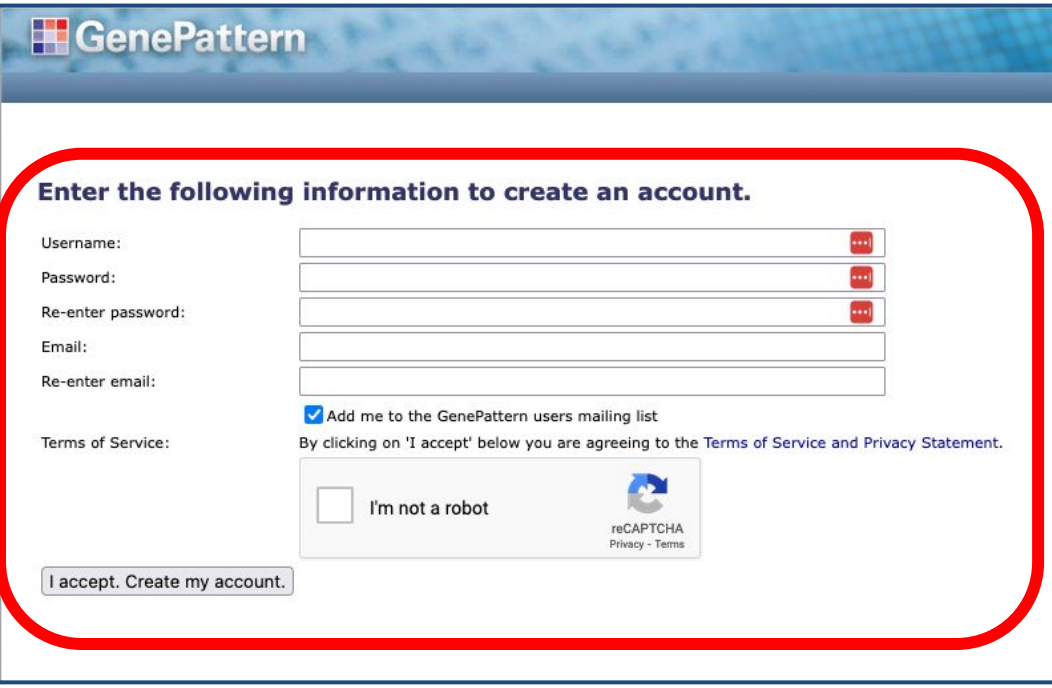
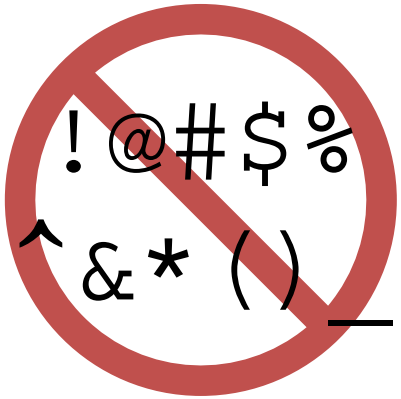
Integrative Genomic Analysis with GenePattern

August 22, 2025

Thorin Tabor **Anthony Castanza** **Alex Wenzel** **Ted Liefeld**

1. Register for an account at:
cloud.genepattern.org





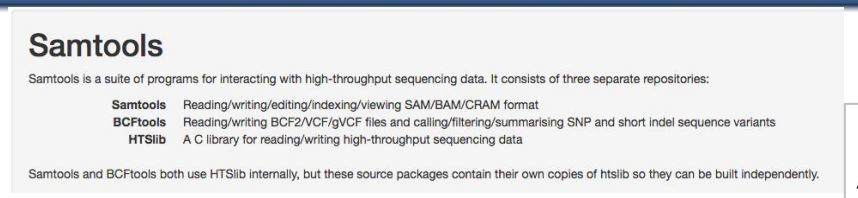
No spaces or special characters in usernames

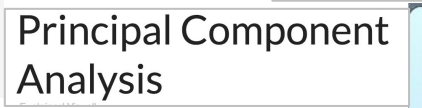
Agenda

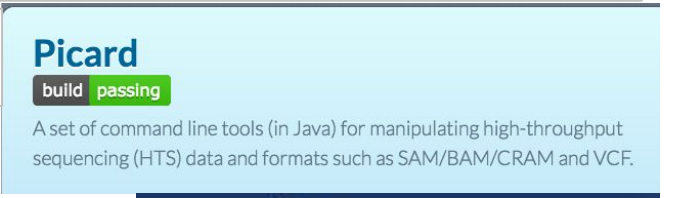
- Overview
- Data Prep
- Differential Analysis
- GSEA
- MSigDB
- **Break**
- Classification & Prediction
- IGV
- Single-cell RNA-Seq Clustering
- GenePattern Copilot
- Closing

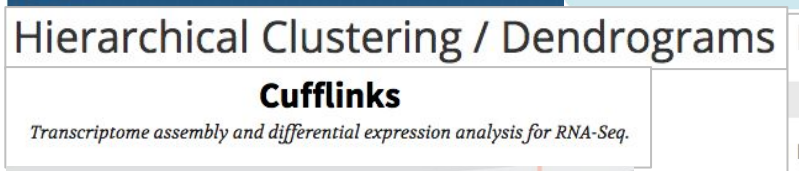
GenePattern Overview

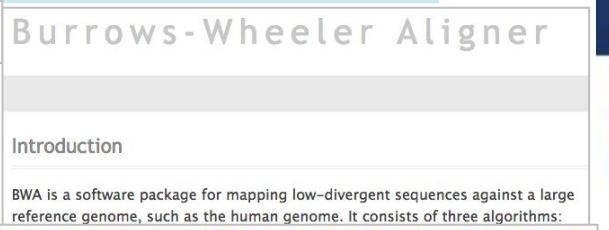
Tools for Bioinformatics

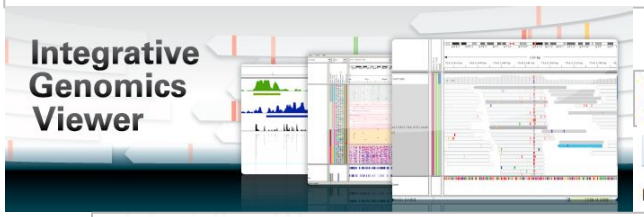


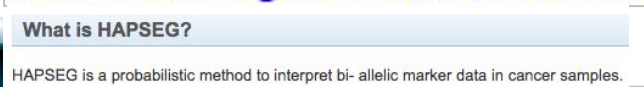




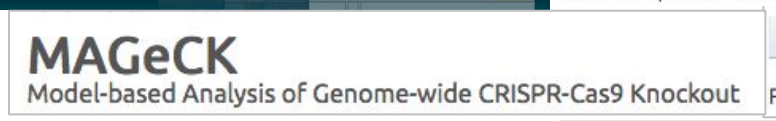


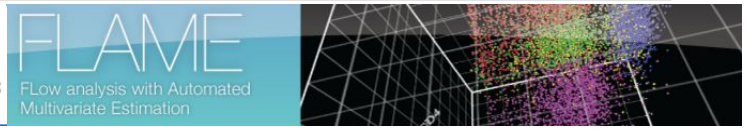






Constellation Map: Downstream visualization and interpretation of gene set enrichment results [version 1; referees: 2 approved]



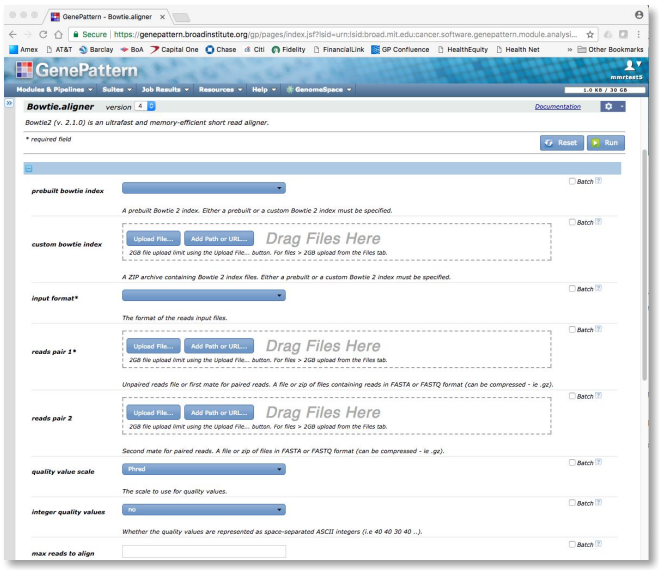


Problems with bioinformatics tool use and interoperability

- Tools are built using different languages
- Each tool has its own installation and operational requirements.
- Tools require (sometimes extensive) Unix knowledge.
- Tools are not designed to communicate with each other.

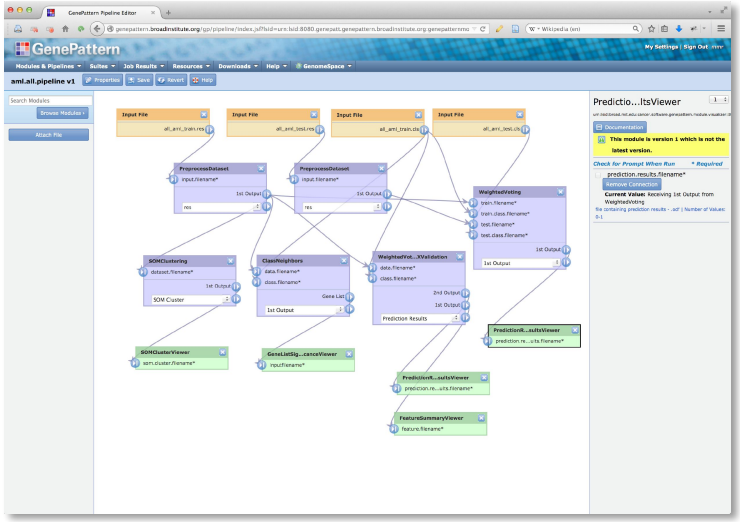
```
bowtie -a --best --strata -S -m 100 -X 400 --chunkmbs 256 --fullref -p 4
Dmel.BDGP5-transcripts \ -1 SRR031714_1.fastq -2 SRR031714_2.fastq |
samtools view -F 0xC -bS - | \ samtools sort -n -
~/Desktop/untreated3-transcriptome
```

Solution : GenePattern capabilities



Wrap tools in a user-friendly interface

Provide a collection of hundreds of ‘omics tools

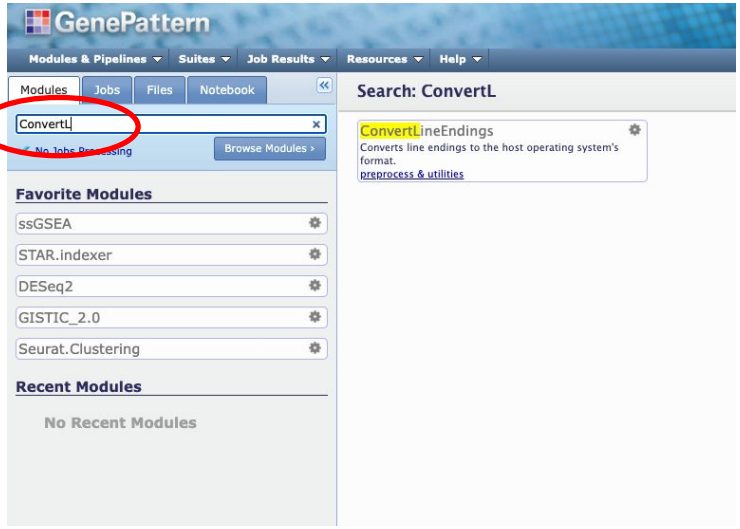


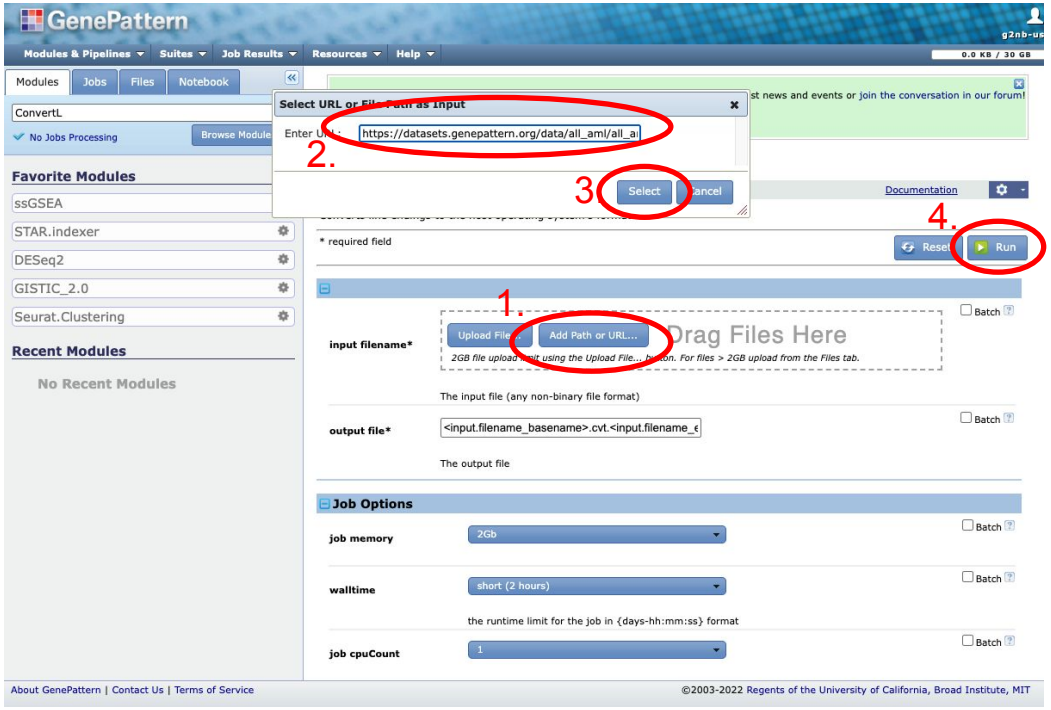
Provide features for reproducibility

- Record and replay of all analyses
- Retain all versions of code
- Chain analyses into “pipelines” that can be shared and published

Exercise #1:

Using Classic GenePattern



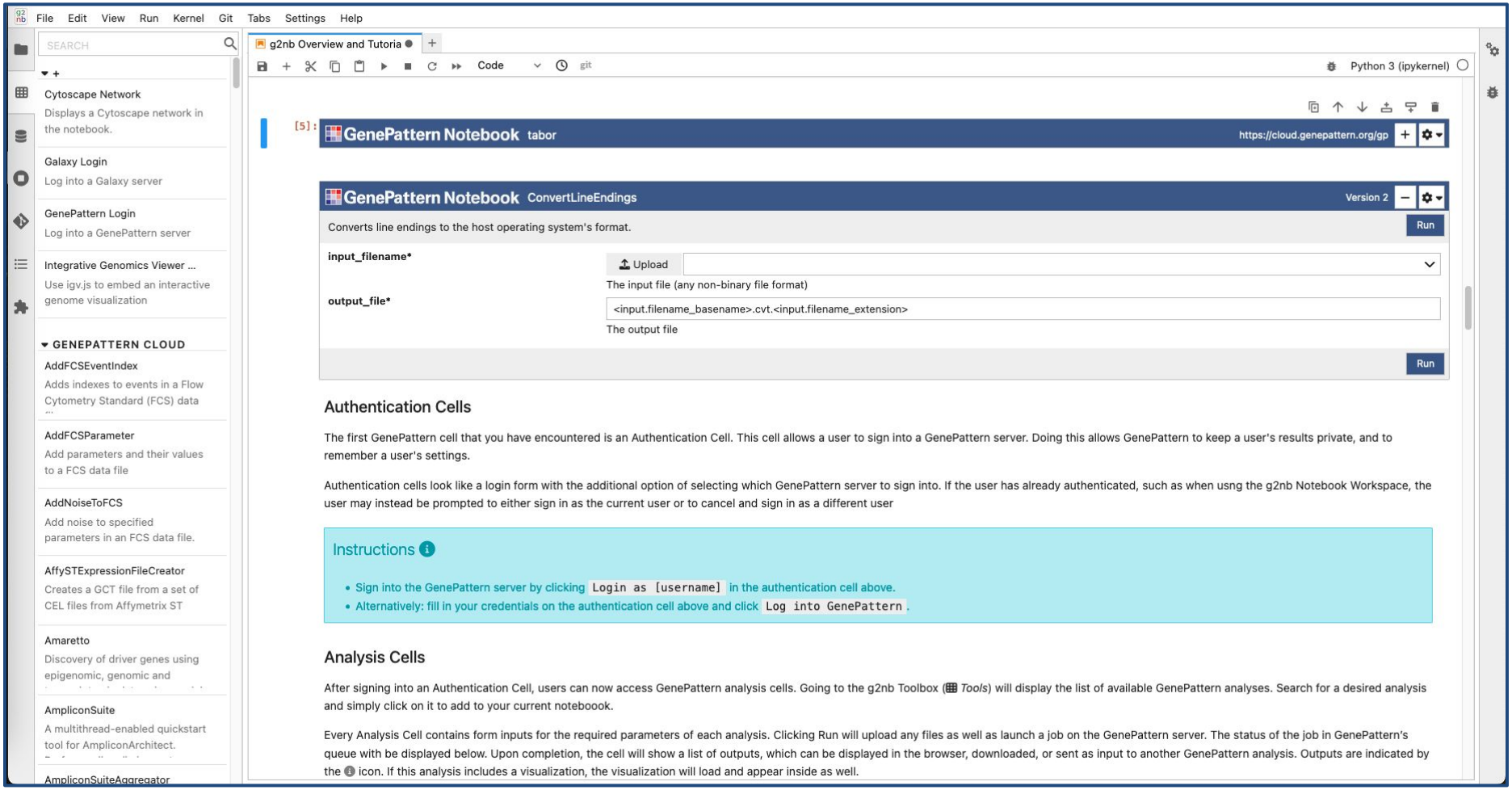


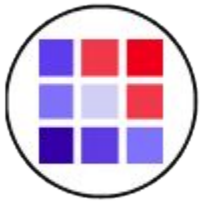
- Go to https://cloud.genepattern.org
- Sign in using your GenePattern username and password.
- On the left, search for ConvertLineEndings and click it in the list when it comes up.
- Click “Add Path or URL” next to the “input filename” parameter and enter: https://datasets.genepattern.org/data/all_aml/all_aml_train.gct

g2nb

The Genomics to Notebook environment

g2nb Workspace





Jupyter

- Popular and well-supported framework for scientific computing
- Ecosystem of available extensions and resources
- Open source

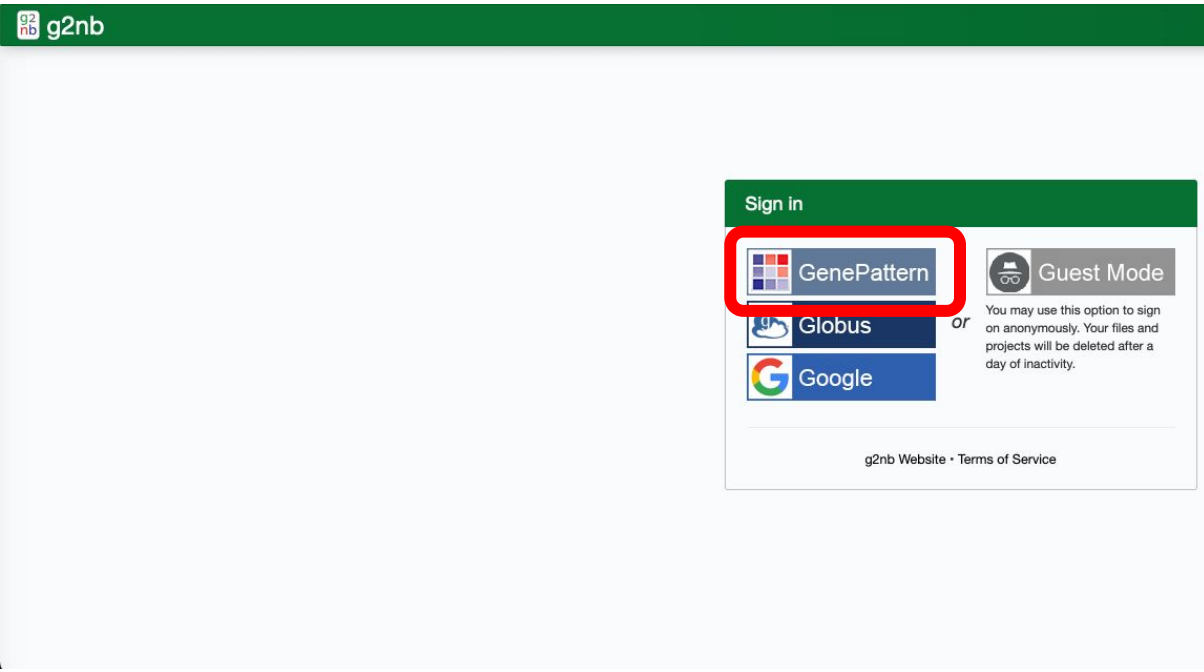
Complete Research Narrative

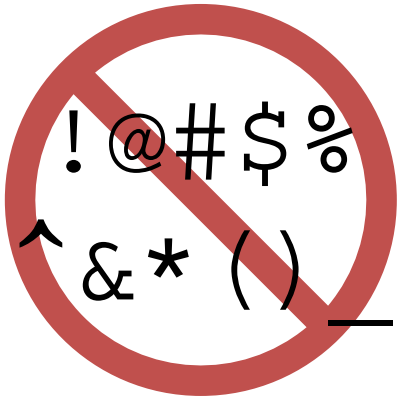
- Leverages the best of Jupyter, GenePattern and other popular bioinformatics platforms
- Interleave text, visualization, graphics and analytical aspects

Exercise #2:

Using g2nb

1. Sign in at:
workspace.g2nb.org





No spaces or special characters in usernames

2. Run the workshop project:
2025-08-22 MSTP Bootcamp