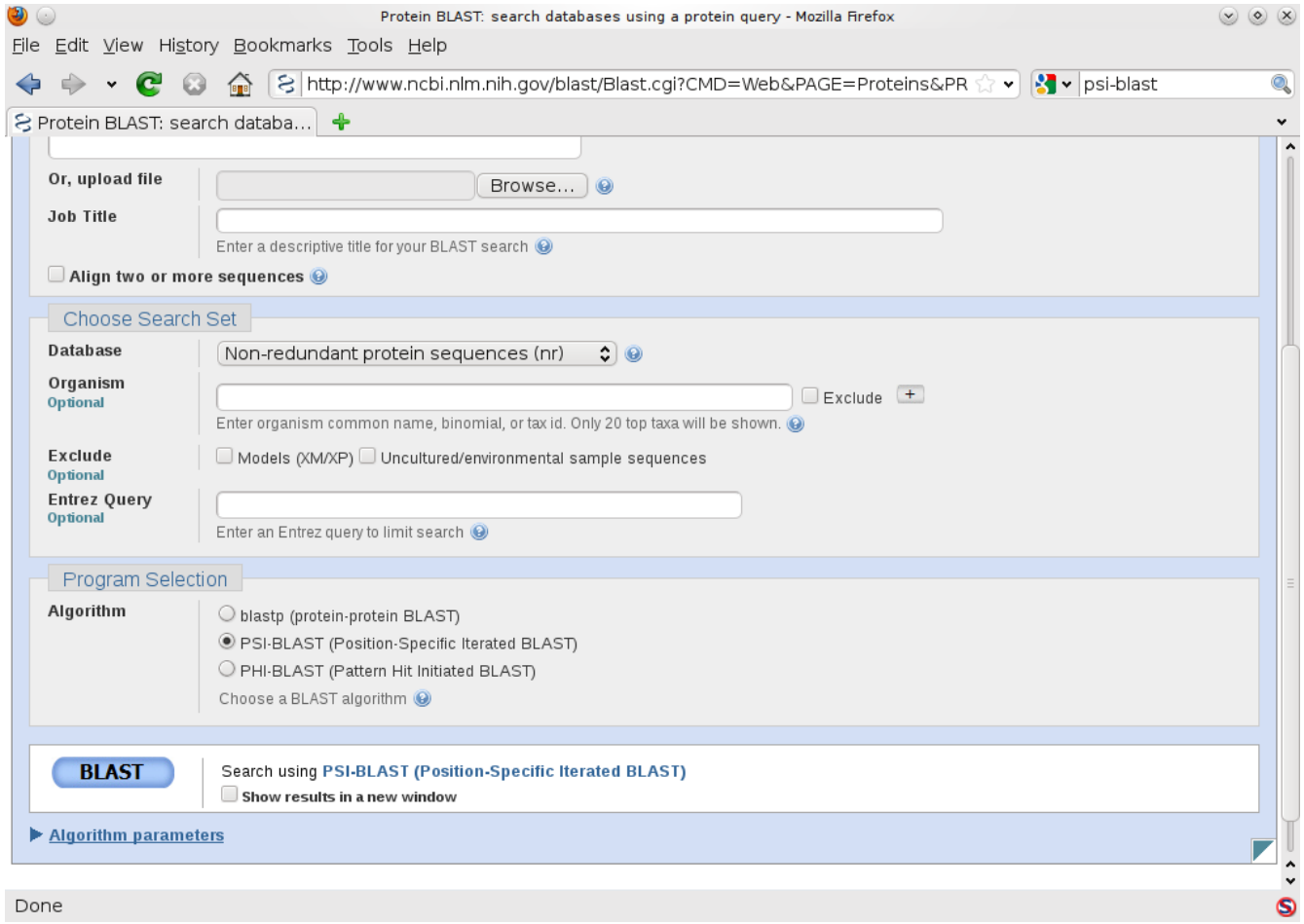


Viewing the alignment and tree in JALVIEW

Related tools

- Multiple Alignment
  - T-Coffee
  - MUSCLE
  - COBALT
- Tree building
  - MrBayes (Bayesian MCMC)
  - PhyML (maximum likelihood)
- Work benches
  - MESQUITE
  - UGENE

Searching with PSI-BLAST



Homework

- Play with CLUSTALX, JALVIEW, and PSI-BLAST
- Read PLoS Comp. Biol. 4:e1000069