

```
#!/usr/bin/env ruby
```

```
require 'bio'
```

BioRuby

片山 俊明, 五斗 進

京大化研バイオインフォマティクスセンター



informatics
BioRuby.org

BioRuby プロジェクトとは

- 国産の優れたオブジェクト指向スクリプト言語 Ruby を使って、Bioinformatics の面倒なルーチンワークを再利用できる形でライブラリ化する
 - 生物学者がいつもの仕事をサラッと片付けられる
 - データベースからエントリを取ってきてパースする
 - Blast をガンガン流して結果を回収する
- 海外の先行プロジェクト BioPerl, BioJava, BioPython 等と比較しての利点
 - Ruby という言語が高機能で使いやすい、短く書ける
 - 開発コミュニティが日本語を話す

Open Bio*

- O|B|F -- Open Bio Foundation
 - バイオインフォマティクスに関するオープンソースコミュニティ
- **BioRuby**
- BioPerl
- BioPython
- BioJava
- BioDAS
- BioMOBY
- EMBOSS
- Ensembl
- OmniGene
- GMOD
- Apollo
- OBDA
- BioCaml
- BioLisp
- BioConductor
- BioPathways
- BioBlog
- BioCyc
- BioDog
-

OBDA

- BioHackathon

- 2002/01 Arizona, 2002/02 Cape Town
- どの Open Bio* ライブラリで作成したデータベースも共通にアクセスできる
- 競合ではなく協調 Open Bio* ハッカーは仲間です:-)

OBDA

- BioHackathon

- 2002/01 Arizona, 2002/02 Cape Town
- どの Open Bio* ライブラリで作成したデータベースも共通にアクセスできる



OBDA

- BioHackathon
 - 2002/01 Arizona, 2002/02 Cape Town
 - どの Open Bio* ライブラリで作成したデータベースも共通にアクセスできる
 - 競合ではなく協調 Open Bio* ハッカーは仲間です:-)
- Open Bio* Sequence Database Access
 - Directory Registry (Stanza)
 - Flat File indexing (DBM, BDB)
 - BioFetch (CGI/HTTP)
 - BioSQL (MySQL, PostgreSQL, Oracle)
 - SOAP (XEMBL based)
 - BioCORBA (BSANE compliant)

OBDA を使ったエントリ取得

- 設定ファイル (Stanza フォーマット)
 - ~/.bioinformatics/seqdatabase.ini
 - /etc/bioinformatics/seqdatabase.ini
 - <http://open-bio.org/registry/seqdatabase.ini>

```
[swissprot]
protocol=biosql
location=db.bioruby.org
dbname=biosql
driver=mysql
biodbname=sp
```

```
[embl]
protocol=biofetch
location=http://bioruby.org/cgi-bin/biofetch.rb
biodbname=embl
```

```
:
```

OBDA を使ったエントリ取得

- 設定ファイル (Stanza フォーマット)
 - ~/.bioinformatics/seqdatabase.ini
 - /etc/bioinformatics/seqdatabase.ini
 - <http://open-bio.org/registry/seqdatabase.ini>

```
[swissprot]
protocol=biosql
location=db.bioruby.org
dbname=biosql
driver=mysql
biodbname=sp
```

```
[emb1]
protocol=biofetch
location=http://bioruby.org/cgi-bin/biofetch.rb
biodbname=emb1
:
```

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

reg = Bio::Registry.new
db = reg.db("swissprot")
entry = db.fetch("TETW_BUTFI")
```


BioRuby にできること

- Bio::Sequence, Bio::Location, Bio::Feature クラス
 - 塩基配列、アミノ酸配列の取り扱い
 - 組成、スプライシング、翻訳、ウィンドウサーチなど
- Bio::DB クラス
 - データベースのパパーザ (現在 20 種くらい)
- Bio::Blast, Bio::Fasta クラス
 - Blast/Fasta を実行し結果をパースするファクトリ
- Bio::PubMed, Bio::Reference クラス
 - 文献データベース検索、BibTeX などのフォーマッタ
- Bio::Registry, Bio::SQL, Bio::Fetch, Bio::FlatFile クラス
 - データベース入出カインターフェイス (OBDA 準拠)
- Bio::Pathway, Bio::Relation クラス
 - グラフ、2 項関係の演算

いまからの課題

- ドキュメント
- SOAP(DAS, XEMBL, 遺伝研), CORBA アクセスの整備
- HMMER, EMBOSS, ClustalW, T-Coffee など対応
- PDB - 構造系のデータベースクラス、メソッドの開発
- PATHWAY, SSDB, KO, GO, InterPro
- BioFetch で Entrez E-utils 対応
- GFF, AGAVE, GAME フォーマット
- 配列データベース系クラスのリファクタリング
- アライメントクラス

BioRuby.org

- 総合情報 <http://bioruby.org/>
 - 開発情報 <http://ura.bioruby.org/>
 - ニュース <http://q--p.bioruby.org/>

 - ソース <ftp://bioruby.org/>
 - CVS <cvs.bioruby.org>

 - M L ja@bioruby.org, dev@bioruby.org
 - 問い合わせ staff@bioruby.org
- presentation by T. Katayama <k@bioruby.org>