

# *Integrating iRODS into scientific workflows*

*Maarten Schermer & Raoul Schram  
Utrecht University*

## **IUGM 2025**

June 18, 2025

## Who we are & what we do

- Part of team of **Research Engineers**
- Goal: **increase the impact of research**
- Write FAIR software to help researchers

## Research means data (lots of it)

- **YoDa** (iRODS): storing, sharing, archiving and publishing data

## Scientific workflow systems

- Software to run complex workflows
- **Workflow**: sequence of steps of computation and data manipulation
- Reproducibility, ease of use, transparency, sharing work
- Examples:
  - CWL (Common Workflow Language)
  - **Galaxy**
  - **Snakemake**
  - Nextflow

## VetGalaxy



Research into canine genetics

- Genetic data (heterogeneous, big files)
- Complex analysis
- Loose collection of shell scripts & programs

→ create a workflow: VetGalaxy

Tools

search tools

browser for iRODS instances

iBridges download from iRODS

downloads files to an iRODS instance!

iBridges upload to iRODS

uploads files to an iRODS instance

PBED compiler

compiles .bed, .bim & .fam into composite .pbed file

PBED decompiler

decompiles composite .pbed file into .bed, .bim & .fam

Consensus Tree Tool

Convert Tool

Quality Control Tool

Lift-Over

Text Manipulation

Convert Formats

Filter and Sort

Join, Subtract and Group

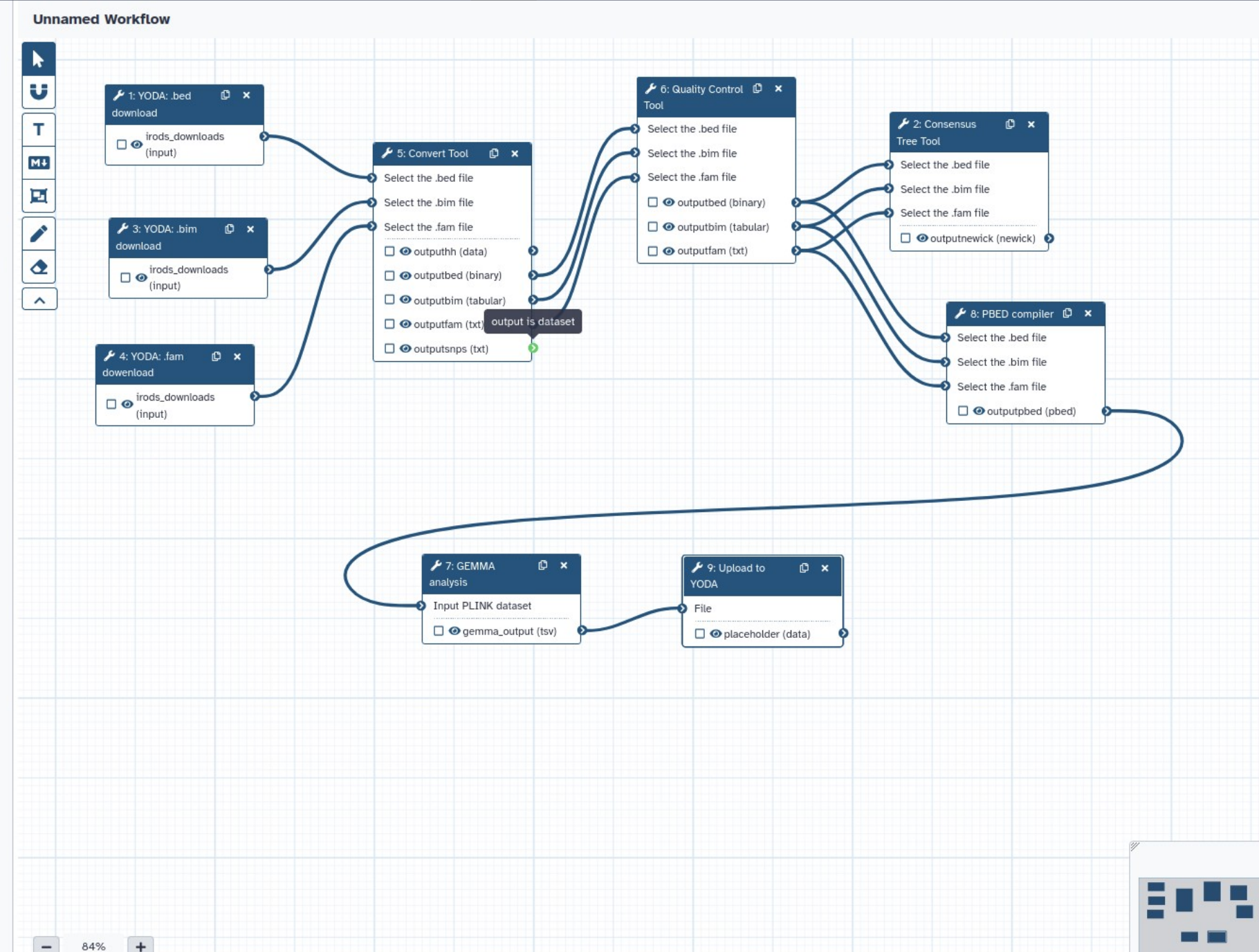
Fetch Alignments/Sequences

Operate on Genomic Intervals

Statistics

Graph/Display Data

Phenotype Association



iBridges upload to iRODS

uploads files to an iRODS instance (Galaxy Version 0.4.0)

Label

Upload to YODA

Add a step label.

Step Annotation

Add an annotation or notes to this step. Annotations are available when a workflow is viewed.

Conditionally skip step?

No

Set to true and connect a boolean parameter that determines whether the step should run. The step runs if the parameter value is true and will be skipped if the parameter value is false.

Tool Parameters

Choose the source type

File

File \*

Data input 'local\_data' (data)

iRODS path (target) \*

/research-test-christine/galaxy-demo

Overwrite?

Yes

Additional Options

Email notification

## Galaxy & iRODS

- Galaxy ObjectStore: data virtualization
- Includes iRODS as optional back-end

“...data belonging to all the users of that Galaxy instance will be stored on the defined persistence media.”

### **VetGalaxy requirements:**

- Integration into workflows
- Personalized access
- Access to different YoDa-instances

## Tool development

- Three iRODS tools: download, upload, browse
- Tools are wrappers
  - Dependencies to install
  - Command(s) to run
  - Input & output hooks to the Galaxy system
  - Limited access to Galaxy configuration
- Modifying Galaxy
  - Personal iRODS configuration

## Dependencies & deployment

- One dependency: **iBridges**
- Published iBridges in **Bioconda**, a Conda-channel specific for biomedical research
- Published iRODS-tools in **Galaxy Toolshed**



Tools

search tools

Upload Data

- Get Data
- VetGalaxy
  - GEMMA analysis GEMMA association tests & kinship analysis
  - iBridges browser for iRODS interactive browser for iRODS instances
  - iBridges download from iRODS downloads files to an iRODS instance!
  - iBridges upload to iRODS uploads files to an iRODS instance
  - PBED compiler compiles .bed, .bim & .fam into composite .pbed file
  - PBED decompiler decompiles composite .pbed file into .bed, .bim & .fam

- Send Data
- Collection Operations
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- Fetch Alignments/Sequences
- Operate on Genomic Intervals
- Statistics

Manage Information

Email address

m.d.schermer@uu.nl

If you change your email address you will receive an activation link in the new mailbox and you have to activate your account by visiting it.

Public name

maarten.schermer

Your public name is an identifier that will be used to generate addresses for information you share publicly. Public names must be at least three characters in length and contain only lower-case letters, numbers, dots, underscores, and dashes (".", "\_", "-").

Address

+ Insert Address

Your iRODS account

Username

m.d.schermer@uu.nl

Required

Data Access Password

.....

Required

Config

{"irods\_host": "its.data.uu.nl", "irods\_port": 1247,"irods\_home": "/nluu12p/home","irods\_user\_name": "m.d.schermer@uu.nl","irods\_zone\_name": "nluu12p","irods\_authentication\_scheme": "pam","irods\_encryption\_algo"

Required

Save

# Galaxy

Workflow Visualize Data Admin Help

## Tools

search tools

Upload Data

### Get Data

#### VetGalaxy

- GEMMA analysis** GEMMA association tests & kinship analysis
- iBridges browser for iRODS** interactive browser for iRODS instances
- iBridges download from iRODS** downloads files to an iRODS instance!
- iBridges upload to iRODS** uploads files to an iRODS instance
- PBED compiler** compiles .bed, .bim & .fam into composite .pbed file
- PBED decompiler** decompiles composite .pbed file into .bed, .bim & .fam

### Send Data

### Collection Operations

### Lift-Over

### Text Manipulation

### Convert Formats

### Filter and Sort

### Join, Subtract and Group

### Fetch Alignments/Sequences

### Operate on Genomic Intervals

### Statistics

### Graph/Display Data

## Active Interactive Tools

Name	Job Info	Created
<input type="checkbox"/> ibridges browser	running	less than a minute ago

## iRODS browser

http://localhost:4002/interactivetool/ep/31u62as57j91p/39nxloh27pctl/?path=/nluu12p/

Server: its.data.uu.nl  
User: m.d.schermer@uu.nl  
Home: /nluu12p/home

/nluu12p/home / research-test-christine / galaxy-demo / data /

/nluu12p/home/research-test-christine/galaxy-demo/data/AFFYMETRIX-AxiomGT1v2.comp\_merge.public\_ids.noDup.bed →  
/nluu12p/home/research-test-christine/galaxy-demo/data/AFFYMETRIX-AxiomGT1v2.comp\_merge.public\_ids.noDup.bim →  
/nluu12p/home/research-test-christine/galaxy-demo/data/AFFYMETRIX-AxiomGT1v2.comp\_merge.public\_ids.noDup.fam →  
/nluu12p/home/research-test-christine/galaxy-demo/data/Dobermanns\_Embarplaat\_2022.bed →  
/nluu12p/home/research-test-christine/galaxy-demo/data/Dobermanns\_Embarplaat\_2022.bim →  
/nluu12p/home/research-test-christine/galaxy-demo/data/Dobermanns\_Embarplaat\_2022.fam →  
/nluu12p/home/research-test-christine/galaxy-demo/data/LUPA-Dobermanns\_Embark\_Lupa174K\_Merge\_AISNPs\_Christina.bed →  
/nluu12p/home/research-test-christine/galaxy-demo/data/LUPA-Dobermanns\_Embark\_Lupa174K\_Merge\_AISNPs\_Christina.bim →  
/nluu12p/home/research-test-christine/galaxy-demo/data/LUPA-Dobermanns\_Embark\_Lupa174K\_Merge\_AISNPs\_Christina.fam →  
/nluu12p/home/research-test-christine/galaxy-demo/data/MDD\_Utrecht2\_1\_Raw.bed →  
/nluu12p/home/research-test-christine/galaxy-demo/data/MDD\_Utrecht2\_1\_Raw.bim →  
/nluu12p/home/research-test-christine/galaxy-demo/data/MDD\_Utrecht2\_1\_Raw.fam →  
/nluu12p/home/research-test-christine/galaxy-demo/data/NEOGEN220\_Utrecht\_University\_CAN230V01\_20240517\_FinalReport.txt →  
/nluu12p/home/research-test-christine/galaxy-demo/data/NEOGEN220\_Utrecht\_University\_CAN230V01\_20240517\_FinalReport.zip →  
/nluu12p/home/research-test-christine/galaxy-demo/data/WGS-CANFAM3-TVD18631.vcf.gz →  
/nluu12p/home/research-test-christine/galaxy-demo/data/WGS-CANFAM3-TVD18631.vcf.gz.tbi →  
/nluu12p/home/research-test-christine/galaxy-demo/data/WGS-CANFAM4-VAN9052\_processed.vcf.gz →  
/nluu12p/home/research-test-christine/galaxy-demo/data/WGS-CANFAM4-VAN9052\_processed.vcf.gz.tbi →  
/nluu12p/home/research-test-christine/galaxy-demo/data/Wisdom\_All\_e3\_data\_with\_X\_Raw.xlsx →

quit

Copy path to clipboard and leave iRODS browser running.  
Export path to Galaxy and quit iRODS browser.  
quit Quit iRODS browser without exporting a path.

442: lupa_170k.fam	👁️ ✎️ 🗑️
441: lupa_170k.bim	👁️ ✎️ 🗑️
440: lupa_170k.bed	👁️ ✎️ 🗑️

## Future work

- Add meta-data capabilities
- Follow future iBridges developments
- Waiting for feedback from VetGalaxy-users

<https://github.com/UtrechtUniversity/galaxy-tools-ibridges>

<https://github.com/UtrechtUniversity/DGK-VetGalaxy>

*Over to Raoul for Snakemake*



**Utrecht  
University**

Sharing science,  
*shaping tomorrow*