



BioRuby & BioJava

Possibilities of use of BioRuby and BioJava
bioinformatics libraries in Ruby and JRuby
scripts.

Andrzej Górecki
Radosław Urbaś

BioRuby & BioJava

Bioinformatics:

- Joins **life sciences**, **information** sciences and **technologies**;
- Applies principles of information sciences and technologies to make the vast, diverse, and **complex** life sciences **data** more understandable and useful;
- Major research efforts in the field include: sequence alignment, gene finding, genome assembly, protein structure alignment, protein structure prediction, prediction of gene expressions, protein-protein interactions, modeling of evolution.



Most recently used languages

- Out of box solutions: *BioLanguage* libraries. **Python**, **Perl**. Being productive means ability for easy parsing heterogeneous biological formats, manipulating biological data with 2D and 3D rendering, extracting automatically biological knowledge using different algorithms.
- **Mathematica** and **Matlab**. Usage of languages from this group is a consequence of reasoning that "Mathematics Is Biology's Next Microscope". Usage of appropriate mathematics can help interpret any kind of data.
- **Ruby** and **JRuby**. Benefits of the first group and code simplicity. Will be discussed later.



Developers' requirements

- Developers need to **focus on** investigating the root of the **problem** – code simplicity, fast code production;
- Many **tasks** and algorithms **are repetitive** – reusability of once developed code;
- **Data** are stored in **various formats** – easy access;

Solution: Bio* libraries

* - insert language name here



BioRuby

- Implementation of an integrated environment for Bioinformatics with Ruby language.
- **Advantages:**
 - code simplicity, acceptable efficiency;
 - a lot of classes for Bioinformatics: database parsers, sequences, algorithms etc.;
 - shell command line, web user interface.
- **Disadvantages:**
 - weak documentation;
 - last version from February 2006.

“Perl : #*@%-ed; Python : Simple, Fun and Indented!; Ruby : Ruby is Perl's younger, prettier sister.”

* footnote from <http://dev.bioruby.org/wiki/en/>



BioRuby – code example

```
#!/usr/bin/env ruby
require 'bio'

seq = Bio::Sequence::NA.new("atgcatgcaaaa")

# original sequence
puts seq
# complementary sequence (Bio::Sequence::NA object)
puts seq.complement
# gets subsequence of positions 3 to 8
puts seq.subseq(3,8)
# shows amino acid names (Array)
puts seq.translate.names
```



BioJava

- Open-source project dedicated to provide Java framework for processing biological data. While using BSF and JRuby it can be used together with Ruby code.
- **Advantages:**
 - classes for operating on databases, a lot of algorithms, xml support, graphical display of biological sequences;
 - no problems with documentation;
 - clear code organized in modular way.
- **Disadvantages:**
 - efficiency;



BioJava – code example

```
// create a DNA SymbolList  
SymbolList symL = DNATools.createDNA("atggccattgaatga");  
  
// transcribe to RNA  
symL = RNATools.transcribe(symL);  
  
// prove that it worked  
System.out.println(symL.seqString());  
  
// translate to protein  
symL = RNATools.translate(symL);  
  
// prove that it worked  
System.out.println(symL.seqString());
```



Conslusions

- BioRuby and BioJava libraries have **similar possibilities**;
- BSF and **JRuby** engine allow to combine **Java and Ruby**;
- **BioJava seems to meet requirements** of bioinformatics.



Questions

Thanks

Any questions?

