



SyBIT

SystemsX.ch

Biology IT

SyBIT

Support for SystemsX.ch Projects

Peter Kunszt

peter.kunszt@systemsx.ch



SystemsX.ch

The Swiss Initiative in Systems Biology

Challenge : Scale Up

- High Throughput Instruments
 - Much larger **data volumes**
 - Increased **data complexity**
- Large Collaborations
 - More people
 - More experiments and measurements



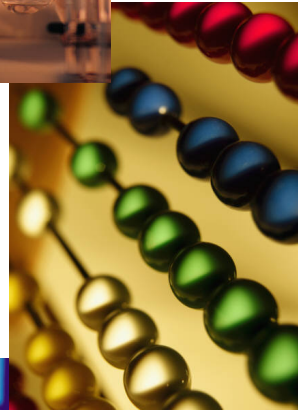
Actors

Data Intensive
Life Science
Research

Researchers

Platforms

IT Resource
providers



3/14/13

Need for an Integrator

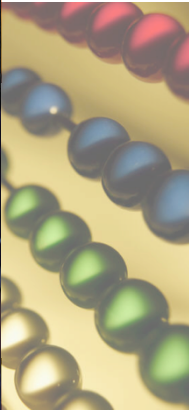
Data Intensive
Systems
Biology
Research

SyBIT
Support

Researchers

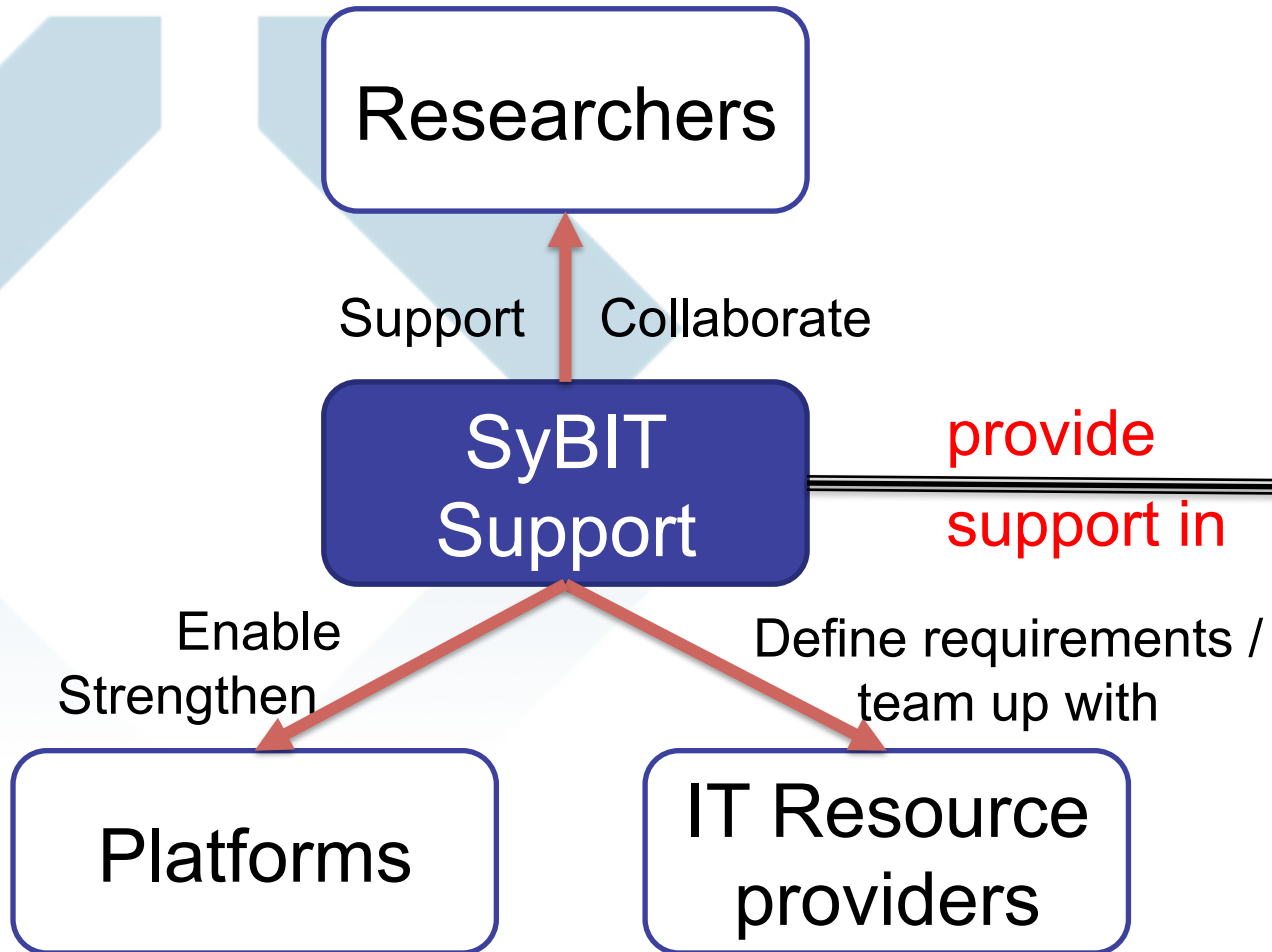
Platforms

IT Resource
providers



3/14/13

Fill the Gap



Tooling
Bioinformatics Services
Data Management
Automation / Workflows
Large Scale analysis
Standards, Integration
Software engineering
Legacy support
Documentation
Education
Collaboration support
Setup, Configuration
Versioning
Commercial software



SyBIT in SystemsX.ch

- SyBIT is the Support project in SystemsX.ch
 - IT : Infrastructure, operations in cooperation with local suppliers – IT depts, lab admins, core facilities
 - Bioinformatics tooling : development, software deployment, support, operations of services
 - supporting 50+ labs all over Switzerland
 - 200+ scientists in direct interaction
 - Distributed team of 25+ people, budget of 2.2Mio/Year

SyBIT Partners

Zurich

- Functional Genomic Center ZH
- ETHZ Institute of Molecular Systems Biology: 4 research labs
- ETHZ Light Microscopy Center

Basel

- BSSE Center for Information Science and Databases
- Biozentrum : Research IT and individual labs
- FMI genomics support and IT support

Lausanne

- Vital-IT of Swiss Institute of Bioinformatics
- EPFL – Bioinformatics and Biostatistics Core Facility



Collaboration Support

- RTD Project Webpages
 - Online content management
 - Maintained, supported
 - Examples: www.cyclix.org www.infectx.org www.cellplasticity.org
- RTD Wiki spaces
 - Supported commercial wiki: Confluence
 - User and group management
 - <http://wiki.systemsx.ch>
- Mailing lists
 - *listname*@sympa.systemsx.ch
 - Web interface at <http://sympa.systemsx.ch>
- Secure subversion code repositories
 - At systemsx02.ethz.ch/svn



Data Publication Support

- Support of international repositories
 - Support with upload to public data repositories at EBI, NCBI, etc
- Custom publication
 - For data where no repositories exist (like microscopy imaging data) we provide public data servers on demand
- All published SystemsX.ch data is linked through www.sybit.net



Infrastructure

- SyBIT works closely with all local infrastructure providers
 - Arc Lémanique: Vital-IT
 - ETH Zurich: Brutus cluster
 - University of Zürich: Functional Genomics Center
 - University of Basel: Biozentrum
- Additional Storage through CSCS
- Good interactions with Network teams and SWITCH



Software

- Bioinformatics support and engineering, standardized tools and analysis
- Parallelization and Cluster-readyness (HPC, automated workflows, processing)
- Data management, large-scale data handling
- Software engineering and refactoring of science codes
- Packaging of software, documentation
- Customizations, Web interfaces, etc



SyBIT Software

- **Open-Source:** All tools and software written by SyBIT is open-source (GNU/BSD/Apache)
 - openBIS, MorphoGraphX, CIFEX, CellClassifier..
- **Community Code:** Contribute software to existing communities if possible
 - QuasR, CellProfiler, BioFormats..
- **Commercial Tool Support:**
 - Matlab, Spotfire, Genohm LIMS ..
- All tools are visible on www.sybit.net





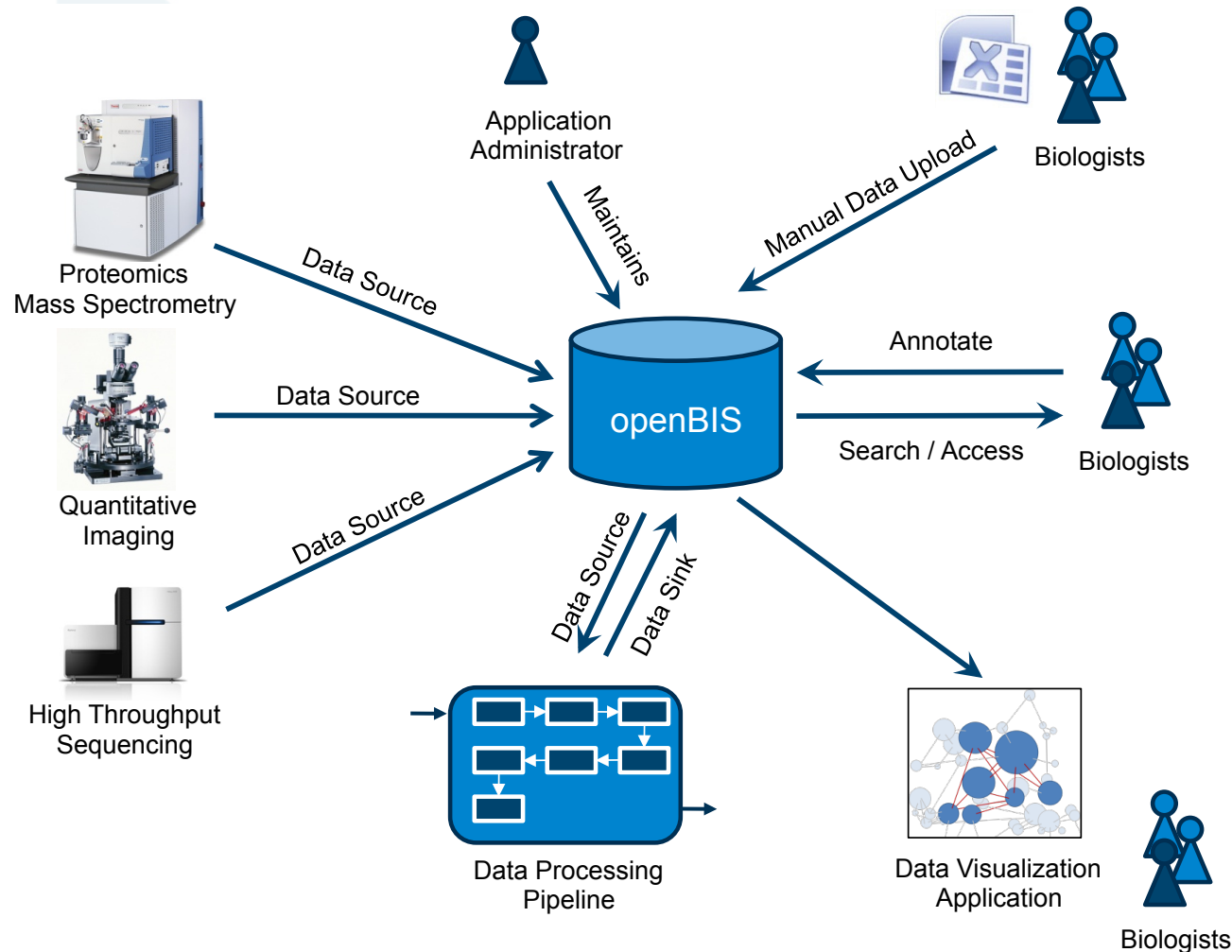
Example 1:

<http://www.cisd.ethz.ch/software/openBIS>

Used in: YeastX, PhosphoNetX, LipidX, InfectX, CINA

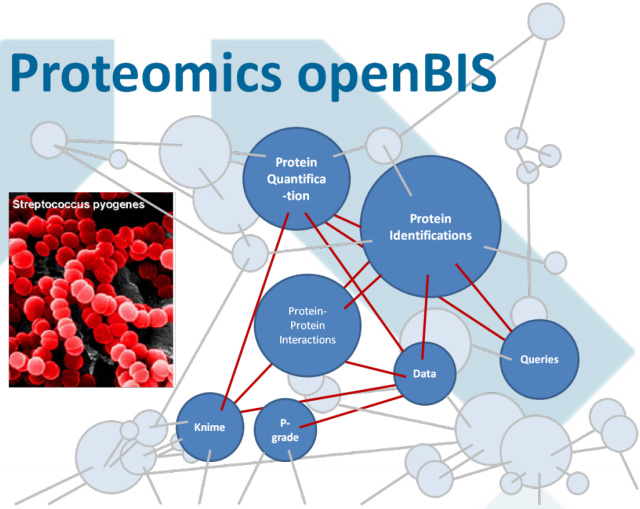


Data Management for Data-Intensive Research

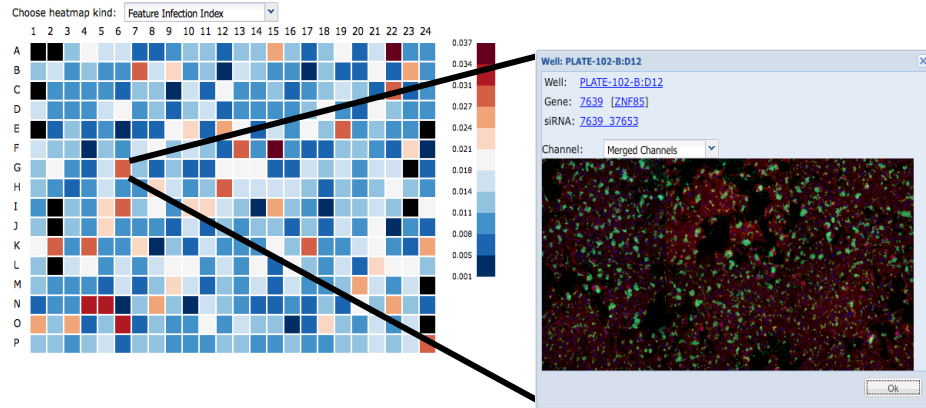


openBIS Variants

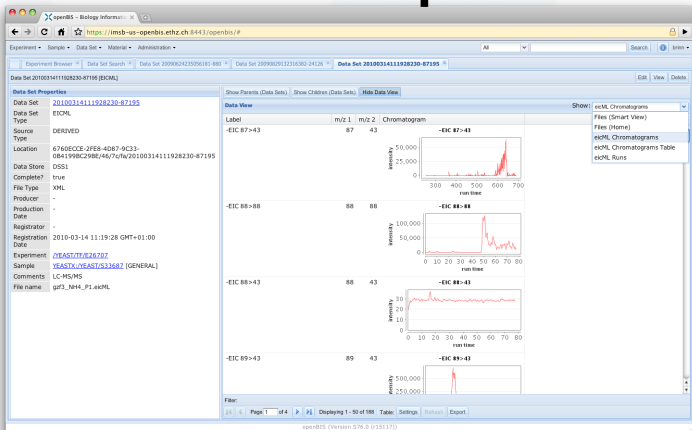
Proteomics openBIS



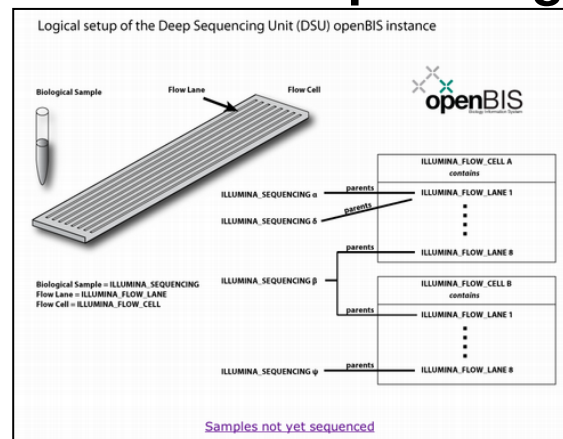
High Content Screening & Imaging



Metabolomics openBIS



Next Gen Sequencing



openBIS as LIMS

Projects	Experiments
Space / Project	Code
• DIANA	FACS
• CELL_SIZE_CONTROL	GROWTH_CURVES
• INDUCIBLE_TRANSCRIPTION (info)	MICROSCOPY
• KRISTINA	HERETOLOGOUS_COMMUNICATION_SYST
• MARIO	GATES
• ROBERT	SHUTTLE_VECTORS
• YEAST_LAB	BACTERIA
	CHEMICALS
	CL_TEMPLATE
	OLIGO
	PLASMIDS
	PROTOCOLS
	YEAST

Results

Materials and methods



Interfaces

Web UI

The screenshot shows the openBIS web interface. At the top, there is a navigation bar with search and menu options. Below that is a table with columns for Gene, Data Set, Well Content, Plate, Well, Well Row, Well Column, File Type, and Well Images. The table lists several entries for gene AAK1 across different data sets and well configurations. To the right of the table, there are four microscopy images showing fluorescence staining of cells.

iPad UI

The screenshot shows the openBIS iPad interface. On the left, there is a list of probes from HT5_PROBE_100 to HT5_PROBE_114. The main area displays the details for HT5_PROBE_105, including its description: "Binding affinity for human cloned 5-hydroxytryptamine 3 receptor". A chemical structure of the probe is shown on the right. Below the description, there is a section for "DESC Binding affinity for human cloned 5-hydroxytryptamine 3 receptor" with a table for data sets.

Custom UI

The screenshot shows a custom web interface. On the left, there is a complex network diagram with nodes representing genes and edges representing relationships. The nodes are color-coded and labeled with gene names like BIOCENTER_ILANDERO, ETHZ_FGZCO, ETHZ_PALALO, etc. On the right, there are two detailed views of gene information for PRK [BSEF-QG9-5893] and PRK [BSEF-QG9-3437], showing various attributes like accession numbers, descriptions, and identifiers.

Custom UI

The screenshot shows a custom web interface. The main part is a table with columns for Plasmid Code, Owner, Owner Number, Plasmid Name, Backbone, and Bacter Antibio Resistal. The table lists 10 plasmids (FRP1 to FRP10) with their respective owners and details. To the right of the table, there is a circular plasmid map showing the arrangement of genes and features on the plasmid.



OpenBIS

- Development Team at ETH Zürich BSSE Department in Basel, Center for Information Science and Databases CISD
 - Bernd Rinn, Chandrasekhar Ramakrishnan, Franz-Josef Elmer, Jakub Straszewski, Manuel Kohler, Antti Luomi, Caterina Barillari, Piotr Kupczyk
- Largely funded by SyBIT, but also directly by ETH and other projects
- Large number of deployments, also outside of SystemsX.ch





Example 2: MorphoGraphX

<http://www.morphographx.org>

Used in: PlantGrowth



MorphGraphX

- **Visualization and processing of 3D biological datasets.**
- Cross-platform, open-source java tool, making use of the GPU if available
- Encourages the development and sharing of specialized image processing algorithms for research purposes,
- Uses several open-source tools from the research community: ITK, VVe, CImg, QGLViewer and Qt.

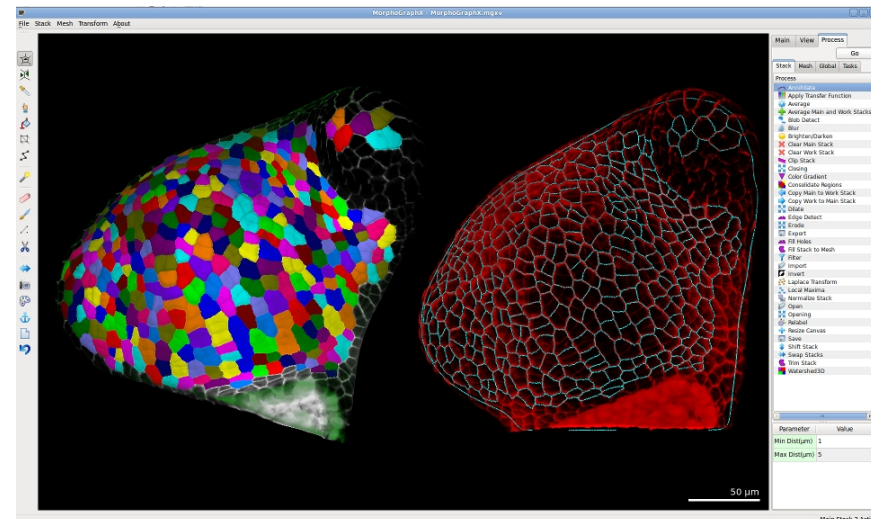
MorphoGraphX

Plant Growth RTD UniBE:

- Richard Smith, Pierre Barbier de Reuille

SyBIT support:

- Thierry Schüpbach (Vital-IT), Adam Srebniak (ETHZ)



Example 3: GDV Genome Data Viewer

<http://gdv.epfl.ch>

Used in: CycliX, LipidX



Genome Data Viewer GDV

- Data Management and Sharing
- Visualization of tracks
- Analysis – start various analyses from the interface
- Using bioscript, which can be embedded in any web interface
 - <http://gdv.epfl.ch/bs>



Genome Data Viewer GDV

Upload new data

Additional actions
when a project is
selected

Search for features

GDV Genome Data Viewer

Tracks | Circles | Help | Admin | Logout

Upload | View | Edit | Share | Delete

Search:

Projects	Name	Color	Created	Assembly	Type
project_CDN	conservation_chr1		13. Jul 2012 11h19	sacCer2	signal
project_ZFP	GC percent		13. Jul 2012 13h02	sacCer2	not determined
Test	Saccharomyces_cerevisiae.EF4.67		13. Jul 2012 10h46	sacCer2	relational
rr	sin		09. Aug 2012 14h15	sacCer2	signal
	transReg_chr1		13. Jul 2012 12h44	sacCer2	features

Shared projects: CST project RDU, ZFP mouscaz R

+ New Project

FAILURE

Download | Edit | Delete

Quickly see whats wrong

Identify tracks fast : set the track color

Utilities

Create new projects

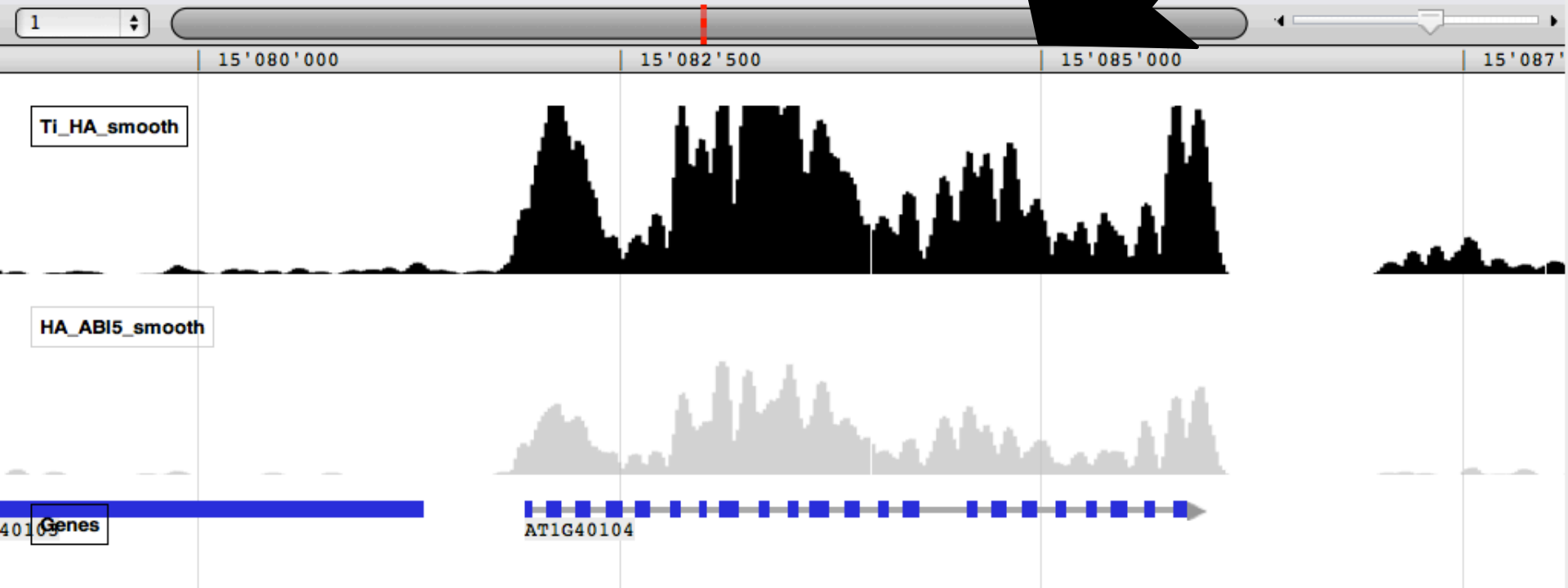
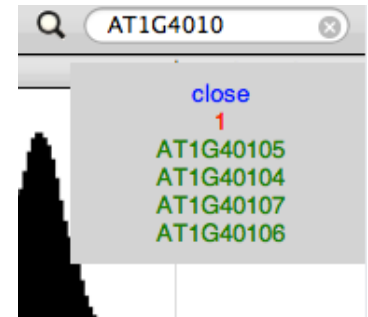


Storage and Management of data, Sharing

Genome Data Viewer GDV

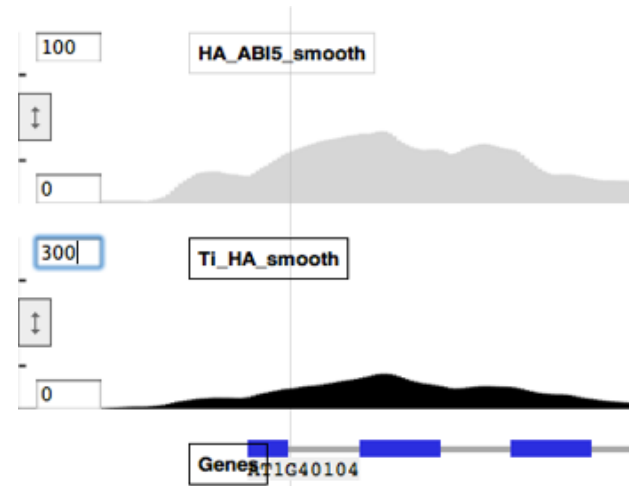
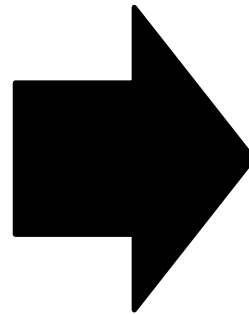
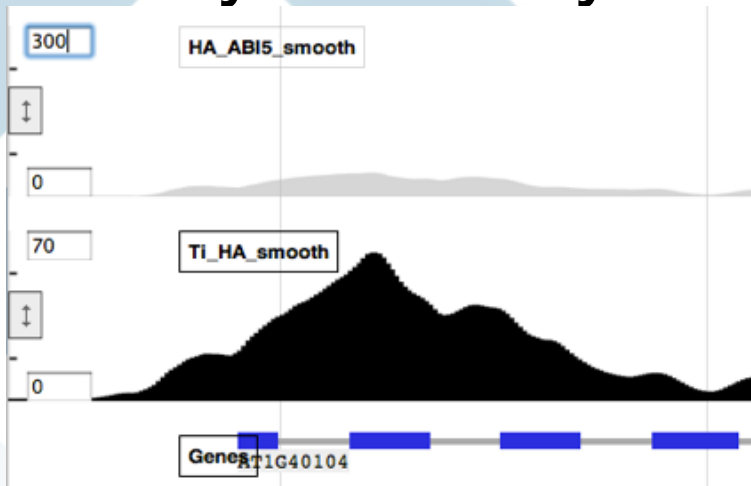
- Visualization of data.

Search for a gene or location

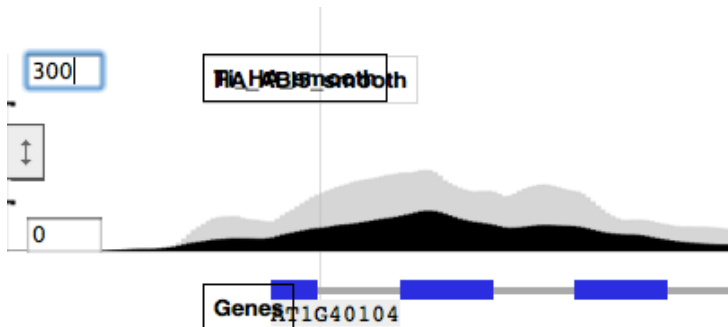


Genome Data Viewer GDV

- Dynamically change track scale



- Superpose tracks



Fetch more Information

- Info panel
- Reflex API

ERF10

ERF10

Links
[Ensembl](#)
[reflex.ws](#)

click

ERF10 AT1G03810

Protein Add About

ERF10 (AT1G03800.1) A. thaliana Edit

No Synonyms

[Sequence](#), [Domains](#), [Structure](#), [Locus](#), [Literature](#)

MTTEKENVTTAVAVKDGGEKSKEVSDKGVKRRKNVTKALAVNDGGEKSKEV

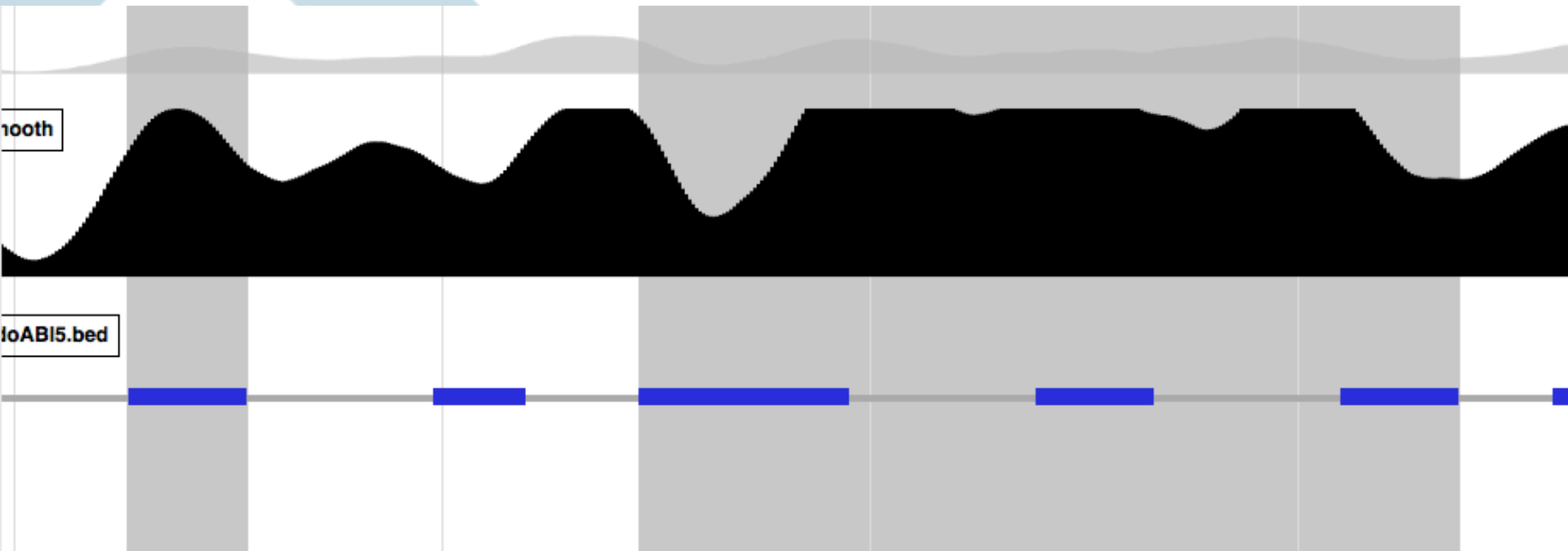
No structure information available

No annotation available



Genome Data Viewer GDV

- Make selections and save them



- Enter a description

1 (15082816,15082887) First peak [Delete](#)

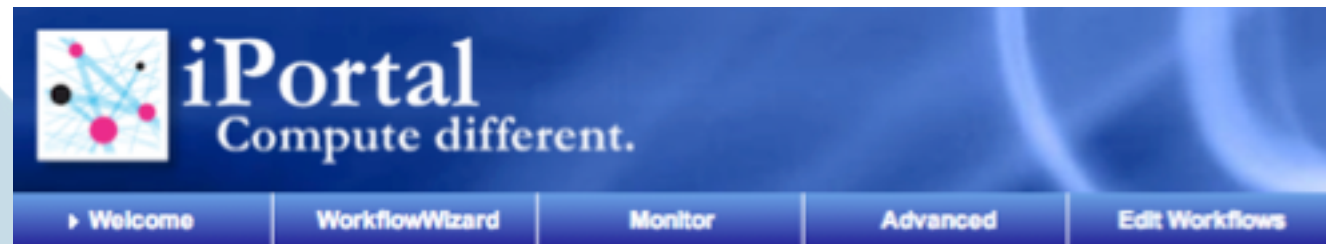
1 (15083115,15083595) Important cluster [Delete](#)



Genome Data Viewer GDV

- Developed at the Bioinformatics and Biostatistics Core Facility BBCF of the EPFL
- Jacques Rougemont, Johan Yarosz, Julien Delafontaine





Example 4: iPortal Proteomics Portal

<https://iportal.ethz.ch/>

Used in: PhosphoNetX



iPortal

Easy automated large-scale proteomics analysis. Wizard to select type of analysis



SEARCH

TPP: Identify peptides/proteins via shotgun mass spectrometry.

Choose



QUANTIFICATION

Quantify protein abundances and relative abundance.

Choose



SWATH

Workflows used with SWATH

Choose



iPortal

Workflow selection

Default	Default peptide/protein identification workflow. Allows to choose search engines and modifications freely. Parameter sweeps possible. Details	Choose
sequestimport_TPP	This workflow takes SEQUEST search output you specify, runs the latest TPP on it and exports the resulting protXML to OpenBIS. Details	Choose
zDONOTUSE_TPP_petunia	Default peptide/protein identification workflow. Choose search engines and modifications freely. Parameter sweeps possible. Details	Choose
zDONOTUSE_sequest_import	This workflow takes SEQUEST search output you specify, runs the latest TPP on it and exports the results to OpenBIS. Details	Choose



iPortal

Data selection (from OpenBIS)

Biological Sample Based Preselection

OpenBIS dataset selection: 0 files are selected.

BIO_EXPERIMENT	CENTROIDED	FILE_NAME	REGISTRATION_DATE	SAMPLE_CODE
E288203	false	130315_LH_cfg_pel_dil10.mzXML	2013/03/28 16:57	130315_LH_CFG_PEL_DIL10
E288203	false	130315_LH_cfg_sol.mzXML	2013/03/28 16:48	130315_LH_CFG_SOL
E288203	false	130315_LH_DigDeAPr_high.mzXML	2013/03/28 16:41	130315_LH_DIGDEAPR_HIGH
E288203	false	130315_LH_DigDeAPr_low.mzXML	2013/03/28 16:34	130315_LH_DIGDEAPR_LOW
E288203	false	130315_LH_sw_pool.mzXML	2013/03/28 01:30	130315_LH_SW_POOL
E288215	false	130308_Lotta_Q99YH8+Medium3.mzXML	2013/03/24 13:06	130308_LOTTA_Q99YH8-MEDIUM3
E288215	false	130308_Lotta_Q99YH8+Medium2.mzXML	2013/03/24 13:02	130308_LOTTA_Q99YH8-MEDIUM2
E288215	false	130308_Lotta_Q99YH8+Medium1.mzXML	2013/03/24 12:58	130308_LOTTA_Q99YH8-MEDIUM1
E288215	false	130308_Lotta_Q9A170+Medium3.mzXML	2013/03/24 12:54	130308_LOTTA_Q9A170-MEDIUM3
E288215	false	130308_Lotta_Q9A170+Medium2.mzXML	2013/03/24 12:50	130308_LOTTA_Q9A170-MEDIUM2
E288215	false	130308_Lotta_Q9A170+Medium1.mzXML	2013/03/24 12:45	130308_LOTTA_Q9A170-MEDIUM1
E288215	false	130308_Lotta_Medium3.mzXML	2013/03/24 12:41	130308_LOTTA_MEDIUM3
E288215	false	130308_Lotta_Medium2.mzXML	2013/03/24 12:36	130308_LOTTA_MEDIUM2

Navigation:



iPortal

Parametrization

BioDB Database Selection

BioDB Type	BioDB Version	BioDB Database
ex_sp	20121201	ex_sp_yeast

Parameter Sets:

Parameter Set Name	Description		
Default - READ ONLY	Default parameters	Show	<input type="checkbox"/>
DefaultXTandem - READ ONLY	Default parameters for running only XTandem	Show	<input type="checkbox"/>
DefaultXTandemOmssa - READ ONLY	Default parameters for running only XTandem and Omssa	Show	<input type="checkbox"/>
Phospho - READ ONLY	Phosphorylation as variable modification on STY	Show	<input type="checkbox"/>
Semi-Tryptic - READ ONLY	Semi-tryptic digestion and oxidation as variable modification on M	Show	<input type="checkbox"/>

COMPACT
 ALL
 DEBUG
 FILTERED

General

FDR for Peptide Identifications	0.01	<input type="button" value="⌵"/>	<input type="button" value="⌶"/>
Fragment Mass Error	0.4	<input type="button" value="⌵"/>	<input type="button" value="⌶"/>
Fragment Mass Error Units	Da	<input type="button" value="⌵"/>	<input type="button" value="⌶"/>
Parent Mass Error	15	<input type="button" value="⌵"/>	<input type="button" value="⌶"/>
Parent Mass Error Unit	ppm	<input type="button" value="⌵"/>	<input type="button" value="⌶"/>
Maximum Missed cleavage sites	1	<input type="button" value="⌵"/>	<input type="button" value="⌶"/>
Enzyme	Trypsin	<input type="button" value="⌵"/>	<input type="button" value="⌶"/>
Static Modifications	Carbamidomethyl (C)	<input type="button" value="⌵"/>	<input type="button" value="⌶"/>
Variable Modifications		<input type="button" value="⌵"/>	<input type="button" value="⌶"/>

Search Engines

Use XTandem	True	<input type="button" value="⌵"/>	<input type="button" value="⌶"/>
Use Omssa	True	<input type="button" value="⌵"/>	<input type="button" value="⌶"/>
Use Myrimatch	True	<input type="button" value="⌵"/>	<input type="button" value="⌶"/>

Save As: Name:

Description:



iPortal Monitoring

zDONOTUSE_Search2Tram1	IPortal_xuboTram1	2013-04-02-202721	loblum	FINISHED	Abort Delete
zDONTUSE_LFQ_110	IPortal_sequestimport_lfq	2013-04-02-202854	loblum	FINISHED	Abort Delete
LFQ	DynMod_MapA15_signoise2	2013-04-02-203334	olgas	RUNNING	
zDONOTUSE_OpenSWATH	IPortal_openswath_xubo	2013-04-03-090309	loblum	ERROR	Abort Rescue Delete
zDONOTUSE_sequest_import	IPortal_sequestimport883	2013-04-03-092518	loblum	FINISHED	Abort Delete
Default	ka_swabs2_frac-174_ribo	2013-04-03-092601	karlassoc	FINISHED	Abort Delete
Default	ka_swabs2_frac-168	2013-04-03-092644	karlassoc	FINISHED	Abort Delete
zDONOTUSE_LibraryCreation	IPortal_ups8_lib	2013-04-03-094749	loblum	FINISHED	Abort Delete
zDONOTUSE_TPP_petunia	20130403_black_MDV_O1212_sp9606_zDONOTUSE	2013-04-03-100832	hehardt	FINISHED	Abort Delete
LFQ	ka_swabs2_frac-168_lfq	2013-04-03-105638	karlassoc	FINISHED	Abort Delete
zDONTUSE_LFQ_110	IPortal_sequestimport883_lfq	2013-04-03-110147	loblum	FINISHED	Abort Delete
zDONOTUSE_LibraryCreation	IPortal_sequestimport883_tram1	2013-04-03-113049	loblum	FINISHED	Abort Delete
zDONOTUSE_TPP_petunia	20130403_black_MDV_O1212_sp9606_zDONOTUSE_XTandem	2013-04-03-150857	hehardt	FINISHED	Abort Delete
zDONTUSE_LFQ_110	20130403_black_MDV_O1212_sp9606_zDONOTUSE_XTandem_LFQ	2013-04-03-160044	hehardt	RUNNING	Abort
zDONOTUSE_LibraryCreation	IPortal_paralell	2013-04-03-163728	loblum	FINISHED	Abort Delete
zDONOTUSE_TPP_petunia	20130403_black_MDV_O1212_sp9606_zDONOTUSE_XTandem_OMSSA	2013-04-04-085340	hehardt	FINISHED	Abort Delete
zDONTUSE_LFQ_110		2013-04-04-111504	wdlski	RUNNING	Abort
zDONOTUSE_OpenSWATH	openSWATH_guot_l1208_003_combined_lib	2013-04-04-145217	petrik	RUNNING	Abort
LFQ	g45_sing_vs_db244_30ppm	v133	karlassoc	FINISHED	Abort Delete

Jobs Show Resource Settings

Name	Total	Ready	Running	Done	Failed
Analyzer	13			10	
ChromatogramExtractor	13			10	
CompressXML	13			10	
FeatureXMLtoTSV	13			10	
Init	1			1	
MProphet	13			10	
ObisDownload	13			13	
ObisUpload	13			10	
Parallelize	13			13	
RTChromatogramExtractor	13			13	
RTNormalizer	13			10	3
RewriteTSVtoFeatureXML	13			10	

Instances

Show: All Running Done Failed

No.	Status	Actions
9	FINISHED	Show Logs
8	ERROR	Show Logs
7	FINISHED	Show Logs
6	FINISHED	Show Logs
5	FINISHED	Show Logs
4	FINISHED	Show Logs
12	FINISHED	Show Logs
3	FINISHED	Show Logs
2	ERROR	Show Logs
11	FINISHED	Show Logs
10	FINISHED	Show Logs
1	ERROR	Show Logs
0	FINISHED	Show Logs

Logs

Sider Hide Debug and Info messages

```
2013-04-03 09:27:15,240 - CRITICAL - exit code of run_app() != 0
tar: output.in*: Cannot stat: No such file or directory
tar: Exiting with failure status due to previous errors
Can not copy the Output file:output.in
```

Stdout

```
None
Loading TrAM file
RT Normalization working on
/dcluster/scratch_xl/shareholder/imb/ra/workflows/20130403_090408_166623/0/1/RTChromatogramExtractor
(ChromatogramExtractor.chrom.ms4)
Progress of loading chromatograms:
-- done [took 0.07 s (CPU), 0.07 s (Wall)] --
nr transitions 56
Will analyse 10 peptides with a total of 56 transitions
rsq: 0.32542 points: 10
rsq: 0.903462 points: 9
Error: Unexpected internal error (WARNING: rsq: 0.90346167180381909 is below limit of
0.9499999999999999. Validate assays for RT-peptides and adjust the limit for rsq or coverage.)
```

gUSE Log

```
-----
WARNING: The option 'usage[mem=8192]' means
8192 MB per processor 'core', not per compute node.
I hope you know what you are doing...
Generic job:
- Exe started at : Wed Apr 3 09:27:05 CEST 2013
- Exe finished at : Wed Apr 3 09:27:15 CEST 2013
- The exit code of the exe: 8
- An error occurred!
- Ran on host: (hostname) : s9097.hpc-net.ethz.ch
- Host info: (uname -a) : Linux s9097.hpc-net.ethz.ch 2.6.32-279.14.1.el6.x86_64 #1 SMP Tue Nov 6 23:43:09
UTC 2012 x86_64 x86_64 GNU/Linux
- Directory list (ls -la) : -----
total 22
drwxr-xr-x 3 loblum hpc 11 Apr 3 09:27
```



iPortal

- Developed in conjunction with EU SCI-BUS project
- Built on LifeRay and gUSE technologies
- Connects to any infrastructure (cluster, grid, cloud)
- Interfaces to OpenBIS
- Public instance will be built on the cloud
- Team at ETH: Lorenz Blum, Béla Hullár, Emanuel Schmid, Witold Wolski



Many more projects done. Many projects to come, also with new RTDs.

Impact through sustainable, reusable tools and services

For more reusable software see

<http://www.sybit.net/>

SyBIT is an ENABLER

Our success is measured by
YOUR successful publications
(please acknowledge us)



Helpdesk / Contact

- helpdesk@systemsx.ch
- For all collaboration services, wiki etc
- For all supported maintained services, error reporting etc

- Peter.Kunszt@systemsx.ch
- With new projects, ideas, high-level questions, how to get started with a new project



Questions?

- www.systemsx.ch
- www.sybit.net

The screenshot shows the SyBIT website homepage. At the top, there is a navigation bar with links for HOME, SYBIT WIKI, SYBIT PROJECTS, TOOLS, EVENTS, BLOG, and ABOUT US. The main content area features a 'Welcome to SyBIT' section with a group photo and a description of the organization. To the right, there is an 'Upcoming Events' section listing 'SyBIT Tech Day' and 'SIB Days'. Below this, there are three columns: 'Recent Blog Posts', 'Learn More About SyBIT', and 'Our Projects', each with a list of links. The footer contains contact information and a copyright notice.

Welcome to SyBIT

SyBIT provides IT support for SystemsX.ch. It was created by the SystemsX.ch executive board to address the computing and IT needs for their research and technology projects. [Read More...](#)

Upcoming Events

- × [SyBIT Tech Day](#)
January 9, 2013
- × [SIB Days](#)
January 28, 2013 - January 29, 2013
Location: Ble/Bienne

Recent Blog Posts

- × [Tech Day on Proteomics](#)
The SyBIT Tech Day on Proteomics was held yesterday in Bern. I'd like to thank...

Learn More About SyBIT

- × [SyBIT projects](#)
- × [Partners](#)
- × [Collaborations](#)
- × [Contact SyBIT](#)

Our Projects

- × [Ongoing projects](#)
- × [Past projects](#)
- × [Suggest a project](#)
- × [Tools](#)

3/14/13

Address: Clausiusstr. 45, 8052 Zurich | T: +41 44 632 83 30/93 37 | Fax: +41 44 632 15 64 | E-mail: [sybit\[at\]aympa.systemsx.ch](mailto:sybit[at]aympa.systemsx.ch) | Wiki: [SyBIT Wiki](#)

Copyright © 2012, SyBIT - [SystemsX.ch](#) Biology IT.

The screenshot shows the SystemsX.ch website homepage. At the top, there is a navigation bar with links for News, Events & Education, and Contact. The main content area features a large image of a colorful molecular structure with the text 'What is Systems Biology?'. Below this, there is a grid of navigation links for News, Partner Institutions, Projects, Education, and Events. The footer contains contact information and the tagline 'We connect... We support... We interact.'

SystemsX.ch
The Swiss Initiative in Systems Biology

News SystemsX.ch Projects
Events & Education Communication Community
Contact Search

What is Systems Biology? →

News → Partner Institutions →
Projects → Education → Events

Search...

SystemsX.ch
The Swiss Initiative in Systems Biology
Clausiusstrasse 45, CH-8052 Zurich
phone +41 44 632 47 75
fax +41 44 632 15 64
email: admin@systemsx.ch

We connect... We support... We interact.

