

Programming for Bioinformatics
BIOL 8803 B
November 16th, 2015

Concepts and such for this week:

Class

Object

Inheritance

Subclass

Passing arguments as a hash

`package` - tell the interpreter what package you are in

`bless` – make a hash reference a full-fledged object

Exercises:

BioPerl is a huge set of Perl classes which allow you to, among many other things, manipulate sequences.

0.) Install BioPerl on your system using CPAN. It's straight-forward (and google-able)

Script #1

- 1.) Use `Bio::SeqIO` to load a genbank file
- 2.) Convert it to a FASTA file
- 3.) Convert it into an EMBL file

Script #2

- 4.) Use the `Bio::Tools::Run::StandAloneBlast` module to BLAST multiple sequences (testSpeciesA.fasta) against a sequence database(testSpeciesB.fasta). The documentations for `StandAloneBlast` can be found at <http://www.bioperl.org/wiki/HOWTO:StandAloneBlast>
- 5.) Use a parser to find the top two best hits for each sequence from the returned BLAST parsing object. The parser will be one of the following: `Bio::Tools::BPbl2seq`, `Bio::Tools::BPpslite`, `Bio::SearchIO` or `Bio::Tools::BPlite` object.

Deliverables

- Code
 - convertgb.pl
 - blastTop2.pl

Additional instructions

For implementation, your code should take options below:

```
convertgb.pl -i [input_file] -f [fasta|embl] -o [output_file]
```

```
blastBest2.pl -i [input_file] -d [sequence_db.fa] -m  
[blast_method] -o [output_file]
```