

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

Every Wed. 10-12pm from Nov. 5th to Nov. 27th in Tupper Building room B-A1 (basement)

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 myself 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

1 lecture

- 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)

2 Lecture

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

- Assemblers
- Running BLAST locally

3 Lecture

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

4 Lecture

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