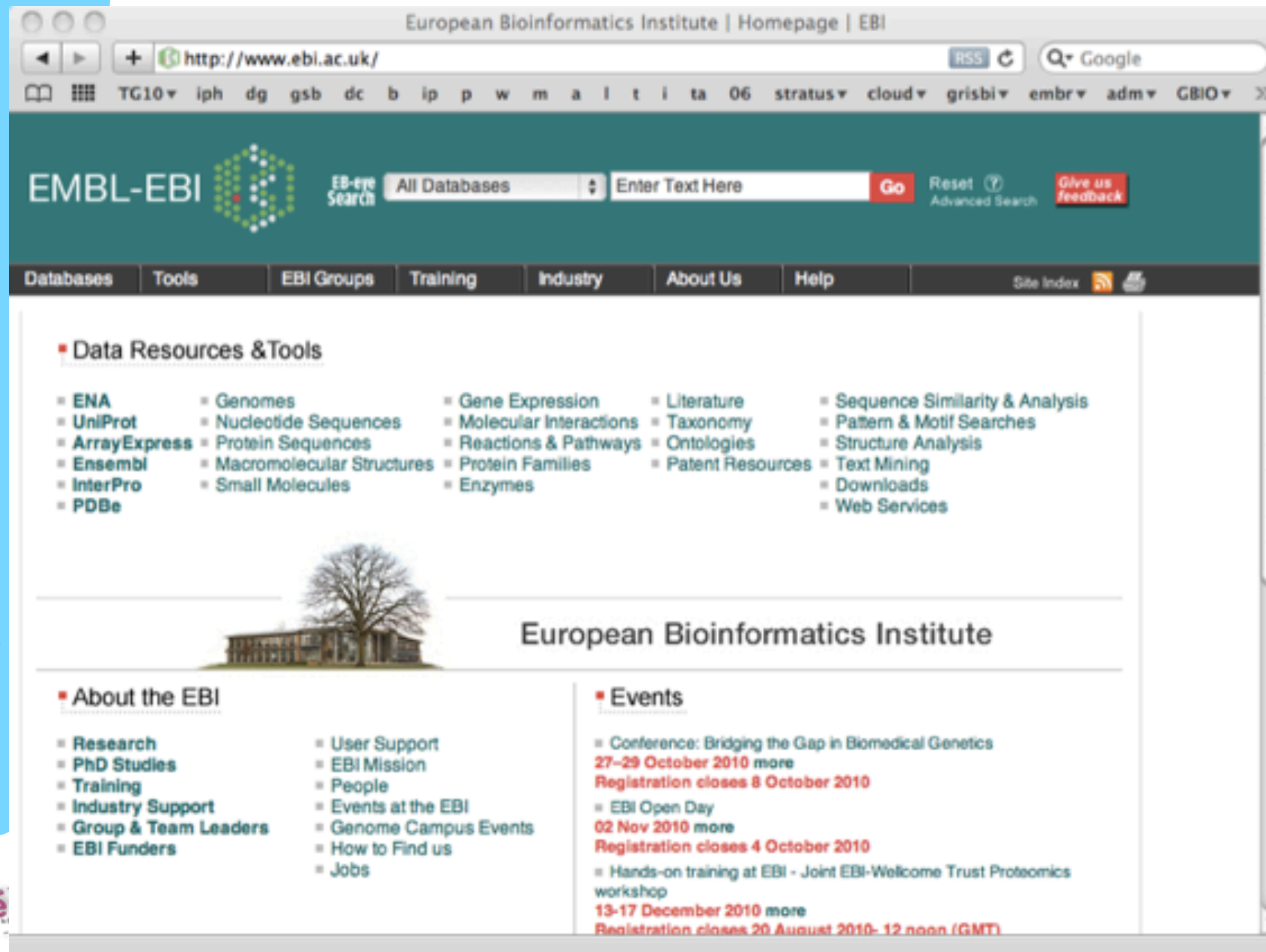


Bioinformatics applications

StratusLab Tutorial (Bordeaux, France)
25-26 October 2012

Infrastructure in Biology



Infrastructure in Biology

The screenshot displays the Galaxy web interface in a browser window. The address bar shows `http://idb-cloud.ibcp.fr:20007/`. The interface includes a top navigation bar with links for **Analyze Data**, **Workflow**, **Shared Data**, **Help**, and **User**. On the left, a sidebar lists tool categories under 'Tools' and 'Databases', including **EMBL**, **ENA**, **UniProt**, **ArrayExpress**, **Ensembl**, **InterPro**, and **PDB**. The main content area features a green status message: 'Hello world! It's running...' and a workflow diagram titled 'WWFSMD? grow noodly appendages...' with the URL `usegalaxy.org`. The workflow diagram shows a sequence of steps: 'Input dataset', 'Filter', 'Join', 'Sort', 'Select first', 'Join two Queries', and 'Group'. A 'History' panel on the right indicates '0 bytes' and provides instructions to click 'Get Data' to start. The footer contains information about project support from NSE, NHGRI, and the Huck Institutes of the Life Sciences, along with dates for EBI Open Day and a hands-on training workshop.

Galaxy

http://idb-cloud.ibcp.fr:20007/

Galaxy

Analyze Data Workflow Shared Data Help User

Tools Options

search tools

Get Data
Send Data
ENCODE Tools
Lift-Over
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Fetch Alignments
Get Genomic Scores
Operate on Genomic Intervals
Statistics
Graph/Display Data
Regional Variation
Multiple regression
Multivariate Analysis
Evolution
Motif Tools
Multiple Alignments
Metagenomic analyses
FASTA manipulation

EMBL-
Databases

Data R
ENA
UniProt
ArrayExp
Ensembl
InterPro
PDB

About t
Research
PhD Stud
Training
Industry Support
Group & Team Leaders
EBI Funders

Events at the EBI
Genome Campus Events
How to Find us
Jobs

EBI Open Day
02 Nov 2010 more
Registration closes 4 October 2010
Hands-on training at EBI - Joint EBI-Wellcome Trust Proteomics workshop
13-17 December 2010 more
Registration closes 20 August 2010- 12 noon (GMT)

History Options

0 bytes

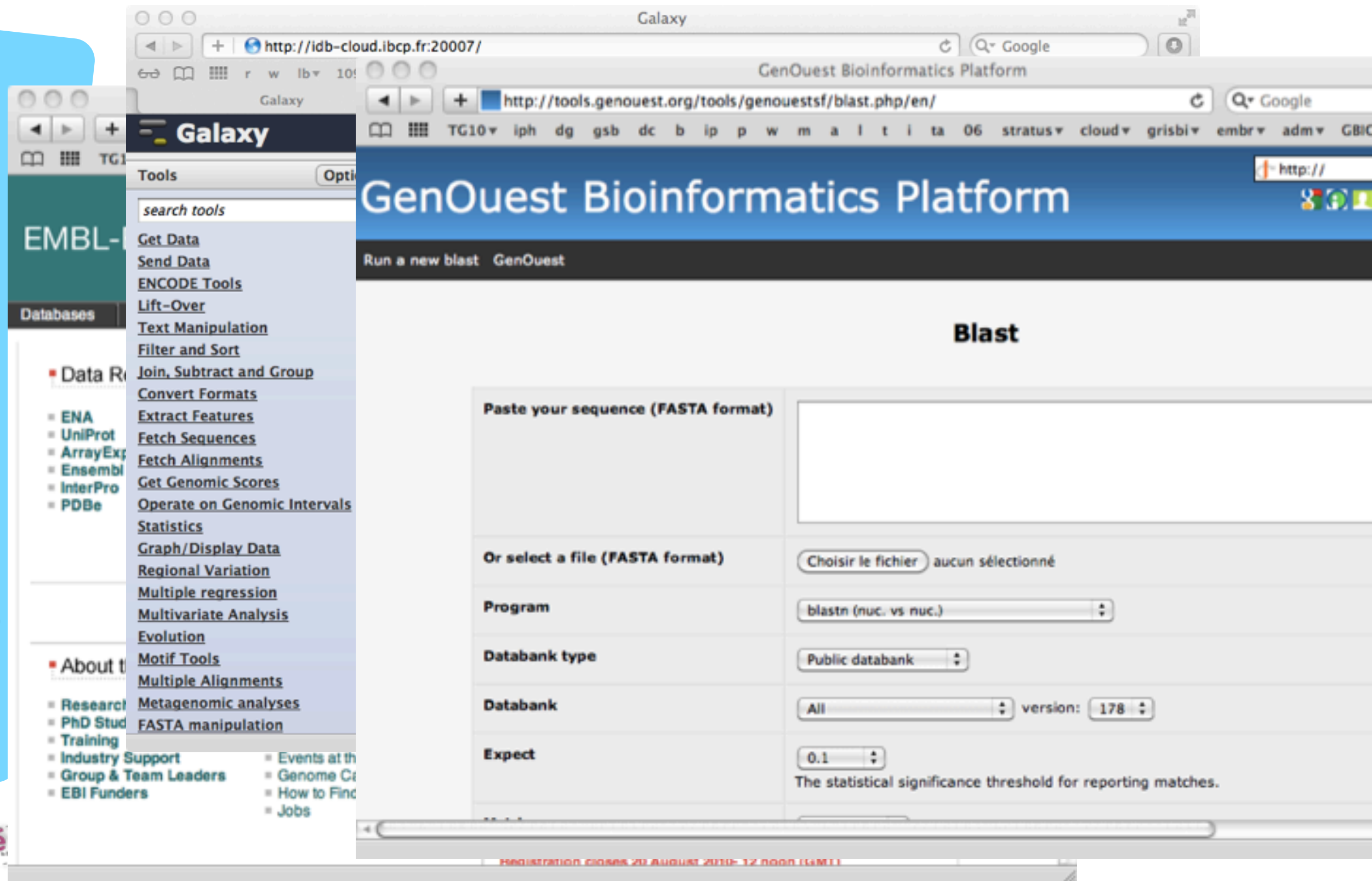
Your history is empty. Click 'Get Data' on the left pane to start

WWFSMD?
grow noodly appendages...

usegalaxy.org

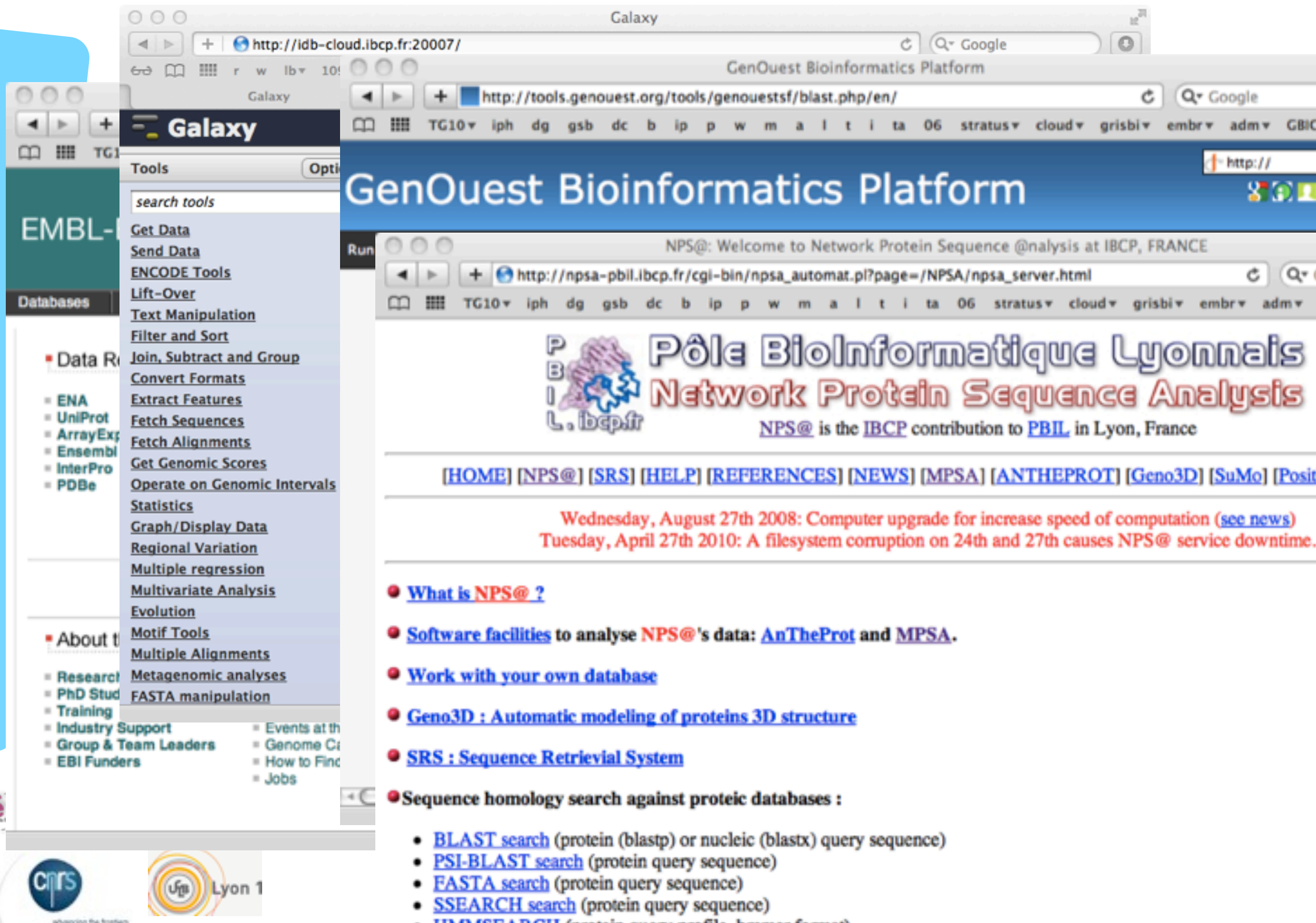
This project is supported in part by [NSE](#), [NHGRI](#), and [the Huck Institutes of the Life Sciences](#).

Infrastructure in Biology



The image displays two overlapping web browser windows. The background window is the GenOuest Bioinformatics Platform, showing a 'Blast' interface with fields for 'Paste your sequence (FASTA format)', 'Or select a file (FASTA format)', 'Program' (set to 'blastn (nuc. vs nuc.)'), 'Databank type' (set to 'Public databank'), 'Databank' (set to 'All'), and 'Expect' (set to '0.1'). The foreground window is the Galaxy platform, showing a 'Tools' menu with various bioinformatics tools like 'Get Data', 'Send Data', 'ENCODE Tools', 'Lift-Over', 'Text Manipulation', 'Filter and Sort', 'Join, Subtract and Group', 'Convert Formats', 'Extract Features', 'Fetch Sequences', 'Fetch Alignments', 'Get Genomic Scores', 'Operate on Genomic Intervals', 'Statistics', 'Graph/Display Data', 'Regional Variation', 'Multiple regression', 'Multivariate Analysis', 'Evolution', 'Motif Tools', 'Multiple Alignments', 'Metagenomic analyses', and 'FASTA manipulation'. The Galaxy window also shows a 'Databases' section with links to ENA, UniProt, ArrayExpress, Ensembl, InterPro, and PDB. The GenOuest window has a top navigation bar with links to various tools and a search bar. The Galaxy window has a top navigation bar with links to various tools and a search bar.

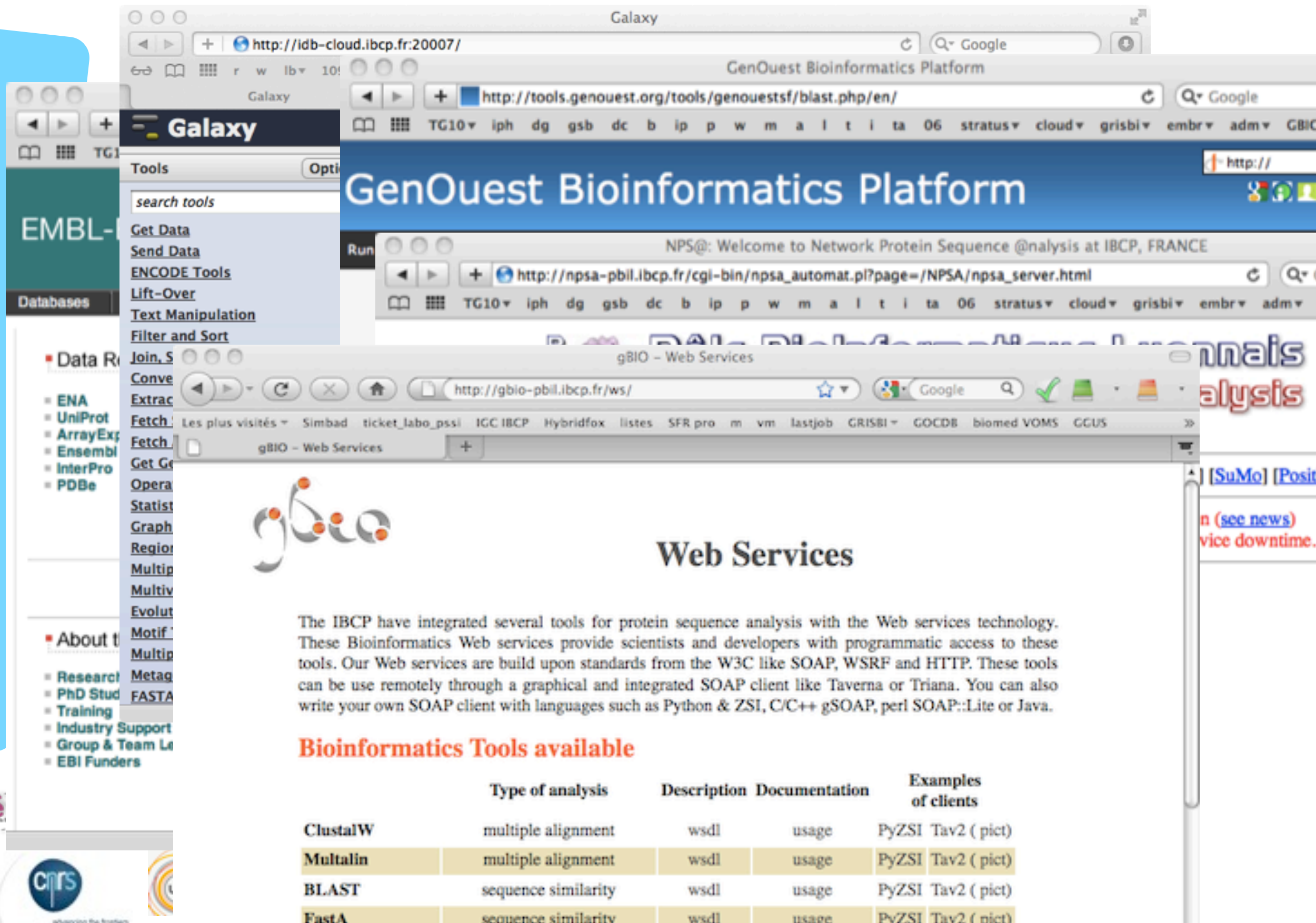
Infrastructure in Biology



The image displays three overlapping web browser windows illustrating biological infrastructure:

- Galaxy (Left):** A web interface for bioinformatics tools. The "Tools" dropdown menu is open, showing categories like "Get Data", "Send Data", "ENCODE Tools", "Lift-Over", "Text Manipulation", "Filter and Sort", "Join, Subtract and Group", "Convert Formats", "Extract Features", "Fetch Sequences", "Fetch Alignments", "Get Genomic Scores", "Operate on Genomic Intervals", "Statistics", "Graph/Display Data", "Regional Variation", "Multiple regression", "Multivariate Analysis", "Evolution", "Motif Tools", "Multiple Alignments", "Metagenomic analyses", and "FASTA manipulation".
- GenOuest Bioinformatics Platform (Top Right):** A web interface for the GenOuest Bioinformatics Platform, showing a search bar and a list of tools.
- NPS@ Network Protein Sequence Analysis (Bottom Right):** A website for the Pôle BioInformatique Lyonnais (PBL) Network Protein Sequence Analysis (NPS@). It includes a welcome message, a list of links (HOME, NPS@, SRS, HELP, REFERENCES, NEWS, MPSA, ANTHEROT, Geno3D, SuMo, Posi), and a list of services:
 - What is NPS@ ?
 - Software facilities to analyse NPS@'s data: AnTheProt and MPSA.
 - Work with your own database
 - Geno3D : Automatic modeling of proteins 3D structure
 - SRS : Sequence Retrieval System
 - Sequence homology search against proteic databases :
 - BLAST search (protein (blastp) or nucleic (blastx) query sequence)
 - PSI-BLAST search (protein query sequence)
 - FASTA search (protein query sequence)
 - SSEARCH search (protein query sequence)
 - UHMSEARCH (protein query sequence)

Infrastructure in Biology



The image displays a collage of web browser windows showcasing various bioinformatics infrastructure platforms. The windows are layered, with the 'gBIO - Web Services' window being the most prominent in the foreground.

Galaxy (EMBL-EBI): A web-based platform for genomic data analysis. The visible URL is <http://idb-cloud.ibcp.fr:20007/>. It features a 'Tools' section with categories like 'Get Data', 'Send Data', 'ENCODE Tools', 'Lift-Over', 'Text Manipulation', and 'Filter and Sort'.

GenOuest Bioinformatics Platform: A platform for genomic data analysis. The visible URL is <http://tools.genouest.org/tools/genouestsf/blast.php/en/>. It includes a search bar and a list of tools.

NPS@: Welcome to Network Protein Sequence @analysis at IBCP, FRANCE: A web service for protein sequence analysis. The visible URL is http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_server.html.

gBIO - Web Services: A web service for protein sequence analysis. The visible URL is <http://gbio-pbil.ibcp.fr/ws/>. The page features the gBIO logo and a section titled 'Web Services'.

Web Services (gBIO - Web Services): The main content of the gBIO window, describing the integration of tools for protein sequence analysis with web services technology. It mentions that these services provide programmatic access to tools like ClustalW, Multalin, BLAST, and FastA.

Bioinformatics Tools available (gBIO - Web Services): A table listing available tools, their types of analysis, descriptions, documentation, and examples of clients.

	Type of analysis	Description	Documentation	Examples of clients
ClustalW	multiple alignment	wsdl	usage	PyZSI Tav2 (pict)
Multalin	multiple alignment	wsdl	usage	PyZSI Tav2 (pict)
BLAST	sequence similarity	wsdl	usage	PyZSI Tav2 (pict)
FastA	sequence similarity	wsdl	usage	PyZSI Tav2 (pict)

Other visible elements:

- EMBL-EBI**: Logo and text visible in the Galaxy window.
- IBCP**: Logo and text visible in the bottom left corner.
- CPRS**: Logo visible in the bottom left corner.
- idé**: Logo visible in the bottom left corner.
- annals analysis**: Text visible in the bottom right corner.
- [SuMo] [Posi]**: Text visible in the bottom right corner.
- n (see news) vice downtime.**: Text visible in the bottom right corner.

Infrastructure in Biology

The image is a collage of several web browser windows displaying different bioinformatics platforms. The windows include:

- Galaxy**: A web interface for Galaxy, showing a search bar and a list of tools like 'Get Data', 'Send Data', 'ENCODE Tools', 'Lift-Over', 'Text Manipulation', and 'Filter and Sort'.
- GenOuest Bioinformatics Platform**: A web interface for GenOuest, showing a search bar and a list of tools like 'TG10', 'iph', 'dg', 'gsb', 'dc', 'b', 'ip', 'p', 'w', 'm', 'a', 'l', 't', 'i', 'ta', '06', 'stratus', 'cloud', 'grisbi', 'embr', 'adm', 'GBIO'.
- NPS@: Welcome to Network Protein Sequence @analysis at IBCP, FRANCE**: A web interface for NPS@, showing a search bar and a list of tools like 'TG10', 'iph', 'dg', 'gsb', 'dc', 'b', 'ip', 'p', 'w', 'm', 'a', 'l', 't', 'i', 'ta', '06', 'stratus', 'cloud', 'grisbi', 'embr', 'adm'.
- gBIO - Web Services**: A web interface for gBIO, showing a search bar and a list of tools like 'Les plus visités', 'Simbad', 'ticket_labo_pssi', 'IGC IBCP', 'Hybridfox', 'listes', 'SFR pro', 'm', 'vm', 'lastjob', 'GRISBI', 'GOCDB', 'biomed VOMS', 'GGUS'.

At the bottom of the collage, there is a table with the following columns: Type of analysis, Description, Documentation, and Examples of clients.

Type of analysis	Description	Documentation	Examples of clients	
ClustalW	multiple alignment	wsdl	usage	PyZSI Tav2 (pict)
Multalin	multiple alignment	wsdl	usage	PyZSI Tav2 (pict)
BLAST	sequence similarity	wsdl	usage	PyZSI Tav2 (pict)
FastA	sequence similarity	wsdl	usage	PyZSI Tav2 (pict)

On the left side of the collage, there is a blue vertical bar with a white line drawing of a plant. At the bottom left, there is a logo for 'idé' and a logo for 'IBCP'. At the bottom right, there is a logo for 'CIRIS'.

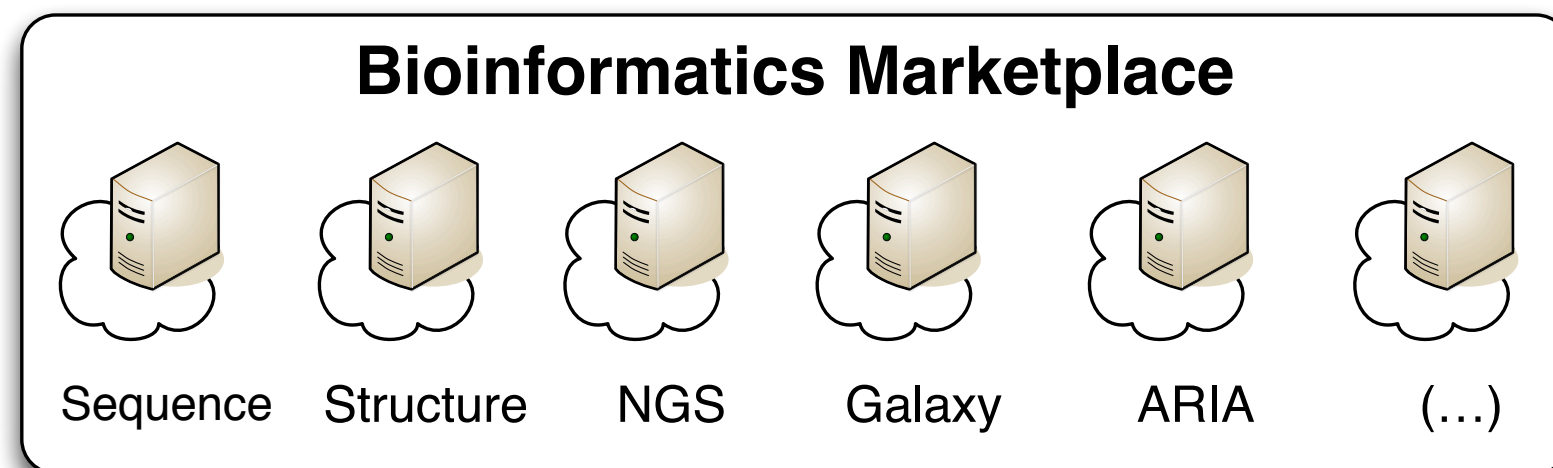
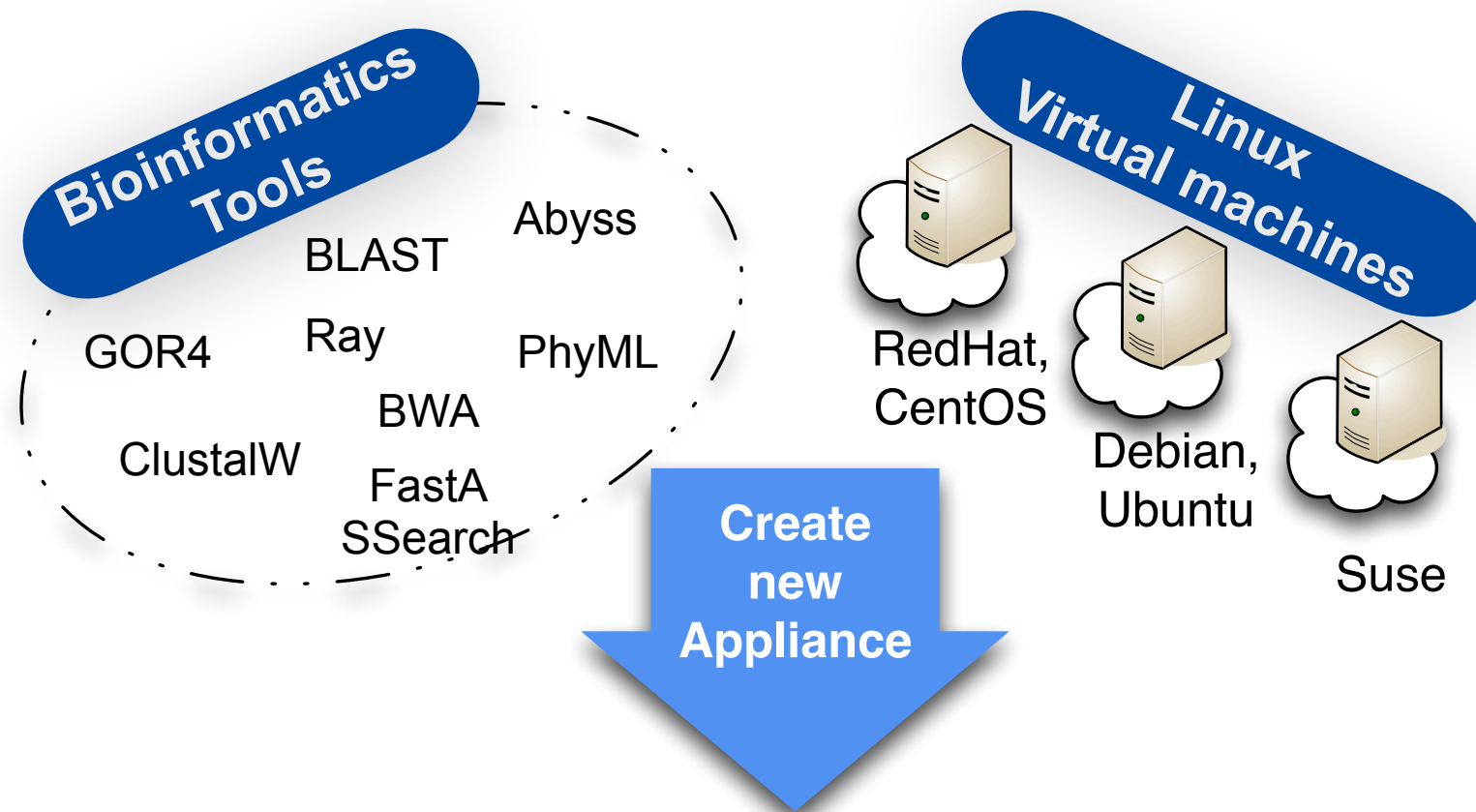
Lot of **tools**
and **web services**
to treat and vizualize lot of **data**

Constraints

- Core facility providers
 - How to deploy easily lot of (incompatible) tools ?
 - To make them connected to public databases ?
 - To limit transfer of huge data ?
 - To provide users with their own computing resources ?
 - With their own isolated storage ?
- Scientific users
 - How easy is it to access/use these tools ?
 - To adapt to your usage ?
 - To get your/other tools deployed on a datacenter ?
 - To combine them ?



Integrate Bioinformatics Tools in Cloud



Appliances for Bioinformatics

StratusLab Marketplace

Metadata

View: [All](#) | [With deprecated](#) | [Deprecated only](#)

Sort by:
Show 10 entries

Search:

biodata v5.7 x86_64

Endorser: christophe.blanchet@ibcp.fr

Identifier: [FtCJFZ7xO5uxKyzThGRX9Ex5cqR](#)

Created: 2012-03-01T12:20:37Z

Biological databases repository appliance built by IDB-IBCP (CNRS, Lyon, France. <http://idee-b.ibcp.fr>). The following databases are installed and available: SwissProt, PROSITE. The appliance can be used to set up your own cloud biological databases repository, and you can add the databases you want with the help of the BioMaj framework. To configure you can use the BioMaj web interface or the command line. To log in with ssh, connect to your VM as root with your SSH authentication key (see general StratusLab documentation for details). To connect to the BioMaj web interface, go to the VM-related URL with a standard web

Bioinformatics Marketplace



Sequence



Structure



NGS



Galaxy



ARIA

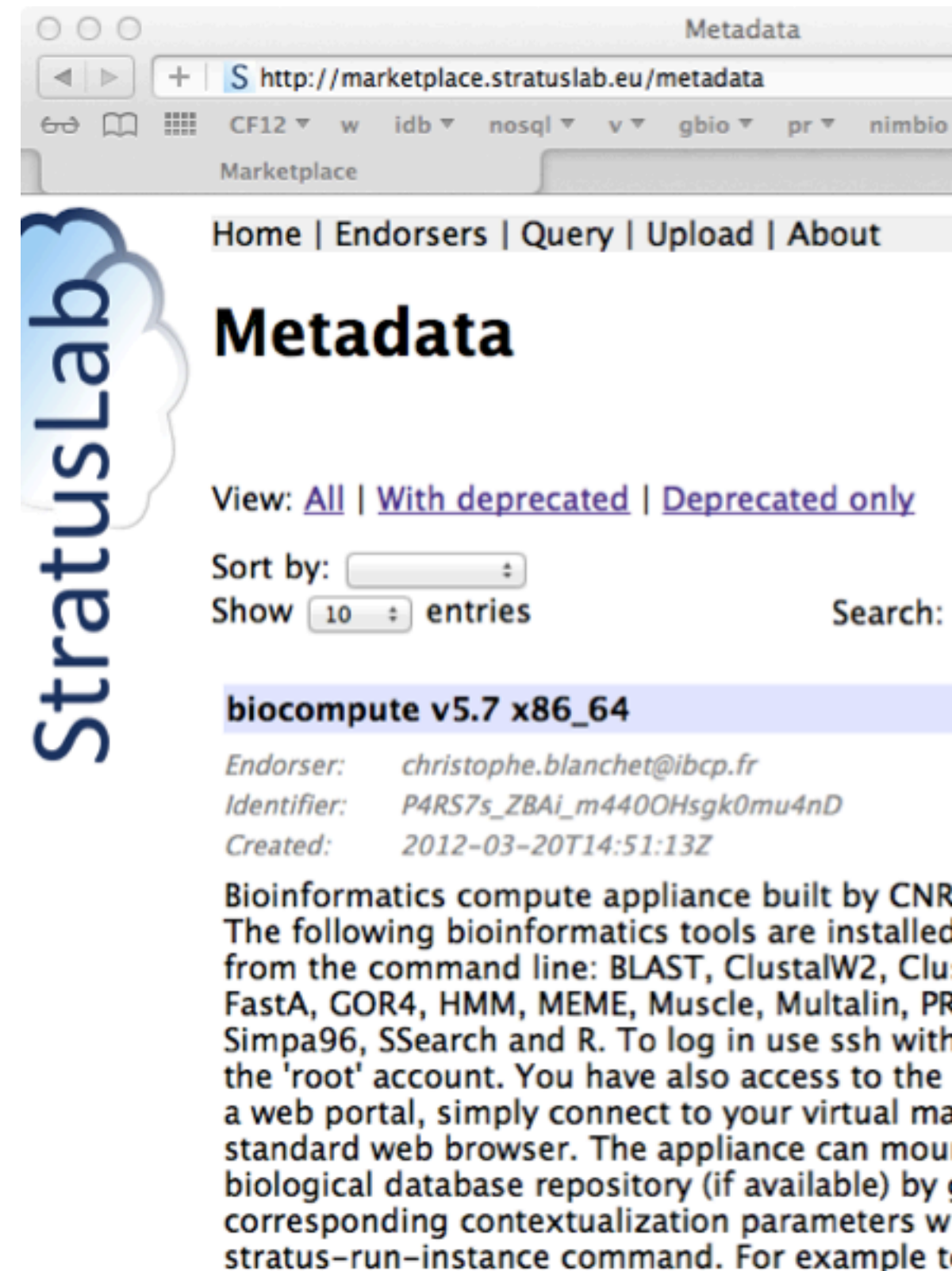


(...)

- Predefined virtual machines
 - small: few GB
 - easy to convert in most formats
- Installed and pre-configured bioinformatics tools
- Several ones already available
 - biodata, biocompute, galaxy, etc.
- with several tools
 - BLAST, Clustalw, ARIA, MEME, HMMer, Abyss, BWA, Ray, etc.

Example of an Appliance

- **‘biocompute’ appliance**
 - pre-installed with bioinformatics tools:
 - BLAST, FastA, SSearch, HMM,
 - ClustalW2, Clustal-Omega, Muscle, Multalin,
 - GOR4, PREDATOR, Simpa96,
 - MEME, R.
 - Deploy **your own** instances
 - **SSH** as root
 - **HTTP** with a standard web browser.
 - Connected to **local public** biological databases
 - Automatically parametrized
 - For example on the IBCP cloud
 - BIO_DB_SERVER=idb-databases.ibcp.fr



The screenshot shows a web browser window displaying the 'Metadata' page of the StratusLab marketplace. The browser's address bar shows the URL 'http://marketplace.stratuslab.eu/metadata'. The page has a navigation bar with links: 'Home | Endorsers | Query | Upload | About'. The main heading is 'Metadata'. Below this, there are filters for 'View' (All, With deprecated, Deprecated only) and 'Sort by'. A 'Show' dropdown is set to '10' entries. A search bar is on the right. The main content area features a highlighted entry for 'biocompute v5.7 x86_64'. The entry details include: 'Endorser: christophe.blanchet@ibcp.fr', 'Identifier: P4RS7s_ZBAi_m440OHsgk0mu4nD', and 'Created: 2012-03-20T14:51:13Z'. A descriptive paragraph follows, stating it is a bioinformatics compute appliance built by CNRS, listing installed tools (BLAST, ClustalW2, Clustal-Omega, FastA, GOR4, HMM, MEME, Muscle, Multalin, PREDATOR, Simpa96, SSearch, and R), and providing instructions for login via SSH and access via a web portal.

StratusLab

Metadata

Home | Endorsers | Query | Upload | About

View: [All](#) | [With deprecated](#) | [Deprecated only](#)

Sort by:

Show entries

Search:

biocompute v5.7 x86_64

Endorser: christophe.blanchet@ibcp.fr

Identifier: P4RS7s_ZBAi_m440OHsgk0mu4nD

Created: 2012-03-20T14:51:13Z

Bioinformatics compute appliance built by CNRS. The following bioinformatics tools are installed from the command line: BLAST, ClustalW2, Clustal-Omega, FastA, GOR4, HMM, MEME, Muscle, Multalin, PREDATOR, Simpa96, SSearch and R. To log in use ssh with the 'root' account. You have also access to the appliance via a web portal, simply connect to your virtual machine with a standard web browser. The appliance can mount a local biological database repository (if available) by using the corresponding contextualization parameters with the stratus-run-instance command. For example to

Bioinformatics Appliances

Protein identification

Endorser: christophe.blanchet@ibcp.fr
Identifier: H6KPqxYIZRdlhPhs2ZKIENiiVyx
Created: 2012-10-23T14:05:07Z

Bioinformatics virtual appliance for protein identification from mass spectrometry data. Contains OMSSA, X!Tandem, PeptideShaker and SearchGUI tools. Constructed by IDB...

[More...](#)

Galaxy portal

Endorser: christophe.blanchet@ibcp.fr
Identifier: OqucGN3bQD9FdlengRIqZ4ZNNHW
Created: 2012-10-11T15:11:59Z

Bioinformatics gateway appliance configured with the GALAXY portal, built by CNRS IBCP-IDB. You have also access to the pre-installed bioinformatics tools through the web...

[More...](#)

Hadoop MapReduce

Endorser: clement.gauthey@ibcp.fr
Identifier: PEIfkAp5mOwULVh1KLsprFcji0s
Created: 2012-10-11T14:42:36Z

This appliance provides an easy way to deploy an Hadoop MapReduce cluster. You just need to run the bash script `hadoop-create-cluster` with a nodes list and an username in parameters and wait few...

[More...](#)

BIO compute node

Endorser: christophe.blanchet@ibcp.fr



Bioinformatics Appliances (...)

ARIA 2.3

Endorser: *christophe.blanchet@ibcp.fr*
Identifier: *N_zDsconV86gvkjZtt7D-ePv4M6*
Created: *2012-10-11T14:01:38Z*

This appliance is part of the StratusLab bioinformatics usecase TOSCANI (TOwards StruCTural Assignment Improvement). The goal is to improve the determination of protein...

[More...](#)

CentOS 6

Endorser: *christophe.blanchet@ibcp.fr*
Identifier: *BI8HibMjBB6uu231adQUkqyGtnI*
Created: *2012-09-27T12:04:15Z*

A minimal installation for CentOS 6.x. Only root account configured. Firewall enabled with SSH and HTTP port open. SELinux disabled. Enhanced StratusLab contextualization used...

[More...](#)

BIO data

Endorser: *christophe.blanchet@ibcp.fr*
Identifier: *FtCJFZ7xO5uxKyzThGRX9Ex5cqR*
Created: *2012-09-24T08:56:26Z*



Biological databases repository appliance built by IDB-IBCP (CNRS, Lyon, France. <http://idee-b.ibcp.fr>). The following databases are installed and available: SwissProt,...

[More...](#)

Mobyle

Endorser: *christophe.blanchet@ibcp.fr*
Identifier: *NaCGZfy9NxClc3ISU158RHrG0ik*
Created: *2012-09-07T14:20:38Z*

This appliance provides cloud users with a fully functional Mobyle portal. Mobyle is a framework and web portal



A bioinformatics web interface

The screenshot displays the IDB Bioinformatics Cloud web interface. A modal dialog titled "Run Instance" is open, allowing users to configure a virtual machine instance. The dialog includes sections for choosing an appliance (currently "ARIA2.3"), filtering by thematic fields or tools (with a dropdown menu open showing options like "Genomics tools", "Molecular structural analysis", "Multiple Sequence Alignment", "Nucleotide and Protein sequence searching", "Public databases", and "Sequence analysis"), and configuring virtual machine settings (Name, Type, Number, and Storage). The background interface shows a sidebar with "Instance them (25.00%)", "Cpu", and "Memory" usage charts, and a "Room for VMs" table on the right. The table lists various VM sizes and their available counts: xsmall (247 / 284), small (145 / 160), medium (69 / 76), large (32 / 34), xlarge (15 / 16), bigmem (2 / 2), xxl (10 / 16), and htc (6 / 8). The top navigation bar includes links for Settings, Home, Help, and Sign out. The bottom status bar indicates an error: "Une erreur lors de l'ouverture de la page. Pour en savoir plus, choisissez Fenêtre > Activité."

Run Instance

Choose the appliance

Select ? ARIA2.3

Filter by:

- thematic fields
- tools

- Genomics tools
- ✓ Molecular structural analysis
- Multiple Sequence Alignment
- Nucleotide and Protein sequence searching
- Public databases
- Sequence analysis

Configure your virt

Name ?

Type ? small (1 CPU, 4GB RAM)

Number ? 1

Storage ?

Run Cancel

Powered by StratusLab

Room for VMs

xsmall	247 / 284
small	145 / 160
medium	69 / 76
large	32 / 34
xlarge	15 / 16
bigmem	2 / 2
xxl	10 / 16
htc	6 / 8

Une erreur lors de l'ouverture de la page. Pour en savoir plus, choisissez Fenêtre > Activité.

A bioinformatics web interface

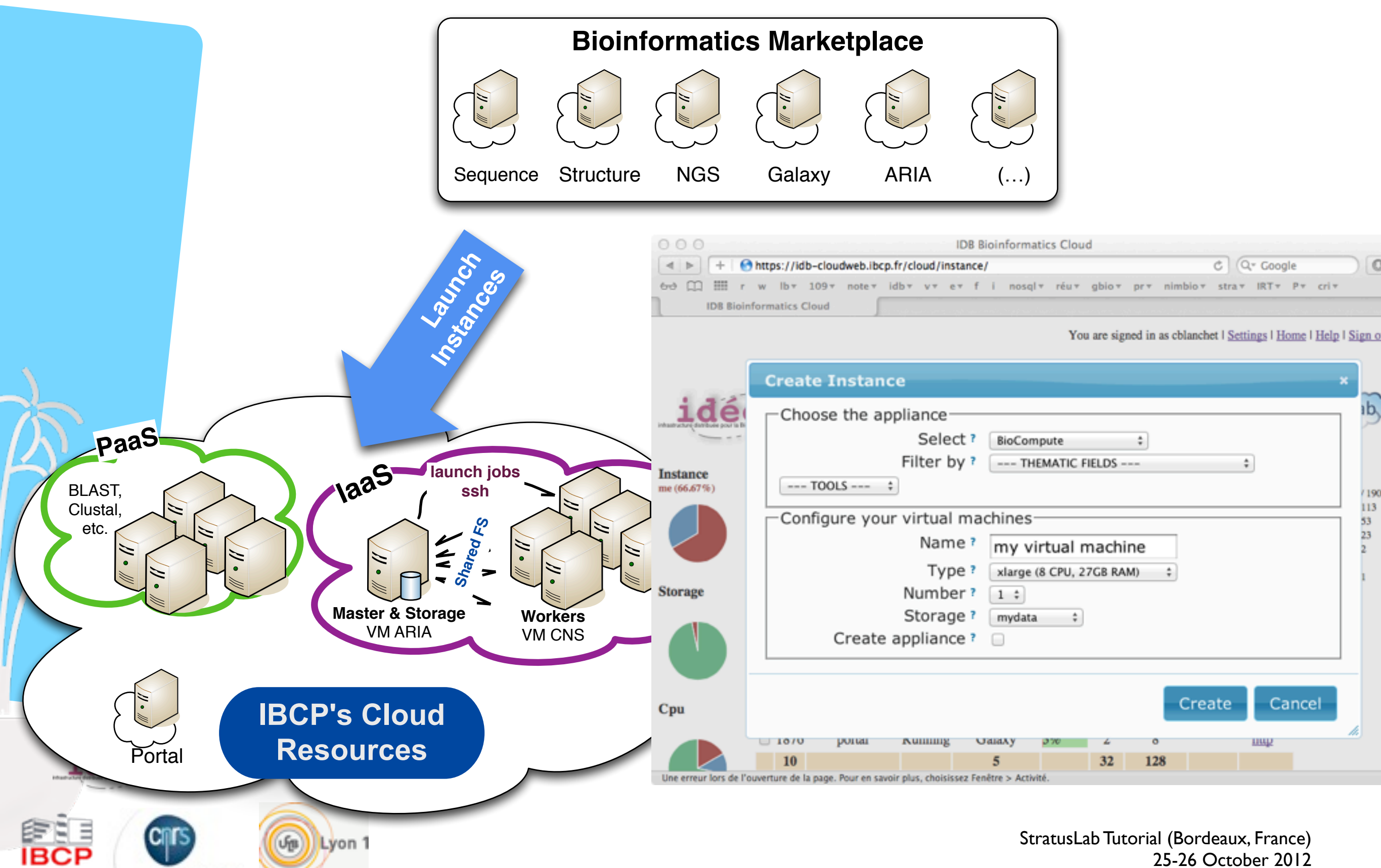
The screenshot displays the IDB Bioinformatics Cloud web interface. The main page features a sidebar with the 'idéeB' logo and navigation links, and a central area with a 'Run Instance' dialog box. The dialog box is titled 'Run Instance' and contains the following sections:

- Choose the appliance:** A dropdown menu showing 'BioCompute'.
- Filter by:** A list of filters including 'thematic fields' and 'tools'.
- Configure your machines:** A section for configuring the instance, including a dropdown for 'small (1 CPU, 4GB RAM)' and a numeric input for '1'.

The background shows the main interface with a list of instances, a 'Shutdown' button, and a 'Run Instance' button. The interface is powered by 'StratusLab'.

At the bottom of the image, there are logos for IBCP, CNRS, and Lyon 1.

Run Bioinformatics Cloud Instances



Manage your Cloud Instances

IDB Bioinformatics Cloud

https://idb-cloudweb.ibcp.fr/cloud/instance/

You are signed in as cblanchet | [Settings](#) | [Home](#) | [Help](#) | [Sign out](#)

idéeB Infrastructure distribuée pour la Biologie

Bioinformatics cloud

Powered by **StratusLab**

Instance me (66.67%)

Storage

Cpu

Shutdown Go New Instance New Storage Show Instances Show Storages Room for VMs

ID	Name	State	Appliance	CPU %	CPU	Mem. (GB)	Storage	Access
1776	cluster	Running	Hadoop GRNet	0%	2	8		ssh
1787	my clustalw node	Running	BioCompute	8%	4	16		ssh http
1789	aria IaaS	Running	ARIA2.3	0%	4	16		ssh
1790	aria IaaS	Running	ARIA2.3	0%	4	16		ssh
1791	aria IaaS	Running	ARIA2.3	0%	4	16		ssh
1792	aria IaaS	Running	ARIA2.3	0%	4	16		ssh
1793	aria IaaS	Running	ARIA2.3	0%	4	16		ssh
1874	Dble lien	Running	Apache AuthN	0%	2	8		ssh http
1875	blast	Running	BioCompute	2%	2	8		ssh http
1876	portal	Running	Galaxy	3%	2	8		http
10			5		32	128		

Room for VMs

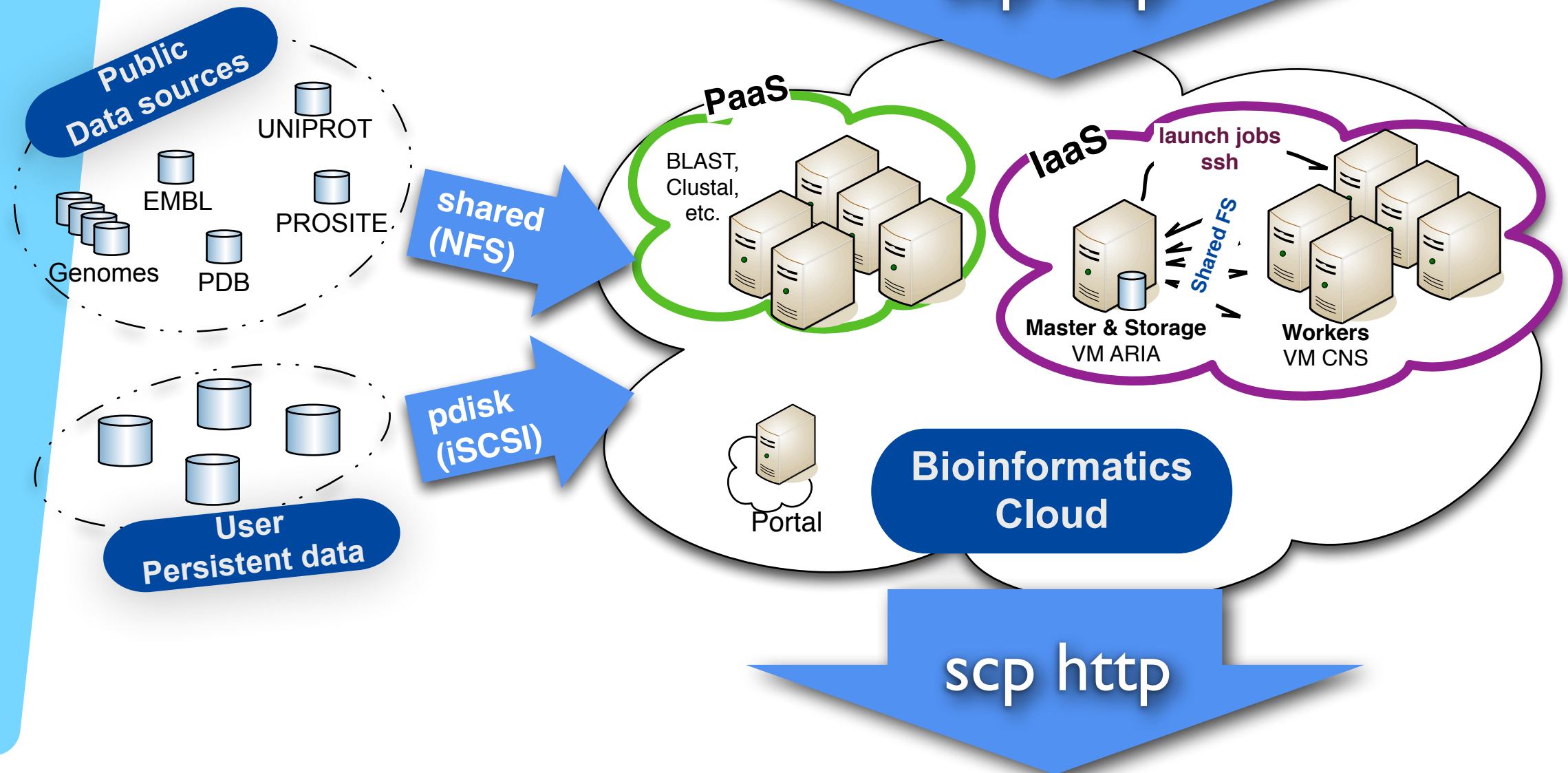
- xsmall 132 / 190
- small 70 / 113
- medium 33 / 53
- large 14 / 23
- xlarge 7 / 12
- bighmem 1 / 1
- xxl 6 / 11
- htc 4 / 4
- htc+ 1 / 1

Une erreur lors de l'ouverture de la page. Pour en savoir plus, choisissez Fenêtre > Activité.

Integrated Bioinformatics Cloud Instances

Upload user data

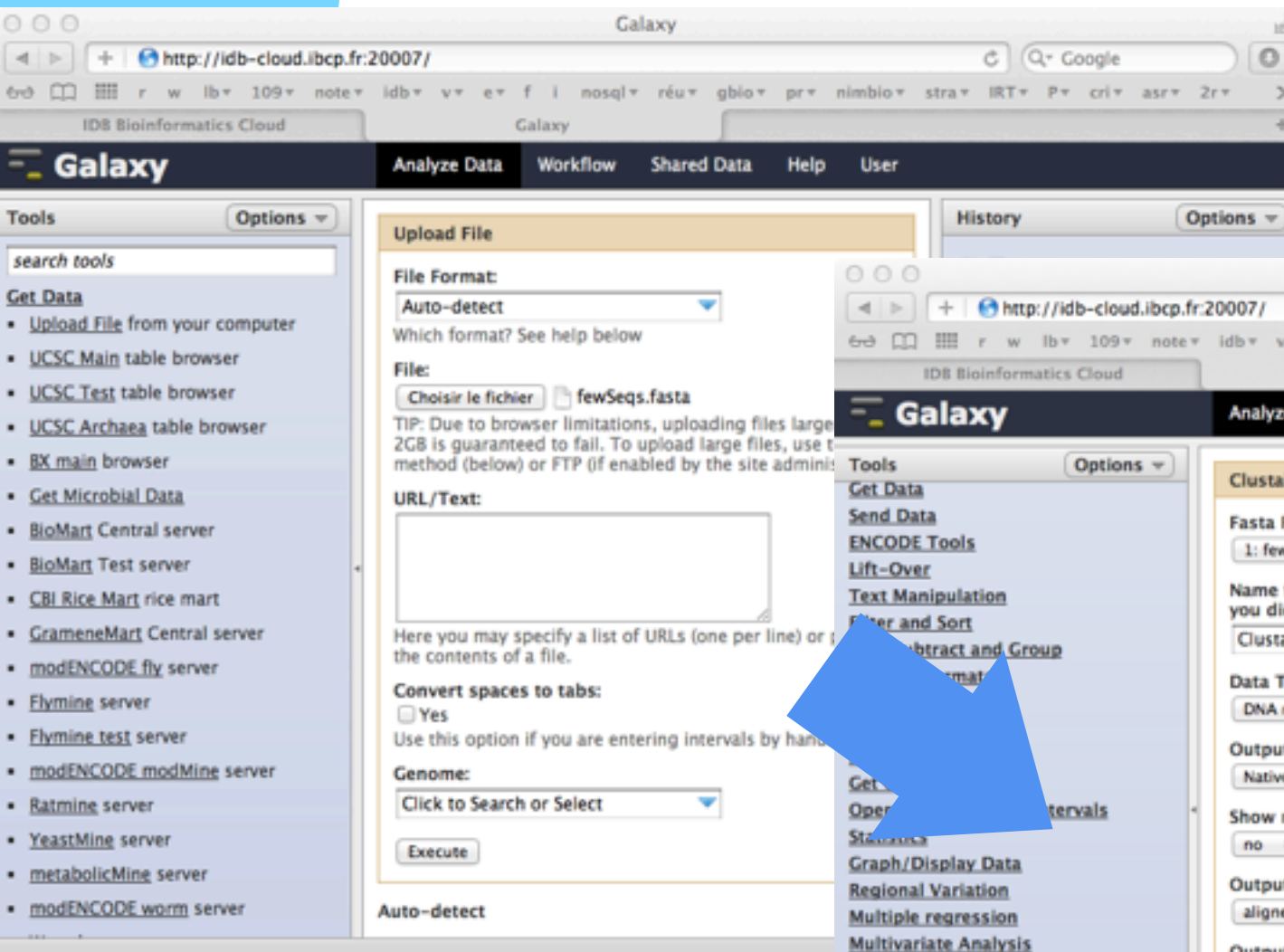
scp http



Get user results

Examples of use

<input type="checkbox"/>	1874	blast	Running	BioCompute	2%	2	8	ssh http
<input type="checkbox"/>	1875	blast	Running	BioCompute	2%	2	8	ssh http
<input type="checkbox"/>	1876	portal	Running	Galaxy	2%	2	8	http
<input type="checkbox"/>	1877	blast machine	Running	BioCompute	2%	4	16	data ssh http



Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

search tools

Get Data

- Upload File from your computer
- UCSC Main table browser
- UCSC Test table browser
- UCSC Archaea table browser
- BX main browser
- Get Microbial Data
- BioMart Central server
- BioMart Test server
- CBI Rice Mart rice mart
- GrameneMart Central server
- modENCODE fly server
- Flymine server
- Flymine test server
- modENCODE modMine server
- Batmine server
- YeastMine server
- metabolicMine server
- modENCODE worm server

Upload File

File Format: Auto-detect

Which format? See help below

File: Choisir le fichier fewSeqs.fasta

TIP: Due to browser limitations, uploading files large 2GB is guaranteed to fail. To upload large files, use the method (below) or FTP (if enabled by the site admin)

URL/Text:

Here you may specify a list of URLs (one per line) or the contents of a file.

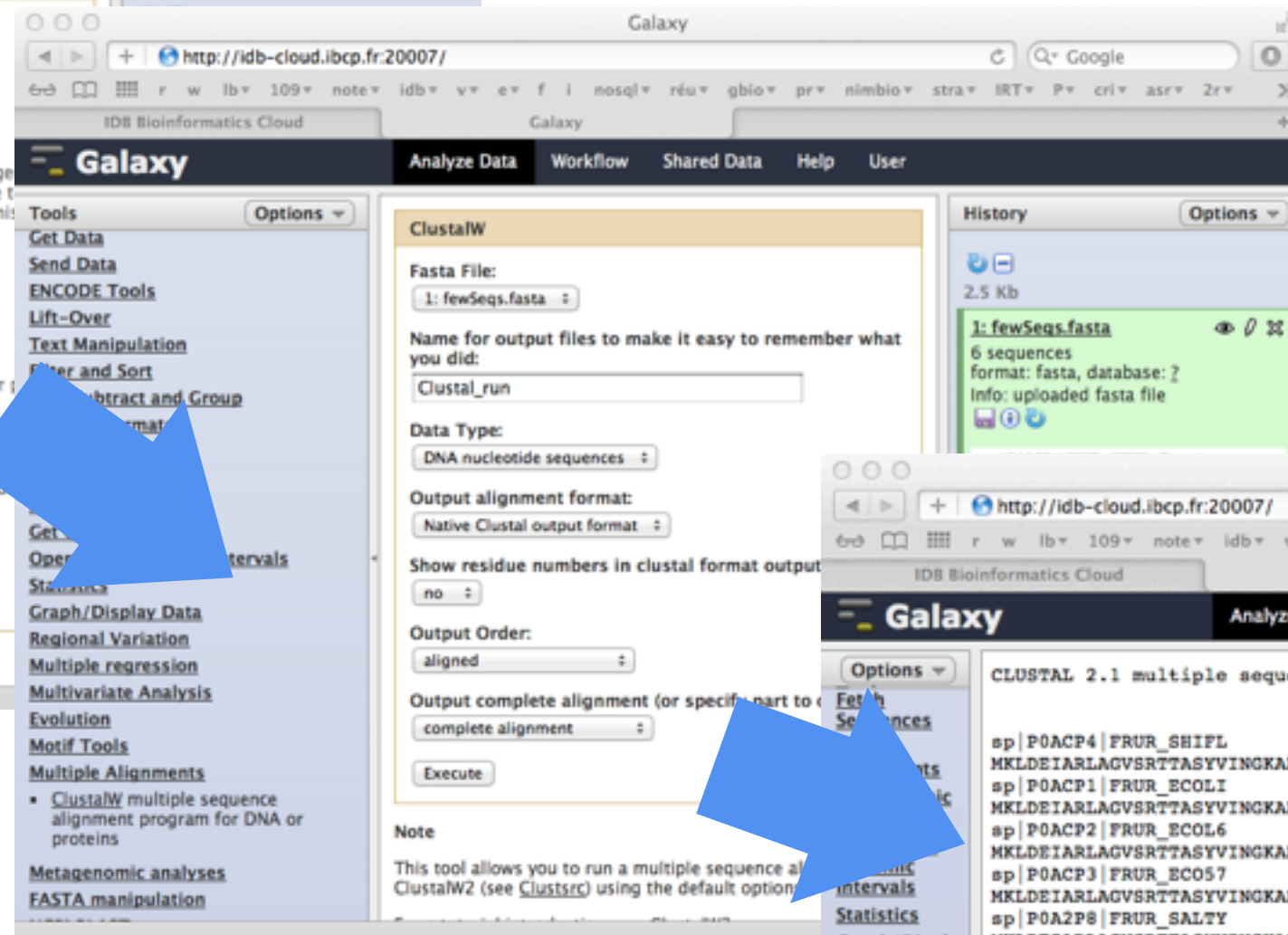
Convert spaces to tabs: ☐ Yes

Use this option if you are entering intervals by hand

Genome: Click to Search or Select

Execute

Auto-detect



Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Get Data

Send Data

ENCODE Tools

Lift-Over

Text Manipulation

Filter and Sort

Extract and Group

Get Data

Open Intervals

Statistics

Graph/Display Data

Regional Variation

Multiple regression

Multivariate Analysis

Evolution

Motif Tools

Multiple Alignments

- ClustalW multiple sequence alignment program for DNA or proteins

Metagenomic analyses

FASTA manipulation

ClustalW

Fasta File: 1: fewSeqs.fasta

Name for output files to make it easy to remember what you did: Clustal_run

Data Type: DNA nucleotide sequences

Output alignment format: Native Clustal output format

Show residue numbers in clustal format output: no

Output Order: aligned

Output complete alignment (or specify part to): complete alignment

Execute

Note

This tool allows you to run a multiple sequence alignment using the ClustalW2 (see Clustal) using the default options

History

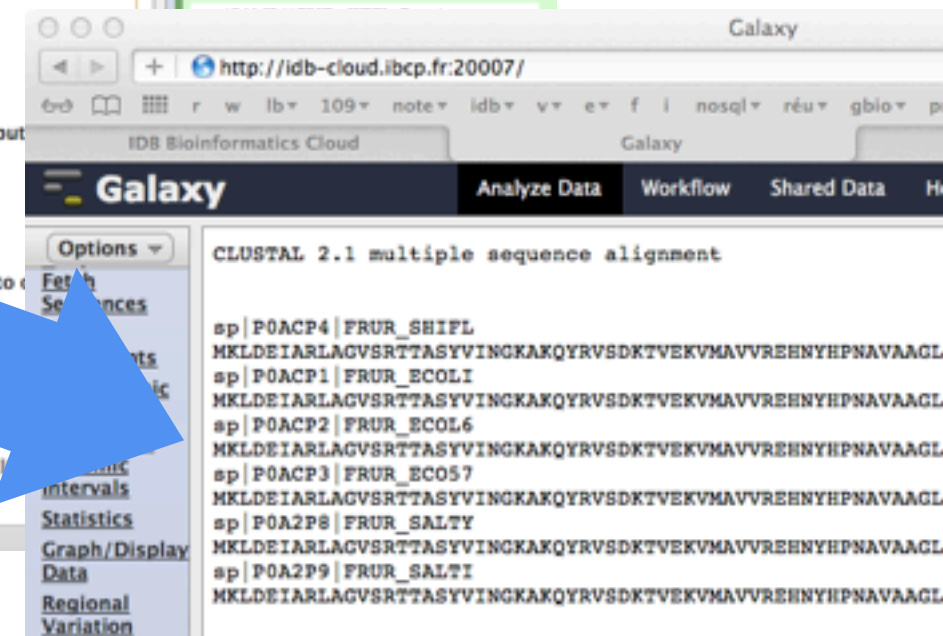
2.5 Kb

1: fewSeqs.fasta

6 sequences

format: fasta, database: ?

Info: uploaded fasta file



Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Get Data

Send Data

ENCODE Tools

Lift-Over

Text Manipulation

Filter and Sort

Extract and Group

Get Data

Open Intervals

Statistics

Graph/Display Data

Regional Variation

CLUSTAL 2.1 multiple sequence alignment

sp|P0ACP4|FRUR_SHIFL

MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL

sp|P0ACP1|FRUR_ECOLI

MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL

sp|P0ACP2|FRUR_ECOL6

MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL

sp|P0ACP3|FRUR_ECO57

MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL

sp|P0A2P8|FRUR_SALTY

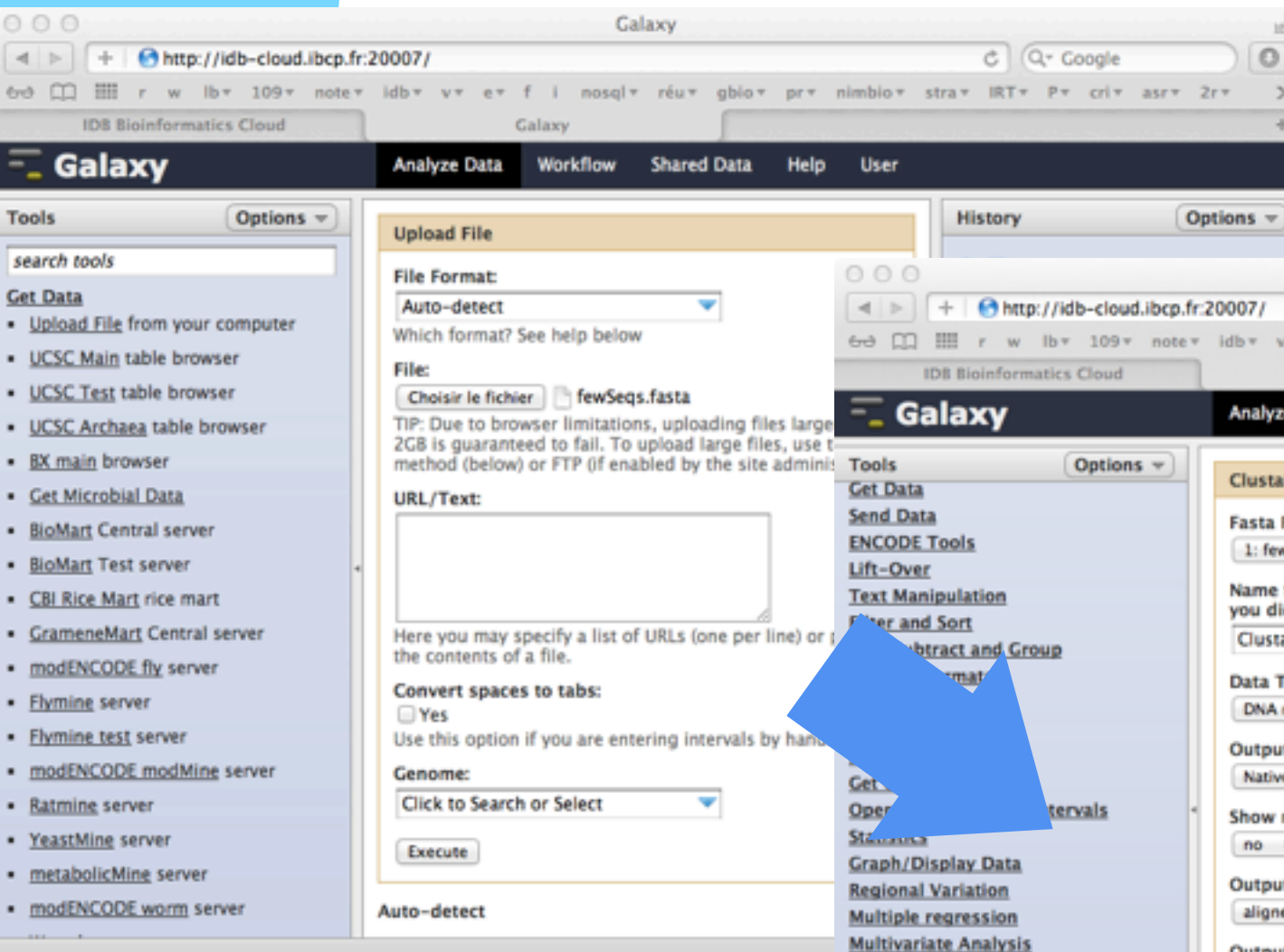
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL

sp|P0A2P9|FRUR_SALTI

MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL

Examples of use

<input type="checkbox"/>	1874	blast	Running	BioCompute	2%	2	8	ssh http://idb-cloud.ibcp.fr:20007/
<input type="checkbox"/>	1875	blast	Running	BioCompute	2%	2	8	ssh http://idb-cloud.ibcp.fr:20007/
<input type="checkbox"/>	1876	portal	Running	Galaxy	2%	2	8	ssh http://idb-cloud.ibcp.fr:20007/
<input type="checkbox"/>	1877	blast machine	Running	BioCompute	2%	4	16	ssh http://idb-cloud.ibcp.fr:20007/



Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

Upload File

File Format: Auto-detect

Which format? See help below

File: Choisir le fichier fewSeqs.fasta

TIP: Due to browser limitations, uploading files large 2GB is guaranteed to fail. To upload large files, use the method (below) or FTP (if enabled by the site administrator)

URL/Text:

Here you may specify a list of URLs (one per line) or the contents of a file.

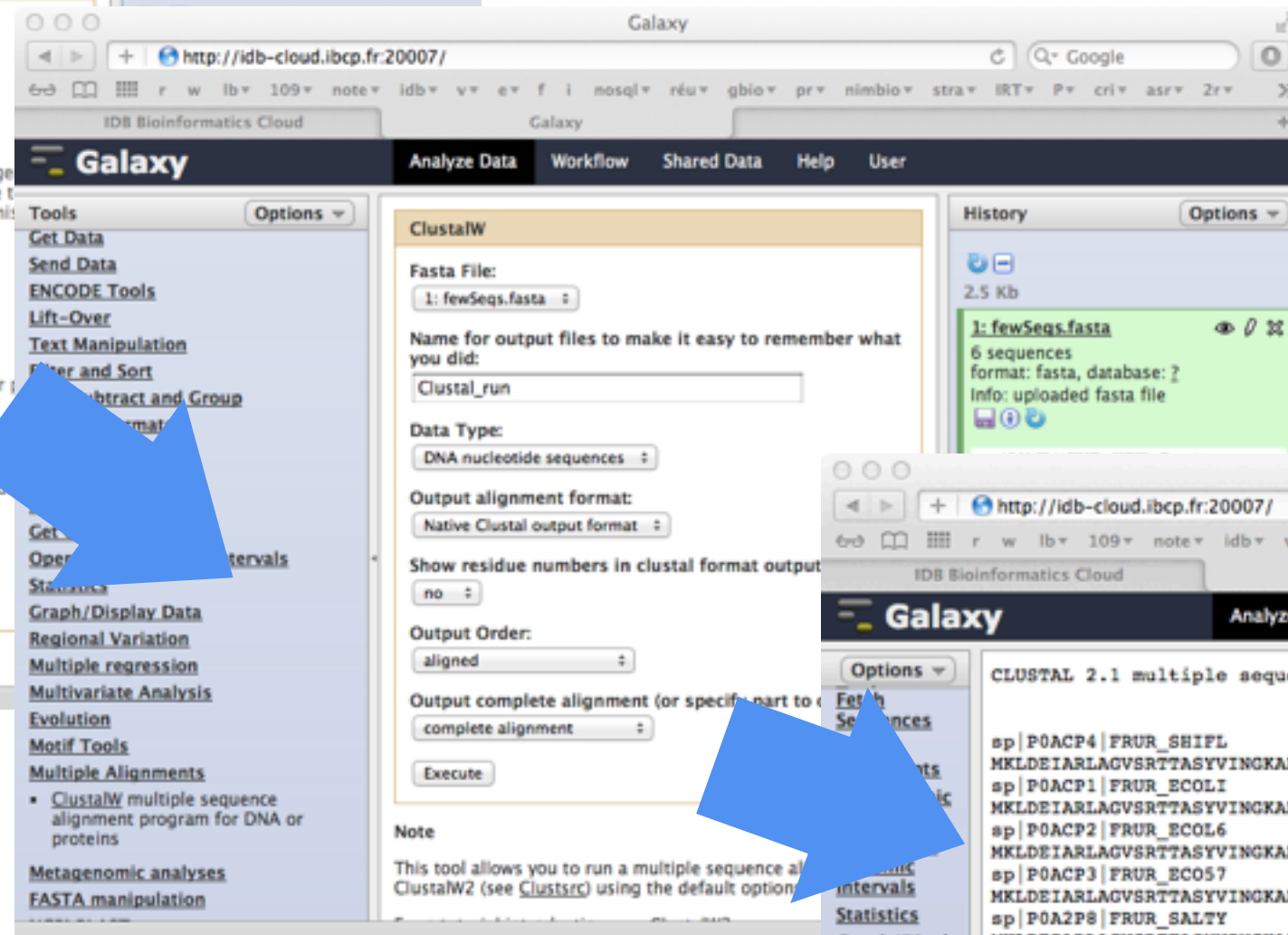
Convert spaces to tabs: ☒ Yes

Use this option if you are entering intervals by hand

Genome: Click to Search or Select

Execute

Auto-detect



Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

ClustalW

Fasta File: 1: fewSeqs.fasta

Name for output files to make it easy to remember what you did: Clustal_run

Data Type: DNA nucleotide sequences

Output alignment format: Native Clustal output format

Show residue numbers in clustal format output: no

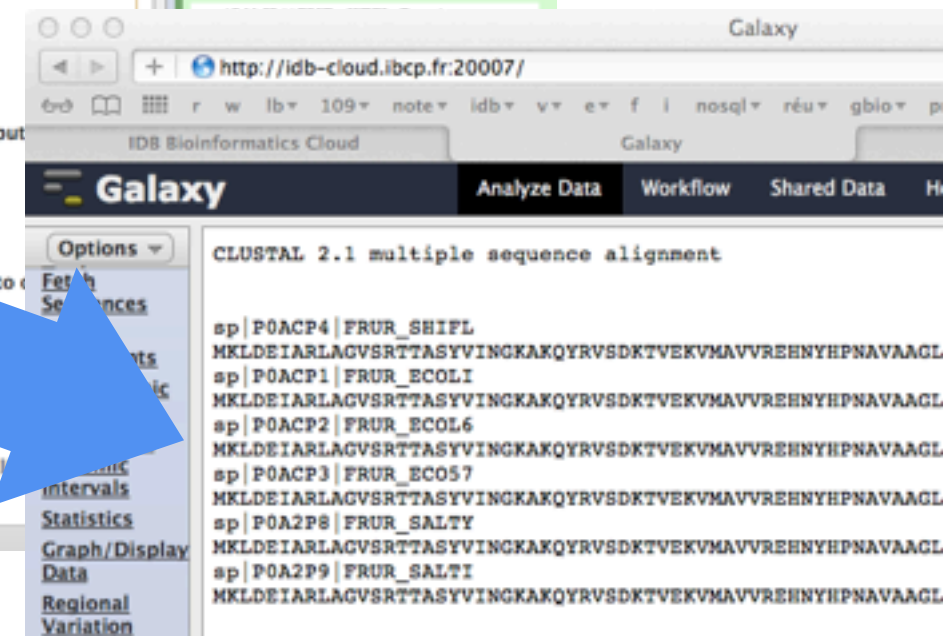
Output Order: aligned

Output complete alignment (or specify part to output): complete alignment

Execute

Note

This tool allows you to run a multiple sequence alignment using the default options of ClustalW2 (see ClustalW2) using the default options



Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

CLUSTAL 2.1 multiple sequence alignment

6 sequences
format: fasta, database: ?
Info: uploaded fasta file

sp|P0ACP4|FRUR_SHIFL
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP1|FRUR_ECOLI
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP2|FRUR_ECOL6
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP3|FRUR_ECO57
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0A2P8|FRUR_SALTY
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0A2P9|FRUR_SALTI
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL

Examples of use

<input type="checkbox"/>	1874	blast	Running	BioCompute	0%	2	8	data http
<input type="checkbox"/>	1875	blast	Running	BioCompute	2%	2	8	ssh http
<input type="checkbox"/>	1876	portal	Running	Galaxy	2%	2	8	http
<input type="checkbox"/>	1877	blast machine	Running	BioCompute	0%	4	16	data ssh

Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

Upload File

File Format: Auto-detect

Which format? See help below

File: Choisir le fichier **fewSeqs.fasta**

TIP: Due to browser limitations, uploading files larger than 2GB is guaranteed to fail. To upload large files, use the method (below) or FTP (if enabled by the site administrator)

URL/Text:

Here you may specify a list of URLs (one per line) or the contents of a file.

Convert spaces to tabs: ☐ Yes

Use this option if you are entering intervals by hand

Genome: Click to Search or Select

Execute

Auto-detect

Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

ClustalW

Fasta File: 1: fewSeqs.fasta

Name for output files to make it easy to remember what you did: Clustal_run

Data Type: DNA nucleotide sequences

Output alignment format: Native Clustal output format

Show residue numbers in clustal format output: no

Output Order: aligned

Output complete alignment (or specify part to output): complete alignment

Execute

Note

This tool allows you to run a multiple sequence alignment using the ClustalW2 (see ClustalW) using the default options

Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

CLUSTAL 2.1 multiple sequence alignment

6 sequences

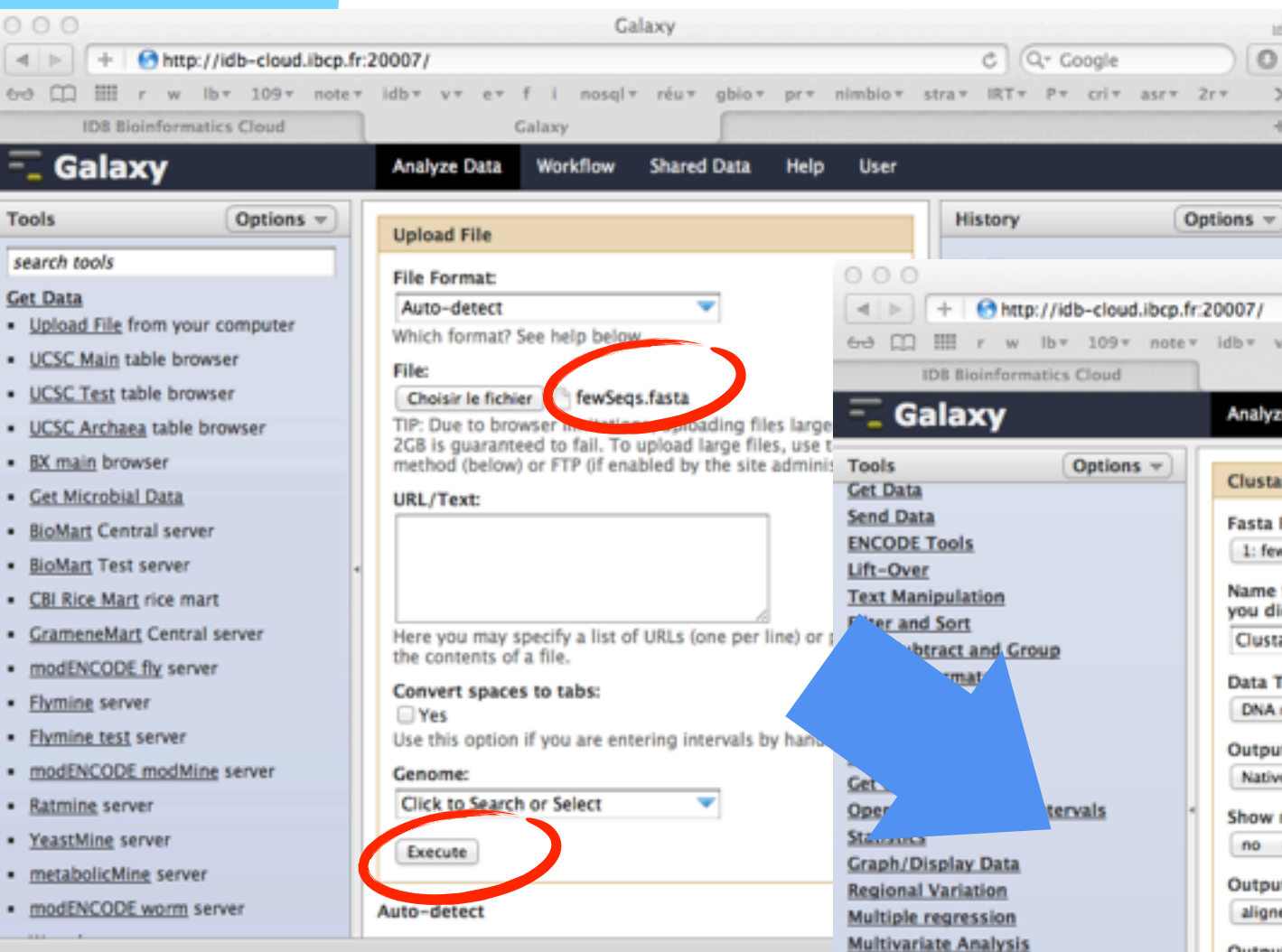
format: fasta, database: ?

Info: uploaded fasta file

sp|P0ACP4|FRUR_SHIFL
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP1|FRUR_ECOLI
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP2|FRUR_ECOL6
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP3|FRUR_ECO57
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0A2P8|FRUR_SALTY
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0A2P9|FRUR_SALTI
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL

Examples of use

<input type="checkbox"/>	1874	blast	Running	BioCompute	0%	2	8	ssh http
<input type="checkbox"/>	1875	blast	Running	BioCompute	2%	2	8	ssh http
<input type="checkbox"/>	1876	portal	Running	Galaxy	2%	2	8	ssh http
<input type="checkbox"/>	1877	blast machine	Running	BioCompute	0%	4	16	ssh http



Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

Upload File

File Format: Auto-detect

Which format? See help below

File: Choisir le fichier **fewSeqs.fasta**

TIP: Due to browser limitations, uploading files larger than 2GB is guaranteed to fail. To upload large files, use the method (below) or FTP (if enabled by the site administrator)

URL/Text:

Here you may specify a list of URLs (one per line) or the contents of a file.

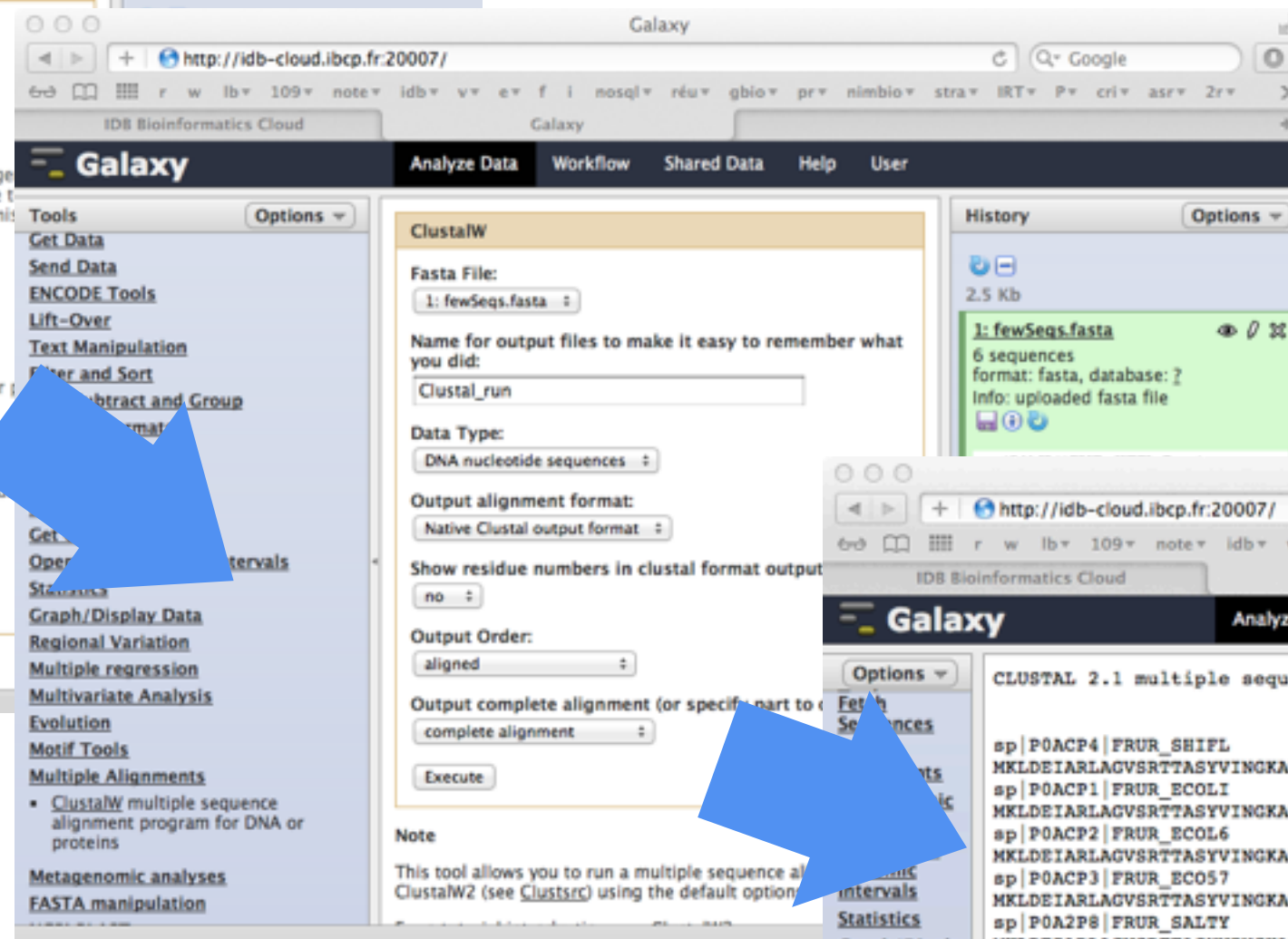
Convert spaces to tabs: ☐ Yes

Use this option if you are entering intervals by hand

Genome: Click to Search or Select

Execute

Auto-detect



Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

ClustalW

Fasta File: 1: fewSeqs.fasta

Name for output files to make it easy to remember what you did: Clustal_run

Data Type: DNA nucleotide sequences

Output alignment format: Native Clustal output format

Show residue numbers in clustal format output: no

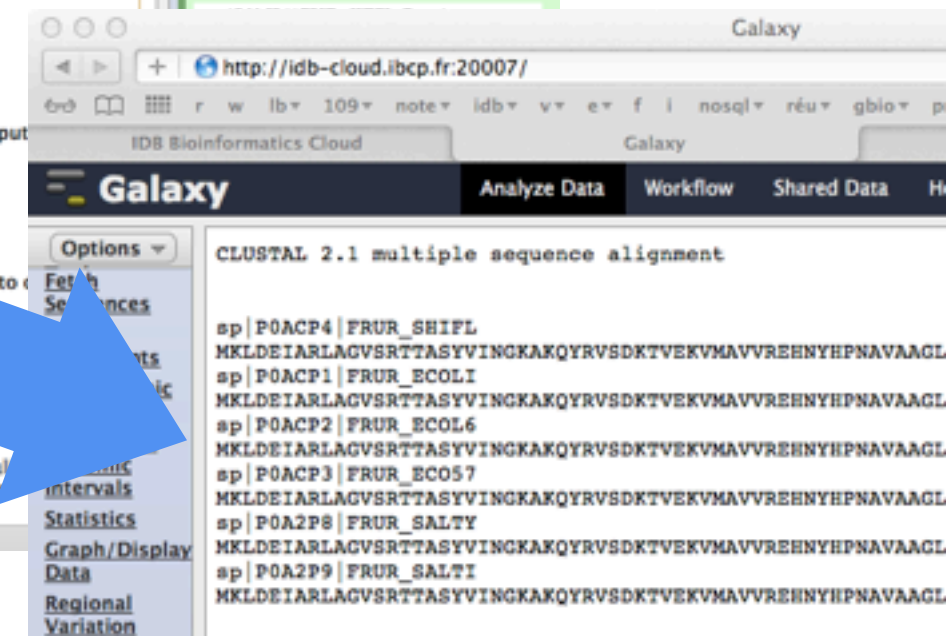
Output Order: aligned

Output complete alignment (or specify part to output): complete alignment

Execute

Note

This tool allows you to run a multiple sequence alignment using the ClustalW2 (see ClustalW) using the default options



Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

CLUSTAL 2.1 multiple sequence alignment

Fasta File: 1: fewSeqs.fasta

Name for output files to make it easy to remember what you did: Clustal_run

Data Type: DNA nucleotide sequences

Output alignment format: Native Clustal output format

Show residue numbers in clustal format output: no

Output Order: aligned

Output complete alignment (or specify part to output): complete alignment

Execute

Note

This tool allows you to run a multiple sequence alignment using the ClustalW2 (see ClustalW) using the default options

Examples of use

<input type="checkbox"/>	1874	blast	Running	BioCompute	2%	2	8	ssh http://idb-cloud.ibcp.fr:20007/
<input type="checkbox"/>	1875	blast	Running	BioCompute	2%	2	8	ssh http://idb-cloud.ibcp.fr:20007/
<input type="checkbox"/>	1876	portal	Running	Galaxy	2%	2	8	ssh http://idb-cloud.ibcp.fr:20007/
<input type="checkbox"/>	1877	blast machine	Running	BioCompute	2%	4	16	ssh http://idb-cloud.ibcp.fr:20007/

Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

Upload File

File Format: Auto-detect

Which format? See help below

File: Choisir le fichier **fewSeqs.fasta**

TIP: Due to browser limitations, uploading files larger than 2GB is guaranteed to fail. To upload large files, use the method (below) or FTP (if enabled by the site administrator)

URL/Text:

Here you may specify a list of URLs (one per line) or the contents of a file.

Convert spaces to tabs: ☐ Yes

Use this option if you are entering intervals by hand

Genome: Click to Search or Select

Execute

Auto-detect

Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

ClustalW

Fasta File: 1: fewSeqs.fasta

Name for output files to make it easy to remember what you did: Clustal_run

Data Type: DNA nucleotide sequences

Output alignment format: Native Clustal output format

Show residue numbers in clustal format output: no

Output Order: aligned

Output complete alignment (or specify part to output): complete alignment

Execute

Note

This tool allows you to run a multiple sequence alignment using the default options of ClustalW2 (see ClustalW2) using the default options

Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

CLUSTAL 2.1 multiple sequence alignment

1: fewSeqs.fasta

6 sequences

format: fasta, database: ?

Info: uploaded fasta file

Execute

CLUSTAL 2.1 multiple sequence alignment

sp|P0ACP4|FRUR_SHIFL
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP1|FRUR_ECOLI
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP2|FRUR_ECOL6
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP3|FRUR_ECO57
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0A2P8|FRUR_SALTY
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0A2P9|FRUR_SALTI
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL

Examples of use

<input type="checkbox"/>	1874	blast	Running	BioCompute	2%	2	8	ssh http://idb-cloud.ibcp.fr:20007/
<input type="checkbox"/>	1875	blast	Running	BioCompute	2%	2	8	ssh http://idb-cloud.ibcp.fr:20007/
<input type="checkbox"/>	1876	portal	Running	Galaxy	2%	2	8	ssh http://idb-cloud.ibcp.fr:20007/
<input type="checkbox"/>	1877	blast machine	Running	BioCompute	2%	4	16	ssh http://idb-cloud.ibcp.fr:20007/

Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

Upload File

File Format: Auto-detect

Which format? See help below

File: Choisir le fichier **fewSeqs.fasta**

TIP: Due to browser limitations, uploading files larger than 2GB is guaranteed to fail. To upload large files, use the method (below) or FTP (if enabled by the site administrator)

URL/Text:

Here you may specify a list of URLs (one per line) or the contents of a file.

Convert spaces to tabs: ☐ Yes

Use this option if you are entering intervals by hand

Genome: Click to Search or Select

Execute

Auto-detect

Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

ClustalW

Fasta File: 1: **fewSeqs.fasta**

Hint: For input files to make it easy to remember what you did:

Clustal_run

Data Type: DNA nucleotide sequences

Output alignment format: Native Clustal output format

Show residue numbers in clustal format output: no

Output Order: aligned

Output complete alignment (or specify part to output): complete alignment

Execute

Note

This tool allows you to run a multiple sequence alignment using the default options of ClustalW2 (see ClustalW2) using the default options

Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

CLUSTAL 2.1 multiple sequence alignment

sp|P0ACP4|FRUR_SHIFL
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP1|FRUR_ECOLI
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP2|FRUR_ECOL6
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP3|FRUR_ECOL57
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0A2P8|FRUR_SALTY
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0A2P9|FRUR_SALTI
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL

Examples of use

<input type="checkbox"/>	1874	blast	Running	BioCompute	2%	2	8	ssh http
<input type="checkbox"/>	1875	blast	Running	BioCompute	2%	2	8	ssh http
<input type="checkbox"/>	1876	portal	Running	Galaxy	2%	2	8	ssh http
<input type="checkbox"/>	1877	blast machine	Running	BioCompute	2%	4	16	ssh http

Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

Upload File

File Format: Auto-detect

Which format? See help below

File: Choisir le fichier **fewSeqs.fasta**

TIP: Due to browser limitations, uploading files larger than 2GB is guaranteed to fail. To upload large files, use the method (below) or FTP (if enabled by the site administrator)

URL/Text:

Here you may specify a list of URLs (one per line) or the contents of a file.

Convert spaces to tabs: ☐ Yes

Use this option if you are entering intervals by hand

Genome: Click to Search or Select

Execute

Auto-detect

Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

ClustalW

Fasta File: **1: fewSeqs.fasta**

Help for input files to make it easy to remember what you did:

Clustal_run

Data Type: DNA nucleotide sequences

Output alignment format: Native Clustal output format

Show residue numbers in clustal format output: no

Output Order: aligned

Output complete alignment (or specify part to output): complete alignment

Execute

Note

This tool allows you to run a multiple sequence alignment using the ClustalW2 (see ClustalW) using the default options

Galaxy

http://idb-cloud.ibcp.fr:20007/

Tools

Options

CLUSTAL 2.1 multiple sequence alignment

sp|P0ACP4|FRUR_SHIFL
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP1|FRUR_ECOLI
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP2|FRUR_ECOL6
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0ACP3|FRUR_ECO57
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0A2P8|FRUR_SALTY
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL
sp|P0A2P9|FRUR_SALTI
MKLDEIARLAGVSRRTASYVINGKAKQYRVSDKTVEKVMVAVREHNYHPNAVAAGL

- Questions and Discussion

Exercise

- Sign in
 - connect to the bioinformatics-devoted cloud
 - <http://idb-cloudweb.ibcp.fr>
 - user 'tbdxnn'
 - Configure you account : 'setttings'
(don't forget to input your SSH pubkey)
- Run instance
 - select an appliance
 - create an instance
 - connect and run tools (ssh or http)
- Store data
 - create a storage
 - use it with an instance



Platform ' Infrastructure Distributed for Biology - IDB ' acknowledges co-funding by the European Community's Seventh Framework Programme (INFSO-RI-261552) and the French National Research Agency's Arpege Programme (ANR-10-SEGI-001)

<http://idee-b.ibcp.fr>

