

MobyNet: service integration over distributed web portals

<http://MobyNet.rpbs.univ-paris-diderot.fr>

Nantes, 21 Octobre 2010

History

- 2004: Institut Pasteur + RPBS -> “Moby” project
– Participatory conception
- Early 2008: Moby@pasteur open
- End 2008: Moby@RPBS open
- 2009: Moby@LIPM, Moby@genouest open
- 2009: MobyNet Project: network of Moby portals
- 2009: IbiSA label
- 2010: Moby@CBS, Moby@Migale, Moby@IPMC, Moby@IGR, ... to appear

1. MobyNet: the project

2. *Moby*

3. *MobyNet specific aspects*

4. *MobyNet today*



Why MobyleNet ?

Nucleic Acids Research Special Issue on Synthetic Biology

Table of Contents

Volume 38 suppl 2 July 2010

[Clear](#) [Get All Checked Abstracts](#)

Editorial: 2010 Web Server Issue

Gary Benson

Editorial

Nucl. Acids Res. (2010) 38(suppl 2): W1-W2 doi:10.1093/nar/gkq262

» [Extract](#) » [Full Text \(HTML\)](#) » [Full Text \(PDF\)](#) » [Screen PDF](#)

OPEN ACCESS [\(CO\)](#) [BY-NC](#)

Articles

Michelle D. Brazas, Joseph T. Yamada, and B. F. Francis Ouellette

Providing web servers and training in Bioinformatics: 2010 update on the Bioinformatics Links Directory

Nucl. Acids Res. (2010) 38(suppl 2): W3-W6 doi:10.1093/nar/gkq253

» [Abstract](#) » [Full Text \(HTML\)](#) » [Full Text \(PDF\)](#) » [Screen PDF](#) » [Web Server Summaries](#)

OPEN ACCESS [\(CO\)](#) [BY-NC](#)

Federico Abascal, Rafael Zardoya, and Maximilian J. Telford

TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations

Nucl. Acids Res. (2010) 38(suppl 2): W7-W13 doi:10.1093/nar/gkq291

» [Abstract](#) » [Full Text \(HTML\)](#) » [Full Text \(PDF\)](#) » [Screen PDF](#) » [Supplementary Data](#)

OPEN ACCESS [\(CO\)](#) [BY-NC](#)

Daniel Glez-Peña, Daniel Gómez-Blanco, Miguel Reboiro-Jato, Florentino Fdez-Riverola, and David Posada

ALTER: program-oriented conversion of DNA and protein alignments

Nucl. Acids Res. (2010) 38(suppl 2): W14-W18 doi:10.1093/nar/gkq321

» [Abstract](#) » [Full Text \(HTML\)](#) » [Full Text \(PDF\)](#) » [Screen PDF](#) » [Supplementary Data](#)

OPEN ACCESS [\(CO\)](#) [BY-NC](#)

Amarendran R. Subramanian, Suvrat Hiran, Rasmus Steinkamp, Peter Meinicke, Eduardo Corel, and Burkhard Morgenstern

DIALIGN-TX and multiple protein alignment using secondary structure information at GOBICS

Nucl. Acids Res. (2010) 38(suppl 2): W19-W22 doi:10.1093/nar/gkq442

[« Previous](#) | [Next Issue »](#)

This Issue

July 2010 38 (suppl 2)



» [Index By Author](#)

» [Cover Image](#)

» [Table of Contents \(PDF\)](#)

» [Editorial: 2010 Web Server Issue](#)

» [Articles](#)

» [Web Services](#)

» [Stand-Alone Programs for High-Throughput Data](#)

» [Front-Matter/Back-Matter](#)

Find articles in this issue containing these words:

[Advance Access](#)

Bioinformatics

ABOUT THIS JOURNAL CONTACT THIS JOURNAL SUBSCRIPTIONS CURRENT ISSUE ARCHIVE SEARCH

Oxford Journals » Life Sciences » Bioinformatics » Volume 26, Issue 20

BRIEFINGS IN BIOINFORMATICS
RANKED 4TH IN BIOCHEMICAL RESEARCH
METHODS WITH A 7.329 IMPACT FACTOR

Table of Contents

Volume 26 Issue 20 October 2010

[Clear](#) [Get All Checked Abstracts](#)

ORIGINAL PAPER

GENOME ANALYSIS

Jianrong Wang, Ahsan Huda, Victoria V. Lunyak, and I. King Jordan

A Gibbs sampling strategy applied to the mapping of ambiguous short-sequence tags
Bioinformatics (2010) 26(20): 2301-2308 doi:10.1093/bioinformatics/btq490

» [Abstract](#) » [FREE Full Text \(HTML\)](#) » [Full Text \(PDF\)](#) » [Supplementary Data](#)

OPEN ACCESS [\(CO\)](#) [BY-NC](#)

APPLICATIONS NOTE

GENOME ANALYSIS

Roy Ronen, Ido Gan, Shira Modai, Alona Sakuchev, Gideon Dror, Eran Halperin, and Noam Shomron

miRNAkey: a software for microRNA deep sequencing analysis
Bioinformatics (2010) 26(20): 2610-2616 doi:10.1093/bioinformatics/btq483

» [Abstract](#) » [Full Text \(HTML\)](#) » [Full Text \(PDF\)](#) » [Supplementary Data](#) » [Permissions](#)

SEQUENCE ANALYSIS

Naohisa Goto, Ppiti Prins, Misuteru Nakao, Raoul Bonnal, Jan Aerts, and Toshiaki Katayama

BioRuby: bioinformatics software for the Ruby programming language
Bioinformatics (2010) 26(20): 2617-2618 doi:10.1093/bioinformatics/btq479

» [Abstract](#) » [FREE Full Text \(HTML\)](#) » [Full Text \(PDF\)](#)

OPEN ACCESS [\(CO\)](#) [BY-NC](#)

Nikos Darzentas

Circloetto: visualizing sequence similarity with Circos
Bioinformatics (2010) 26(20): 2620-2621 doi:10.1093/bioinformatics/btq484

» [Abstract](#) » [Full Text \(HTML\)](#) » [Full Text \(PDF\)](#) » [Permissions](#)

I. V. Kulakovskiy, V. A. Boeva, A. V. Favorov, and V. J. Makeev

Deep and wide digging for binding motifs in ChIP-Seq data
Bioinformatics (2010) 26(20): 2622-2623 doi:10.1093/bioinformatics/btq485

» [Abstract](#) » [Full Text \(HTML\)](#) » [Full Text \(PDF\)](#) » [Supplementary Data](#) » [Permissions](#)

STRUCTURAL BIOINFORMATICS

D. P. Kloos, B. A. Wallace, and Robert W. Janes

2Struc: the secondary structure server
Bioinformatics (2010) 26(20): 2624-2625 doi:10.1093/bioinformatics/btq486

» [Abstract](#) » [FREE Full Text \(HTML\)](#) » [Full Text \(PDF\)](#) » [Supplementary Data](#)

OPEN ACCESS [\(CO\)](#) [BY-NC](#)

Marco Baccini, Valerio Mariani, Jürgen Haas, Stefan Scheuber, Andreas D. Schenk, Torsten Schwede, and Ansgar Philippson

OpenStructure: a flexible software framework for computational structural biology
Bioinformatics (2010) 26(20): 2630-2633 doi:10.1093/bioinformatics/btq487

» [Abstract](#) » [FREE Full Text \(HTML\)](#) » [Full Text \(PDF\)](#) » [Supplementary Data](#)

OPEN ACCESS [\(CO\)](#) [BY-NC](#)

[« Previous](#)

This Issue

October 2010 20 (20)



» [Index By Author](#)

» [Front Matter \(PDF\)](#)

» [Table of Contents \(PDF\)](#)

» [Back Matter \(PDF\)](#)

» [ORIGINAL PAPER](#)

» [GENOME ANALYSIS](#)

» [SEQUENCE ANALYSIS](#)

» [STRUCTURAL BIOINFORMATICS](#)

» [GENE EXPRESSION](#)

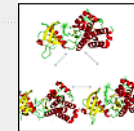
» [SYSTEMS BIOLOGY](#)

» [APPLICATIONS NOTE](#)

Welcome to BMC Bioinformatics published by BioMed Central

BMC Bioinformatics is an open access journal publishing original peer-reviewed research articles in all aspects of the development, testing and novel application of computational and statistical methods for the modeling and analysis of all kinds of biological data, as well as other areas of computational biology. BMC Bioinformatics (ISSN 1471-2105) is indexed/tracked/covered by PubMed, MEDLINE, BIOSIS, CAS, EMBASE, Scopus, ACM, CABI, Thomson Reuters (ISI) and Google Scholar.

Featured articles [Open Access](#)



Predicting protein-small molecule binding

BMC Bioinformatics 2010, 11:365

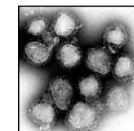
Protein-small molecule interactions can now be predicted with greater accuracy by a new algorithm on the web server IBIS that utilizes inference-by-homology to rank predicted binding sites based on their biological relevance.



Raising the MetaBar for data storage

BMC Bioinformatics 2010, 11:358

Contextual data can now be integrated onto environmental sequence datasets through a web-based software tool, MetaBar, that barcodes samples and stores relevant data such as geographical location and sampling time in an online database.



Automated typing of influenza virus

BMC Bioinformatics 2010, 11:366

Rapid and automated typing and subtyping of the influenza virus is now possible with a new algorithm "FluTyper" designed to analyse MALDI mass spectra from proteolytic digests of the whole influenza virus and antigens.

Why MobyNet ?

Bioinformatics

Nucleic Acids Research Web Server Issue

2004 ~ 137 papers

“from BLAST services to three-dimensional protein structure prediction.”

R. Gautier, A.-C. Camproux, and P. Tufféry
SCit: web tools for protein side chain conformation analysis
Nucl. Acids Res. 2004 32: W508-W511.

2009 ~120 papers

“metagenomics, molecular network and pathway analysis and biological text mining.”

2010 ~ 120 papers “Synthetic biology.”

Peter Schmidtke, Vincent Le Guilloux, Julien Maupetit, and Pierre Tufféry
fpocket: online tools for protein ensemble pocket detection and tracking
Nucleic Acids Research Advance Access published on May 16, 2010.

Table of Contents

Volume 38 suppl 2

Clear Get

Editorial:

Gary Benson
Editorial
Nucl. Acids Res.

» Extract » Full Text » OPEN ACCESS

Articles

Michelle D. Easton
Providing vLinks Direc
Nucl. Acids Res.

» Abstract » Full Text » OPEN ACCESS

Federico Abati
TranslatorX
Nucl. Acids Res.

» Abstract » Full Text » OPEN ACCESS

Daniel Glez-Posada
ALTER: prc
Nucl. Acids Res.


» Abstract » Full Text » OPEN ACCESS

Amarendran R. Subramanian, Suvrat Hiran, Rasmus Steinkamp, Peter Meinicke, Eduardo Corel, and Burkhard Morgenstern


DIALIGN-TX and multiple protein alignment using secondary structure information at GOBICS
Nucl. Acids Res. (2010) 38(suppl 2):W19-W22 doi:10.1093/nar/fgq442

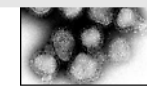
STRUCTURAL BIOINFORMATICS

D. P. Klose, B. A. Wallace, and Robert W. Janes
Z2Struct: the secondary structure server
Bioinformatics (2010) 26(20):2624-2625 doi:10.1093/bioinformatics/btq480

» Abstract » FREE Full Text (HTML) » Full Text (PDF) » Supplementary Data
OPEN ACCESS 

Marco Biaschi, Valerio Mariani, Jürgen Haas, Stefan Scheuber, Andreas D. Schenk, Torsten Schwede, and Ansgar Philippson
OpenStructure: a flexible software framework for computational structural biology
Bioinformatics (2010) 26(20):2630-2633 doi:10.1093/bioinformatics/btq481

» Abstract » FREE Full Text (HTML) » Full Text (PDF) » Supplementary Data
OPEN ACCESS 



Rapid and automated typing and subtyping of the influenza virus is now possible with a new algorithm 'FluTyper' designed to analyse MALDI mass spectra from proteolytic digests of the whole influenza virus and antigens.

arch
al
well
CM,

I with
iS
ling

al
line

Why MobyNet ?

1. Lots of new methods each year: $k * 100$ new services . year

New themes, new skills, new data types, ...

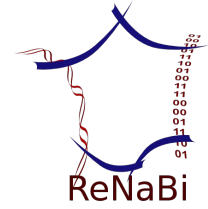
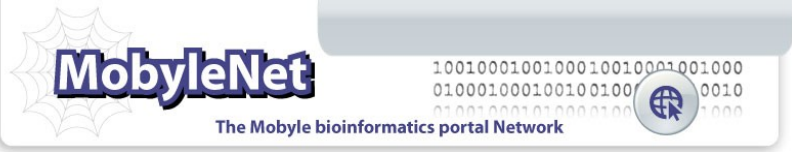
==> *Need for resource discovery facilities*

2. Evolution towards complex treatments, multi-scale analyses, knowledge integration

==> *Need for complex treatments facilities (pipelines, workflows, protocols, ...)*

3. Biologists and bioinformaticians do not share the same usage of the tools

==> **Avoid programmatic access to tools**



Why MobyNet ?

Possible strategy 1:

Setup reference **bioinformatics Centers** (national, worldwide):
Integrate databanks + tools at one place

But: *uneasy to deal with increasing Data, methods, approaches, ...*
Surprisingly, no offer (yet ?) for Integrated workflows through the Web.

The screenshot shows the EMBL-EBI website interface. At the top, there is a search bar with 'All Databases' selected and a search input field. Below the search bar is a navigation menu with links for Databases, Tools, EBI Groups, Training, Industry, About Us, Help, and Site Index. The main content area is divided into several sections:

- Data Resources & Tools:** A grid of links including ENA, UniProt, ArrayExpress, Ensembl, InterPro, PDBe, Genomes, Nucleotide Sequences, Protein Sequences, Macromolecular Structures, Small Molecules, Gene Expression, Molecular Interactions, Reactions & Pathways, Protein Families, Enzymes, Literature, Taxonomy, Ontologies, Patent Resources, Sequence Similarity & Analysis, Pattern & Motif Searches, Structure Analysis, Text Mining, Downloads, and Web Services.
- European Bioinformatics Institute:** A central banner featuring a tree and the institute's name.
- About the EBI:** Links for Research, PhD Studies, Training, Industry Support, Group & Team Leaders, EBI Funders, User Support, EBI Mission, People, Events at the EBI, Genome Campus Events, How to Find us, and Jobs.
- Events:** A list of upcoming events with dates and registration information, such as 'Small Molecule Bioactivity Resources' (14-18 February 2011) and 'In silico Systems Biology: Network Reconstruction, Analysis and Network-based Modelling' (23-26 May 2011).
- Research Highlights:** A section titled 'Third mosquito genome sequenced' dated Sept 30, 2010, describing the sequencing of the Southern house mosquito, Culex quinquefasciatus.
- EBI Hosted Project Websites:** A list of various databases and tools including 1000 Genomes, BioCatalogue, BioSapiens, CREATE, E-MeP, EGA, ELIXIR, EMMA, EMBRACE, EMERALD, ENFIN, FELICS, IMPACT, INSDC, LRG, SLING, Microme, SPINE, SYBARIS, and SYMBIOmatics.
- Latest News:** A section with a RSS feed icon.

Why MobyNet ?

Possible strategy 2:

Setup a network of interoperable Services as **web-services**:
Elaborate complex treatment flows

But: *Unlike bioinformaticians, biologists do not like/know how to deal with programmatic access to treatment. Prefer visual access to resource.*

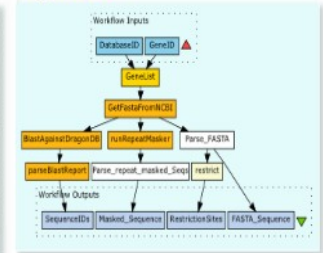


Moby 1.0 Code Released
Start the [Moby Dashboard](#) to explore the Moby Central Registry

Tired of This?

Protocol
Create a gene list in Excel
Go to NCBI
Retrieve FASTA for each gene
DragonDB Blast each sequence
Copy/parse IDs into a spreadsheet
Run Repeat Masker on each sequence
copy/paste masked sequence into Excel
Run MacVector cut each seq with EcoRI

Try This!



Worldwide Distribution of MOBY Services



Archives

- » June 2008
- » May 2008
- » February 2008
- » June 2007
- » April 2006

Categories

- » General (24)
- » MOBY Clients (1)
- » Moby for Java (2)
- » News (3)

Downloads, Docs, and Stats

- » Code Releases
- » CVS Code Repository
- » General Docs
- » Currently 'dead' Services

RSS Feeds
 feeds are updated hourly...

- » MOBY Services
- » MOBY Objects

Meta

- » Log in
- » WordPress

Why MobyNet ?

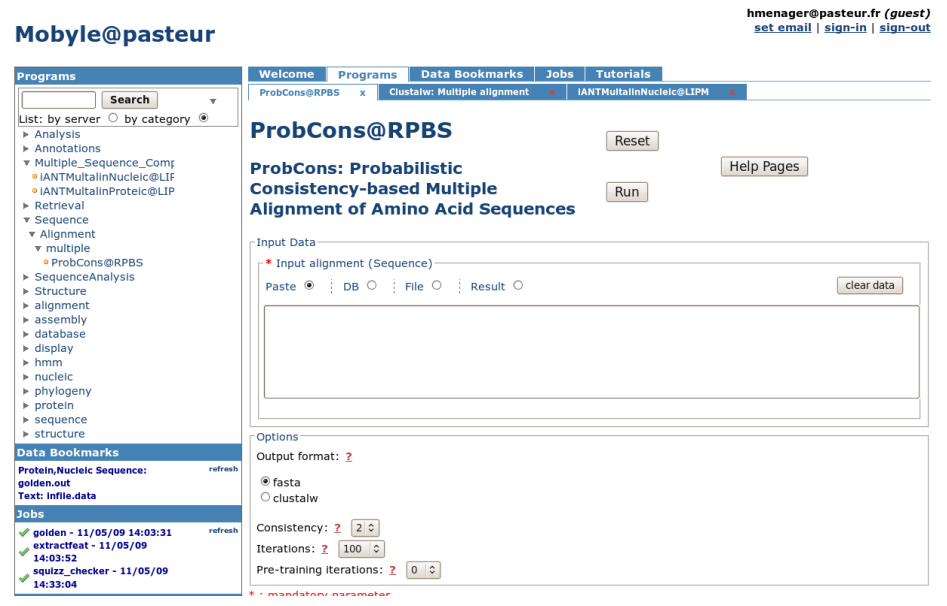
Possible strategy 3:

Setup resources available through web.

Make web servers interoperable so as to allow complex protocols.

Make resources available for programmatic access for Bioinformaticians.

*Keep skills close to services.
Make easy to aggregate even small contributions.*



MobyNet@pasteur hmenager@pasteur.fr (guest)
[set_email](#) | [sign-in](#) | [sign-out](#)

Programs | Welcome | Programs | Data Bookmarks | Jobs | Tutorials

ProbCons@RPBS x Clustalw: Multiple alignment x IANTMultalinNucleic@LIPM x

ProbCons@RPBS

ProbCons: Probabilistic Consistency-based Multiple Alignment of Amino Acid Sequences

Reset Help Pages Run

Input Data

* Input alignment (Sequence)

Paste DB File Result

clear data

Options

Output format: ?

fasta
 clustalw

Consistency: ?

Iterations: ?

Pre-training iterations: ?

* mandatory parameter

Programs

Search

List: by server by category

- Analysis
- Annotations
- Multiple_Sequence_Comparisons
 - IANTMultalinNucleic@LIP
 - IANTMultalinProtein@LIP
- Retrieval
- Sequence
 - Alignment
 - multiple
 - ProbCons@RPBS
- SequenceAnalysis
- Structure
- alignment
 - assembly
 - database
 - display
 - hmm
 - nucleic
 - phylogeny
 - protein
 - sequence
 - structure

What is MobyNet ?

- **An ensemble of web portals**
 - User-centered interface using Web
 - Based on *MobyNet*
 - Service deployment as web-services still possible
 - Services close to skills : “*specialized*” nodes
- **A network of interoperable portals**
 - Seamless invocation of remote tools
 - Workflows possible (workflow concept integrated to MobyNet design)
 - Enable a large spectrum of applications (aggregating nodes)
- **A confidence network**
 - Control over service import/export
 - Redundancy
 - Availability of services
- **Opened project**
 - Open source
 - New nodes

What is MobyNet not ?

- **An integrated computational resource**
 - No grid, no cloud computing considerations
 - standard bandwidth for data -> limit over data size affordable by a web portal.
 - Each node might have a distinct architecture for calculation (flexible).
- **Data sharing space**
 - No data sharing: each node manages its banks and workspace independently.
- **Frozen vision of online service setup**
 - Each portal is customizable
 - Themes, services, data-types can evolve freely (new types, new categories of services).
- **A 100% operational project (some features still in development)**

1. MobyNet: the project

2. Moby

- **Concepts**
- Moby Portal
- Web services
- Moby Workflows
- Moby vs Galaxy

3. MobyNet specific aspects

4. MobyNet today

Nantes, 21 Octobre 2010

- **Moby**
 - Moby core server
- **Moby companions**
 - BMID (interface designer)
 - BCBBPipe (workflows)
 - Charlyse (workflows with indexation of analyses)
 - PlayMoby (BioMOBY web-services)
 - Moby Opal (Opal web-services)

What is Moby ?

- a bioinformatics analysis Web portal.
- provides an access to bioinformatics command line tools via a web interface.



- emphasis on:
 - **usability** for biologists,
 - **extensibility** for developers,
 - **maintainability** for server administrators



Moby design for users :

Moby

Programs

- 2D_structure
- alignment
- assembly
- database
- display
- genetics
- hmm
- nucleic
- phylogeny
- sequence
- rpbs

Data Bookmarks

Sequence: ABCD
Sequence: CBM20.fasta

Jobs

- ✔ toppred - 06/14/10 17:00:17
- ✔ toppred - 06/14/10 17:02:00
- ✔ clustalw-multialign - 06/15/10 09:14:02
- ✔ protdist - 06/15/10 16:00:47
- ✔ protdist - 06/15/10 16:11:58
- ✔ clustalw-multialign - 06/16/10 09:23:54
- ✔ clustalw-multialign - 06/16/10 09:24:33
- ✔ clustalw-multialign - 06/16/10 09:27:47
- ✔ clustalw-multialign - 06/16/10 09:28:09
- ✔ clustalw-multialign - 06/16/10 09:33:18
- ✔ clustalw-multialign - 06/16/10 09:34:05
- ✔ clustalw-multialign - 06/16/10 09:35:00
- ✔ clustalw-multialign - 06/16/10 09:35:48
- ✔ clustalw-multialign - 06/17/10 09:11:57

Welcome | Programs | Data Bookmarks | Jobs | Tutorials

clustalw-multialign ✖

Clustalw: Multiple alignment

version 2.0.11

Do full multiple alignment

Sequences File (a file containing several sequences) (-infile)

paste db upload

General settings

Toggle Slow/Fast pairwise alignments (-quicktree)

Slow Fast
Protein or DNA (-ttype)
 Automatic Protein DNA

Multiple Alignments parameters?

Gap opening penalty (-gapopen)

Gap extension penalty (-gapext)

No end gap separation penalty (-endgaps?)

Gap separation penalty range (-gapdist?)

Delay divergent sequences : % ident. for delay (-maxdiv)?

File for new guide tree (-newtree)

File for old guide tree (-usetree)

paste upload

Enter your data below:

hmenager@pasteur.fr (guest)
set email | sign-in | OpenId sign-in | activate | sign-out
refresh workspace

Moby

Programs

- 2D_structure
- alignment
- assembly
- database
- display
- genetics
- hmm
- nucleic
- phylogeny
- sequence
- rpbs

Data Bookmarks

Sequence: ABCD
Sequence: CBM20.fasta

Jobs

- ✔ toppred - 06/14/10 17:00:17
- ✔ toppred - 06/14/10 17:02:00
- ✔ clustalw-multialign - 06/15/10 09:14:02
- ✔ protdist - 06/15/10 16:00:47
- ✔ protdist - 06/15/10 16:11:58
- ✔ clustalw-multialign - 06/16/10 09:23:54
- ✔ clustalw-multialign - 06/16/10 09:24:33
- ✔ clustalw-multialign - 06/16/10 09:27:47
- ✔ clustalw-multialign - 06/16/10 09:28:09
- ✔ clustalw-multialign - 06/16/10 09:33:18
- ✔ clustalw-multialign - 06/16/10 09:34:05
- ✔ clustalw-multialign - 06/16/10 09:35:00
- ✔ clustalw-multialign - 06/16/10 09:35:48
- ✔ clustalw-multialign - 06/17/10 09:11:57

Welcome | Programs | Data Bookmarks | Jobs | Tutorials

clustalw-multialign - 06/17/10 15:55:09 ✖

✔ http://rita.sis.pasteur.fr:85/data/jobs/clustalw-multialign/T14274755194902

results

Output parameters

Alignment file

```

CLUSTAL 2.0.10 multiple sequence alignment

sp|P22832|AMYG_ASPSH      -MSFRSLLALSGLVCSGLAS--VISKRATLDSWLSNEATVARTAILNINIG
sp|P23176|AMYG_ASPKA      -MSFRSLLALSGLVCSGLAN--VISKRATLDSWLSNEATVARTAILNINIG
sp|P36914|AMYG_ASPOR      MVSFSSCLRALALGSSVLAQPVLQRATGLDWTLSSTEANFSRQAILNINIG
sp|P26827|CDGT_THETU      -MKKTFKLLILVLM.SLTLVFGLTAPIQAASDTAVSNVNVNYS-TDVIYQIV
tr|Q92394|Q92394_9HETE    -MAPVRSRSLAGALLASLGLVAGLSPAENRSQS-----IYQVV
tr|Q1KLC8|Q1KLC8_THEFU    -MGVRRSLAALLAALLGCAT-----
    
```

Tree file

```

(
  (
    sp|P36914|AMYG_ASPOR:0.15853,
    (
      sp|P22832|AMYG_ASPSH:0.00803,
      sp|P23176|AMYG_ASPKA:0.00762)
    :0.14492)
  :0.29501,
  tr|Q1KLC8|Q1KLC8_THEFU:0.41547,
  /
)
    
```

Standard output

```

CLUSTAL 2.0.10 Multiple Sequence Alignments

Sequence format is Pearson
Sequence 1: sp|P36914|AMYG_ASPOR      612 aa
Sequence 2: sp|P22832|AMYG_ASPSH    639 aa
Sequence 3: sp|P23176|AMYG_ASPKA    639 aa
Sequence 4: sp|P26827|CDGT_THETU    639 aa
Sequence 5: tr|Q92394|Q92394_9HETE   639 aa
Sequence 6: tr|Q1KLC8|Q1KLC8_THEFU   639 aa
    
```

parameters

hmenager@pasteur.fr (guest)
set email | sign-in | OpenId sign-in | activate | sign-out
refresh workspace

Mobyle design for developers :

AJAX Dialog

Form.py
Automatic forms

job-submit

job-status

job-kill

Program
definition
s

MobyleJob
parameter
validation

WF engine
JobFacade

Mobyle Core

Job

RunnerFather
command line
building
Jobs status

RunnerChild

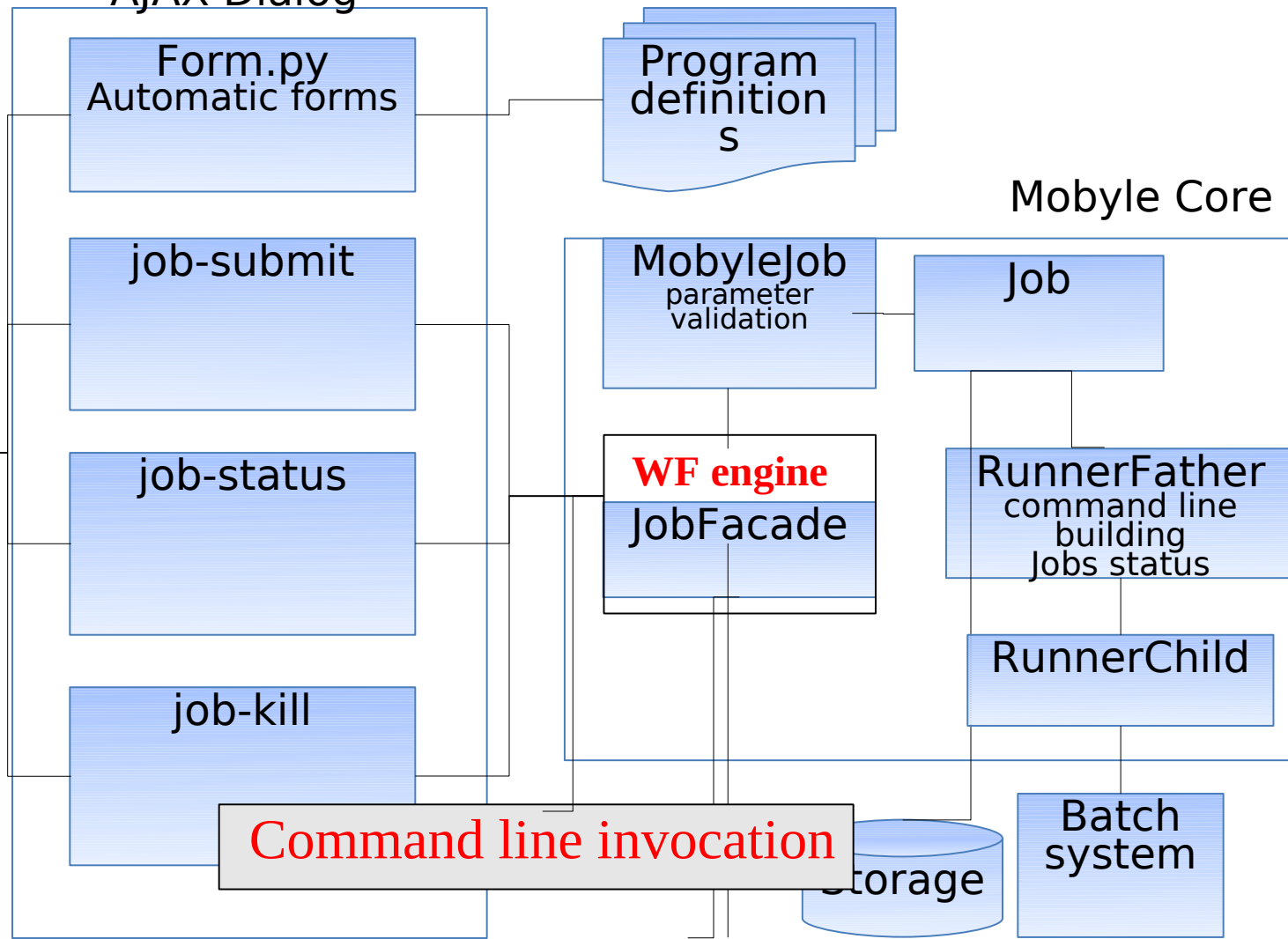
Command line invocation

Storage

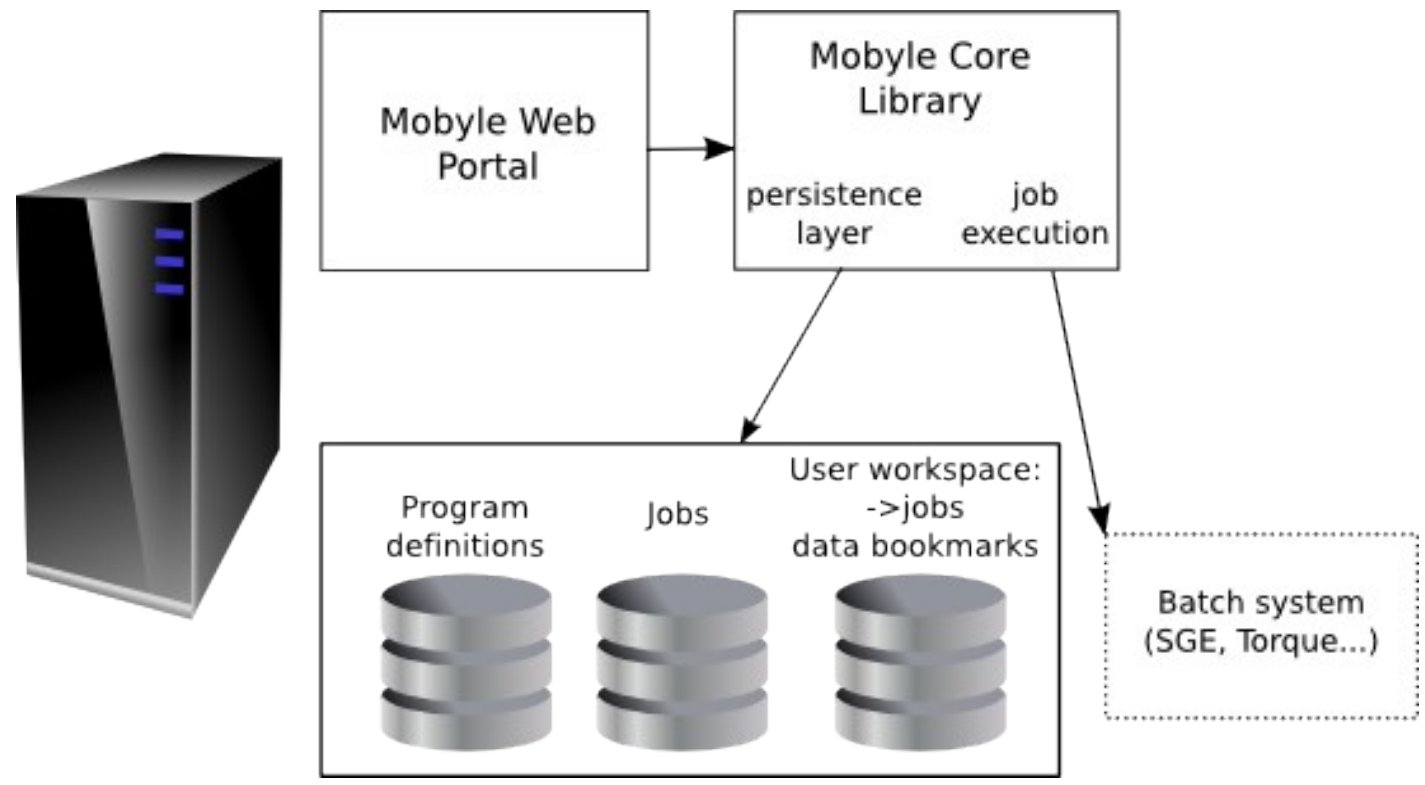
Batch
system

Browser
Portal.py

MobyleNet



Moby design for administrators and service providers :





<program> **Moby service publication: xml / relax NG**

```
<head>
  <name>fpocket2</name>
  <version>1.0</version>
  <doc>
    <title>fpocket2</title>
    <description>
      <text lang="en">
        A program for pocket identification
        in protein structures
      </text>
    </description>
    <comment>
      <text lang="en">
        fpocket is a very fast open source protein pocket (cavity)
        detection algorithm based on Voronoi tessellation. As the
        algorithm is very fast it can be used on a large scale level
        (PDB size for instance).
      </text>
    </comment>
    <reference>
      fpocket: online tools for protein ensemble pocket detection
      and tracking. Schmidtke P, Le Guilloux V, Maupetit J, Tuffery
      P. Nucleic Acids Res. 2010 May 16.
    </reference>
    <reference>
      Fpocket: an open source platform for ligand pocket detection.
      Le Guilloux V, Schmidtke P, Tuffery P.
      BMC Bioinformatics. 2009 Jun 2;10:168.
    </reference>
    <authors>Schmidtke P, Le Guilloux V, Maupetit J, Tuffery P.</authors>
    <doclink>http://fpocket.sourceforge.net/</doclink>
  </doc>
  <category>Structure:Pockets</category>
  <command>wfpocket</command>
</head>
```



```
<program>
<head>
  <name>fpocket2</name>
  <version>1.0</version>
  <doc>
    <title>fpocket2</title>
    <description>
      <text lang="en">
        A program for pocket identification
        in protein structures
      </text>
    </description>
    <comment>
      <text lang="en">
        fpocket is a very fast open source
        detection algorithm based on Voronoi
        tessellation. As the algorithm is very fast it can be
        used on a large scale level (PDB size for
        instance).
      </text>
    </comment>
    <reference>
      fpocket: online tools for protein
      and tracking. Schmidtke P, Le Guill
      P. Nucleic Acids Res. 2010 May 16.
    </reference>
    <reference>
      Fpocket: an open source platform for
      protein pocket detection. Le Guilloux V, Schmidtke P, Tuffery
      P. BMC Bioinformatics. 2009 Jun 2;10:100.
    </reference>
    <authors>Schmidtke P, Le Guilloux V, Maupetit J, Tuffery P.</authors>
    <doclink>http://fpocket.sourceforge.net/</doclink>
  </doc>
  <category>Structure:Pockets</category>
  <command>wfpacket</command>
</head>
```

Moby@RPBS

(guest)
[set email](#) | [sign-in](#) | [sign-out](#)

Contact us

Programs

fpocket Search or

All

- Structure
- Pockets
 - HMDPocket
 - fpocket
 - fpocket2
 - hpocket
 - mdpocket

Data Bookmarks

Jobs

Tutorials

- [How to use Moby? A step by step tutorial](#)
- [Registration information](#)
- [Sequence formats](#)
- [Alignment formats](#)

Welcome Programs Data Bookmarks Jobs Tutorials

fpocket2 x

fpocket2

Reset

Help Pages

Run

fpocket is a very fast open source protein pocket (cavity) detection algorithm based on Voronoi tessellation. As the algorithm is very fast it can be used on a large scale level (PDB size for instance).

Demonstration mode

Test the service with server sample data (input parameters will be discarded)

Input parameters

Job label

Query PDB (PDB) ?

Paste ; DB ; File

clear data

On-line creation/editition of interface: BMID <http://exon.niaid.nih.gov/BMID/>

```
clustalw -infile=infile.data -align -type=protein -usetree=usetree.data&& perl /common/bin/aln-color.pl *.aln *.phy *.nxs
```

Moby XML

```
<?xml?>
<?xml-stylesheet href="http://www.ebi.ac.uk/ontology/ontology.xsl" type="text/xsl"?>
<!--
<title>ClustalW</title>
<version>1.83</version>
<description>Multiple Alignments</description>
<author>Des Higgins</author>
<reference>
  Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving
  </reference>
<category>Sequence Alignment</category>
<doclink>http://www.ebi.ac.uk/clustalw/faq.html</doclink>
</head>
<command>clustalw</command>
<parameters>
  + <parameter iscommand="1" ishidden="1" type="String"></parameter>
  + <parameter ismandatory="1" isimple="1" type="Sequence"></parameter>
  + <parameter ismandatory="1" type="Excl"></parameter>
  + <parameter isimple="1" type="Switch"></parameter>
  + <parameter type="Paragraph"></parameter>
  + <parameter type="Paragraph"></parameter>
  + <parameter type="Paragraph"></parameter>
  + <parameter type="Paragraph"></parameter>
  - <parameter type="Paragraph">
    - <paragraph>
      <name>outputparam</name>
      <prompt>Output parameters</prompt>
      <group-2</group>
      - <parameters>
        + <parameter type="OutFile"></parameter>
        + <parameter type="Excl"></parameter>
        + <parameter type="Switch"></parameter>
        + <parameter type="Excl"></parameter>
        + <parameter type="Switch"></parameter>
      </parameters>
    </paragraph>
  </parameter>
  + <parameter type="Results"></parameter>
  + <parameter type="Results"></parameter>
  + <parameter type="Results"></parameter>
  + <parameter type="Results"></parameter>
  + <parameter type="Results"></parameter>
  + <parameter type="Results"></parameter>
  + <parameter type="Paragraph"></parameter>
  + <parameter type="Paragraph"></parameter>
</parameters>
</?xml?>
```

BMID (BCBB Moby Interface Designer) - BETA vers

Design Preview Source

BMIDcontrols

Paragraph Sequence Choice List Integer Float Boolean String Text Horizontal Layout

clear all load templates save as new draft edit delete all

Parameters Panel

program properties

Header

Drag items here

db

Database

Choose a database

query

Query (Entry name or Accession number)

ac

Search with Accession number only (-a)

id

Search with entry name only (-l)

golden_out

Browse...

OR

1. MobyNet: the project

2. Moby

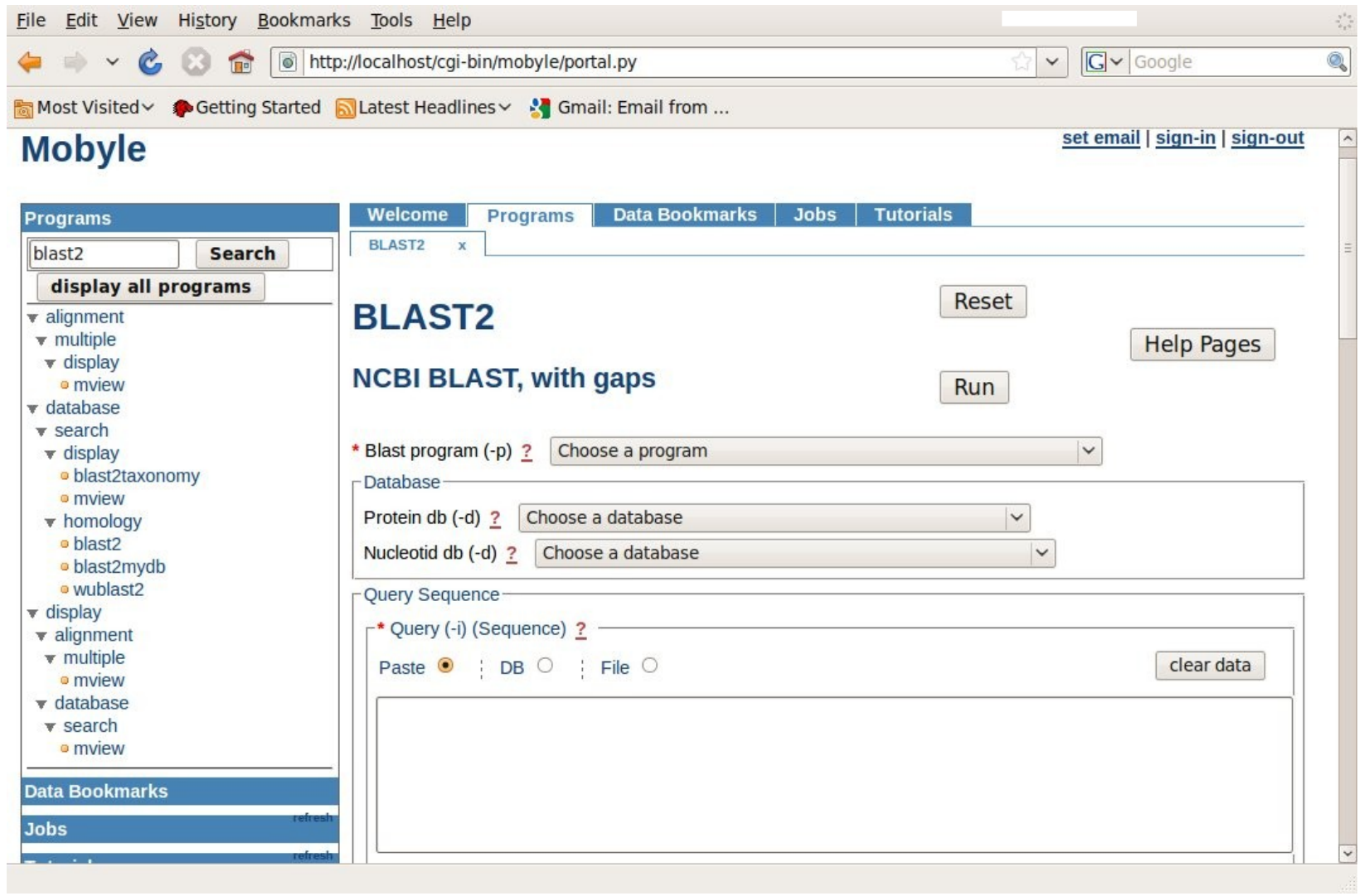
- Concepts
- **Moby Portal**
- Web services
- Moby Workflows
- Moby vs Galaxy

3. MobyNet specific aspects

4. MobyNet today

Nantes, 21 Octobre 2010

Moby user side



The screenshot shows a web browser window with the URL `http://localhost/cgi-bin/moby/portal.py`. The page title is "Moby" and it includes navigation links for "set email", "sign-in", and "sign-out".

The main content area features a navigation menu with "Welcome", "Programs", "Data Bookmarks", "Jobs", and "Tutorials". The "Programs" section is active, showing a search bar with "blast2" and a "Search" button. Below the search bar is a "display all programs" button and a tree view of program categories:

- alignment
 - multiple
 - display
 - mview
- database
 - search
 - display
 - blast2taxonomy
 - mview
 - homology
 - blast2
 - blast2mydb
 - wublast2
 - display
 - alignment
 - multiple
 - mview
 - database
 - search
 - mview

Below the tree view are sections for "Data Bookmarks" and "Jobs", each with a "refresh" button.

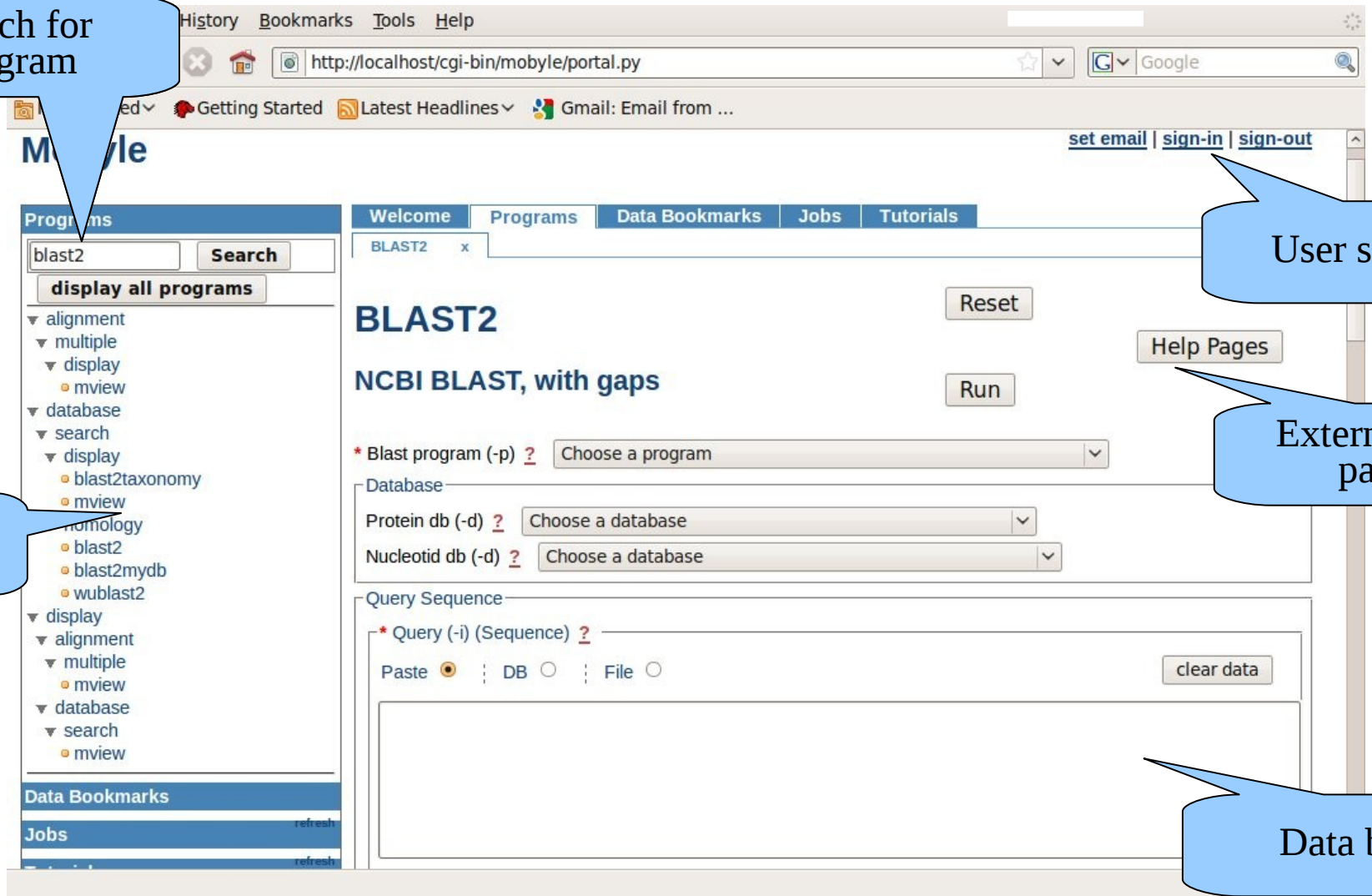
The main tool configuration area is titled "BLAST2" and "NCBI BLAST, with gaps". It includes a "Reset" button and a "Help Pages" button. The configuration fields are:

- * Blast program (-p) ?
- Database
 - Protein db (-d) ?
 - Nucleotid db (-d) ?
- Query Sequence
 - * Query (-i) (Sequence) ?
 - Input options: Paste (selected), DB, File
 - clear data button

Moby user side

Search for program

Program menu



History Bookmarks Tools Help

http://localhost/cgi-bin/moby/portal.py

Getting Started Latest Headlines Gmail: Email from ...

set email | sign-in | sign-out

Programs

blast2 Search

display all programs

- alignment
 - multiple
 - display
 - mview
- database
 - search
 - display
 - blast2taxonomy
 - mview
 - homology
 - blast2
 - blast2mydb
 - wublast2
 - display
 - alignment
 - multiple
 - mview
 - database
 - search
 - mview

User session

External help pages

Data box



Databox

* Query (-i) (Sequence) ?

Paste | DB | File

clear data

MKFLILLFNILCLFPVLAADNHGVPQGAS

copy/paste

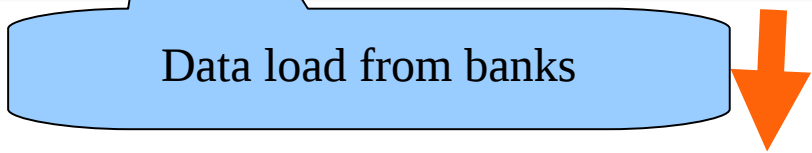
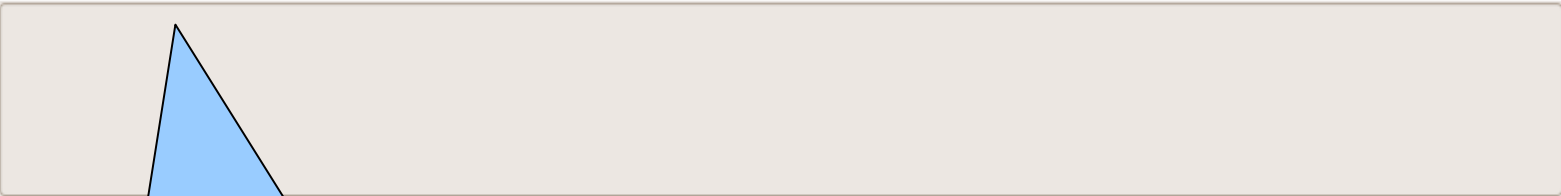
Databox

* Query (-i) (Sequence) ?

Paste | DB | File

edit data clear data

uniprot 104K_THEPA add



* Query (-i) (Sequence) ?

Paste | DB | File

edit data clear data

<> add

```
ID 104K_THEPA Reviewed; 924 AA.  
AC P15711; Q4N2B5;  
DT 01-APR-1990, integrated into UniProtKB/Swiss-Prot.  
DT 01-APR-1990, sequence version 1.  
DT 10-JUN-2008, entry version 41.  
DE 104 kDa microneme/rhoptry antigen precursor (p104).  
GN OrderedLocusNames=TP04_0437;
```

Databox

* Query (-i) (Sequence) ?

Paste | DB | File | Result

edit data clear data

Results bookmarks: golden.out select

Bookmark

* Query (-i) (Sequence) ?

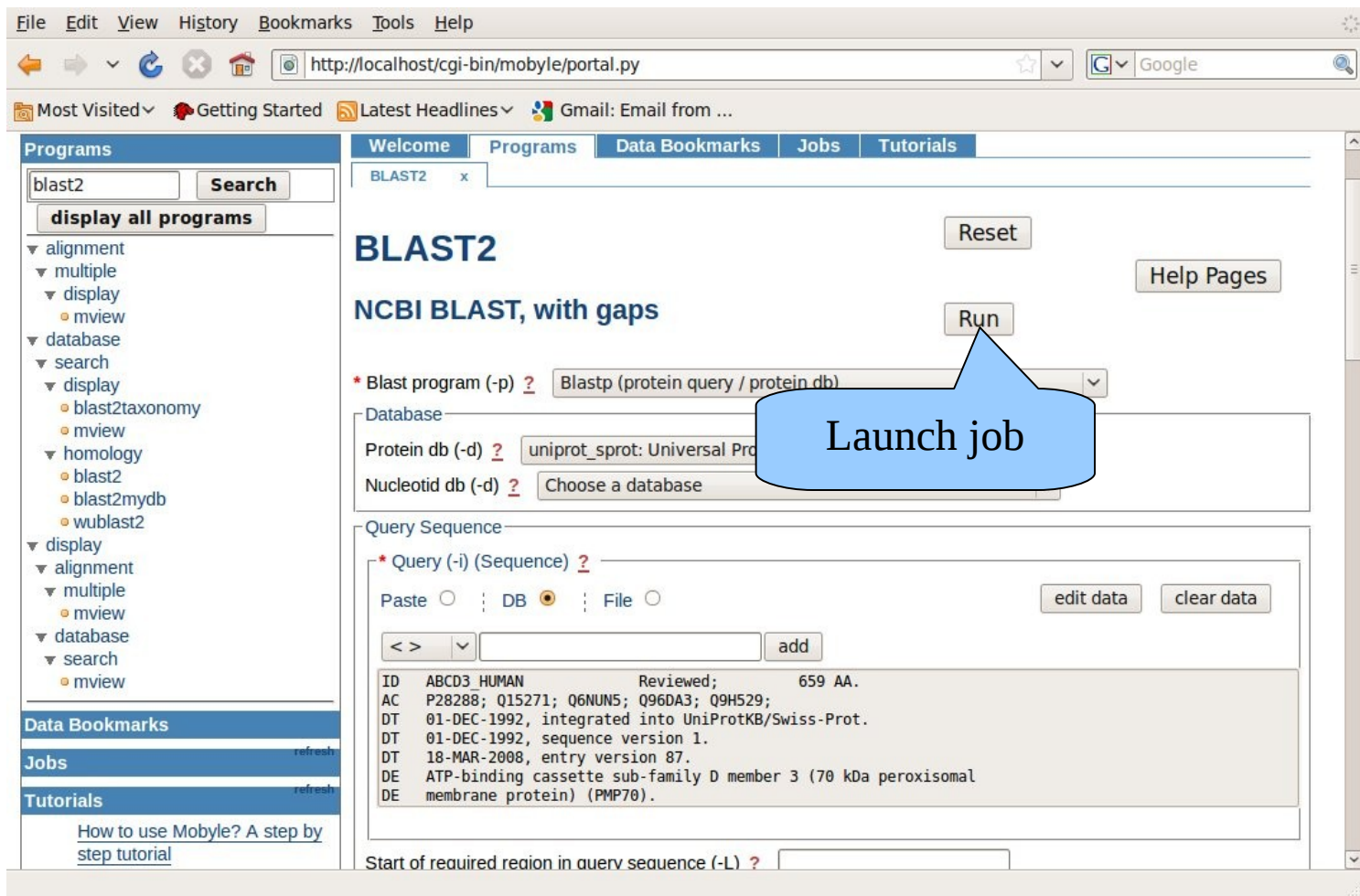
Paste | DB | File | Result

edit data clear data

Results bookmarks: < > select

```
ID ABCD1_MOUSE Reviewed; 736 AA.  
AC P48410; Q9QY41; Q9QZ32;  
DT 01-FEB-1996, integrated into UniProtKB/Swiss-Prot.  
DT 01-FEB-1996, sequence version 1.  
DT 10-JUN-2008, entry version 72.  
DE ATP-binding cassette sub-family D member 1 (Adrenoleukodystrophy  
DE protein) (ALDP).
```

Lancement

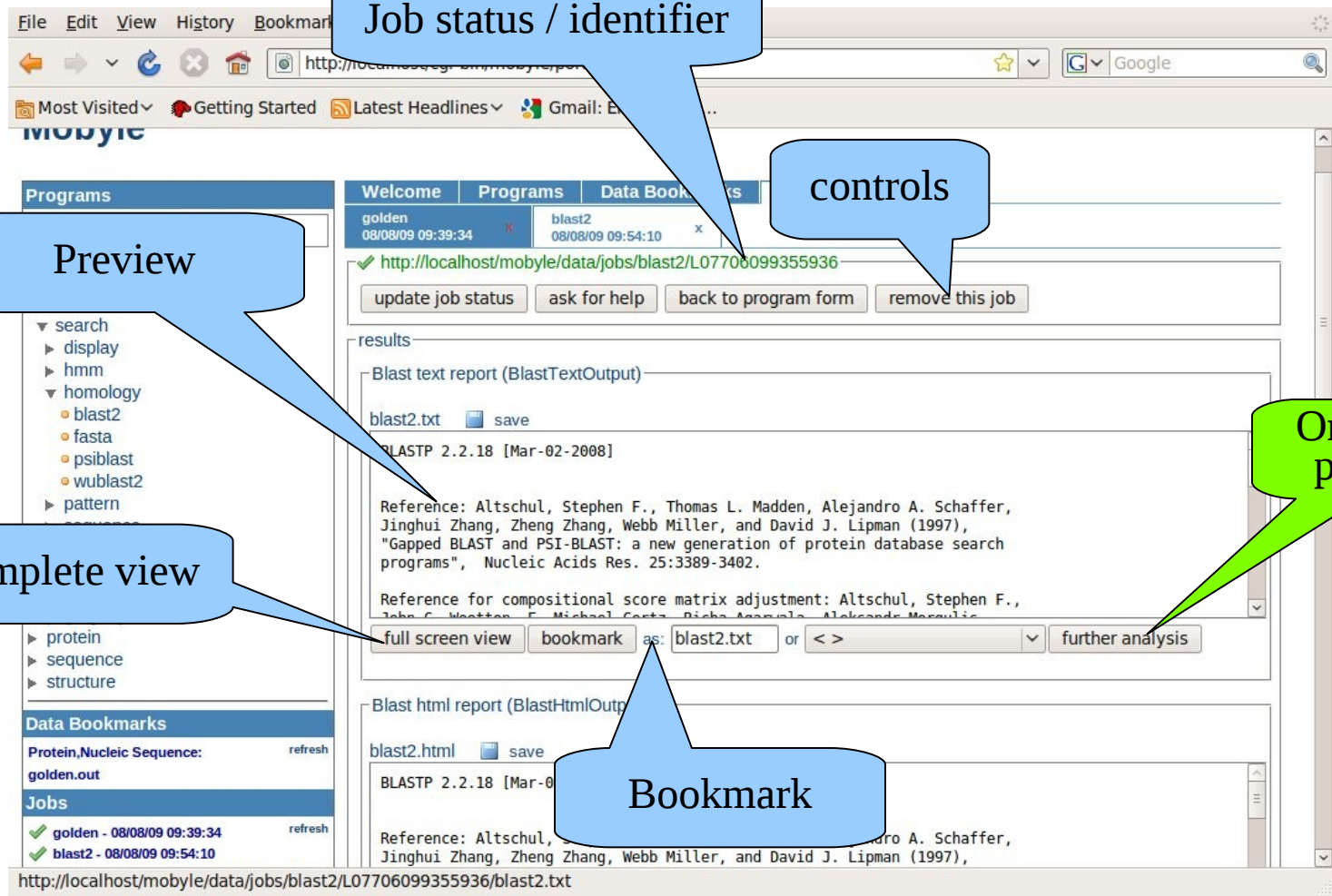


The screenshot shows a web browser window at `http://localhost/cgi-bin/moby/portal.py`. The page title is "BLAST2" and the main heading is "NCBI BLAST, with gaps". A blue callout bubble with the text "Launch job" points to the "Run" button. The interface includes a left sidebar with a "Programs" menu, a top navigation bar with "Welcome", "Programs", "Data Bookmarks", "Jobs", and "Tutorials", and a main content area with input fields for "Blast program (-p)", "Database", "Protein db (-d)", "Nucleotid db (-d)", and "Query Sequence".

Programs

- blast2
-
- alignment
 - multiple
 - display
 - mview
- database
 - search
 - display
 - blast2taxonomy
 - mview
 - homology
 - blast2
 - blast2mydb
 - wublast2
 - display
 - alignment
 - multiple
 - mview
 - database
 - mview

Résultats



The screenshot shows a web browser window with the URL `http://localhost/moby/data/jobs/blast2/L07706099355936/blast2.txt`. The page has a navigation menu on the left with categories like 'search', 'display', 'hmm', 'homology', 'blast2', 'fasta', 'psiblast', 'wublast2', and 'pattern'. The main content area shows job details for 'golden' and 'blast2', a 'Blast text report (BlastTextOutput)' with a 'blast2.txt' file, and a 'Blast html report (BlastHtmlOutput)' with a 'blast2.html' file. Callouts point to various elements: 'Job status / identifier' points to the job ID, 'controls' points to buttons like 'update job status', 'ask for help', 'back to program form', and 'remove this job', 'Preview' points to the 'blast2.txt' file, 'Complete view' points to the 'full screen view' button, and 'Bookmark' points to the 'bookmark' button.

On the fly pipeline

Bookmark



Moby: embedding applets

The screenshot shows the Moby portal interface in a Mozilla Firefox browser. The address bar shows the URL: `http://moby.repbs.univ-paris-diderot.fr/cgi-bin/portal.py`. The page title is "Moby portal".

JME editor

Draw compound using JME, get its description in the smiles or mol format. Load compound from its mol description, graphically edit it using JME.

Buttons: **Reset**, **Run**, **Help Pages**

SMILES/MOL data

Use this applet to edit your SMILES/MOL data

Buttons: CLR, NEW, DEL, 123, D-R, +/-, UDO, JME

Buttons: To SMILES input, To MOL input, Draw MOL input

SMILES data (Smiles_structure)

Paste (selected), File, clear data

MOL data (MOL_structure)

Paste (selected), File, clear data

Tutorials

- How to use Moby? A step by step tutorial
- Registration information
- Sequence formats
- Alignment formats

* : mandatory parameter

Data viewers



Jalview
version 2.0.11
Multiple Alignment Viewer

File Edit Select View Format Colour Calculate Help

Sequence 1 ID: ABCD4_HUMAN

Sequence: abcd
Alignment: musc
Sequence: abcd

References:
Andrew M. Waterhouse, James B. Procter, David M. A. Martin, Michèle Clamp, and Geoffrey J. Barton
Jalview Version 2 - a multiple sequence alignment editor and analysis workbench
Bioinformatics Advance Access published on January 16, 2009, DOI 10.1093/bioinformatics/btp033

Author(s): Andrew M. Waterhouse, James B. Procter, David M. A. Martin, Michèle Clamp and Geoffrey J.

VARNa: Visualization Applet for RNA
version 3.6
A Java lightweight component and applet for drawing the RNA secondary structure

File Edit View History Bookmarks Tools Help

Sequence: 165
DNA Sequence: abcd
Sequence: abcd
Sequence: sequ
Sequence: sequ

Jobs

Archaeopteryx
version 0.955 beta
Phylogenetic Tree Viewer

File Edit View History Bookmarks Tools Help

Tools: View as Text, Font Size, Options, Type, Help

Phylogram
 Dynamic Hiding
 Show Internal Data
 Taxonomy colorize
 Collapse branches
 Display Data

Node Name
 Bootstrap Value
 Taxonomy Name
 Protein/Gene Name
 Protein/Gene Acc
 Confidence Value
 Event

(Click on Node to):
[treeview] [treeview] [treeview]

Zoom: In Y, Out Y, Fit

Reverse to Supp/Tree:
Order Subtree
Uncollapse All

Search:

Sequence: 165
DNA Sequence: abcd
Sequence: abcd
Sequence: sequ
Sequence: sequ

Jobs

Author(s): Mira V Han and Christian M Zmasek

1. MobyNet: the project

2. Moby

- Concepts
- Moby portal
- **Web services**
- Moby Workflows
- Moby vs Galaxy

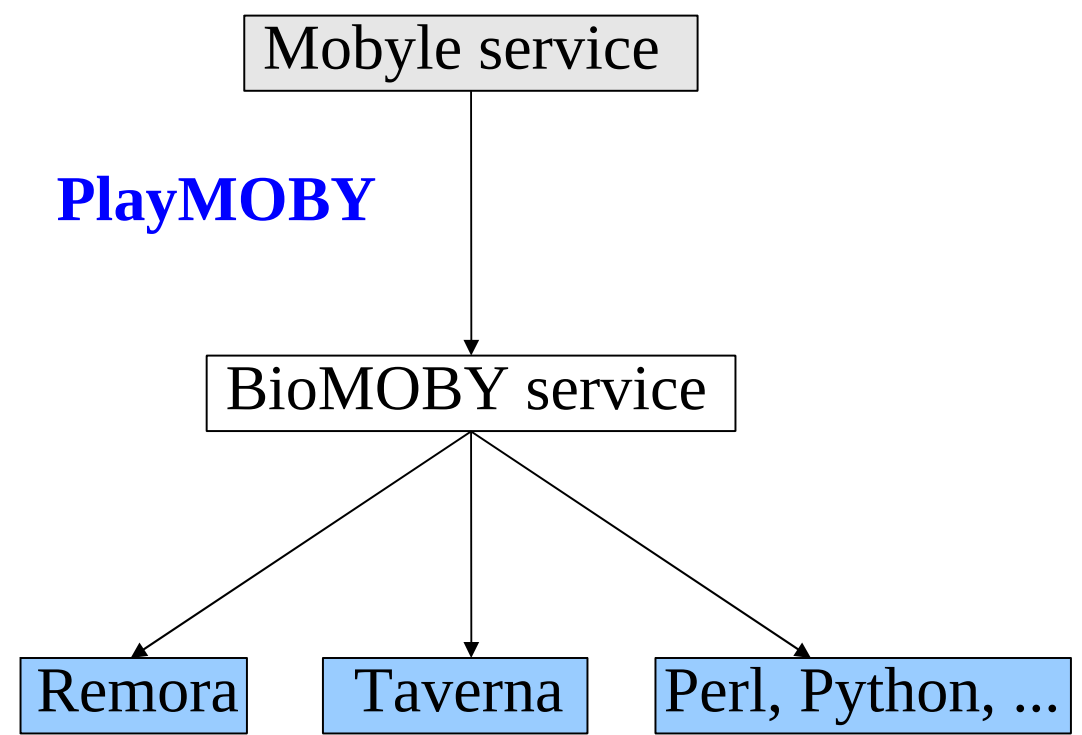
3. MobyNet specific aspects

4. MobyNet today

Nantes, 21 Octobre 2010

PlayMOBY: Interoperability through BioMOBY

Sébastien Letort, Sébastien Carrere, Jérôme Gouzy



<http://lipm-bioinfo.toulouse.inra.fr/download>

Licence CeCILL

PlayMOBY: Trusted network and QoS

- Monitoring
 - Report Qbios compatible
- <http://gforge.inria.fr/projects/qbios/>
F. Moreews, C. Caron et al.
 (cf ReNaBi BioWorkFlow project)

location: <http://www.legoo.org/biomoby/playmoby/cgi/dispatcher.cgi>

LegooBlast	biomoby	OK	2009.02.16-02:00
LegooBlastFromXml	biomoby	OK	2009.02.16-02:00
LegooGetKnowledgeBaseRelationships	biomoby	OK	2009.02.16-02:00
LegooNicknamesSearch	biomoby	OK	2009.02.16-02:00
LegooPrimerSearch	biomoby	OK	2009.02.16-02:02
MensGetExpressionPattern	biomoby	OK	2009.02.16-02:02

location: http://narcisse.toulouse.inra.fr/multi/web_services/playmoby/cgi/dispatcher.cgi

NarcisseBlast	biomoby	OK	2009.02.15-23:35
NarcisseGetAnnotationsFromLocus_tags	biomoby	OK	2009.02.15-23:35
NarcisseGetCDSFromGenesLocus_tags	biomoby	OK	2009.02.15-23:35
NarcisseGetCDSFromLocus_tags	biomoby	OK	2009.02.15-23:35
NarcisseGetCDSFromRegions	biomoby	OK	2009.02.15-23:35
NarcisseGetDownstreamFromLocus_tags	biomoby	OK	2009.02.15-23:35
NarcisseGetGenesFromLocus_tags	biomoby	OK	2009.02.15-23:35
NarcisseGetGenesFromRegions	biomoby	OK	2009.02.15-23:35
NarcisseGetLocus_tags	biomoby	OK	2009.02.15-23:35
NarcisseGetLocus_tagsFromLegooRegions	biomoby	OK	2009.02.15-23:35
NarcisseGetProteinsFromGenesLocus_tags	biomoby	OK	2009.02.15-23:35
NarcisseGetProteinsFromLocus_tags	biomoby	OK	2009.02.15-23:35
NarcisseGetProteinsFromRegions	biomoby	OK	2009.02.15-23:35
NarcisseGetSequencesFromLegooRegions	biomoby	OK	2009.02.15-23:35
NarcisseGetSliceFromRegions	biomoby	OK	2009.02.15-23:35
NarcisseGetTranscriptsFromGenesLocus_tags	biomoby	OK	2009.02.15-23:36
NarcisseGetTranscriptsFromLocus_tags	biomoby	OK	2009.02.15-23:36
NarcisseGetUpstreamFromLocus_tags	biomoby	OK	2009.02.15-23:36
NarcissePsiBlast	biomoby	OK	2009.02.15-23:36

1. MobyNet: the project

2. Moby

- Concepts
- Moby portal
- Web services
- **Moby Workflows**
- Moby vs Galaxy

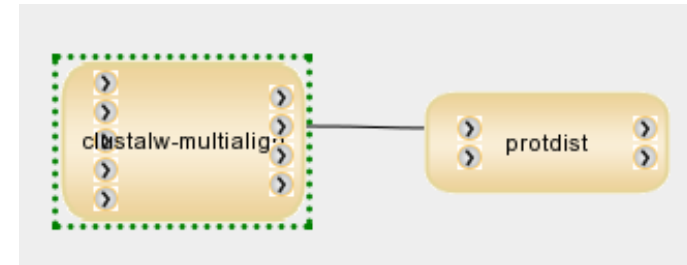
3. MobyNet specific aspects

4. MobyNet today

Nantes, 21 Octobre 2010



BCBBPipe (Coll. NIAID / NIH)



```
<workflow>
  <head>
    <name>toto</name>
    <version>1.1</version>
    <doc>
      <title>test workflow</title>
      <description>this workflow is a test one</description>
    </doc>
  </head>
  <flow>
    <task id="1" service="clustalw-multialign" suspend="False">
      <description>Run a clustalw</description>
    </task>
    <task id="2" service="protdist" suspend="True">
      <description>Run a protdist</description>
    </task>
    <link toTask="1" fromParameter="1" toParameter="infile"/>
    <link toTask="1" fromParameter="2" toParameter="outputformat"/>
    <link fromTask="1" toTask="2" fromParameter="aligfile" toParameter="infile"/>
    <link fromTask="2" fromParameter="outfile" toParameter="3"/>
  </flow>
  <parameters>
    <parameter id="1">
      </parameter>
    </parameters>
</workflow>
```



BCBB Workflow Designer (NIAID / NIH): <http://exon.niaid.nih.gov/mobyWorkflow/>

test

▼ MyWorkflow(s)

new

test_wf rename del


▼ Programs

find reset

All Apps

- All apps
- display
 - abiview
 - cirdna
 - feature_table
 - lindna
 - protein
 - nucleic
 - information
 - sequence
 - phylogeny
 - alignment
 - database
 - search
 - homology
 - blast2
 - fasta
 - phiblast
 - psiblast
 - wublast2
 - display
 - sequence
 - golden
 - assembly
 - hmm

save reset Edge style: Bezier ▼



```
graph LR; golden --> psiblast; psiblast --> abiview;
```

Pause before running this application

submit reset delete

► Header

Browse...

OR

Sequence File (-i)

Start of required region in query (-S)

End of required region in query (-H)

Protein database (-d)

Choose a database ▼

► Scoring option

► Filtering and masking options

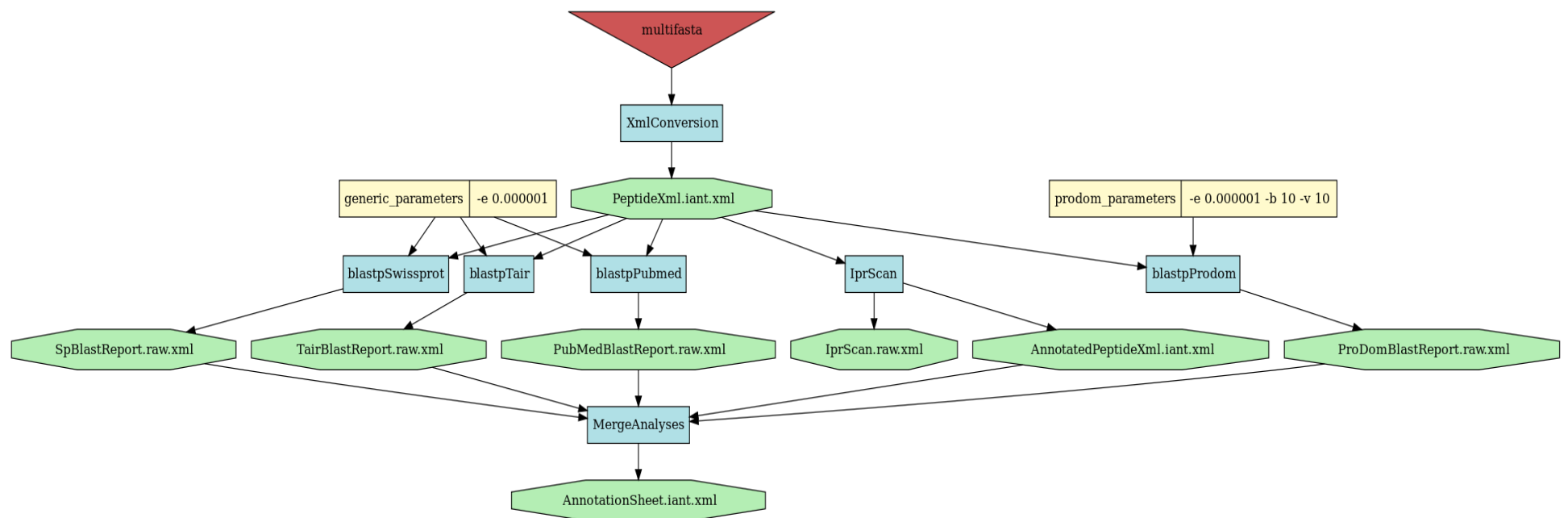
▼ Selectivity options

- Expected value (-e)
- Word Size (-W)
- Multiple hits window size (-A)

Show job status

Charlyse (LIPM)

In situ workflows





Charlyse (LIPM) Workflow meta data management

Logged as: Not connected
 CharLyse

[Workflows](#)
[New account](#)
[Login](#)

NAME	DESCRIPTION	PARAMETERS
HelianthusBacAnalysis	Full BAC analysis. Foreach BAC contigs, predict genes/mRNA/CDS/proteins using Eugene, microsatellites using SputNik and enzyme restriction sites using EMBOSS restrict tool. Foreach protein predicted by Eugene, automatic annotation and similarity searches are done (Against Swiss-Prot, PubMed, TAIR, V.vinifera and HeliaGene Peptides.	<input type="text" value="Contigs"/> <input type="text" value="FASTA_NA_multi"/>
test	Moby Workflow Example, implementing 2 blast (Enzyme and SwissProt) the a merge	<input type="text" value="InputSequence1"/> <input type="text" value="FASTA_AA"/> <input type="text" value="InputSequence2"/> <input type="text" value="FASTA_AA"/>
test	Moby Workflow Example, implementing 2 blast (Enzyme and SwissProt) the a merge	<input type="text" value="InputSequence"/> <input type="text" value="FASTA_AA"/>
UC1: PeptideAnalyses	chargement d'un lot de peptides de plante, annotation fonctionnelle	<input type="text" value="PeptideSequence"/> <input type="text" value="FASTA_AA"/>
UC2: GenomicAnalyses	chargement de sequences genomiques, recherche de sites de restrictions, de microsat et prediction de genes	<input type="text" value="GenomicSequence"/> <input type="text" value="FASTA_NA_multi"/>
UC3: SeqCleanAndAssembly	chargement de lectures, netooyage de vecteur et assemblage	<input type="text" value="RawSequences"/> <input type="text" value="FASTA_NA_multi"/> <input type="text" value="RawSequencesQuality"/> <input type="text" value="FASTA_Base_Quality_multi"/>

HelianthusBacAnalysis

Full BAC analysis. Foreach BAC contigs, predict genes/mRNA/CDS/proteins using Eugene, microsatellites using SputNik and enzyme restriction sites using EMBOSS restrict tool. Foreach protein predicted by Eugene, automatic annotation and similarity searches are done (Against Swiss-Prot, PubMed, TAIR, V.vinifera and HeliaGene Peptides.

Workflow's name:

Workflow's description:

A multifasta with all contigs for a single BAC

Blast parameters

HeliaGene

Author(s): Sebastien.Carrere@toulouse.inra.fr

Charlyse (LIPM) Indexation of analyses (EZLeucine), annotation

Session : XRQ_262F10 (lipm_PV_vRXGLDJ) 20100930

Workflow summary:

NAME	TITLE	DESCRIPTION	
XRQ_262F10	HelianthusBacAnalysis	Full XRQ_262F10 (from filevalidated_contigs.262F10.fasta) BAC analysis. Foreach BAC contigs, predict genes/mRNA/CDS/proteins using Eugene, microsatellites using Sputnik and enzyme restriction sites using EMBOSS restrict tool. Foreach protein predicted by Eugene, automatic annotation and similarity searches are done (Against Swiss-Prot, PubMed, TAIR, V.vinifera and HeliGene Peptides).	
NAME	DESCRIPTION	PARAMETERS	VALUE
Contigs	A multifasta with all contigs for a single BAC		/www-charlyse/site/data/tmp/infiles/of5tu4vMHC_.fasta
GenePrediction	Merge of all contigs Eugene gene prediction with evidences		Eugene.gff3
Graphics	Merge of all contigs Eugene gene prediction graphical display		Eugene.png
Protein	Merge of all contigs Eugene prediction protein sequences		Eugene.prot
mRNA	Merge of all contigs Eugene prediction mRNA sequences		Eugene.mrna
CDS	Merge of all contigs Eugene prediction CDS sequences		Eugene.ods
Gene	Merge of all contigs Eugene prediction gene sequences		Eugene.gene
SputnikPrediction	Merge of all contigs Sputnik micro satellites prediction		Sputnik.iant.xml
RestrictionSites	Merge of all contigs Restriction enzyme cleavage sites report		Restriction.xls

Enlarge



Related files:

Contigs	A multifasta with all contigs for a single BAC	Download
CDS	Merge of all contigs Eugene prediction CDS sequences	Download
Gene	Merge of all contigs Eugene prediction gene sequences	Download
GenePrediction	Merge of all contigs Eugene gene prediction with evidences	Download
mRNA	Merge of all contigs Eugene prediction mRNA sequences	Download
Graphics	Merge of all contigs Eugene gene prediction graphical display	Download
Protein	Merge of all contigs Eugene prediction protein sequences	Download
PeptideXml	Peptide XML File	Download
IprscanReport	InterPro scan output	Download
FinalSheet	Annotated Peptide with embedded analyses	Download
RestrictionSites	Merge of all contigs Restriction enzyme cleavage sites report	Download
SputnikPrediction	Merge of all contigs Sputnik micro satellites prediction	Download
PeptideXml	Peptide XML File	Download

See : All Files

1. *MobyNet: the project*

2. **Moby**

- Concepts
- Moby Core
- Web services
- Moby Workflows
- **Moby vs Galaxy**

3. *MobyNet specific aspects*

4. *MobyNet today*

Nantes, 21 Octobre 2010

Moby vs Galaxy ?

Goecks et al., *Genome Biology* 2010, 11:R86

Table 1
Comparing Galaxy to other genomic workbenches

Galaxy functionality	Description	GenePattern comparison	Moby comparison
Making computation accessible			
Unified, web-based tool interface	All tool interface share same style and use web components; tool interfaces are generated from tool configuration file	Same functions as Galaxy	Same functions as Galaxy
Simple tool integration	Tool developers can integrate tools by writing a tool configuration file and including tool file in Galaxy configuration file	Similar but not as flexible tool configuration file; easy installation of selected tools via a web-based interface	Remote services can be added using a server configuration file
Integrated datasources	Transparent access to established data warehouses	No similar functions	No similar functions
Ensuring reproducibility			
Automatic metadata	Provenance, inputs, parameters, and outputs for each tool used; analysis steps grouped into histories	Same functions as Galaxy	Same functions as Galaxy
User tags	Can apply short tags to histories, datasets, workflows, and pages; tags are searchable and facilitate reuse	No similar functions	No similar functions
User annotations	Can add descriptions or notes to histories, datasets, workflows, workflow steps, and pages to aid in understanding analyses	Cannot annotate a history but can annotate a workflow (pipeline) with an external document	No similar functions
Creating and running workflows	Can create, either by example or from scratch, a workflow that can be repeatedly used to perform a multi-step analysis	Same functions as Galaxy, although editor is form-based rather than graphical	In development
Workflow metadata	Automatic documentation is generated when a workflow is run; users can also tag and annotate workflows and workflow steps	Same functions as Galaxy for generating automatic metadata; cannot annotate workflow steps	In development
Promoting transparency			
Sharing model	Datasets, histories, workflows, and Pages can be shared at progressive levels and published to Galaxy's public repositories; datasets have more advanced sharing options, including groups	Can share analyses and workflows with individuals or groups	No similar functions
Item reuse, display framework and public repositories	Shared or published items displayed as webpages and can be imported and used immediately; public repositories can be searched; archives of analyses and workflows for sharing between servers are under development	Can create an archive of an analysis or workflow and share that with others; author information is included in archive	Can create an archive of an analysis and share that with others
Pages with embedded items	Can create custom webpages with embedded Galaxy items; each page can document a complete experiment, providing all details and supporting reuse of experiment's outputs	Microsoft Word plugin enables users to embed analyses and workflows in Word documents	No similar functions
Coupling between analysis workspace and publication workspace	Can import and immediately start using any shared, published, or embedded item without leaving web browser or Galaxy	Can run embedded analyses and save results in Microsoft Word documents	No similar functions

A summary of Galaxy's functionality and how Galaxy's functionality compares to the functionality of two other genomic workbenches, GenePattern and Moby. Galaxy's novel functionality includes (but is not limited to) integrated datasources, user annotations, a graphical workflow editor, Pages with embedded items, and coupling the workspaces for analysis and publication using an open, web-based model.

Goecks et al. *Genome Biology* 2010 11:R86 doi:10.1186/gb-2010-11-8-r86

[OPEN DATA](#)

Thoughts about using Galaxy and Moby pipelines for NGS

Sophie Créno, Institut Pasteur

- User needs:
 - ➔ Tracability of analyses
 - ✓ Reusability and sharing of pipelines
 - ✓ Easiness to create pipelines
 - ➔ Workspace shared by/accessible to several users
- Requirements due to data:
 - ➔ Do not upload
 - ➔ Do not duplicate
 - ➔ Accessible by several persons

Moby vs Galaxy ?

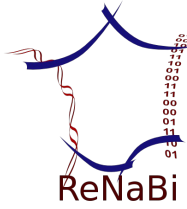
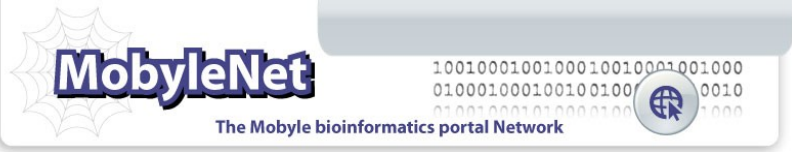
Moby pros:

- Friendly portal
- Diversity of services
- Flexible link towards web-services
- Remote invocation of services

Moby cons:

- No workgroup
- No data sharing
- Weak data annotation/indexing
- Large data not yet supported

1. *MobyNet: the project*
2. *Moby*
3. **MobyNet specific aspects**
4. *MobyNet today*



MobyNet: a network of Moby portals

1. Seamless remote invocation for user
2. Trusted network
3. Security, robustness

MobyNet: user side

Moby@pasteur

hmenager@pasteur.fr (guest)
[set email](#) | [sign-in](#) | [sign-out](#)

Programs

List: by server by category

- ▶ Analysis
- ▶ Annotations
- ▼ Multiple_Sequence_Compare
 - IANTMultalinNucleic@LIP
 - IANTMultalinProteic@LIP
- ▶ Retrieval
- ▼ Sequence
 - ▼ Alignment
 - multiple
 - ProbCons@RPBS
- ▶ SequenceAnalysis
- ▶ Structure
- ▶ alignment
- ▶ assembly
- ▶ database
- ▶ display
- ▶ hmm
- ▶ nucleic
- ▶ phylogeny
- ▶ protein
- ▶ sequence
- ▶ structure

Data Bookmarks

Protein,Nucleic Sequence: refresh
golden.out

Text: infile.data

Jobs

- ✓ golden - 11/05/09 14:03:31 refresh
- ✓ extractfeat - 11/05/09 14:03:52
- ✓ squlzz_checker - 11/05/09 14:33:04

Welcome | **Programs** | Data Bookmarks | Jobs | Tutorials

ProbCons@RPBS x | Clustalw: Multiple alignment x | IANTMultalinNucleic@LIPM x

ProbCons@RPBS

ProbCons: Probabilistic Consistency-based Multiple Alignment of Amino Acid Sequences

Input Data

* Input alignment (Sequence)

Paste | DB | File | Result

Options

Output format: ?

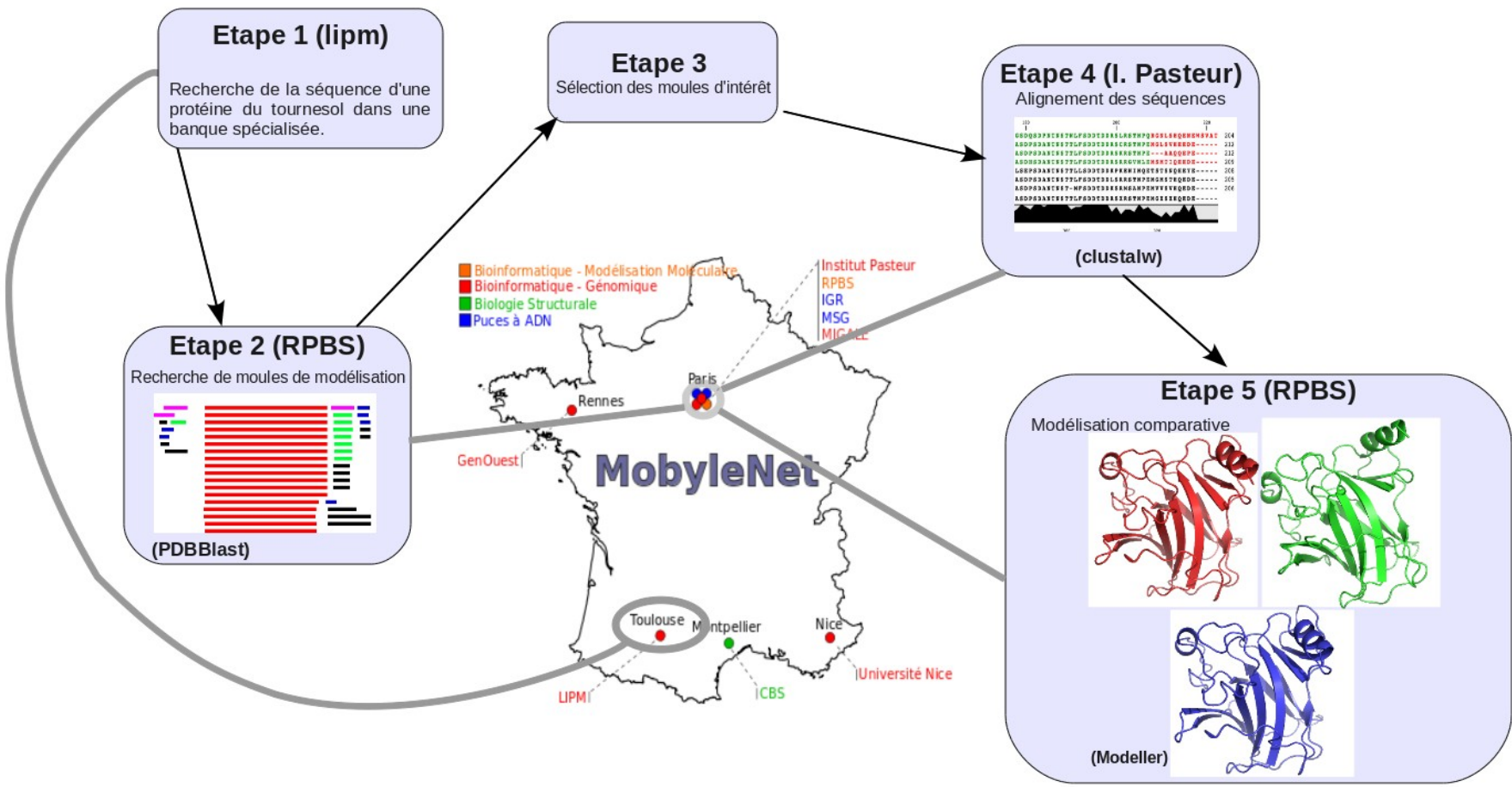
fasta
 clustalw

Consistency: ?

Iterations: ?

Pre-training iterations: ?

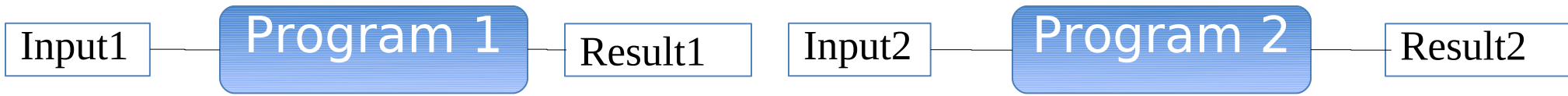
* : mandatory parameter



MobyNet: trusted network

- **Service Import/Export** on demand:
 - Administrator selects “exported” services.
 - Administrator selects “imported” services.
- **Data**
 - User space: stored on the “client” portal
 - Jobs: stored on the execution portal

MobyNet: Datatypes issues



- Type defined by:
 - Biotype: protein, drug, nucleic, ...
 - Data-type: (class, superclass) e.g. sequence
 - Data-Format: e.g. Fasta
- result1 as input2 ?
 - Data-types and Biotypes must be compatible
 - Data-Formats must be compatible (i.e. interconversion is possible)

Dictionary of datatypes over MobyNet nodes.

<http://MobyNet.rpbs.univ-paris-diderot.fr>













(not public yet)

← → ☆ http://mobylenet.rpbs.univ-paris-diderot.fr:8080/mobytypes/bytype

asadin | Sign out

MobyNet
The Moby bioinformatics portal Network

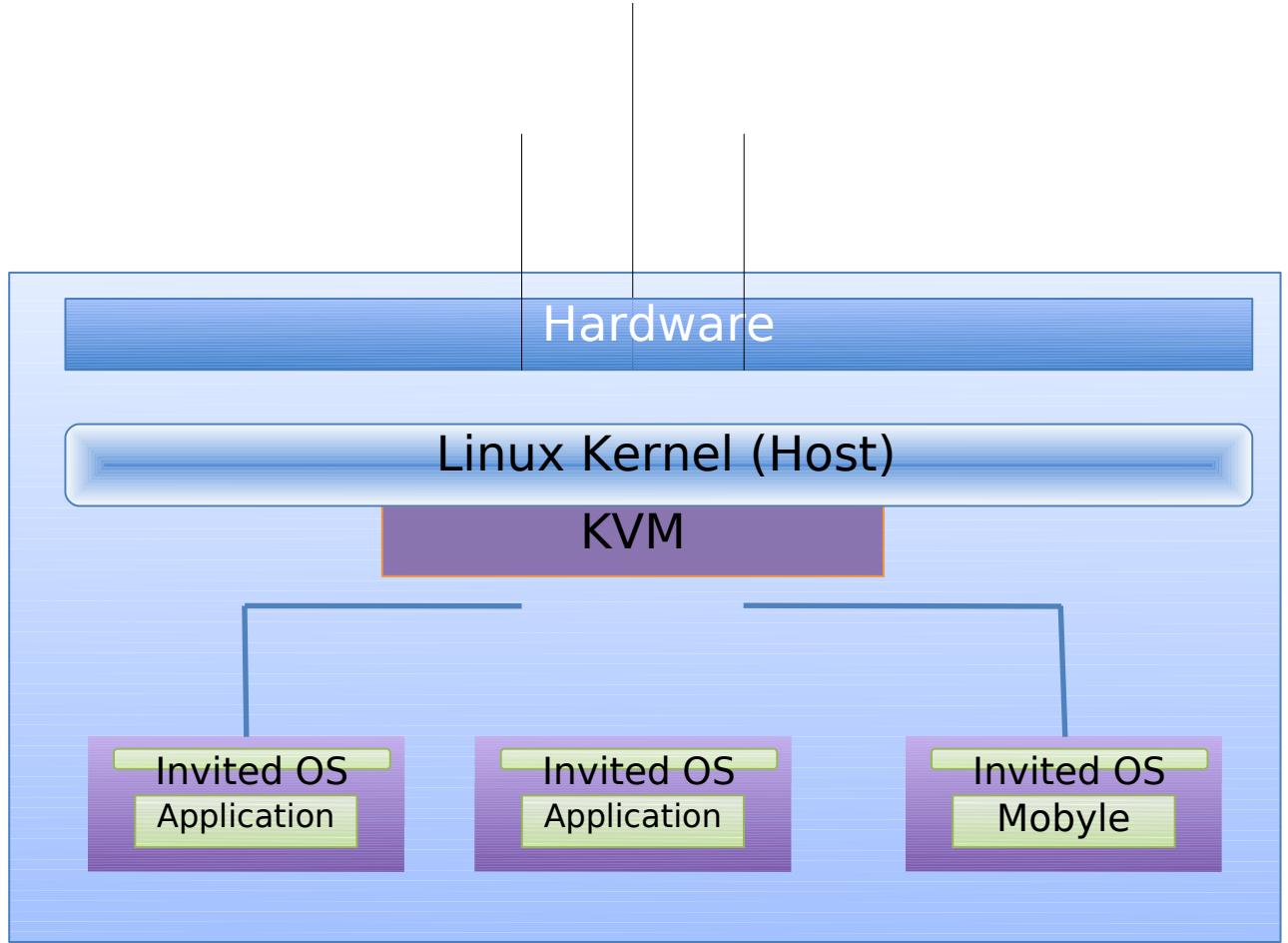
Project News Documentation <Intra>

Biotype	superclass	class	Info	nb_utilisations	used_by
★ Protein	AbstractText	3DAlignment		1	rpbs.Modeller
None	AbstractText	3DAlignment		1	rpbs.ProFitv2_6
★ Protein	AbstractText	3DAlignment		1	rpbs.Ali2Modeller
None	AbstractText	3DSequence		1	rpbs.HMMEncode
None	Report	ACCproReport		2	rpbs.ACCpro rpbs.QMean
None	AbstractText	AVPReport		1	rpbs.avp
Protein	AbstractText	AaindexData		6	genouest.pepwindowall genouest.pepwindow lipm.pepwindow lipm.pepwindowall rpbs.pepwindow rpbs.pepwindowall
None	Binary	AbiTraceFile		2	lipm.abiview rpbs.abiview
None	AbstractText	AceAssembly		2	genouest.cap3 lipm.cap3
None	None	AceAssembly		2	lipm.Cap3 lipm.Cap3FromXml
None	AbstractText	AceAssembly		1	lipm.IANTautoSNP
None	None	Alignment		69	genouest.hmmsearch genouest.clustalw-sequence genouest.dialign genouest.hmmalign

Mobylenet: Easy deployment through KVM virtualization

Advantages:

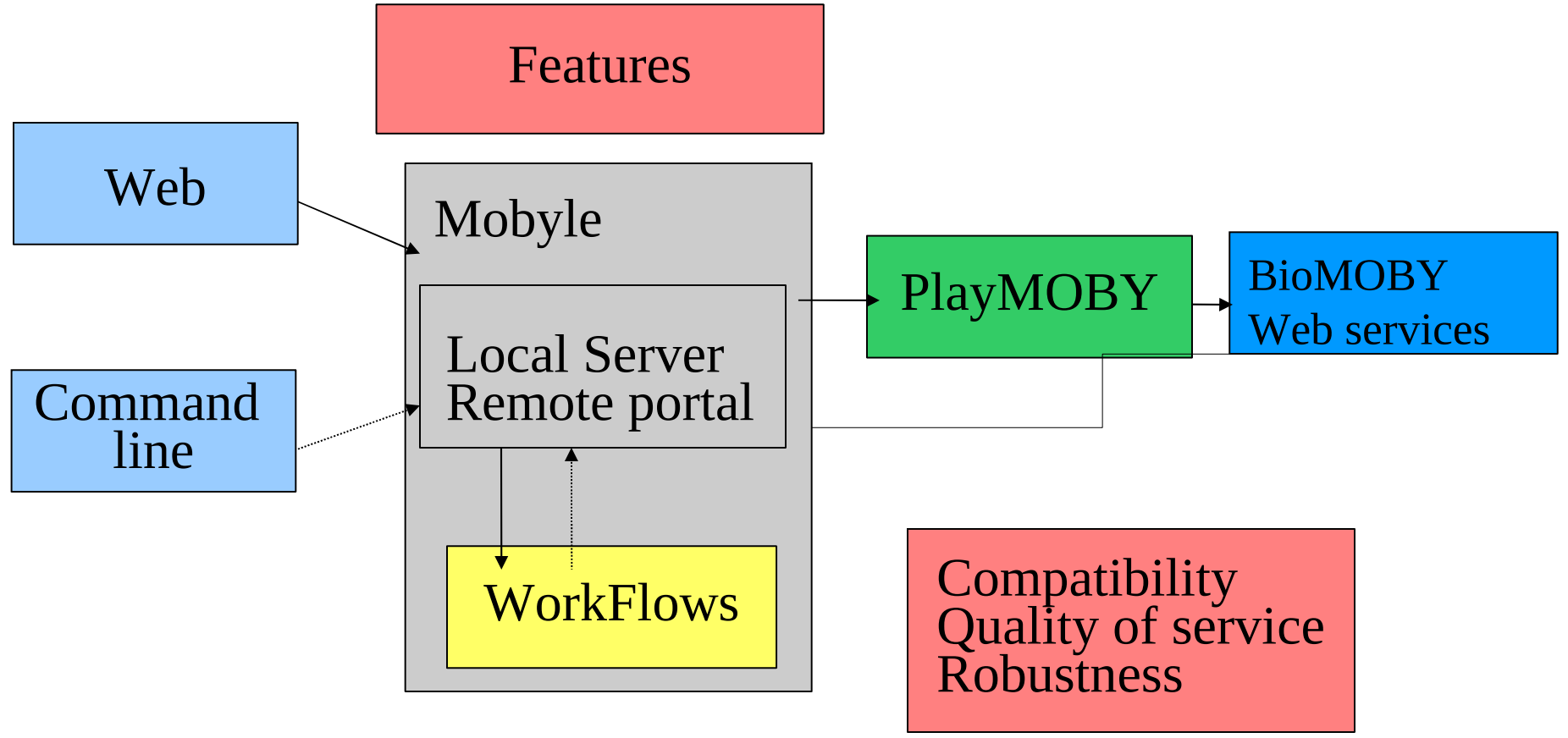
- Flexible administration
- Make testing easy
- **Better security**
 - Service isolation
 - Host isolation
- Easier administration at the level of the network (supervisor)
- **High availability of services**



ISO images to appear on mobylenet.rpbs.univ-paris-diderot.fr

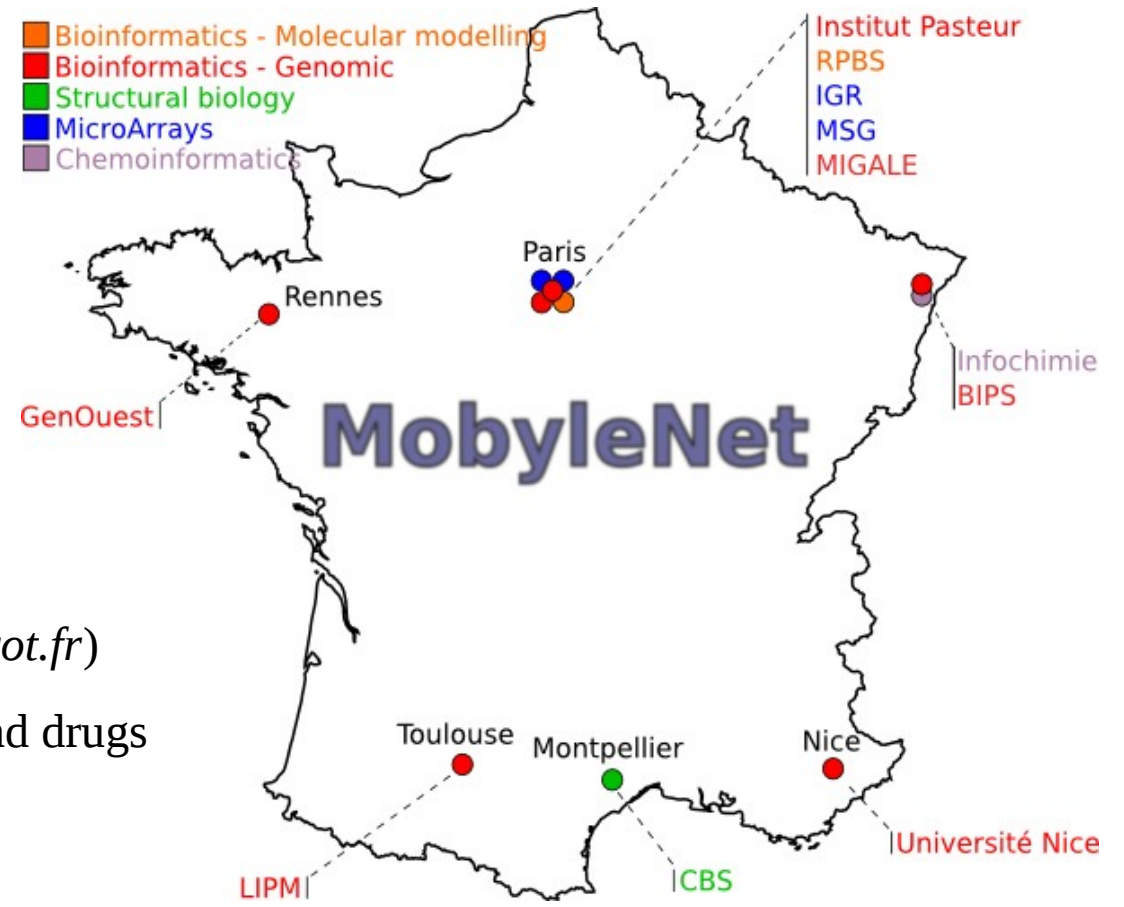
1. *MobyNet: the project*
2. *Moby*
3. *MobyNet specific aspects*
4. **MobyNet today**

“Complex” project

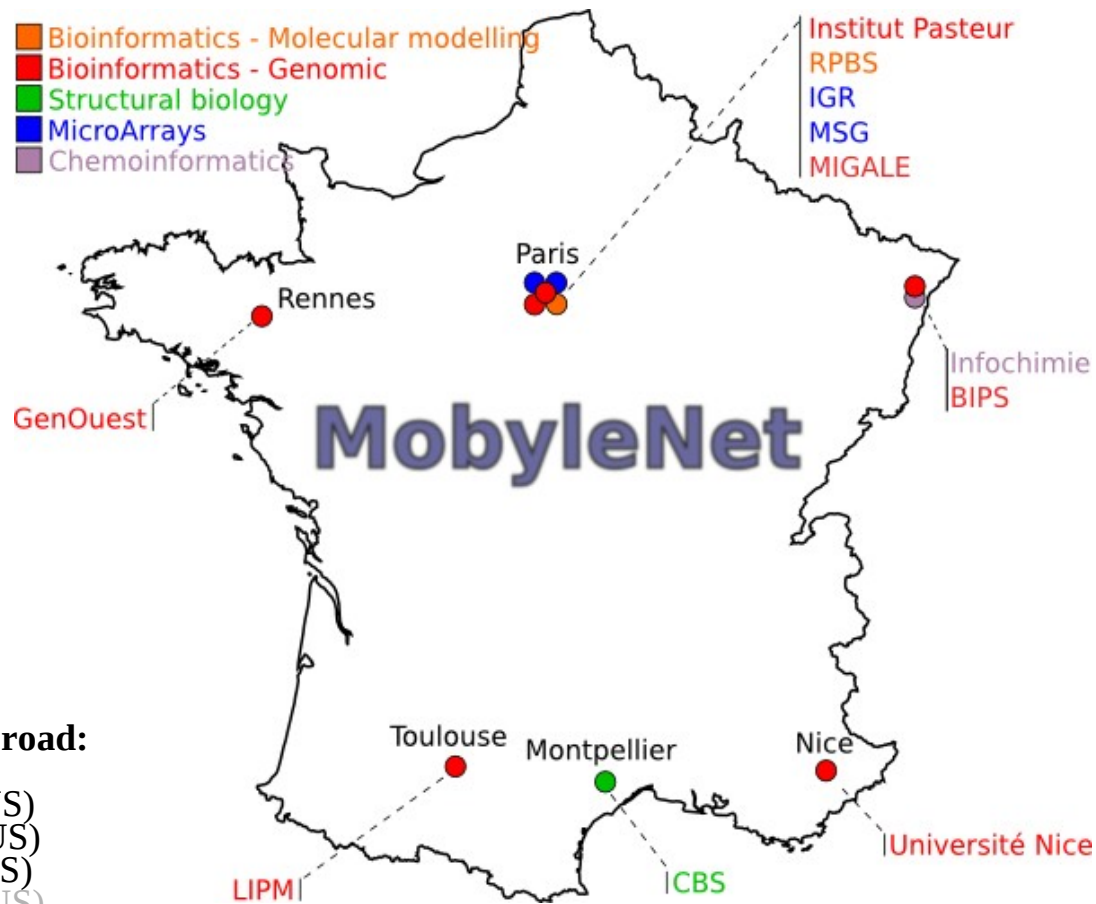


MobyNet' 2010 summary

- ✓ Moby deployed on 8 sites
in production on 4
- ✓ Interoperability is effective –
inter-site pipelining possible
(<http://MobyNet.rpbs.univ-paris-diderot.fr>)
- ✓ Service spectrum from genomics to 3D and drugs



Presently involves 11 nodes nodes



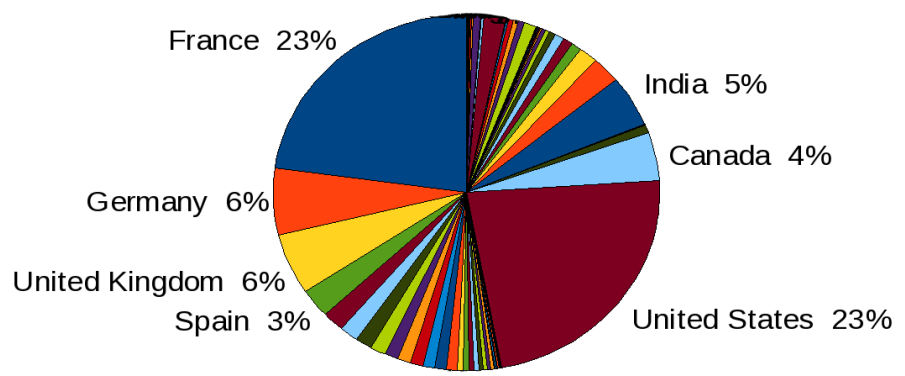
(Known) Moby servers abroad:

- NIH/NIAID (Washington, US)
- NIH/HELIX (Washington, US)
- NCSU (Caroline du Nord, US)
- PACIFIC BIOSCIENCES (US)
- CGIAR/CIP (Perou)
- UNAM (Mexique, Mexico)
- UNINA (Naples, Italy)
- ICHEC (Irlande)
- USAL (Salamanque, Spain)
- AG-RESEARCH (New Zealand)

....

Mobyle / MobyleNet

- Several hundreds on-line services (Emboss, Phylip, 3D modeling, small compounds, ...)
- Since 2008, over 1 000 000 jobs by over 40 000 users worldwide



Moby / MobyNet: What next ?

- Moby 1.0 in test
 - release schedule: before end 2010
 - Workflows, viewers, command line invocation

- Large scale application test: early 2011

- Moby 2.0 (Workgroup, annotation, ...) requires core rewriting: end 2011

Thanks !

C. Letondal
H. Ménager
B. Néron
N. Joly
S. Larroudé
B. Caudron

S. Letort
S. Carrere
J. Gouzy

O. Sallou

J. Maupetit
P. Tufféry

A. Saladin

Thanks !

C. Letondal
H. Ménager
B. Néron
N. Joly
S. Larroude
B. Caudron

S. Letort
S. Carrere
J. Gouzy

O. Sallou

J. Maupetit
P. Tufféry

A. Saladin

V. Gopalan
Y. Huyen

GenOuest: O. Colin, O. Sallou
CBS: G. Labesse ,JL Pons
Nice: R. Christen, D. Douguet
ENS: S. LeCrom, L. Jourdren
IGR: Ph. Dessen, H. Ripoché
Migale: F. Samson
InfoChimie: A. Varnek



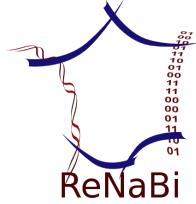
Ceci n'est pas une plate-forme (?)



1001000100100010010001001000
010001000100100100 0010
010010001010000100 1000



The Moby bioinformatics portal Network





1001000100100010010001001000
010001000100100100 0010
010010001010000100 1000



The Moby bioinformatics portal Network

