BioRuby

Yoshinori K. Okuji
Toshiaki Katayama
Mitsuteru S. Nakao

Bioinformatics Center, Institute for Chemical Research, Kyoto University, Japan

July 19, 2001
Background

1. *Too many* DB formats exist (GenBank, SWISS-PROT, KEGG, etc.).

2. The interfaces for tools are often *unfriendly* for batch processing (FASTA, BLAST, CLUSTALW, etc.).

3. Object-oriented design is very useful for an *unified interface*, regardless of underlying data objects!
Overview of Ruby

• Interpreted scripting language for quick and easy object-oriented programming.

• Complete, full, pure object-oriented language: *All data are objects.*

• Simple, straight-forward, extensible, and portable.

• Very popular, especially in Japan. In part, because the author “Matz” is Japanese.

• Totally free.
Hello World

Ruby
puts 'Hello'

C
main(){puts("Hello");}

Perl
print "Hello\n";

Python
print "Hello"

Tcl
puts "Hello"

BASH
echo Hello
Serious Example

# Perl
package Foo;
sub new {
    my $self = shift;
    bless { 'foo' => 0 }, $self;
}
sub add {
    my $self = shift;
    my $arg = shift;
}
sub output {
    my $self = shift;
    print $self->{'foo'}, "\n";
}

package main;
$foo = new Foo;
$foo->output;
$foo->add(10);
$foo->output;

# Ruby
class Foo
    def initialize
        @foo = 0
    end
    def add(arg)
        @foo += arg
    end
    def output
        puts @foo
    end
end

foo = Foo.new
foo.output
foo.add 10
foo.output
What is BioRuby

- The aim is to provide "open source" resources for bioinformatics.
- Based on the pure object-oriented language Ruby.
- Creating applications, as well as libraries.
- Licensed under the term of LGPL.
Why BioRuby

- Ruby has **good readability**, so your code won’t be write-only code.

- Ruby is **easily extensible** by Ruby itself (i.e. over-riding methods) and by dynamically loadable C modules.

- Ruby’s object-oriented design support is straightforward and complete.

- We love Ruby! Programming with Ruby is *quite delightful*!
Current Status

**DB parsers**

GenBank, GenBank location, DDBJ, KEGG/GENES, KEGG/GENOME, PROSITE, TRANSFAC, LITDB, and Gene Ontology classes.

**Data modules**

Amino Acids, Nucleic Acids, Codon Table, and KEGG organisms modules.

**Tool support**

FASTA and SSEARCH.

**Misc**

DBGET module, Sequence class, modules for Hierarchical Clustering and Smith-Waterman algorithm alignment, Matrix class for Bioinformatics, etc.
BioRuby Genome Browser is an example of BioRuby applications, which supports GIF, PNG and SVG formats.
Staff

“Yoshinori K. Okuji” o@bioruby.org

“Toshiaki Katayama” k@bioruby.org

“Mitsuteru S. Nakao” n@bioruby.org

Feel free to join us!
http://bioruby.org/
staff@bioruby.org