FMI

Friedrich Miescher Institute
for Biomedical Research

Setting up a Galaxy Instance as a Service

Nikolay Vazov
Jochen Bick
Hans-Rudolf Hotz

You are given the task
to set up a Galaxy instance for others
(i.e. as a core service in your institute)
and you are not really familiar
with Galaxy.

this is not an ordinary training session

...more like a workshop, with presentations

this is not an ordinary training session

...more like a workshop, with presentations

Hans-Rudolf Hotz	'10 rules'
Jochen Bick	starting problems and experiences
Nikolay Vazov	how can we make simple things more complicated

...and a panel discussion at the end

though we are flexible

What this training session is not about:

- writing tool wrappers

- how to use galaxy

this is not an ordinary training session

'10 rules' for Setting up a Galaxy Instance as a Service

Hans-Rudolf Hotz (hrh@fmi.ch)
Friedrich Miescher Institute for Biomedical Research
Basel, Switzerland

6. July 2015

some info about me:

FMI

Friedrich Miescher Institute
for Biomedical Research

some info about me:

Friedrich Miescher Institute

317 employees
(incl. 95 PhD students, 99 Post Docs)

Epigenetics
(7 research groups)

Cancer
(9 research groups)

Neurobiology
(8 research groups)

Technology Platforms
Computational Biology – Cell Sorting – Imaging and Microscopy – *C. elegans*
Functional Genomics – Histology – Mass Spectrometry – Protein Structure

- **funded by the Novartis Research Foundation**
- **affiliated institute of Basel University**

FMI
Friedrich Miescher Institute
for Biomedical Research

The Computational Biology platform is providing support for....

FMI

Friedrich Miescher Institute
for Biomedical Research

The Computational Biology platform is providing support for....

the “average” lab scientist, using computers to:

- draw plasmids
- do BLAST searches
- use Excel

the “modern” lab scientist, using computers to:

- analyze NGS data with R/Bioconductor scripts

FMI
Friedrich Miescher Institute
for Biomedical Research

The Computational Biology platform is providing support for....

the “average” lab scientist, using computers to:

- draw plasmids
- do BLAST searches
- use Excel

? →

the “modern” lab scientist, using computers to:

- analyze NGS data with R/Bioconductor scripts

FMI
Friedrich Miescher Institute
for Biomedical Research

The Computational Biology platform is providing support for....

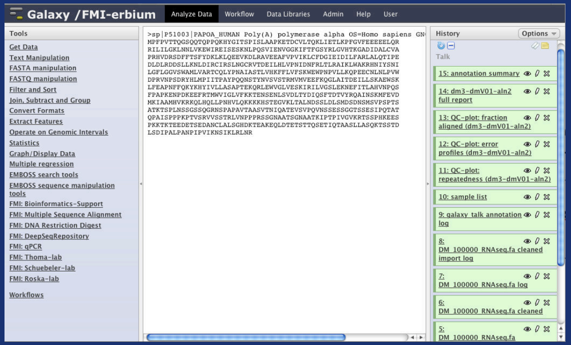
the “average” lab scientist, using computers to:

- draw plasmids
- do BLAST searches
- use Excel

→

the “modern” lab scientist, using computers to:

- analyze NGS data with R/Bioconductor scripts


Galaxy /FMI-erbium
Analyze Data
Workflow
Data Libraries
Admin
Help
User
Tools
Get Data
Text Manipulation
FASTA manipulation
FASTQ manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Operate on Genomic Intervals
Statistics
Graph/Display Data
Multiple regression
EMBOSS search tools
EMBOSS sequence manipulation tools
FMI: Bioinformatics-Support
FMI: Multiple Sequence Alignment
FMI: DNA Restriction Digest
FMI: DeepSeqRepository
FMI: qPCR
FMI: Thoma-lab
FMI: Schuebeler-lab
FMI: Roska-lab
Workflows
History
Options
Talk
15: annotation summary
14: dm3-dmV01-aln2 full report
13: QC-plot: fraction aligned (dm3-dmV01-aln2)
12: QC-plot: error profiles (dm3-dmV01-aln2)
11: QC-plot: repeatedness (dm3-dmV01-aln2)
10: sample list
9: galaxy_talk annotation log
8: DM_100000_RNAseq.fa cleaned import log
7: DM_100000_RNAseq.fa log
6: DM_100000_RNAseq.fa cleaned
5: DM_100000_RNAseq.fa

```
>sp|P51003|PAPOA_HUMAN Poly(A) polymerase alpha OS=Homo sapiens GN
MPFPVTTQGSQQTQPPQKHYGITSPISLAAPKETDCVLTQKLIETLKPFGVFEEEEELQR
RILILGKLNNLVKEWIREISESKNLPQSVIENVGGKIFTFGSYRLGVHTKGADIDALCVA
PRHVDRSDFFTSFYDKLKLQEEVKDLRAVEEAFVPVIKLCFDGIEIDILFARLALQTIPE
DLDLRDDSLLKNLDIRCIRSLNGCRVTDEILHLVPNIDNFRLTLRAIKLWAKRHNIYSNI
LGFLGGVSWAMLVARTCQLYPNAIASTLVHKFFLVFSKWEWPNPVLLKQPEECNLNLPVW
DPRVNPSDRYHLMPIITPAYPQQNSTYNVSVSTRMVMVEEFKQGLAITDEILLSKAEWSK
LFEAPNFFQKYKHYIVLLASAPTEKQRLEWVGLVESKIRILVGSLEKNEFITLAHVNPQS
FPAPKENPDKEEFRTMWVIGLVFKKTENSENLSVDLTYDIQSFTDTVYRQAINSKMFEVD
MKIAAMHVKRKQLHQLLPNHVLQKKKKHSTEGVKLTALNDSSLDLSMDSDNSMSVPSPTS
ATKTSPLNSSGSSQGRNSPAPAVTAASVTNIQATEVSVPQVNSSESSGGTSSESIPQTAT
QPAISPPPKPTVSRVVSSTRLVNPPPRSSGNAATSGNAATKIPTPIVGVKRTSSPHKEES
PKKTKTEEDETSEDANCLALSGHDKTEAKEQLDTETSTTQSETIQTAASLLASQKTSSTD
LSDIPALPANPIPVIKNSIKLRLNR
```

http://galaxyproject.org/

FMI
Friedrich Miescher Institute
for Biomedical Research

why are we using Galaxy

- open source software / no license fee
- it provides a standard set of Bioinformatics tools
- we can add our own scripts and tools
- in addition to the ~15 core developers, there is a huge world wide community
- a local installation is simple to set up
- it is flexible (you can adjust it to your needs)

FMI
Friedrich Miescher Institute
for Biomedical Research

why are we using Galaxy

- open source software / no license fee
- it provides a standard set of Bioinformatics tools
- we can add our own scripts and tools
- in addition to the ~15 core developers, there is a huge world wide community
- a local installation is simple to set up
- it is flexible (you can adjust it to your needs)

in use at the FMI since early 2008 (2007)

FMI
Friedrich Miescher Institute
for Biomedical Research

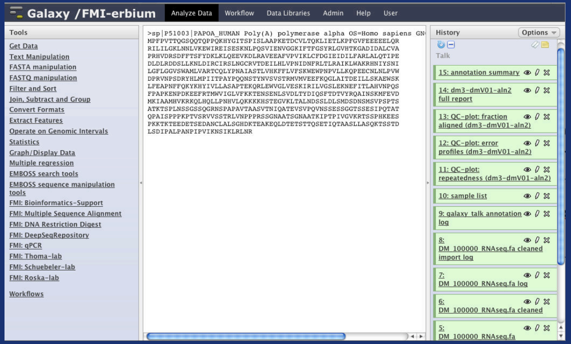
Galaxy as a stepping stone

the “average” lab scientist, using computers to:

- draw plasmids
- do BLAST searches
- use Excel

the “modern” lab scientist, using computers to:

- analyze NGS data with R/Bioconductor scripts


Galaxy /FMI-erbium
Analyze Data
Workflow
Data Libraries
Admin
Help
User
Tools
Get Data
Text Manipulation
FASTA manipulation
FASTQ manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Operate on Genomic Intervals
Statistics
Graph/Display Data
Multiple regression
EMBOSS search tools
EMBOSS sequence manipulation tools
FMI: Bioinformatics-Support
FMI: Multiple Sequence Alignment
FMI: DNA Restriction Digest
FMI: DeepSeqRepository
FMI: qPCR
FMI: Thoma-lab
FMI: Schuebeler-lab
FMI: Roska-lab
Workflows
History
Options
Talk
15: annotation summary
14: dm3-dmV01-aln2 full report
13: QC-plot: fraction aligned (dm3-dmV01-aln2)
12: QC-plot: error profiles (dm3-dmV01-aln2)
11: QC-plot: repeatedness (dm3-dmV01-aln2)
10: sample list
9: galaxy_talk annotation log
8: DM_100000_RNAseq.fa cleaned import log
7: DM_100000_RNAseq.fa log
6: DM_100000_RNAseq.fa cleaned
5: DM_100000_RNAseq.fa

```
>sp|P51003|PAPOA_HUMAN Poly(A) polymerase alpha OS=Homo sapiens GN
MPFPVTTQGSQQTQPPQKHYGITSPISLAAPKETDCVLTQKLIETLKPFGVFEEEEELQR
RILILGKLNNLVKEWIREISESKNLPQSVIENVGGKIFTFGSYRLGVHTKGADIDALCVA
PRHVDRSDFFTSFYDKLKLQEEVKDLRAVEEAFVPVIKLCFDGIEIDILFARLALQTIPE
DLDLRDDSLLKNLDIRCIRSLNGCRVTDEILHLVPNIDNFRLTLRAIKLWAKRHNIYSNI
LGFLGGVSWAMLVARTCQLYPNAIASTLVHKFFLVFSKWEWPNPVLLKQPEECNLNLPVW
DPRVNPSDRYHLMPIITPAYPQQNSTYNVSVSTRMVMVEEFKQGLAITDEILLSKAEWSK
LFEAPNFFQKYKHYIVLLASAPTEKQRLEWVGLVESKIRILVGSLEKNEFITLAHVNPQS
FPAPKENPDKEEFRTMWVIGLVFKKTENSENLSVDLTYDIQSFTDTVYRQAINSKMFEVD
MKIAAMHVKRKQLHQLLPNHVLQKKKHSTEGVKLTALNDSSLDLSMDSDNSMSVPSPTS
ATKTSPLNSSGSSQGRNSPAPAVTAASVTNIQATEVSVPQVNSSESSGGTSSESIPQTAT
QPAISPPPKPTVSRVVSSTRLVNPPPRSSGNAATSGNAATKIPTPIVGVKRTSSPHKEES
PKKTKTEEDETSEDANCLALSGHDKTEAKEQLDTETSTTQSETIQTAASLLASQKTSSTD
LSDIPALPANPIPVIKNSIKLRLNR
```

http://galaxyproject.org/

FMI

Friedrich Miescher Institute
for Biomedical Research

Galaxy as a teaching tool

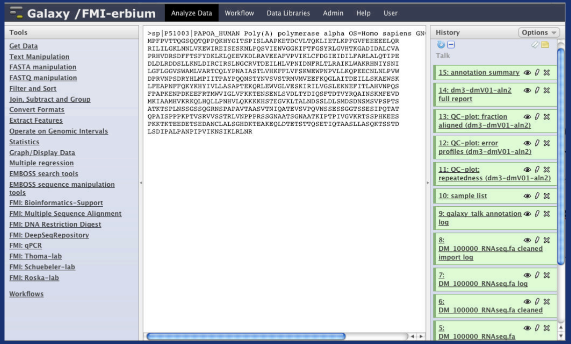
for learning Bioinformatics....

the “average” lab scientist, using computers to:

- draw plasmids
- do BLAST searches
- use Excel

the “modern” lab scientist, using computers to:

- analyze NGS data with R/Bioconductor scripts


Galaxy /FMI-erbium
Analyze Data
Workflow
Data Libraries
Admin
Help
User
Tools
Get Data
Text Manipulation
FASTA manipulation
FASTQ manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Operate on Genomic Intervals
Statistics
Graph/Display Data
Multiple regression
EMBOSS search tools
EMBOSS sequence manipulation tools
FMI: Bioinformatics-Support
FMI: Multiple Sequence Alignment
FMI: DNA Restriction Digest
FMI: DeepSeqRepository
FMI: qPCR
FMI: Thoma-lab
FMI: Schuebeler-lab
FMI: Roska-lab
Workflows
History
Options
Talk
15: annotation summary
14: dm3-dmV01-aln2 full report
13: QC-plot: fraction aligned (dm3-dmV01-aln2)
12: QC-plot: error profiles (dm3-dmV01-aln2)
11: QC-plot: repeatedness (dm3-dmV01-aln2)
10: sample list
9: galaxy_talk annotation log
8: DM_100000_RNAseq.fa cleaned import log
7: DM_100000_RNAseq.fa log
6: DM_100000_RNAseq.fa cleaned
5: DM_100000_RNAseq.fa

```
>sp|P51003|PAPOA_HUMAN Poly(A) polymerase alpha OS=Homo sapiens GN
MPFPVTTQGSQQTQPPQKHYGITSPISLAAPKETDCVLTQKLIETLKPFGVFEEEEELQR
RILILGKLNNLVKEWIREISESKNLPQSVIENVGGKIFTFGSYRLGVHTKGADIDALCVA
PRHVDRSDFFTSFYDKLKLQEEVKDLRAVEEAFVPVIKLCFDGIEIDILFARLALQTIPE
DLDLRDDSLLKNLDIRCIRSLNGCRVTDEILHLVPNIDNFRLTLRAIKLWAKRHNIYSNI
LGFLGGVSWAMLVARTCQLYPNAIASTLVHKFFLVFSKWEWPNPVLLKQPEECNLNLPVW
DPRVNPSDRYHLMPIITPAYPQQNSTYNVSVSTRMVMVEEFKQGLAITDEILLSKAEWSK
LFEAPNFFQKYKHYIVLLASAPTEKQRLEWVGLVESKIRILVGSLEKNEFITLAHVNPQS
FPAPKENPDKEEFRTMWVIGLVFKKTENSENLSVDLTYDIQSFTDTVYRQAINSKMFEVD
MKIAAMHVKRKQLHQLLPNHVLQKKKKHSTEGVKLTALNDSSLDLSMDSDNSMSVPSPTS
ATKTSPLNSSGSSQGRNSPAPAVTAASVTNIQATEVSVPQVNSSESSGGTSSESIPQTAT
QPAISPPPKPTVSRVVSSTRLVNPPPRSSGNAATSGNAATKIPTPIVGVKRTSSPHKEES
PKKTKTEEDETSEDANCLALSGHDKTEAKEQLDTETSTTQSETIQTAASLLASQKTSSTD
LSDIPALPANPIPVIKNSIKLRLNR
```

http://galaxyproject.org/

... which is more than pressing a red button

FMI

Friedrich Miescher Institute
for Biomedical Research

the FMI Galaxy Server

single (dedicated) multi-core box

- 16 cpu (four quad-core Intel X7350)
- 128GB RAM
- python 2.6.5
- 34 TB local attached storage

connected to a MySQL database

external authentication

FMI
Friedrich Miescher Institute
for Biomedical Research

the FMI Galaxy Server

users

- 238 registered users
- 20 'heavy users'
- 30 'ocassional users'

jobs

~ 500 jobs/month

- NGS and MA analysis

FMI

Friedrich Miescher Institute
for Biomedical Research

disclaimer

I will talk about my experience over the last 7 years

FMI

Friedrich Miescher Institute
for Biomedical Research

disclaimer

I will talk about my experience over the last 7 years

- this might not be up-to-date
- our Galaxy server is heavily adjusted to our need

FMI

Friedrich Miescher Institute
for Biomedical Research

10 rules for setting up a galaxy instance as a service

FMI

Friedrich Miescher Institute
for Biomedical Research

10 rules for setting up a galaxy instance as a service

Check: what are you actually asked for

Check: what resources do you have / need

Follow the suggestion on the wiki

Set up only what you have been asked to

Know the tools you offer

Prevent data duplication

Set up 'reports'

Offer training

Keep the Galaxy software (and you) up to date

Adjust your Galaxy server to changes in requirements

1) Check: what are you actually asked for

1) Check: what are you actually asked for

- talk to the person(s) who contacted you

    - why Galaxy ?

Friedrich Miescher Institute
for Biomedical Research

1) Check: what are you actually asked for

- talk to the people who will use your service

    - using Galaxy for what?
    - do they know use.galaxy.org ?
    - alternatives ?

FMI

Friedrich Miescher Institute
for Biomedical Research

1) Check: what are you actually asked for

- talk to the people who will use your service

    - using Galaxy for what?

    - do they know use.galaxy.org ?

    - alternatives ?

→ define the tools

1) Check: what are you actually asked for

- talk to the people who will use your service
  - using Galaxy for what?
  - do they know use.galaxy.org ?
  - alternatives ?

→ define the tools

→ use the toolshed

→ develop your own tools

1) Check: what are you actually asked for

- talk to potential users

  - what are they using now ?
  - do they know use.galaxy.org ?
  - are there overlaps with the initial request ?

FMI

Friedrich Miescher Institute
for Biomedical Research

1) Check: what are you actually asked for

- is Galaxy the right tool ?

  - Galaxy is not Bioinformatician

  - Galaxy might be too 'big' for the task

  - Galaxy is not (yet) a LIMS

  (- Galaxy is not good for 1000 of repeated jobs)

  - another system is already in place

1) Check: what are you actually asked for

Visibility

- internal web site
- public web site

Access

- everybody can create an account
- accounts are created for the users
- external authentication

2) Check: what resources do you need / have ?

FMI

Friedrich Miescher Institute
for Biomedical Research

2) Check: what resources do you need / have ?

- hardware
  - cpu / memory
  - storage
    - fast (local) storage
    - slow (network) storage

2) Check: what resources do you need / have ?

- people / knowledge
  - system administration for the Galaxy serer
  - Bioinformatics background

select the option(s) which fits your requirements and resources

- use Main (usegalaxy.org)
- use another public galaxy server

  https://wiki.galaxyproject.org/PublicGalaxyServers

- install galaxy locally

- use galaxy on the cloud

- get 'SlipStream' galaxy appliance

https://wiki.galaxyproject.org/BigPicture/Choices

3) Follow the suggestion on the wiki

3) Follow the suggestion on the wiki

read them first......before you set up a production server

https://wiki.galaxyproject.org/Admin

https://wiki.galaxyproject.org/Admin/Config/Performance/ProductionServer

3) Follow the suggestion on the wiki

read them first……before you set up
a production server

https://wiki.galaxyproject.org/Admin

https://wiki.galaxyproject.org/Admin/Config/Performance/ProductionServer

switching to a database server

SQLite → PostgreSQL

Friedrich Miescher Institute
for Biomedical Research

3) Follow the suggestion on the wiki

make sure you are admin

https://wiki.galaxyproject.org/Admin/Interface

3) Follow the suggestion on the wiki

remove 'deleted' datasets

https://wiki.galaxyproject.org/Admin/Config/Performance/Purge%20Histories%20and%20Datasets

→ **set up a cron job**

3) Follow the suggestion on the wiki

setup 'Trackster"

https://wiki.galaxyproject.org/VisualizationSetup

3) Follow the suggestion on the wiki

check other installations:

https://wiki.galaxyproject.org/Community/Deployments

Galaxy Community Log Board:
(a place to share how you addressed a particular task)

https://wiki.galaxyproject.org/Community/Logs

FMI
Friedrich Miescher Institute
for Biomedical Research

3) Follow the suggestion on the wiki

check other installations:

https://wiki.galaxyproject.org/Community/Deployments

Galaxy Community Log Board:
(a place to share how you addressed a particular task)

https://wiki.galaxyproject.org/Community/Logs

→ and add your stuff as well

FMI
Friedrich Miescher Institute
for Biomedical Research

4) Set up only what you have been asked to

Friedrich Miescher Institute
for Biomedical Research

4) Set up only what you have been asked to

.....at least in the beginning:
don't confuse your clients with too many tools

4) Set up only what you have been asked to

.....at least in the beginning:
don't confuse your clients with too many tools

offer group/user specific tools

https://wiki.galaxyproject.org/UserDefinedToolboxFilters
https://wiki.galaxyproject.org/Admin/Config/Access%20Control

4) Set up only what you have been asked to

.....at least in the beginning:
don't confuse your clients with too many tools

offer group/user specific tools

https://wiki.galaxyproject.org/UserDefinedToolboxFilters
https://wiki.galaxyproject.org/Admin/Config/Access%20Control

→ good for testing

4) Set up only what you have been asked to

.....but make sure you have:

production server / development server

4) Set up only what you have been asked to

.....but make sure you have:

production server / development server

and the production server backed-up
(including the database server)

5) Know the tools you offer

Friedrich Miescher Institute
for Biomedical Research

5) Know the tools you offer

First, make sure you know how to use galaxy

https://wiki.galaxyproject.org/Learn

5) Know the tools you offer

First, make sure you know how to use galaxy

https://wiki.galaxyproject.org/Learn

Second, understand the tools you offer

- **can you execute them on the command line**

6) Set up 'reports'

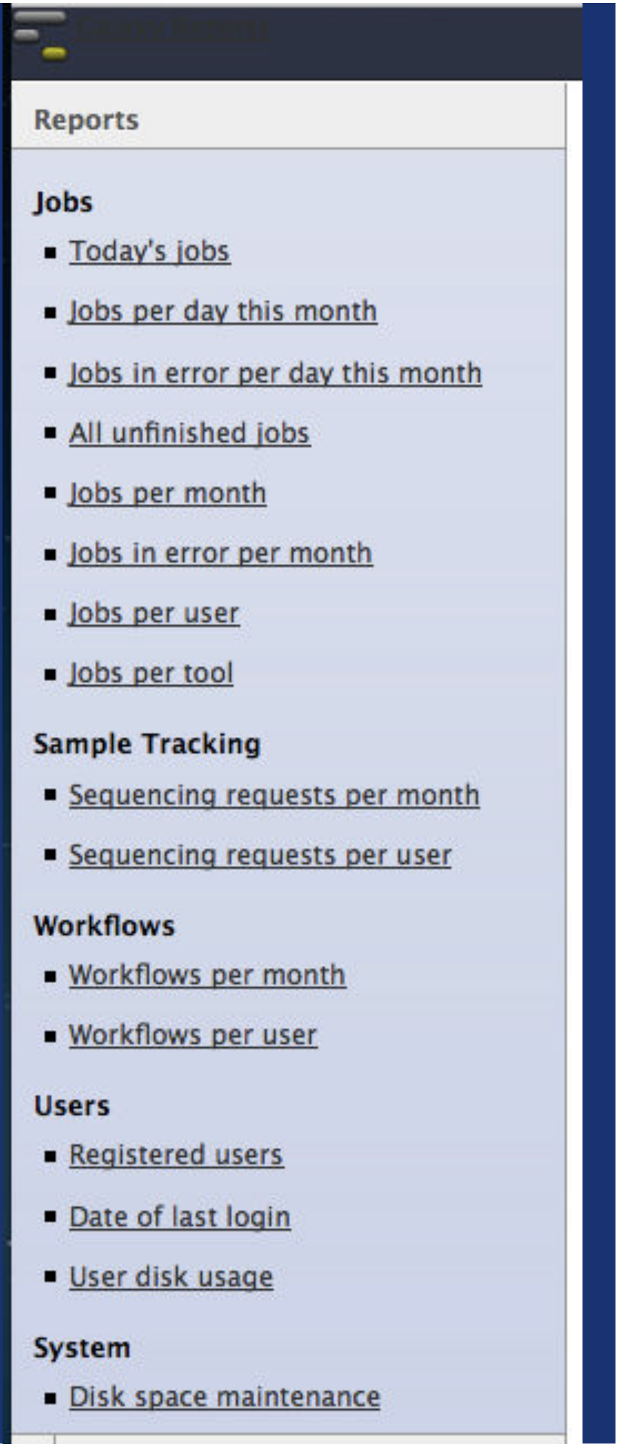
Friedrich Miescher Institute
for Biomedical Research

6) Set up 'reports'

https://wiki.galaxyproject.org/Admin/UsageReports

6) Set up 'reports'

it is a second web site



Reports

Jobs

- Today's jobs
- Jobs per day this month
- Jobs in error per day this month
- All unfinished jobs
- Jobs per month
- Jobs in error per month
- Jobs per user
- Jobs per tool

Sample Tracking

- Sequencing requests per month
- Sequencing requests per user

Workflows

- Workflows per month
- Workflows per user

Users

- Registered users
- Date of last login
- User disk usage

System

- Disk space maintenance

6) Set up 'reports'

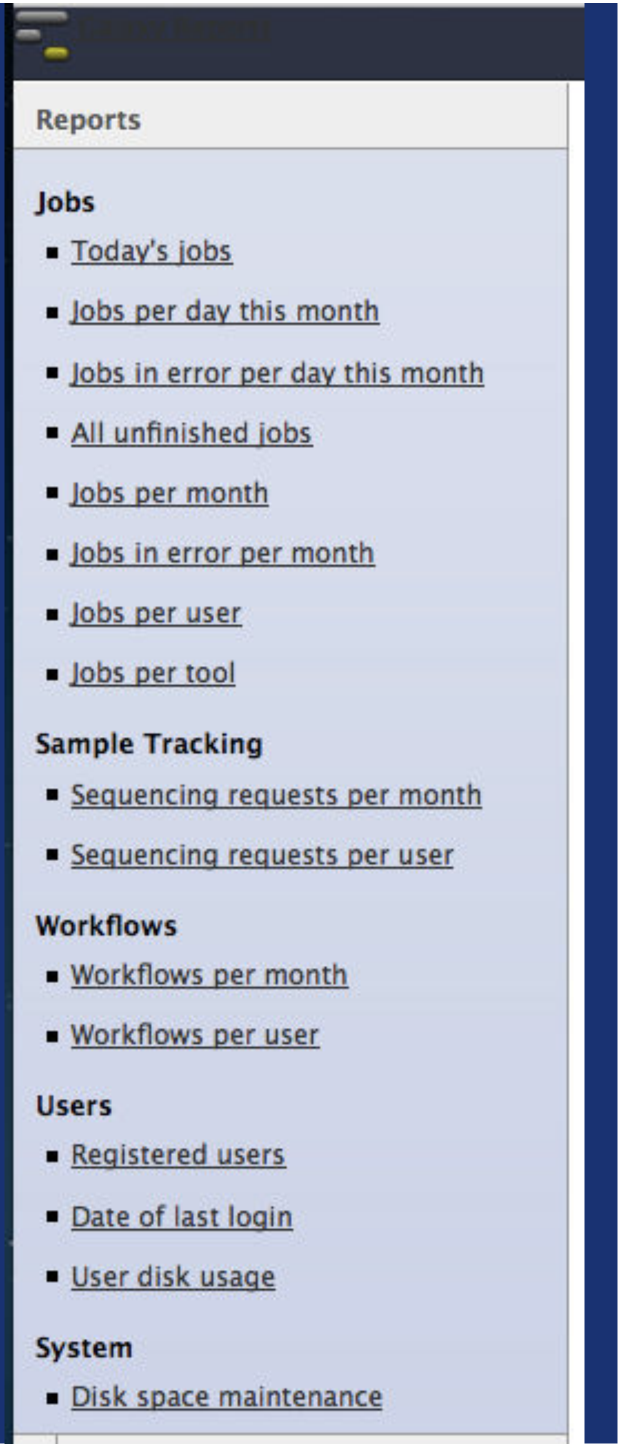
Today's jobs

All unfinished jobs

Jobs per tool

Jobs per user

User disk usage



Galaxy Reports

Reports

Jobs

- Today's jobs
- Jobs per day this month
- Jobs in error per day this month
- All unfinished jobs
- Jobs per month
- Jobs in error per month
- Jobs per user
- Jobs per tool

Sample Tracking

- Sequencing requests per month
- Sequencing requests per user

Workflows

- Workflows per month
- Workflows per user

Users

- Registered users
- Date of last login
- User disk usage

System

- Disk space maintenance

6) Set up 'reports'

....learn about the database

Friedrich Miescher Institute
for Biomedical Research

6) Set up 'reports'

....learn about the database:

- execute queries which are not covered by 'reports'

6) Set up 'reports'

....learn about the database:

- execute queries which are not covered by 'reports'

- *fix* the database

FMI
Friedrich Miescher Institute
for Biomedical Research

6) Set up 'reports'

....learn about the database:

- execute queries which are not covered by 'reports'

- *fix* the database

I have not recommended this to you

FMI

Friedrich Miescher Institute
for Biomedical Research

7) Prevent data duplication

Friedrich Miescher Institute
for Biomedical Research

7) Prevent data duplication

a quote from a bioinformatics mailing list:

“We don’t want to use Galaxy because it produces to much data”

7) Prevent data duplication

a quote from a bioinformatics mailing list:

> *"We don't want to use Galaxy because it produces to much data"*

Galaxy can help you reducing the storage requirements

7) Prevent data duplication

use 'Data Libraries'

https://wiki.galaxyproject.org/Admin/DataLibraries

FMI

Friedrich Miescher Institute
for Biomedical Research

7) Prevent data duplication

use 'Data Libraries'

https://wiki.galaxyproject.org/Admin/DataLibraries

- *'Link to files without copying into Galaxy,*
- enable *'Upload files from filesystem paths'*

https://wiki.galaxyproject.org/Admin/DataLibraries/UploadingLibraryFiles

FMI

Friedrich Miescher Institute
for Biomedical Research

7) Prevent data duplication

promote history sharing

promote Galaxy 'pages'

7) Prevent data duplication

promote history sharing

promote Galaxy 'pages'

allow user to see the full path of datasets
(`expose_dataset_path = True`)

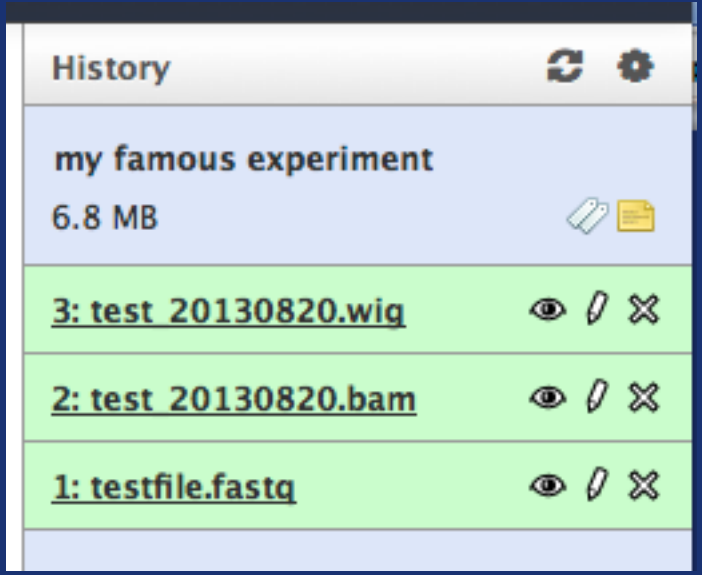
7) Prevent data duplication

use 'external' data

7) Prevent data duplication

a simple NGS analysis:

fastq file → BAM file → wig file



History

my famous experiment

6.8 MB

3: test_20130820.wig

2: test_20130820.bam

1: testfile.fastq

FMI

Friedrich Miescher Institute
for Biomedical Research

7) Prevent data duplication

storing data outside of Galaxy:

- raw data (fastq) files are in central/group specific repositories
- the Galaxy ‘aligner’ knows the location of the fastq files and stores the BAM file again in a group specific repository and creates just a 'log file' as history item
- the Galaxy ‘count’ tool uses the 'log file' as input

7) Prevent data duplication

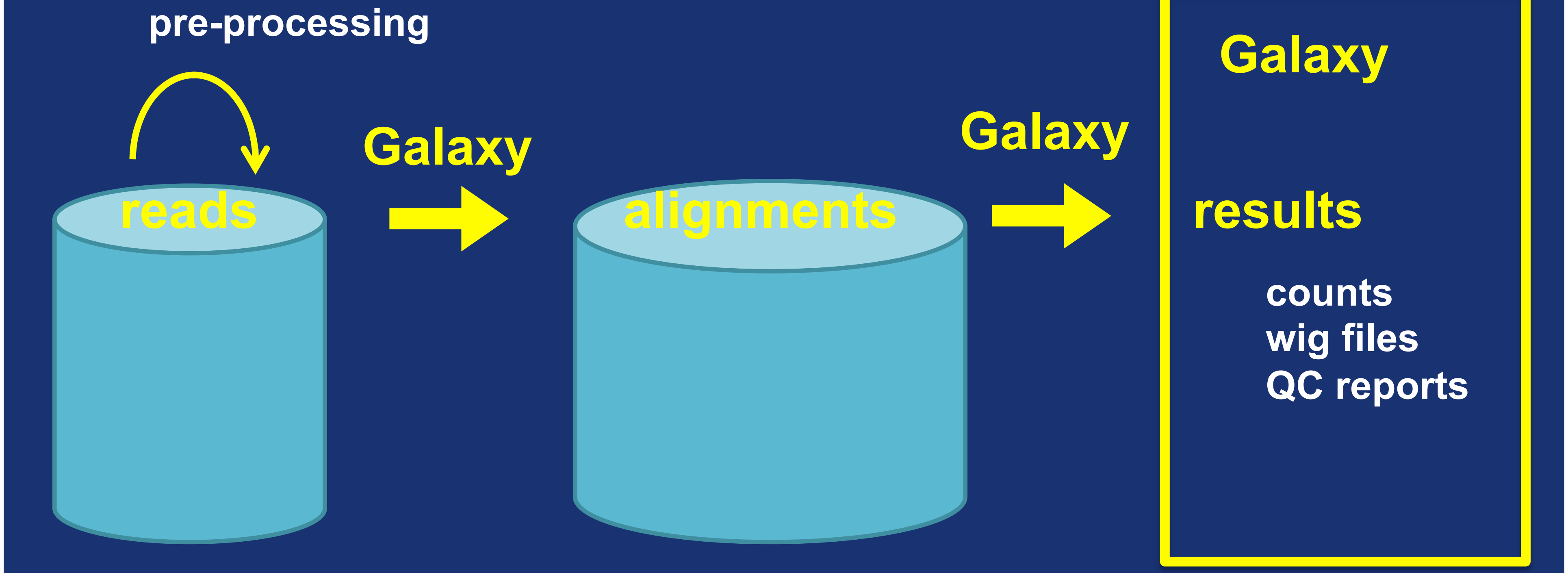
storing data outside of Galaxy:

- raw data (fastq) files are in central/group specific repositories
- the Galaxy ‘aligner’ knows the location of the fastq files and stores the BAM file again in a group specific repository and creates just a 'log file' as history item
- the Galaxy ‘count’ tool uses the 'log file' as input

this is not really best (Galaxy) practice, but it allows to collaborate with non-Galaxy users ....and reproducibility is still guaranteed

FMI
Friedrich Miescher Institute
for Biomedical Research

7) Prevent data duplication



pre-processing

reads

Galaxy

alignments

Galaxy

Galaxy

results

counts
wig files
QC reports

group specific storage
(accessible on the command line)

8) Offer training

Friedrich Miescher Institute
for Biomedical Research

8) Offer training

individual training

run training courses

https://wiki.galaxyproject.org/Teach

8) Offer training

individual training

run training courses

https://wiki.galaxyproject.org/Teach

→ stress testing for the server

Friedrich Miescher Institute
for Biomedical Research

9) Keep your server (and you) up to date

Friedrich Miescher Institute
for Biomedical Research

9) Keep your server up to date

unless you have a very good reason, make sure your are running the latest (or at least a recent) code version

- it is easier for others to help you
- the reported issue might already be fixed in the current release

9) Keep your server up to date

unless you have a very good reason, make sure your are running the latest (or at least a recent) code version

- it is easier for others to help you
- the reported issue might already be fixed in the current release

find a balance between updates (with new or different features) and continuity

we do 3 update per year

FMI

Friedrich Miescher Institute
for Biomedical Research

9) Keep your server up to date

doing an update is easy

9) Keep your server up to date

doing an update is easy ....most of the time

9) Keep your server up to date

doing an update is easy ....most of the time

- announce the down time one week in advance
- install a new server
- update the 'new server' from last time
- update the development server
- update the production server

9) Keep your server up to date

doing an update is easy ....most of the time

- announce the down time one week in advance
- install a new server
- update the 'new server' from last time
- update the development server
- update the production server

goal: minimize the down time

FMI

Friedrich Miescher Institute
for Biomedical Research

9) Keep yourself up to date

Friedrich Miescher Institute
for Biomedical Research

9) Keep yourself up to date

- read the DevNewsBriefs

  https://wiki.galaxyproject.org/DevNewsBriefs

- follow the mailing lists

  https://wiki.galaxyproject.org/MailingLists

- join the "Galaxy Admins"

  https://wiki.galaxyproject.org/Community/GalaxyAdmins

  BOF: Tuesday, 7 July, 18:20 Franklin Room, JICCC

- go to GCC2016

- form regional communities

9) Keep yourself up to date

take the time to look at new features

10) Adjust your Galaxy server to changes in requirements

Friedrich Miescher Institute
for Biomedical Research

10) Adjust your Galaxy server

you don't need a cluster to set up different queues

10) Adjust your Galaxy server

you don't need a cluster to set up different queues

→ job.conf.xml

`job_conf.xml.sample_advanced`

https://wiki.galaxyproject.org/Admin/Config/Jobs

FMI

Friedrich Miescher Institute
for Biomedical Research

10) Adjust your Galaxy server

you can change the hardware

10) Adjust your Galaxy server

you can change the hardware, as long as you keep the ‘database/’ directory and the SQL DB in sync.

10) Adjust your Galaxy server

you can change the hardware, as long as you keep the ‘database/’ directory and the SQL DB in sync.

old recipe:

- make a copy of the SQL DB
- copy the complete galaxy directory to the new server (make sure you keep the path)
- point the new galaxy server to the MySQL DB copy and start it
  - -> due to the higher Python version, news eggs were downloaded
  - -> all python code was re-compiled
- test the new server (while the old one is still in use)
- stop the old server
- rsync ~/galaxy_dist/database/files/
- point the new galaxy server to the ‘live’ MySQL DB and re-start it

10) Adjust your Galaxy server

you can change the database server

Friedrich Miescher Institute
for Biomedical Research

10) Adjust your Galaxy server

you can change the database server

MySQL → PostgreSQL

10) Adjust your Galaxy server

you can change the database server

MySQL → PostgreSQL

recipe will be posted on

https://wiki.galaxyproject.org/Community/Logs

FMI
Friedrich Miescher Institute
for Biomedical Research

Friedrich Miescher Institute
for Biomedical Research

Acknowledgment

Computational Biology

- Michael Stadler / Christian Hundsrücker

Functional Genomics

- Tim Roloff

IT Support

- Stefan Grzybeck

....and all the people from the "Galaxy"

hrh@fmi.ch
@hrhotz

FMI
Friedrich Miescher Institute
for Biomedical Research