

# “Lightweight” HPC Client

*Leveraging Galaxy Tool configs*

Samuel Lampa, UPPMAX, Uppsala University  
[samuel.lampa@gmail.com](mailto:samuel.lampa@gmail.com)  
[twitter.com/smllmp](http://twitter.com/smllmp)

# Background

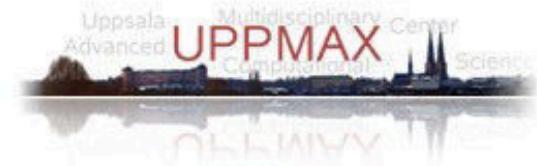
- Want to re-use tool wrapping work already done in Galaxy
- Sysadmins with ... strict ... policies
- Only accepted authentication: SSH
- :/

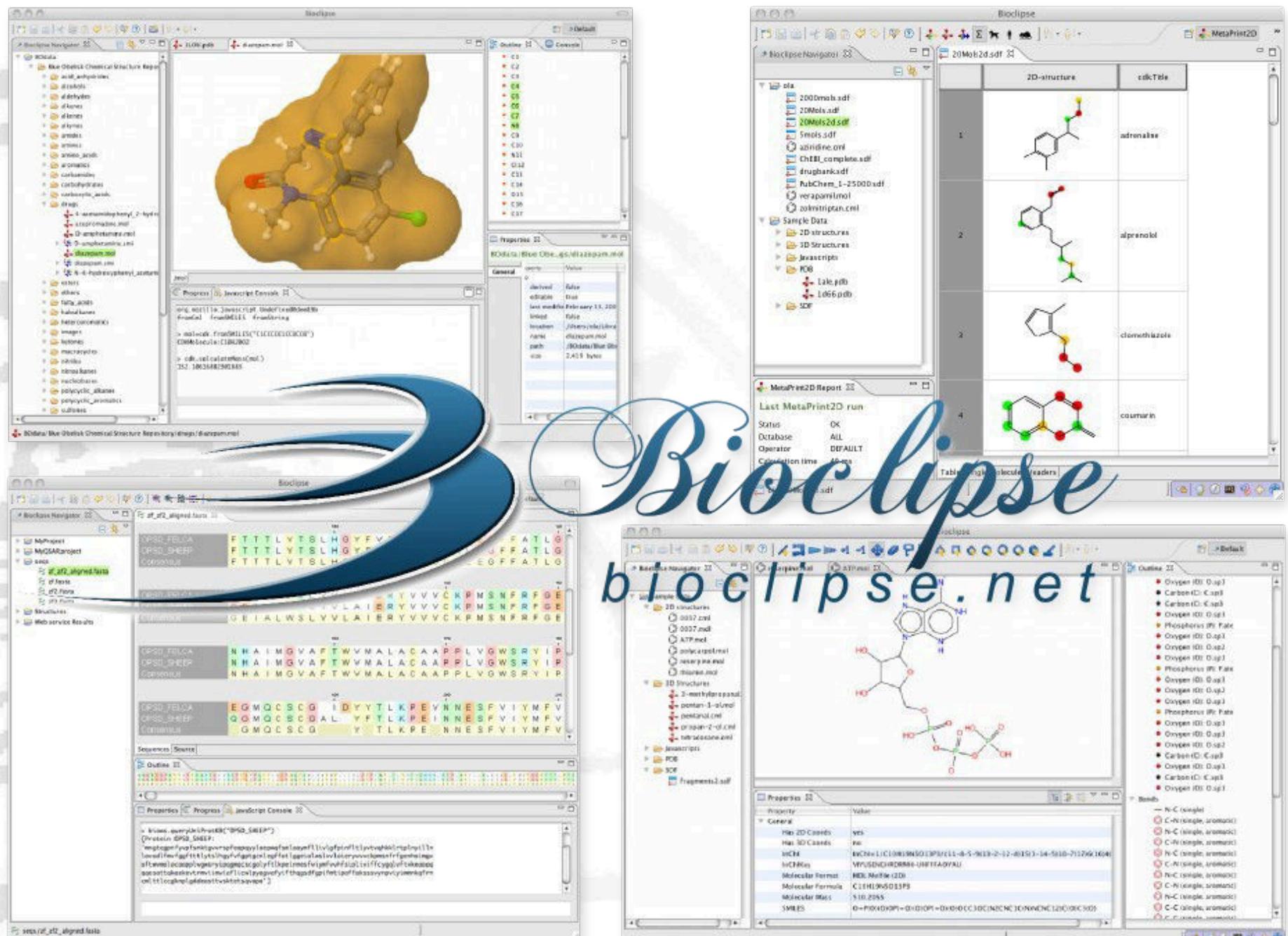
# Solution

- Use SSH!
- Use (Galaxy) Toolconfigs!
- :)



UPPSALA  
UNIVERSITET

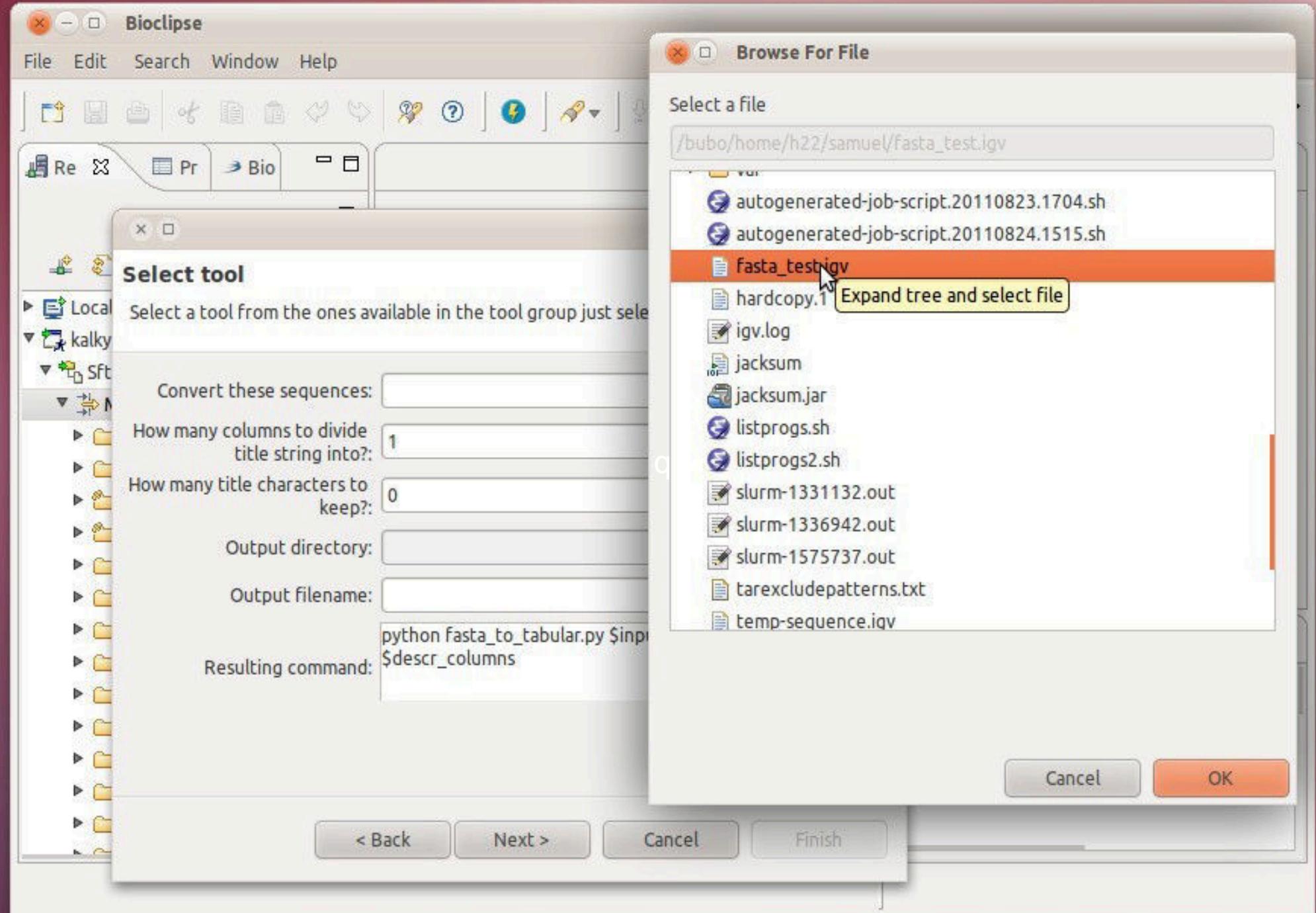


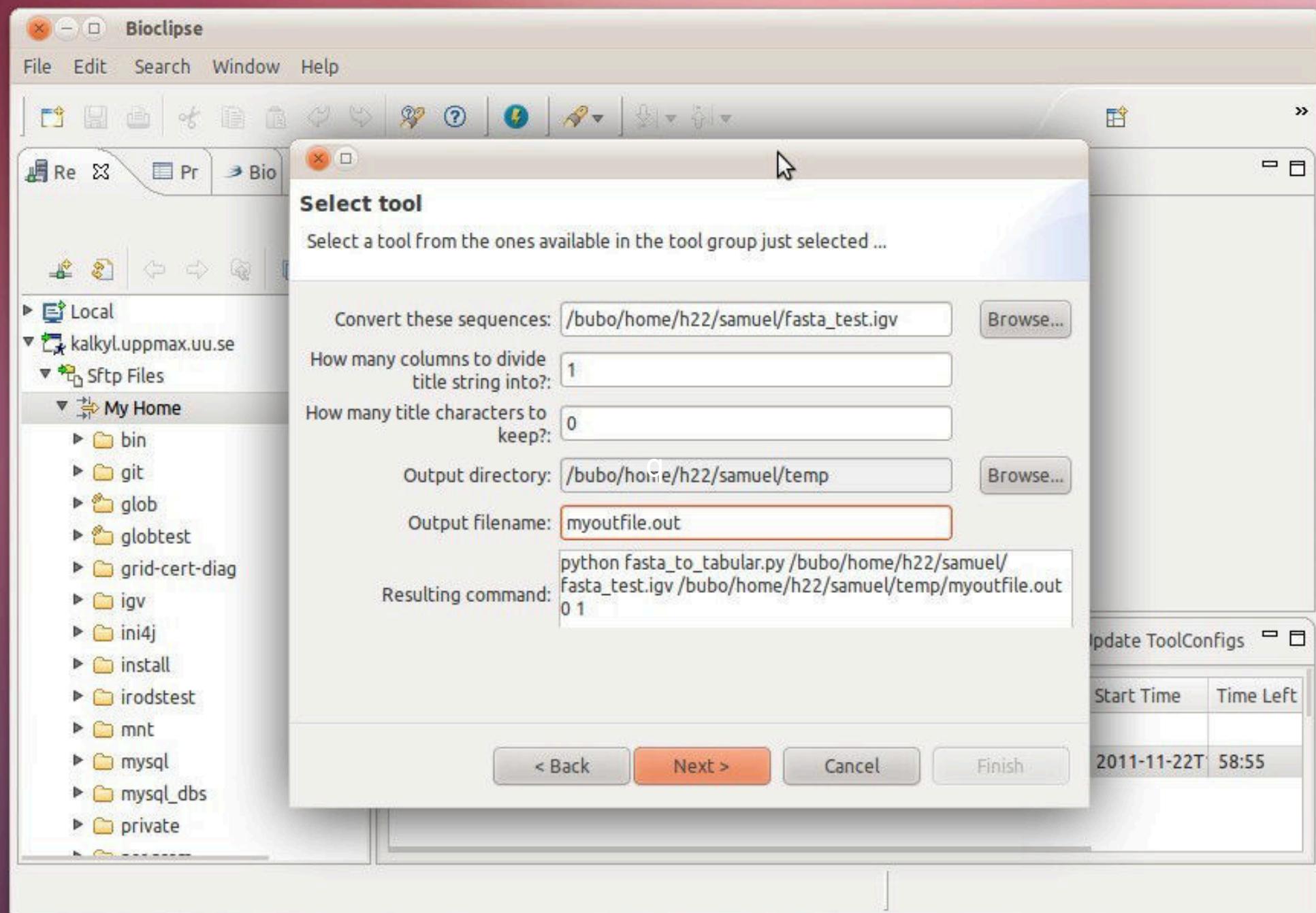


# Bioclipse

[bioclipse.net](http://bioclipse.net)

UPPNE





**Bioclipse**

File Edit Search Window Help

Re Pr Bio

Local kalkyl.uppmax.uu.se Sftp Files My Home bin git glob globtest grid-cert-diag igv ini4j install irodstest mnt mysql mysql\_dbs private

**Configure SBATCH parameters**

Set the parameters for the SBATCH job batch script to be sent to the SLURM resource manager

HPC Module to load:

Project to account: uppmax\_staff

Partition (type of job): core

No of Nodes: core

No of CPUs: node  
devel

Running time (d-hh:mm:ss):

Activate --qos=short option?: no

Job name: Untitled

```
#SBATCH -A uppmax_staff
#SBATCH -p core
#SBATCH -N [noofnodes]
#SBATCH -n [noofcpus]
#SBATCH -t [runtime]
#SBATCH -J Untitled

module load bioinfo-tools
[modulename]

perl #if $pileup_type.type_select ==
"six" #pileup_parser.pl $input "3" "5"
```

< Back Next > Cancel Finish

Date ToolConfigs

Start Time	Time Left
2011-11-22T	58:55

**Bioclipse**

File Edit Search Window Help

IGV Editor X

**IGV Genome Viewer**

Human hg19 chr4 chr4:34,897,325-96,612,848 Go

72 mb

40 mb 50 mb 60 mb 70 mb 80 mb 90 mb 100 mb

p15.33 p15.1 p12 q13.1 q21.1 q22.2 q25 q27 q28.3 q31.3 q33 q35.1

RefSeq Genes RAP2 BEND4 CORIN SGCB PPAT LPHN3 STAP1 GC USO1 BMP3 PTPN13 GRID2 EIF4E CXOC4

chr4:36,399,251 245.1 OT 294M

**Submit new jobs**

**View Running Jobs**

JobInfo X Refresh Submit New Job Update ToolConfigs

Job ID	Partition	Job Name	User	Project	Status	Start Time	Time Left	Nodes
1591866	devel	_interactive	samuel	uppmax_staff	R	2011-11-22T	58:55	1

Browse files

- private
- program
- projects
- R
- refcards
- sandboxes
- scripts
- src
- sw
- temp
- test
- tmp
- uppnex-ga
- usr
- var
- autogenerated-job-script.201
- autogenerated-job-script.201
- fasta\_test.igv
- hardcopy.1
- igv.log
- jacksum
- jacksum.iac

**[MOCKUP IMAGE]**

Bioclipse

File Edit Search Window Help

Galaxy via SSH-tunneled browser

Galaxy Analyze Data Workflow Shared Data Admin Help User

Tools Options

- 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

Configure and start batch jobs

Job runner parameters

Native job runner options

Project b2010074

Max job time limit 3-00:00:00

Format: (days-hh:mm:ss)  
This specifies the maximum running time for a job, helps the job manager decide on better scheduling strategies for jobs.

Partition

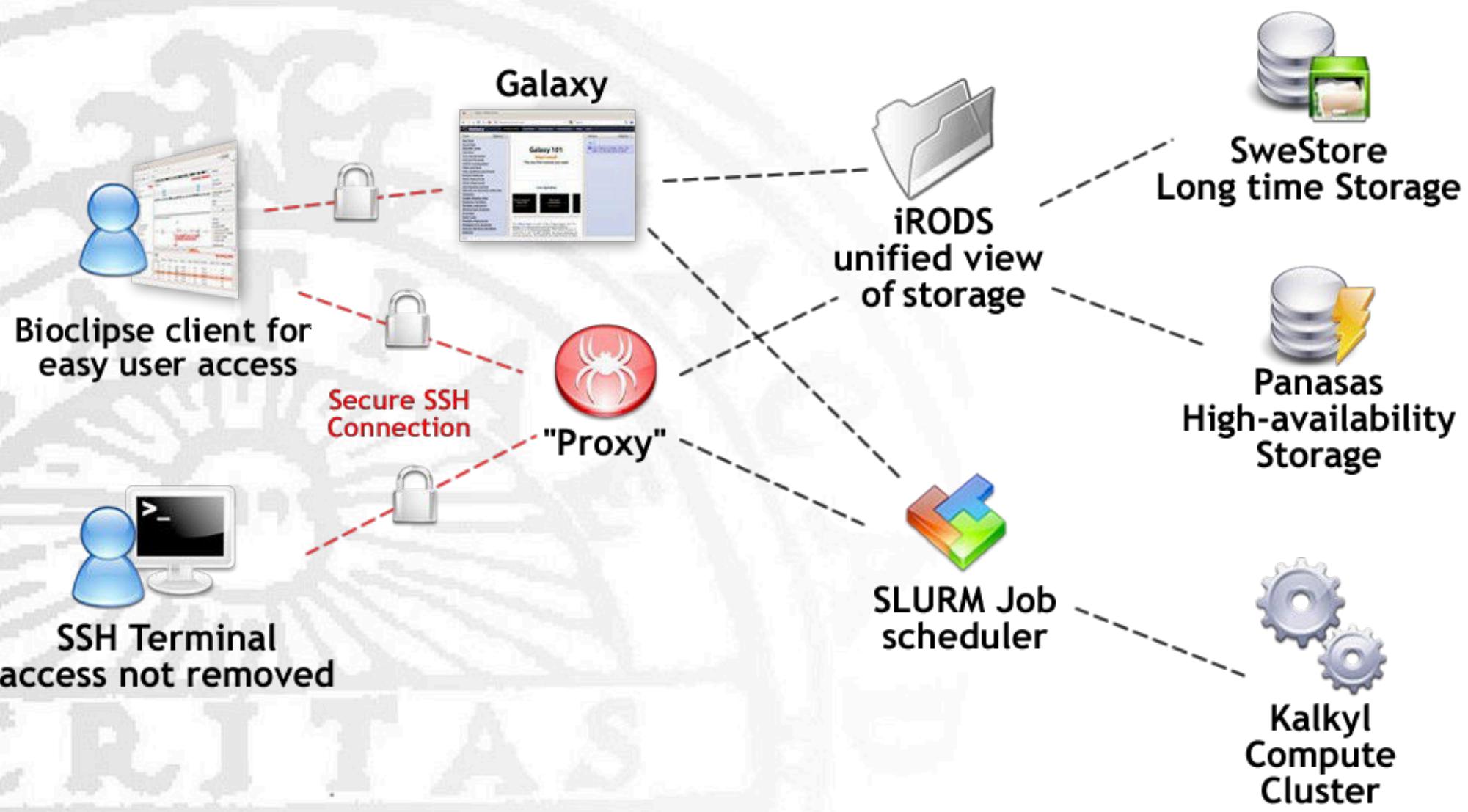
Node  Core  Devel

Specifies the job granularity: whole nodes or single CPU cores per job (task).

JobInfo X Refresh Submit New Job Update ToolConfigs

Job ID	Partition	Job Name	User	Project	Status	Start Time	Time Left	Nodes
autogenerated-job-script.201								
autogenerated-job-script.201								
fasta_test.igv								
hardcopy.1								
igv.log								
jacksum								
jacksum.iac								
1591866	devel	_interactive	samuel	uppmx_staff	R	2011-11-22T 58:55		1

**[MOCKUP IMAGE]**



# Take home messages / questions

- “light-weight” helper tool, using normal HPC
- Teaches (normal) HPC usage
- Ways to make tool configs more declarative?
- We will (try to) use Galaxy in the normal way too!



UPPSALA  
UNIVERSITET



# Status: **Proof of concept**

Youtube Demo:  
**[tinyurl.com/hpcclient](http://tinyurl.com/hpcclient)**

Samuel Lampa, UPPMAX, Uppsala University  
[samuel.lampa@gmail.com](mailto:samuel.lampa@gmail.com)  
[twitter.com/smlmp](http://twitter.com/smlmp)