BIOL5705: Gene Sequence Analysis

Title

Computational methods for detecting gene homology and phylogenetic tree construction.

Lecture Layout

There are 4 lectures and each lecture is 2 hours in length. I will use 1 – 1.5 hours for lecture and you can work on the assignment and ask questions during the remaining time.

Lecture Location & Time

Wednesdays 1:30-3:30 in 4202 Mona Campbell Building

Jan. 16th – Feb. 6th

Grading/Evaluation

Grading will be based on 4 equally weighted assignments.

Assignments will be due at the start of the next lecture (or one week later for the last assignment). Students are encouraged to ask each other and me for help, but assignments should be done individually.

Assignments are to be emailed to morgan.g.i.langille@gmail.com

Course Materials

All lecture materials including presentations and assignments are available online at: http://morganlangille.com/biol5705.html

Outline

1st lecture (Jan. 16st)
- • What is homology?
- • Orthologs, paralogs, etc.
- • Local vs global alignment
- • E values, bit scores, "coverage", identity vs similarity
- • Different BLAST flavours (blastn, blastp, tblastn, etc.)
- • Databases (nr, refseq, etc.)
- • Blast (Web)

2nd Lecture (Jan 23rd)

- HMMs (HMMer & PFAM website)
- PSSMs (PSI-BLAST)
- RNA secondary structures (stochastic free grammars)
- Genome Alignment (Mauve)
- Mappers (Bowtie)
- Assemblers
- Running BLAST locally

3rd Lecture (Jan. 30th)

- Multiple Alignment (Clustalw, Muscle, T-COFFEE, HMMAlign)
- Editing Alignments (Gblocks)
- Viewing Alignments (Jalview)

4th Lecture (Feb. 6th)

- Why we build trees?
  Tree file formats
  Parsimony
  Neighbour Joining (distance based)
  ML (PhyML, RAxML, FAST-Tree)
  Bayes (Mr. Bayes, Beast)
  http://www.phylogeny.fr/