Bio::Phylo

Rutger Vos

University of British Columbia
What it is

• Bio::Phylo is a set of perl5 libraries for handling phylogenetic data:
  – Models nodes, trees, sets of trees; character sequences, matrices, sets of matrices; taxa, sets of taxa.
  – Includes BioPerl-like extensible IO architecture.
  – On http://www.cpan.org
  - (Includes ad-hoc methods for biodiversity calculations and treedrawing.)
Why?

- Historically: small, easy (-er) install
- Test bed for:
  - “best practices”
  - performance optimizations (caching, data models)
  - integration with CIPRES et al.
- “Narrower but deeper” than BioPerl
How it works

• Bio::Phylo is written in OO perl5, follows “perl best practices”:
  – exception hierarchies,
  – inside-out objects,
  – separate setters and getters
  – UIDs for all objects
  – (additional syntax sugar for array-like operations)
How it should work

• Bio::Phylo functionality should be exposed to other libraries, hence:
  – Should adopt BioPerl interfaces (already done where possible, but BioPerl doesn’t do character matrices that well)
  – Should be able to exchange data “through the wire” (webservices, CORBA, CIPRES)
  – Should have more explicit “bookkeeping” of sets of trees + matrices + taxa (CDAT)
Relation to other projects

• Bio::Phylo forms the “tree-aware” part of the CIPRES/CORBA stack (see Mark’s talk)

• Bio::Phylo will become part of the Bio::CDAT architecture (see Aaron’s talk)

• Bio::Phylo should become part (somehow) of BioPerl (see Bio* bootcamp)