

# GENOMESPACE



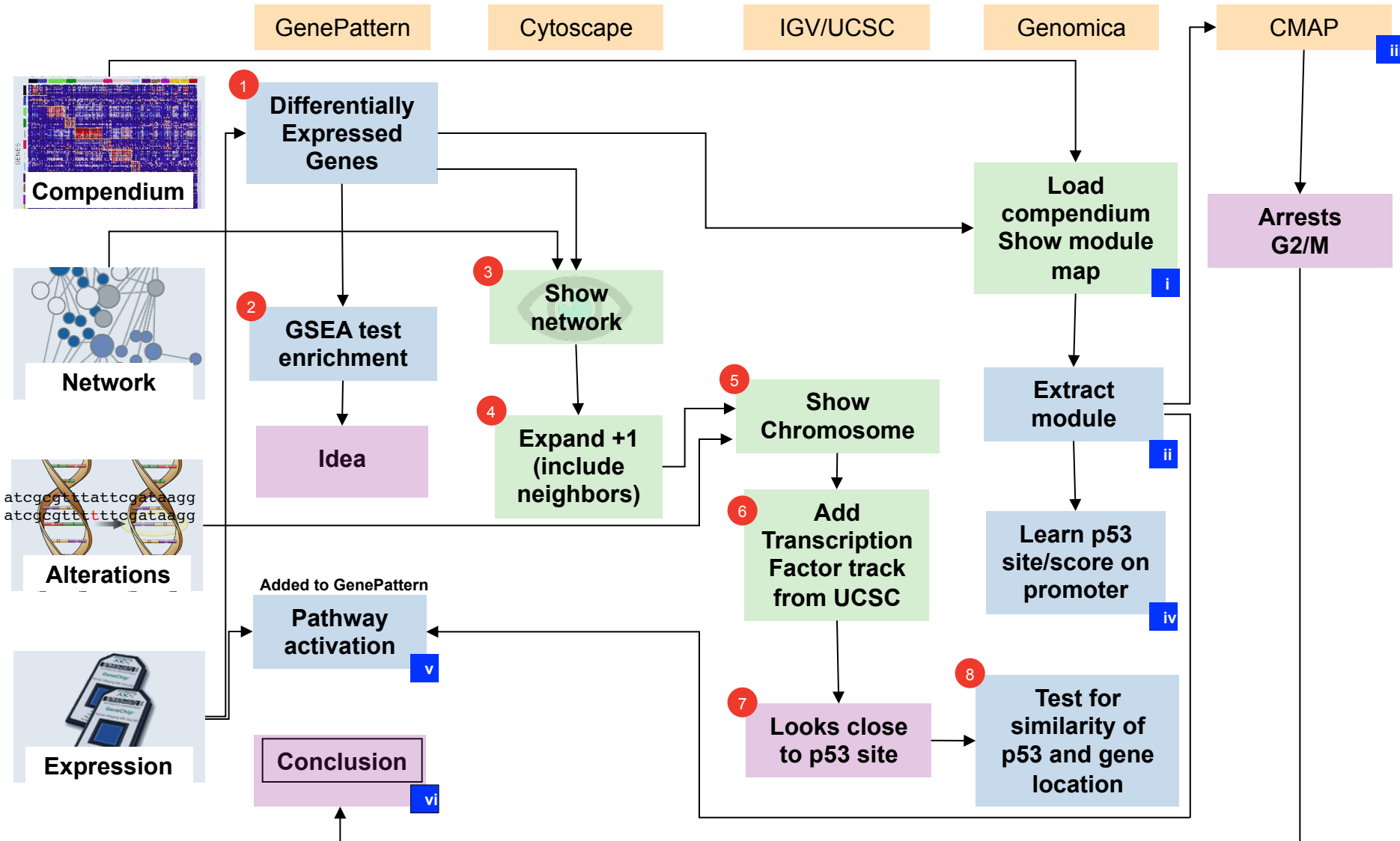
# GENOMESPACE



## Outline

- Introduction to GenomeSpace
- GenomeSpace Tools and Recipes
- GenomeSpace User Interface
- Integrative analysis exercise
- Other GenomeSpace Tools
- GenomeSpace development
- Q and A

# The vision: Integrative Translational Genomics



# Online community to share diverse computational tools

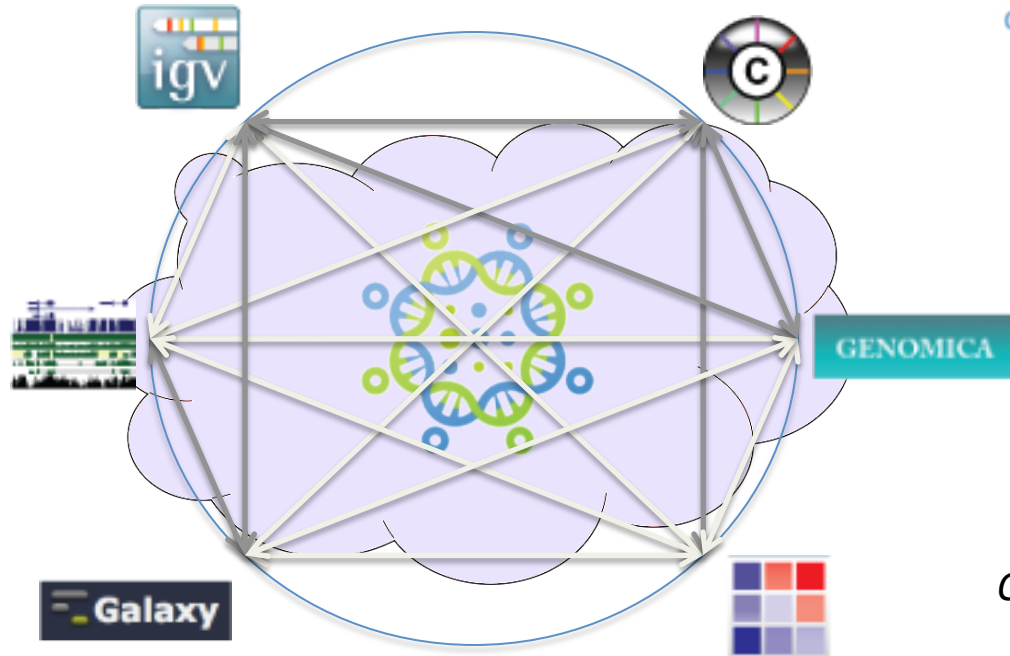
## GENOMESPACE



### Seed Tools

Cytoscape  
Galaxy  
GenePattern  
Genomica  
IGV  
UCSC Browser

*Outreach to new tools*



### Driving Biological Projects

lincRNAs  
Cancer stem cells  
Patient Stratification  
*Outreach to new DBPs*

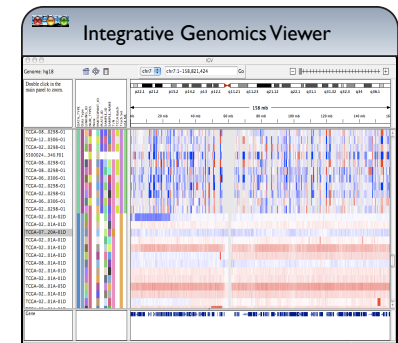
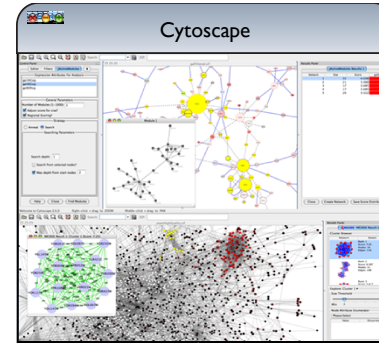
[www.genomespace.org](http://www.genomespace.org)



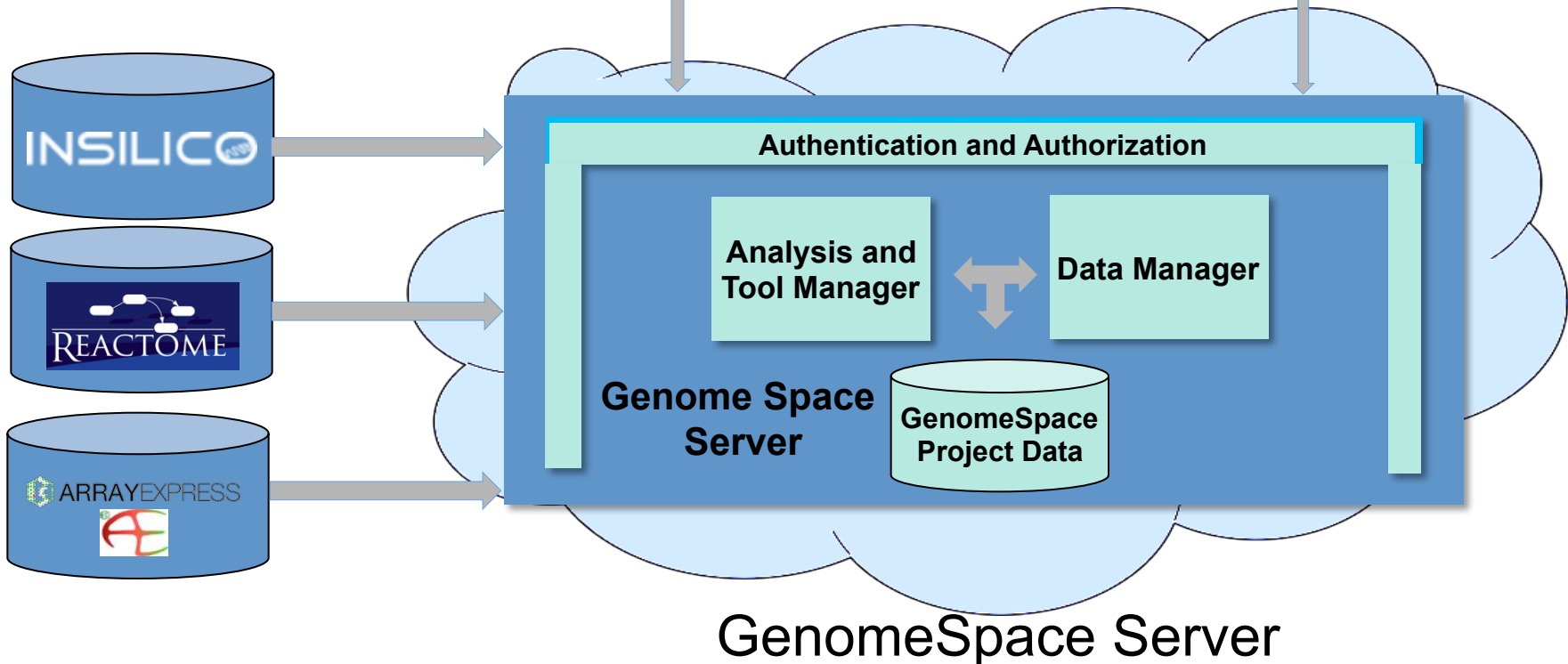
# **GenomeSpace: a connection layer between integrative analysis tools**

- Support for all types of resource: Web-based, desktop, etc.
- Automatic conversion of data formats between tools
- Easy access to data from any location
- Ease of entry into the environment

# GenomeSpace Components



## GenomeSpace-Enabled Tools



# Register

The image is a screenshot of a web browser displaying the GenomeSpace website. The browser's address bar shows 'www.genomespace.org'. The website has a blue header with the GenomeSpace logo on the left and a 'User Login' link on the right. Below the header is a navigation bar with links: 'What is GenomeSpace?', 'Tools', 'Documentation', 'Developers', 'Support', and 'About'. A search bar is also present. The main content area features a large banner with the text 'GENOMESPACE BETA' and 'Frictionless connection of bioinformatics tools'. Below this text are two buttons: 'Register' (blue) and 'User Login' (green). The 'Register' button is circled in red. To the right of the buttons is a graphic showing various bioinformatics data visualizations. Below the banner is a 'STATUS' section indicating that all systems are operating normally. Further down is a 'WHAT'S NEW' section with a 'News Highlights' tab selected, showing an article about GenomeSpace at the BOSC-Broad Interoperability Hackathon. On the right side of the page, there is a social media section for '@genomespace' with a 'Mentions' tab, displaying tweets about the VIZBI workshop and the BOSC Interoperability Hackathon.

www.genomespace.org

GenomeSpace: Home

User Login

What is GenomeSpace? Tools Documentation Developers Support About

**GENOMESPACE BETA**  
Frictionless connection of bioinformatics tools

**Register** **User Login**

STATUS 03.18.13 07:01AM All systems are operating normally.

WHAT'S NEW

News Highlights GenomeSpace Team Blog

GenomeSpace at BOSC-Broad Interoperability Hackathon

GenomeSpace will be one of the featured tools at the BOSC/Broad interoperability hackathon on Sunday and Monday, April 7 and 8 at the [Broad Institute](#). It occurs just prior to [Bio-IT World](#) in Boston and is a warm up event for [BOSC 2013](#) occurring July 19-20th in Berlin. The goal of the hackathon will be developing open source tools to help enhance bioinformatics interoperability. See <http://www.broadinstitute.org/events/interoperability-hackathon-2013/home> for more information as

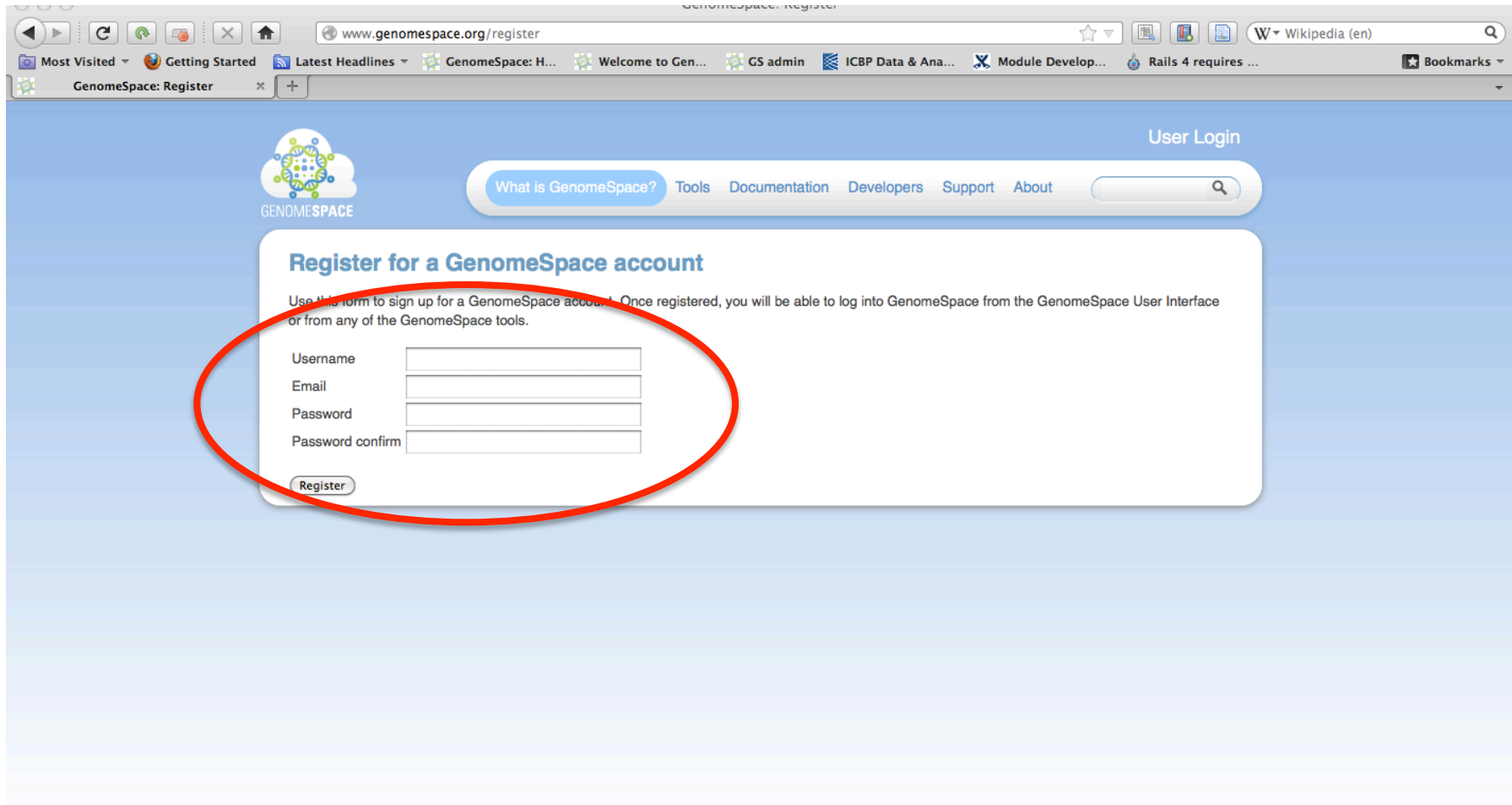
@genomespace Mentions

genomespace Learn GenomeSpace at VIZBI: Visualizing Biological Data, March 19-22 - workshop space still available [lanyrd.com/2013/vizbi/](#) #bioinformatics  
12 days ago · reply · retweet · favorite

genomespace Join GenomeSpace and others at the Bioinformatics Interoperability Hackathon in Boston, April 7-8. #bioinformatics  
[groups.google.com/forum/#!forum/interoperability-hackathon-2013](#)

www.genomespace.org

# Register



The screenshot shows a web browser window with the address bar displaying `www.genomespace.org/register`. The browser's bookmark bar includes links such as "Most Visited", "Getting Started", "Latest Headlines", "GenomeSpace: H...", "Welcome to Gen...", "GS admin", "ICBP Data & Ana...", "Module Develop...", "Rails 4 requires ...", and "Bookmarks". The page header features the GenomeSpace logo, a navigation menu with links like "What is GenomeSpace?", "Tools", "Documentation", "Developers", "Support", and "About", and a "User Login" link. The main content area is titled "Register for a GenomeSpace account" and contains a registration form. A red circle highlights the form, which includes fields for "Username", "Email", "Password", and "Password confirm", along with a "Register" button. The form is set against a light blue background with a subtle grid pattern.

www.genomespace.org/register

Most Visited Getting Started Latest Headlines GenomeSpace: H... Welcome to Gen... GS admin ICBP Data & Ana... Module Develop... Rails 4 requires ... Bookmarks

GenomeSpace: Register

GENOMESPACE

User Login

What is GenomeSpace? Tools Documentation Developers Support About

## Register for a GenomeSpace account

Use this form to sign up for a GenomeSpace account. Once registered, you will be able to log into GenomeSpace from the GenomeSpace User Interface or from any of the GenomeSpace tools.

Username

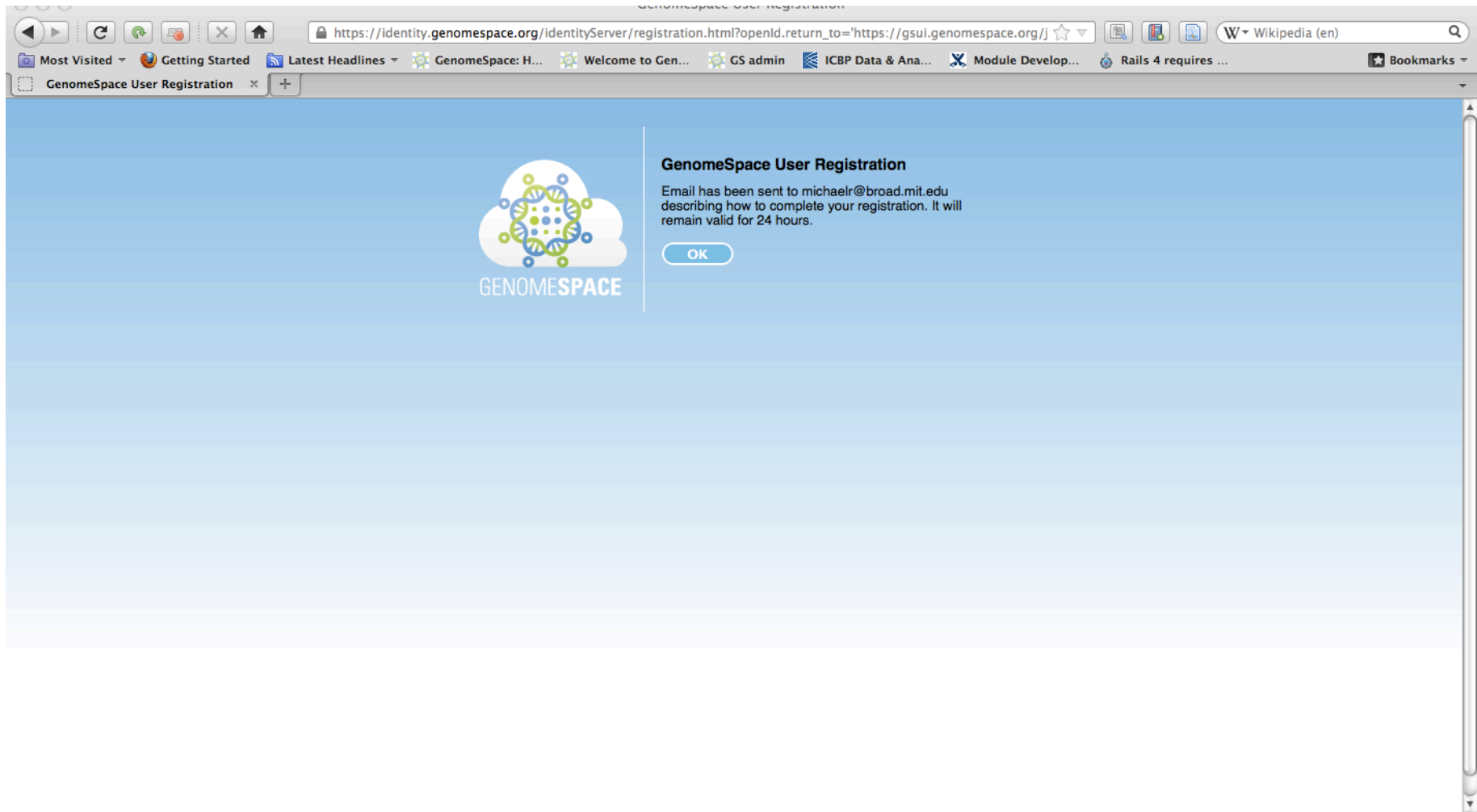
Email

Password

Password confirm

Register

# Register



# Register

The screenshot shows a Gmail interface with a browser window at the top. The browser's address bar displays the URL: <https://mail.google.com/mail/u/0/?ui=2&shva=1#inbox/13d7f233a77b6490>. The browser's tab bar includes several open tabs, with 'GenomeSpace: H...' and 'Welcome to Gen...' being relevant. The Gmail interface shows the 'Mail' view with a search bar and a list of folders on the left, including 'Inbox (14)', 'Starred', 'Important', 'Sent Mail', 'Drafts (45)', 'Circles', '[Gmail] Sent Mail', '2 Do 4 J', '2 Do 4 T', 'Archive/Classes', 'B1Tst0re', 'Boomerang', 'Boomerang-Out...', 'Broad-E', '2012-06-25 R...', '2013-05-07 R...', 'Broad-wide Soft...', 'Cancer Portals', 'Deleted Messages', and 'Events'. The main content area displays an email from 'genomespacenotifier@gmail.com' to 'michaelr' at '4:14 PM (3 minutes ago)'. The email body contains the following text: 'Hello. To complete the registration of the GenomeSpace account for you, please follow this link: <https://identity.genomespace.org/identityServer/usermanagement/register/pendingUuid/7368493c-062f-4acb-8575-dd99a781418a>'. This link is circled in red. Below the link, it says: 'You can also copy and paste it into your browser's address bar. After 24 hours the link will expire, and you will need to request registration again. If you did not request a GenomeSpace account, or if the request was made in error, you should do nothing and let the temporary registration expire normally.' The right sidebar shows the contact 'genomespacenotifier' with an 'Add to circles' button and a 'Show details' link. The bottom of the page includes a footer with '20% full Using 5.1 GB of your 25 GB', '©2013 Google - Terms of Service - Privacy Policy - Program Policies', 'Powered by Google', and 'Last account activity: 1 minute ago Open in 1 other location Details'.

GenomeSpace user registration

genomespacenotifier@gmail.com 4:14 PM (3 minutes ago)

Hello. To complete the registration of the GenomeSpace account for you, please follow this link:

<https://identity.genomespace.org/identityServer/usermanagement/register/pendingUuid/7368493c-062f-4acb-8575-dd99a781418a>

You can also copy and paste it into your browser's address bar.

After 24 hours the link will expire, and you will need to request registration again.

If you did not request a GenomeSpace account, or if the request was made in error, you should do nothing and let the temporary registration expire normally.

Click here to [Reply](#), [Reply to all](#), or [Forward](#)

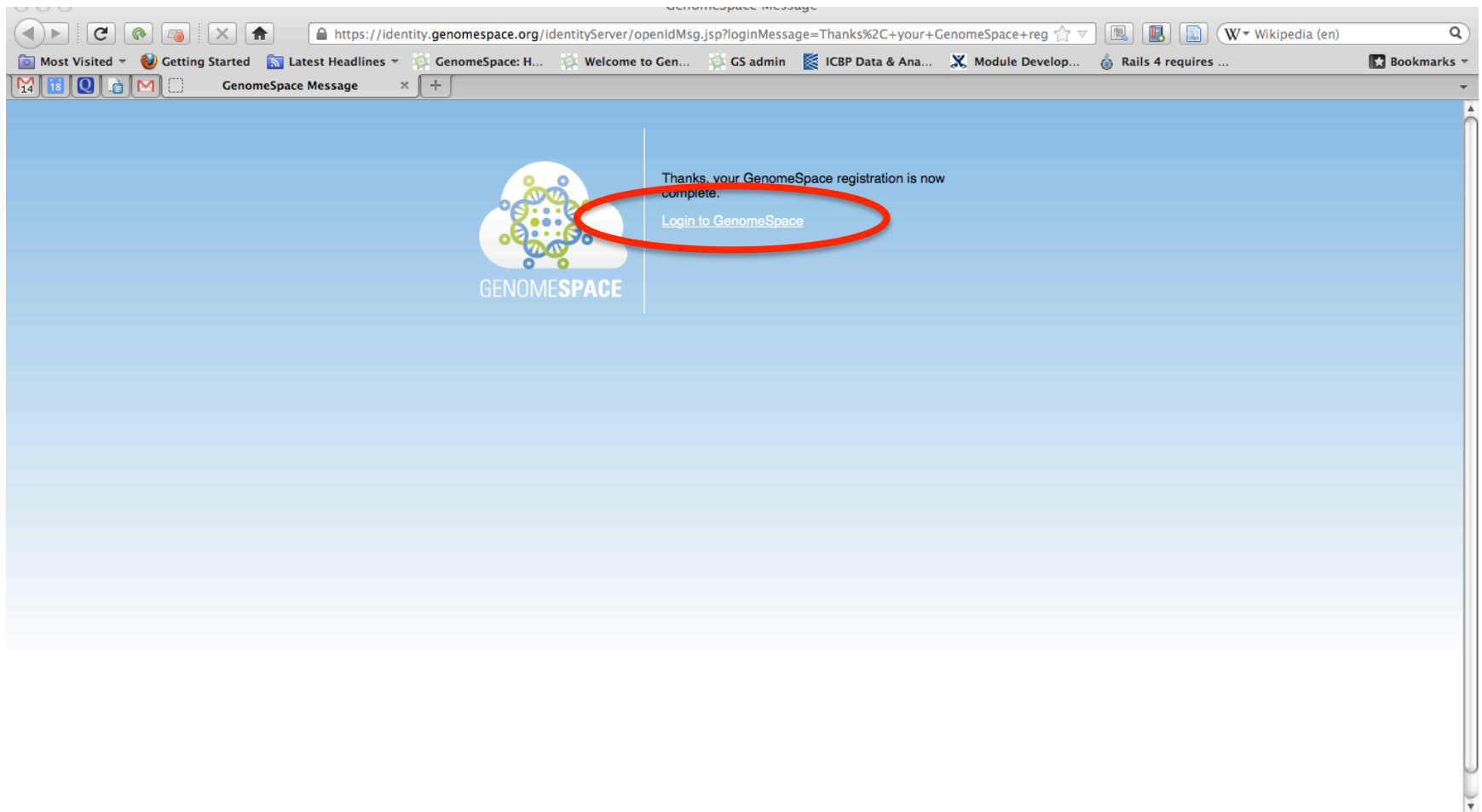
20% full Using 5.1 GB of your 25 GB

©2013 Google - [Terms of Service](#) - [Privacy Policy](#) - [Program Policies](#)

Powered by Google

Last account activity: 1 minute ago Open in 1 other location [Details](#)

# Login






# Login

GenomeSpace OpenID Login

https://identity.genomespace.org/identityServer/openIdProvider?openid.ns=http%3A%2F%2Fspecs.openid.net%2Fauth' Wikipedia (en)

Most Visited Getting Started Latest Headlines GenomeSpace: H... Welcome to Gen... GS admin ICBP Data & Ana... Module Develop... Rails 4 requires ... Bookmarks

GenomeSpace OpenId Login



**OpenID Login**

USERNAME:

PASSWORD:

[Register new GenomeSpace user](#)  
[Forgot your password?](#)

# GenomeSpace UI

Browser window showing the GenomeSpace UI at <https://gsui.genomespace.org/jsui/gsui.html>.

The interface includes a navigation bar with the following menu items: File, Launch, View, Manage, Recipes, Help.

The main content area displays a grid of application icons for launching various tools:

- Cistrome
- Cytoscape
- Galaxy
- GenePattern
- Genomica
- geWorkbench
- IGV
- InSilicoDB
- UCSC Table Browser
- ArrayExpress
- Gitools

Below the application grid, there is a section for Directories, showing a list of folders and their details:

Directories	Filename	Owner	Size	Last Modified
Home				
mmr	mmr	mmr		
Shared to mmr	Shared to mmr	System		
Public	Public	System		

# Tools and Recipes

Focus on Kitchen Skills

# Agenda

- Review of GenomeSpace tools in the first exercises
- Basic recipes for using GenomeSpace
  - Launching tools
  - Uploading data to GenomeSpace
  - Sending data to tools

# GenomeSpace Tools



ArrayExpress



geWorkbench



Galaxy



Gitools



Cistrome



IGV



Cytoscape



InSilicoDB



GenePattern



Genomica



UCSC Table Browser



ISAcreeator

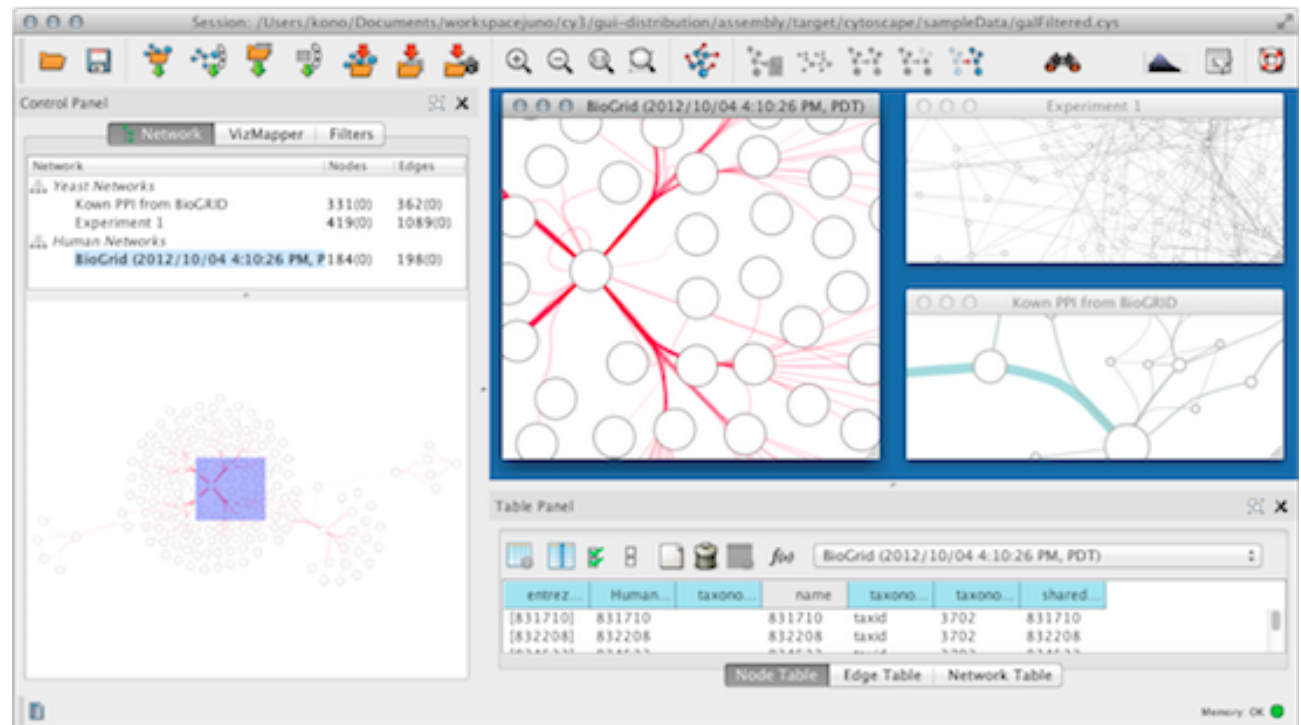


MSigDB



# Cytoscape

Cytoscape is an open-source bioinformatics software platform for visualizing molecular interaction networks and biological pathways, and integrating these networks with annotations, gene expression profiles, and other state data.





# Galaxy

Galaxy is an open-source, scalable framework for tool integration that allows users to analyze multiple alignments, compare genomic annotations, and profile metagenomic samples, among many possible analyses; workflows allow the linking together of analyses.

The screenshot displays the Galaxy web interface in a browser window. The address bar shows <https://main.g2.bx.psu.edu>. The navigation bar includes links for **Analyze Data**, **Workflow**, **Shared Data**, **Visualization**, **Cloud**, **Help**, and **User**, along with a **Using 0%** status indicator. A yellow banner at the top of the main content area states: "Galaxy will be down on Thursday, March 14, to relocate hardware to a new server room. All jobs running at the time of the shutdown will be terminated."

The interface is divided into three main sections:

- Tools (Left Panel):** A sidebar with a search bar and a list of tool categories: [Get Data](#), [Send Data](#), [ENCODE Tools](#), [Lift-Over](#), [Text Manipulation](#), [Convert Formats](#), [FASTA manipulation](#), [Filter and Sort](#), [Join, Subtract and Group](#), [Extract Features](#), [Fetch Sequences](#), [Fetch Alignments](#), [Get Genomic Scores](#), and [Operate on Genomic Intervals](#).
- History (Right Panel):** A list of recent jobs. The top job is "imported: Exons vs Repeats" (416.4 MB). Below it are several jobs using the "GenomeSpace importer" tool on various datasets, including "10: GenomeSpace importer on CCLE Expression Entrez 2012-04-06.res", "9: GenomeSpace importer on SNPs.bed.bed", "8: GenomeSpace importer on all aml test.res", "7: top 5 exons", "2: UCSC Main on Human: rmsk (genome)", and "1: UCSC Main on Human: rmsk (genome)".
- Live Quickies (Bottom):** A row of four featured workflows: "Imported fastQ population: quickie # 14", "454 Mapping: Single End: Galaxy quickie # 15", "Uploading Data using FTP: Galaxy quickie # 17", and "Managing account histories: Galaxy quickie # 19".

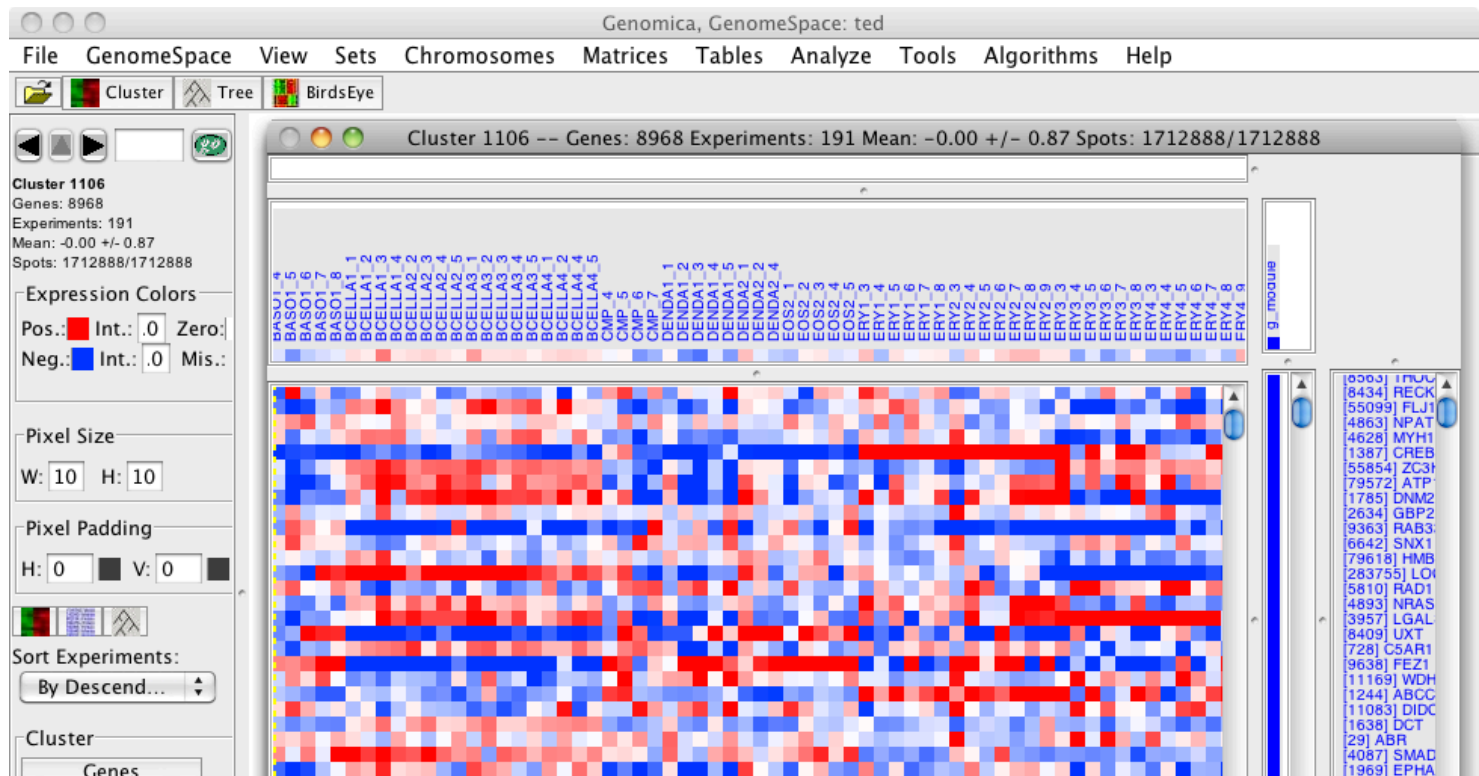
The central area features a large graphic with the text "Galaxy is hiring" and "OSLO", surrounded by various scientific and data-related terms like "transparent", "research", "accessible", "workflows", "reproducible", "biology", "data", "analysis", "visualization", "informatics", "science", "NISC", "UIO", and "University of Oslo".





# Genomica

Genomica is an analysis and visualization tool for genomic data that can integrate gene expression data, DNA sequence data, and gene and experiment annotation information.





# GenePattern

GenePattern is a powerful genomic analysis platform that provides access to more than 150 tools for gene expression analysis, proteomics, SNP analysis, flow cytometry, RNA-seq analysis, and common data processing tasks. A web-based interface provides easy access to these modules and allows for the creation of multi-step analysis pipelines that enable reproducible *in silico* research.

The screenshot shows the GenePattern web interface in a browser window. The address bar displays `genepattern.broadinstitute.org/gp/pages/index.jsf`. The page features a navigation bar with links for **Modules & Pipelines**, **Suites**, **Job Results**, **Resources**, **Downloads**, **Help**, and **GenomeSpace**. On the left, a sidebar titled **Modules & Pipelines** includes a search bar and radio buttons for **category**, **suite**, and **all**. Below these are links for **open all** and **close all**, followed by a list of **Recently Used** modules such as Annotation, caBIG, Clustering, Data Format Conversion, Flow Cytometry, Gene List Selection, GENE-E, IGV, Image Creators, Missing Value Imputation, MIT\_701X, Pathway Analysis, Prediction, Preprocess & Utilities, Projection, Proteomics, RNA-seq, RNAi, Sequence Analysis, SNP Analysis, Statistical Methods, Survival Analysis, Uncategorized, Visualizer, and pipeline. The main content area displays a message: **3/8/13 GSEALeadingEdgeViewer updated for compatibility with Java 7.** Below this is a **Welcome to GenePattern** section with the heading **Analyzing genomic data in GenePattern** and the question **what do you want to do?** It provides instructions to click a **protocol** or **Quick Start** for analysis guidance. A section titled **Protocols for running common analyses in GenePattern:** lists **Run an Analysis in GenePattern** (preprocessing and visualization) and **Differential Expression Analysis** (finding differentially expressed genes). On the right, a **Recent Jobs** sidebar shows a list of jobs including **ConvertLineEndings**, **PreprocessDataset**, and **ExtractColumnName** with their respective IDs, dates, and file names.



# ArrayExpress

**ArrayExpress** is a repository of over 30,000 functional genomics experiments comprising nearly 1 million assays. Users can query and retrieve data in a number of different formats including the [MIAME](#) and [MINSEQE](#) standards.

The screenshot shows the ArrayExpress website in a web browser. The browser's address bar displays [www.ebi.ac.uk/arrayexpress/](http://www.ebi.ac.uk/arrayexpress/). A cookie notice is visible at the top, stating that cookies are used for site functionality and providing links to the Cookie and Privacy statements. The main navigation bar includes links for Services, Research, Training, Industry, and About us. The ArrayExpress logo is prominently displayed, followed by a search bar with a 'Search' button and an 'Advanced' link. Below the search bar, a navigation menu includes Home, Experiments, Arrays, Submit, Help, and About ArrayExpress. The main content area features the heading 'ArrayExpress - functional genomics data' and a description of the database. To the right, a 'Data Content' section provides statistics on the repository's data.

ArrayExpress - functional genomics data

ArrayExpress is a database of functional genomics experiments that can be queried and the data downloaded. It includes gene expression data from microarray and high throughput sequencing studies. Data is collected to [MIAME](#) and [MINSEQE](#) standards. Experiments are submitted directly to ArrayExpress or are imported from the NCBI GEO database.

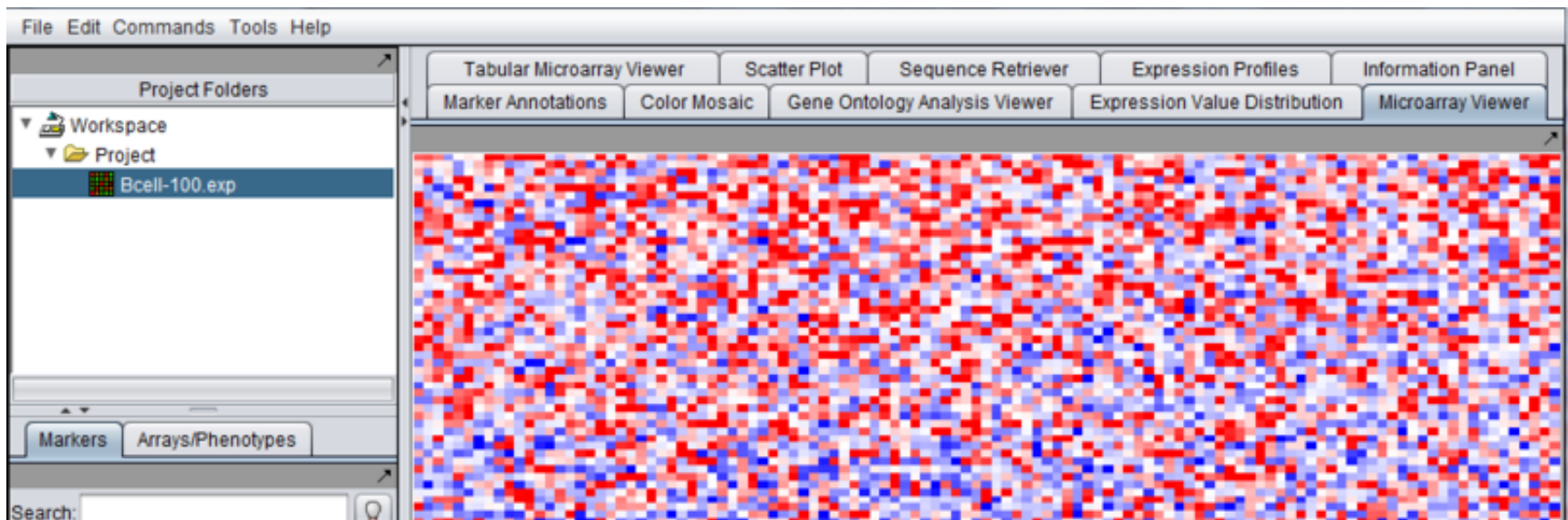
**Data Content**  
Updated today at 06:00

- 36287 experiments
- 1050421 assays
- 13.52 TB of archived data



# geWorkbench

geWorkbench is an open-source bioinformatics platform that offers a comprehensive and extensible collection of tools for the management, analysis, visualization, and annotation of biomedical data. For microarrays, there are tools for filtering and normalization, basic statistical analyses, clustering, network reverse engineering, as well as many common visualization tools





# Cistrome

In addition to the standard Galaxy functions, Cistrome has 29 ChIP-chip- and ChIP-seq-specific tools in three major categories, from preliminary peak calling and correlation analyses, to downstream genome feature association, gene expression analyses, and motif discovery.

The screenshot displays the Galaxy / Cistrome web interface in a browser window. The address bar shows `cistrome.org/ap/root`. The navigation bar includes links for **Analyze Data**, **Workflow**, **Shared Data**, **Lab**, **Visualization**, **Help**, and **User**. A status bar on the right indicates **Using 0 bytes**.

On the left, the **Tools** panel is visible, featuring a search bar and two categories of tools: **CISTROME TOOLBOX** (with links for [Import Data](#), [Data Preprocessing](#), [Gene Expression](#), [Integrative Analysis](#), and [Liftover/Others](#)) and **GALAXY TOOLBOX** (with a link for [Get Data](#)).

The main content area displays a system error message titled **Galaxy/Cistrome! is experiencing problems**, dated **Feb. 6th, 2013**. The message states: "We are currently investigating why jobs on Galaxy/Cistrome are failing. We hope to find the cause of the problem and fix it soon. We will update this page when that happens. Thanks for your bug reports and for your patience in this matter!" It includes an **UPDATE** section: "UPDATE: We believe that we found the cause of the problem. Everything should be back and functional. **\*\*Please report the bugs if you are still experiencing problems\*\***" and a concluding statement: "Our server has a reboot at the beginning of this month, and that cause some problem with cistrome. Now it fixed after restart the cistrome. Sorry for the inconvenience."

On the right, the **History** panel shows **Unnamed history** with **0 bytes** and a message: "Your history is empty. Click 'Get Data' on the left pane to start".



# Gitools

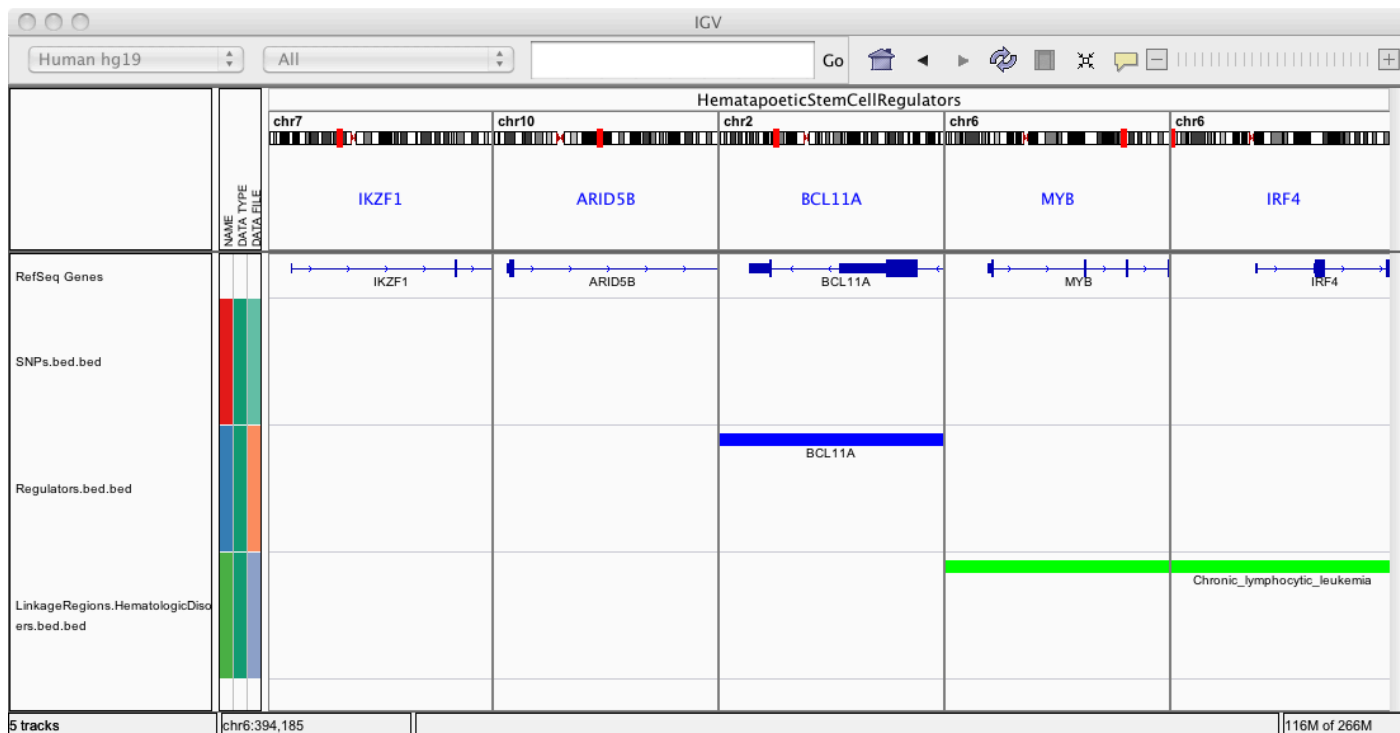
- Gitools is a framework for analysis and visualization of genomic data using interactive heatmaps.





# Integrative Genomics Viewer (IGV)

The **Integrative Genomics Viewer (IGV)** is a high-performance visualization tool for interactive exploration of large, integrated genomic datasets. It supports a wide variety of data types, including array-based and next-generation sequence data, and genomic annotations.







# InSilicoDB

InSilico DB is a web-based genomics data manager containing thousands of curated public datasets. The datasets can be exported to analysis tools and GenomeSpace.

The screenshot displays the InSilicoDB web interface. The top navigation bar includes the InSilicoDB logo, a link to 'Select and export your results to analysis tools', a 'Samples basket' icon, the user email 'liefeld@broadinstitute.org', and a help icon. Below the navigation bar, a search bar contains 'gse14990'. The main content area is divided into a left sidebar with filters and a right section for dataset details.

**Filters:**

- DataSets source:**
  - ☒ My safe [2/2]
  - ☒ Public [6/6]
- Curation:**
  - ☒ Manually curated
  - ☐ To curate
- Platforms:**
  - ☒ MicroArray [13/13]
  - ☒ High Throughput Sequencing [13/13]
- Measurement type:**
  - ☒ MicroArray [1/1]
  - ☒ High Throughput Sequencing [3/3]

**Dataset Details:**

Dataset: **GSE14990** MYC regulation of a "poor prognosis" metastatic cancer cell state

Buttons: Edit/Show clinical annotation, Download, Export, Twitter, Email

Dataset	Sharing	Relevance	#Samples	M
GSE14990 MYC regulation of a "poor prognosis" metastatic cancer cell state	Public	100%	15	



# UCSC Table Browser

The Table Browser allows you to retrieve data associated with a track in text format, to calculate intersections between tracks, and to retrieve DNA sequence covered by a track. After you select the options for your output file, you can opt to send your output file to your GenomeSpace cloud storage.

**Table Browser**

Use this program to retrieve the data associated with a track in text format, to calculate intersections between tracks, and to retrieve DNA sequence covered by a track. For help in using this application see [Using the Table Browser](#) for a description of the controls in this form, the [User's Guide](#) for general information and sample queries, and the OpenHelix Table Browser [tutorial](#) for a narrated presentation of the software features and usage. For more complex queries, you may want to use [Galaxy](#) or our [public MySQL server](#). To examine the biological function of your set through annotation enrichments, send the data to [GREAT](#). Send data to [GenomeSpace](#) for use with diverse computational tools. Refer to the [Credits](#) page for the list of contributors and usage restrictions associated with these data. All tables can be downloaded in their entirety from the [Sequence and Annotation Downloads](#) page.

clade:  genome:  assembly:

group:  track:

table:

region: ☒ genome ☐ position

identifiers (names/accessions):

filter:

intersection:

correlation:

output format:  Send output to ☐ [Galaxy](#) ☐ [GREAT](#) ☐ [GenomeSpace](#)

output file:  (leave blank to keep output in browser)

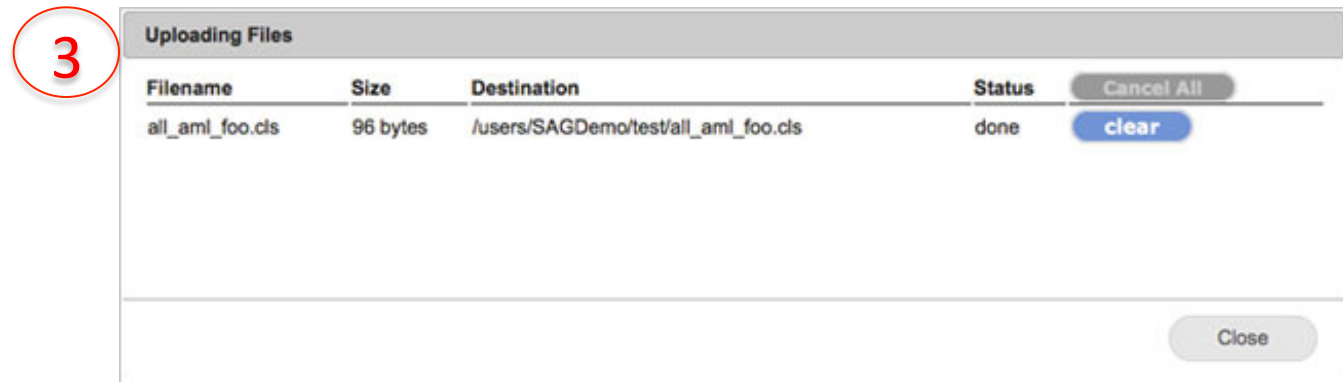
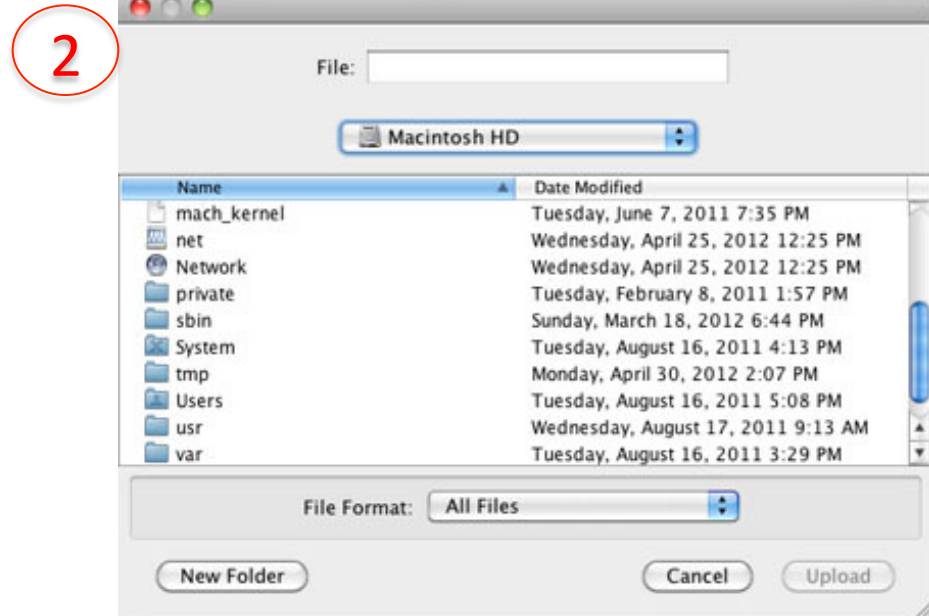
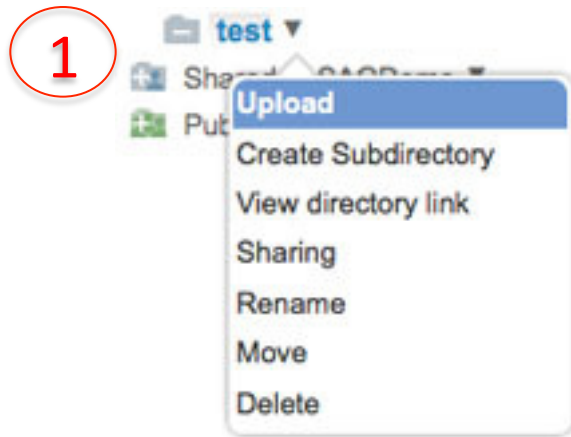
file type returned: ☒ plain text ☐ gzip compressed

To reset all user cart settings (including custom tracks), [click here](#).

# Basic GenomeSpace recipes

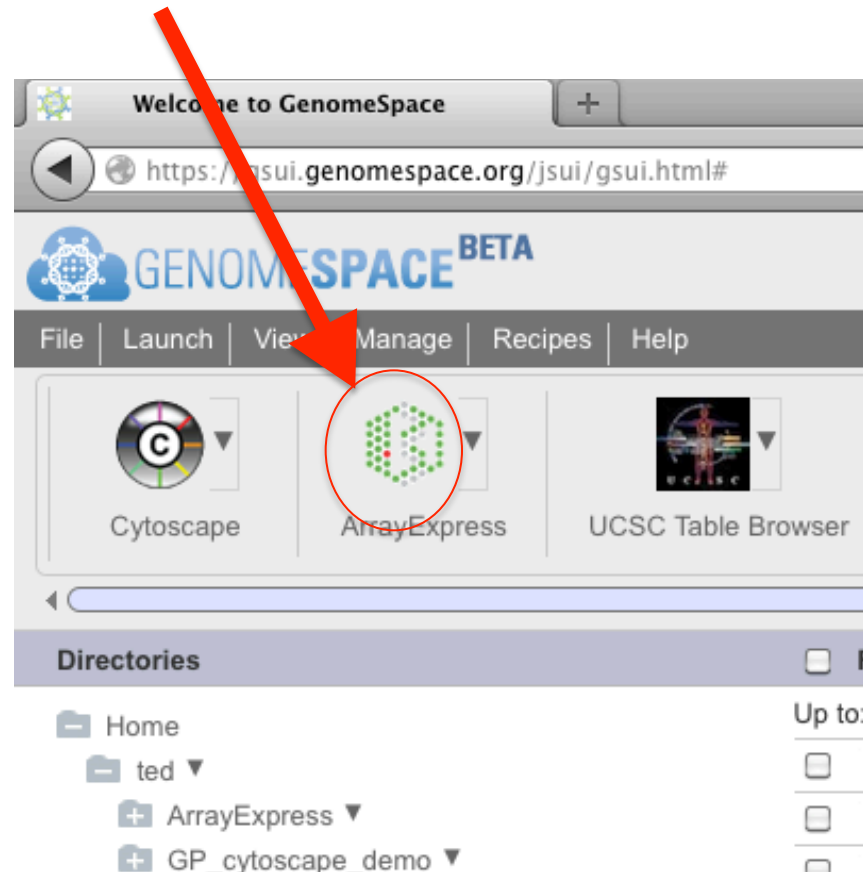
- Uploading data
- Launching tools
- Transitioning across tools

# Uploading Data



# Launching tools

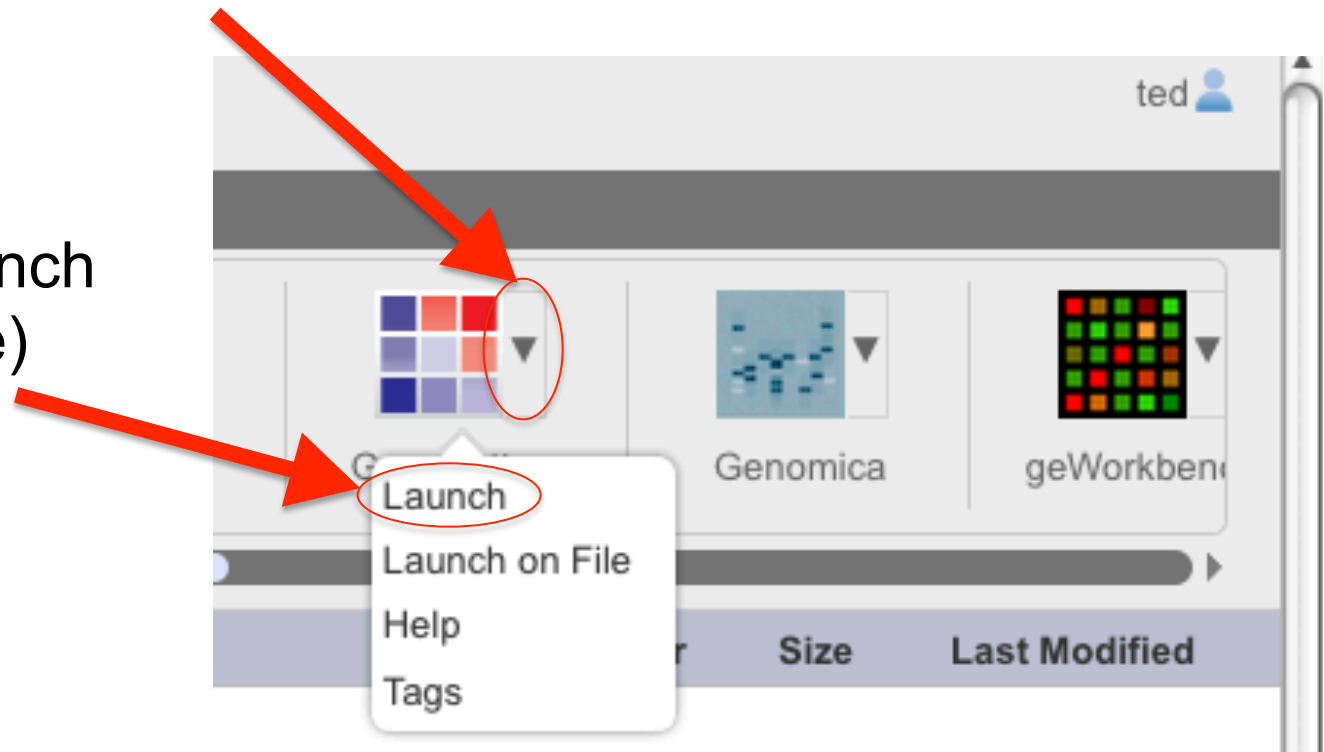
Click on the tool's icon



# Launching tools

Open the tool's context menu

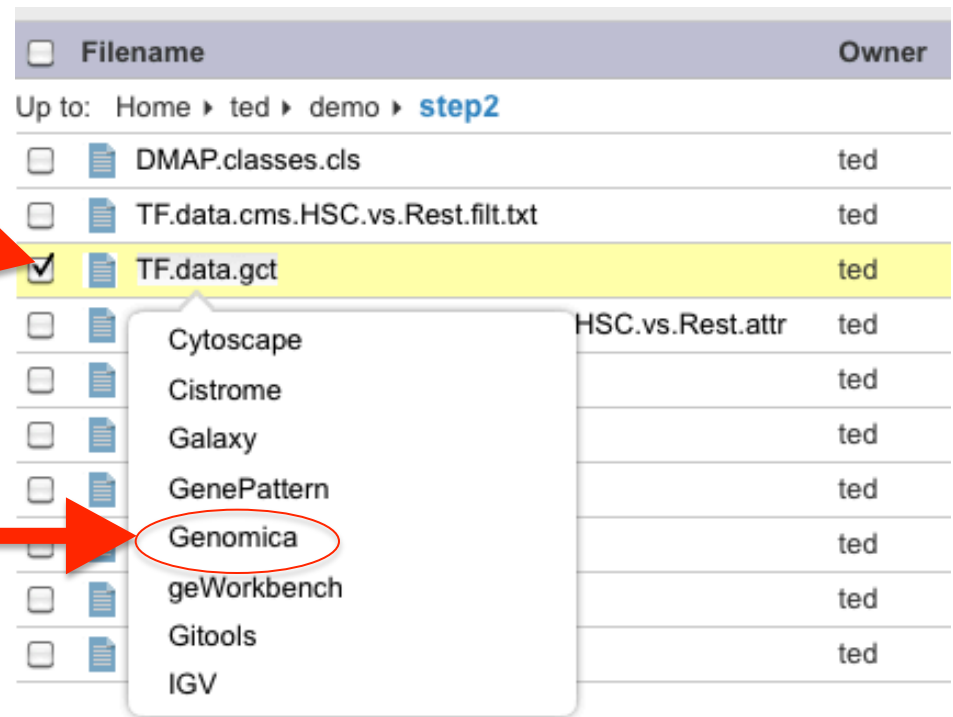
Then click on Launch  
(or Launch on File)



# Launching Tools

Click the checkbox for one (or more) files

Then click on one of the files to get the Launch menu and pick Your tool





# Launching tools

Then click the Launch button

Click and drag a File onto a tool icon

Launch GenePattern

Currently Selected Files

- /ted/demo/step2/TF.data.gct [remove]

Drop files here

Launch Close

Filename	Owner	Size	Last Modified
Up to: Home > ted > demo > step2			
<input type="checkbox"/> DMAP.classes.cls			
<input type="checkbox"/> TF.data.cms.HSC.H3.Rest.filt.txt			
<input type="checkbox"/> TF.data.gct			
<input type="checkbox"/> TF.data.genesymbol.comp.man			
<input type="checkbox"/> TF.data.genesymbol.gct			
<input type="checkbox"/> TF.data.genesymbol.geneset.tab			
<input type="checkbox"/> TF.data.tab	ted	1.31 Mb	12 months ago
<input type="checkbox"/> all_aml_test.res	ted	1.83 Mb	12 months ago
<input type="checkbox"/> all_aml_test.slice.gos.res	ted	433 bytes	5 months ago
<input type="checkbox"/> all_aml_train.comp.marker.attr	ted	2.02 Mb	12 months ago

Owner	Size	Last Modified
ted	664 bytes	12 months ago
ted	1 Kb	12 months ago
ted	1.31 Mb	12 months ago
ted	134 Kb	12 months ago
ted	1.31 Mb	12 months ago
ted	4 Kb	10 months ago
ted	1.31 Mb	12 months ago
ted	1.83 Mb	12 months ago
ted	433 bytes	5 months ago
ted	2.02 Mb	12 months ago

and Harvard | Funded by the National Human Genome Research Institute (NHGRI)

GSE14990GPL570\_RNA\_FRMAGENE\_16835

NP\_minidbp

demo

- step1
- step2
- step3
- step4

galaxy\_export

working dir

©2012 The Broad Institute of MIT and Harvard | Funded by the National Human Genome Research Institute (NHGRI)

# Transitioning across tools

## 1. Launch Genomica

- Load (shared) data from GenomeSpace
- Save it back to a new folder

## 2. Launch GenePattern on your data

- Do a simple processing step
- Save it back to GenomeSpace
- Send it to IGV

## 3. Visualize the processed data IGV

# Launch Genomica

- Using one of the options you saw earlier
  - Click on the icon
  - or use the context menu
  - or use the launch menu
- Load data from GenomeSpace

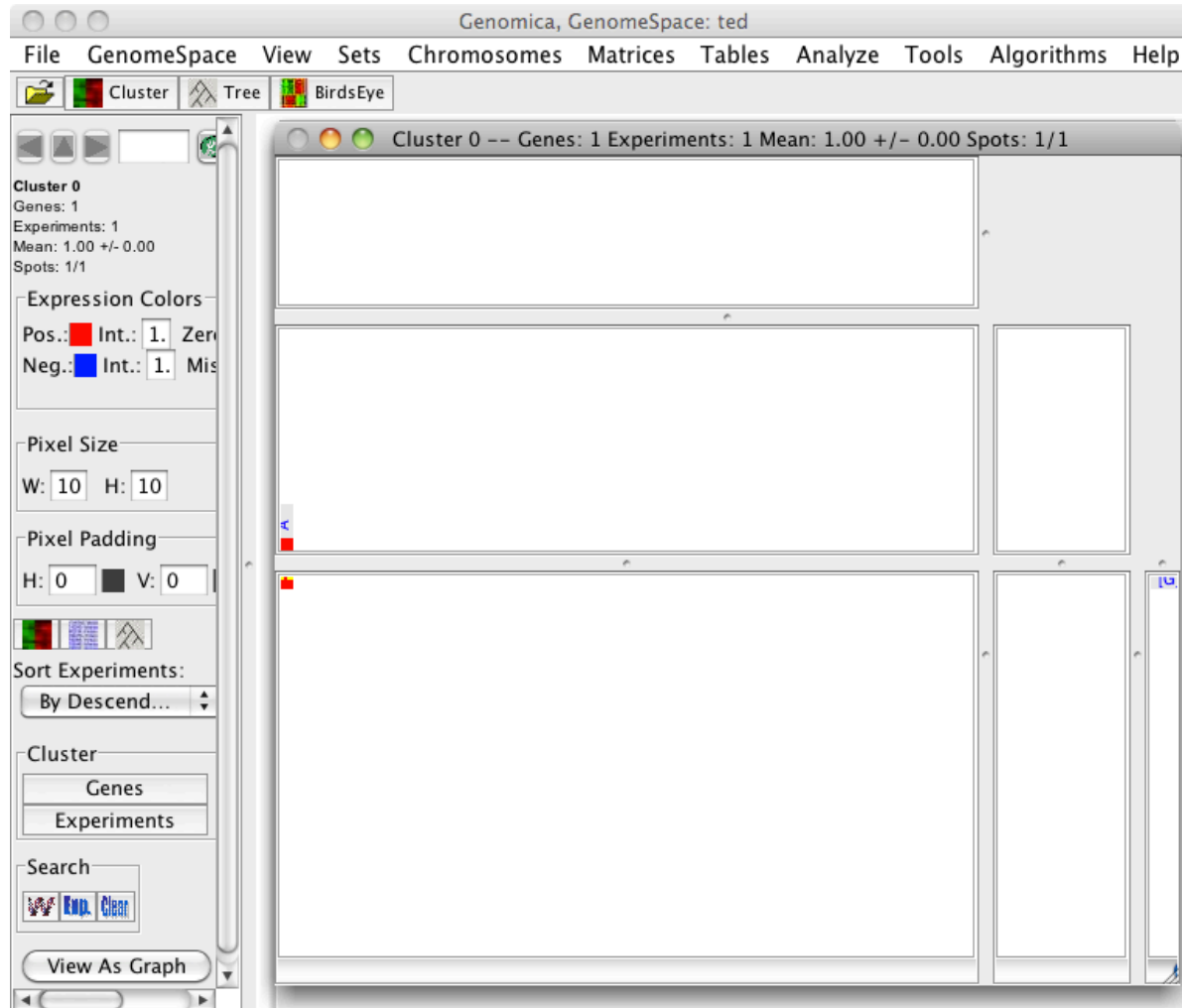
Home ▶ Public ▶ SharedData ▶ Demos ▶ Scenario ▶ step3 ▶ **80\_module.gxp**

Or

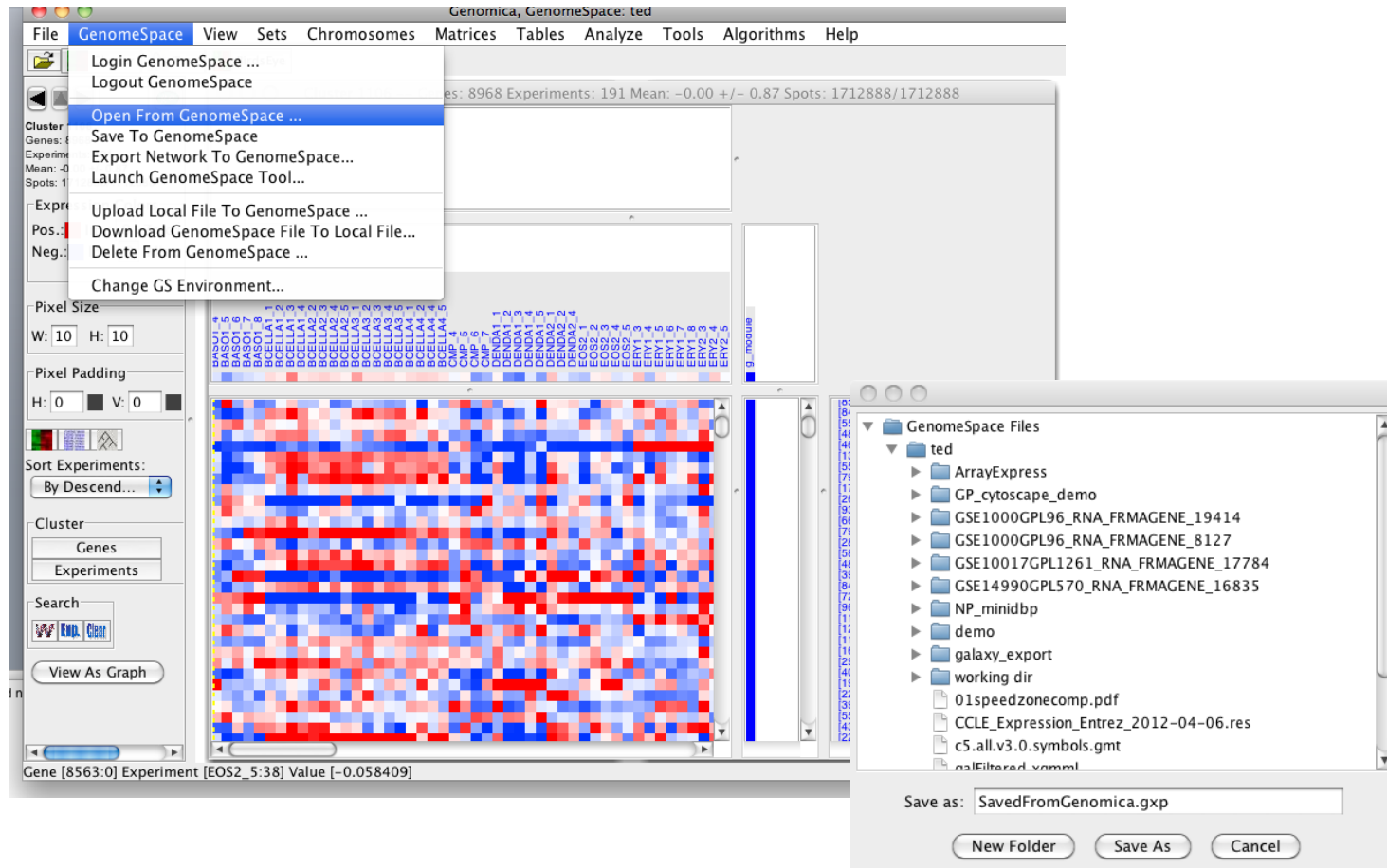
Home ▶ Shared to <your id> ▶ mmr ▶ FGED ▶ **80\_module.gxp**

# Loading into Genomica

Home ► Shared to <your id> ► mmr ► FGED ► **80\_module.gxp**



# Saving Back to GenomeSpace



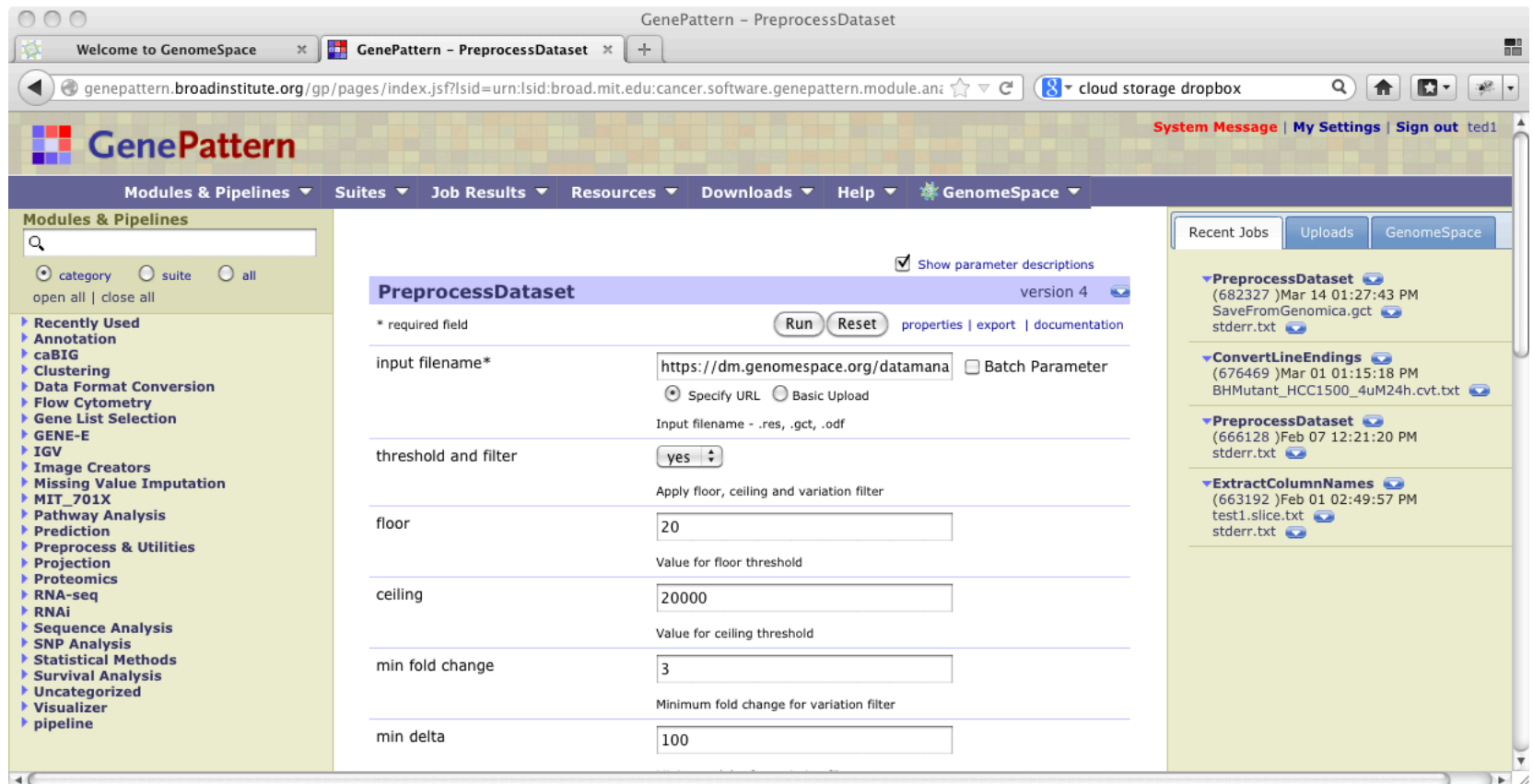
# Launching GenePattern

- You can do this from within Genomica or also from the GenomeSpace interface
- Select “PreprocessDataset” in the send to module



# Process the data

- Run PreprocessDataset with default parameters



The screenshot displays the GenePattern web interface for the PreprocessDataset module. The browser window shows the URL `genepattern.broadinstitute.org/gp/pages/index.jsf?lsid=urn:lsid:broad.mit.edu:cancer.software.genepattern.module.anz`. The interface includes a top navigation bar with links for System Message, My Settings, and Sign out. Below this is a menu bar with options like Modules & Pipelines, Suites, Job Results, Resources, Downloads, Help, and GenomeSpace. A left sidebar lists various modules and pipelines, with 'Preprocess & Utilities' highlighted. The main content area shows the 'PreprocessDataset' module configuration for version 4. It includes a 'Run' button and a 'Reset' button. The configuration fields are as follows:

Parameter	Value
input filename*	<code>https://dm.genomespace.org/datamana</code>
threshold and filter	yes
floor	20
ceiling	20000
min fold change	3
min delta	100

On the right side, there is a 'Recent Jobs' panel showing a list of completed jobs, including 'PreprocessDataset' and 'ConvertLineEndings'.

# Save the result

Use the context menu for the file on  
either the job result page ...

The screenshot shows the GenePattern web interface for Job 682330. The browser address bar shows the URL: [genepattern.broadinstitute.org/gp/jobResults/682330?openVisualizers=true](http://genepattern.broadinstitute.org/gp/jobResults/682330?openVisualizers=true). The page header includes the GenePattern logo and navigation links: [System Message](#), [My Settings](#), and [Sign out](#). The main navigation bar contains: [Modules & Pipelines](#), [Suites](#), [Job Results](#), [Resources](#), [Downloads](#), [Help](#), and [GenomeSpace](#).

Below the navigation bar, a message states: "To see options click [context menu icon] icon. For Input Parameters click [document icon] icon." There is a checkbox for "Show Execution Logs".

The job results section for "682330. PreprocessDataset" shows the following details:

- submitted: Mar 14 01:30:40 PM, completed: Mar 14 01:31:04 PM
- size: 1.2 MB
- input.filename: <https://dm.genomespace.org/dm/genomespace.org/datamanager/dataformat/gct>
- [SaveFromGenomica.gct](#) (1.2 MB) (Last modified: Thu Mar 14 13:30:45 EDT 2013)
- [stderr.txt](#) (1.0 KB) (Last modified: Thu Mar 14 13:30:52 EDT 2013)

At the bottom, a red message states: "These job results are scheduled to be purged from the server on 3/21/13 11:00 PM". Below this is a link: [Return to Modules & Pipelines Start](#).



# Save the result

...or the context menu for the file on the GenePattern home page.

The screenshot shows the GenePattern web interface in a browser window. The address bar displays `genepattern.broadinstitute.org/gp/pages/index.jsf`. The page features a navigation bar with links for **Modules & Pipelines**, **Suites**, **Job Results**, **Resources**, **Downloads**, **Help**, and **GenomeSpace**. On the right, there are links for **My Settings**, **Sign out**, and a user identifier **ted1**.

The main content area includes a green notification box stating: **3/8/13 GSEALeadingEdgeViewer** updated for compatibility with Java 7. Below this is a yellow banner with the text: **Welcome to GenePattern** and **Analyzing genomic data in GenePattern**.

A sidebar on the left lists various modules and pipelines, including **Recently Used**, **Annotation**, **caBIG**, **Clustering**, **Data Format Conversion**, **Flow Cytometry**, **Gene List Selection**, **GENE-E**, **IGV**, **Image Creators**, **Missing Value Imputation**, **MIT\_701X**, **Pathway Analysis**, **Prediction**, **Preprocess & Utilities**, **Projection**, **Proteomics**, **RNA-seq**, **RNAi**, **Sequence Analysis**, **SNP Analysis**, **Statistical Methods**, **Survival Analysis**, **Uncategorized**, **Visualizer**, and **pipeline**.

On the right side, there is a panel titled **Recent Jobs** with tabs for **Uploads** and **GenomeSpace**. A red arrow points to the **GenomeSpace** tab. Below the tabs, a list of recent jobs is shown, including **PreprocessDataset** and **ConvertLineEndings**. A red circle highlights the context menu icon (three dots) next to the first **PreprocessDataset** entry.

The main content area also includes a section titled **what do you want to do?** with two bullet points:

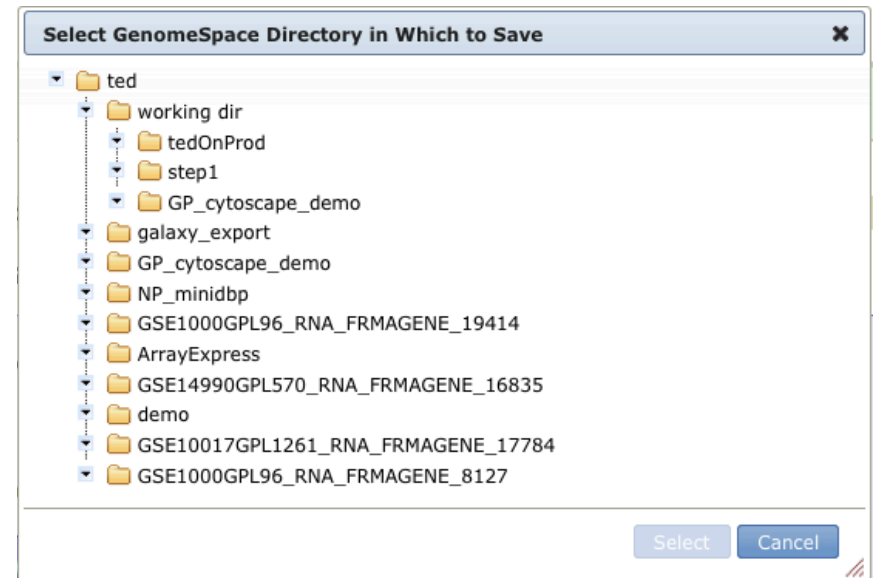
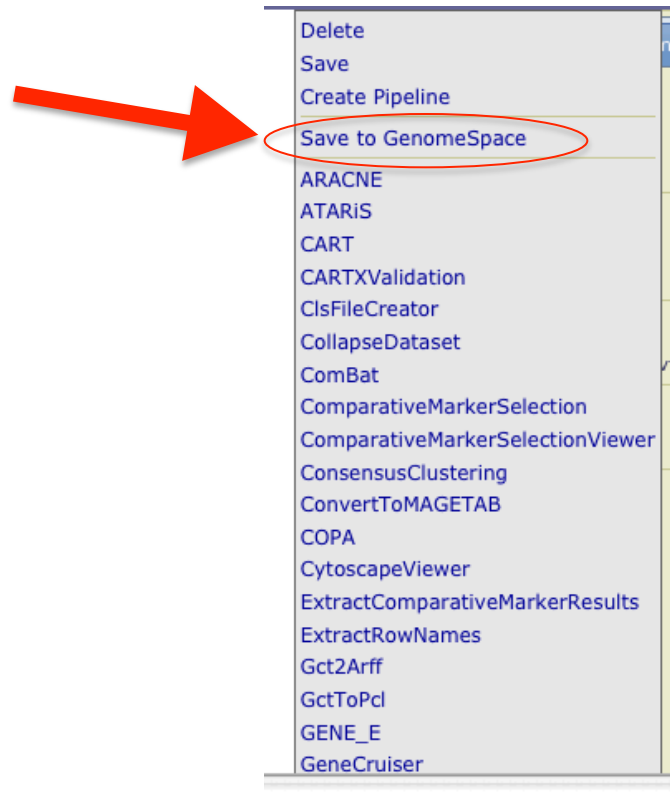
- Click a **protocol** to run an analysis. GenePattern guides you step by step.
- Click **Quick Start** for instructions on how to run any module in GenePattern.

Below this, there is a section titled **Protocols for running common analyses in GenePattern:** with two entries:

- Run an Analysis in GenePattern**: Learn how to run an analysis in GenePattern by preprocessing gene expression data and visualizing the resulting data as a heat map.
- Differential Expression Analysis**: Find genes that are significantly differentially expressed between classes of samples.

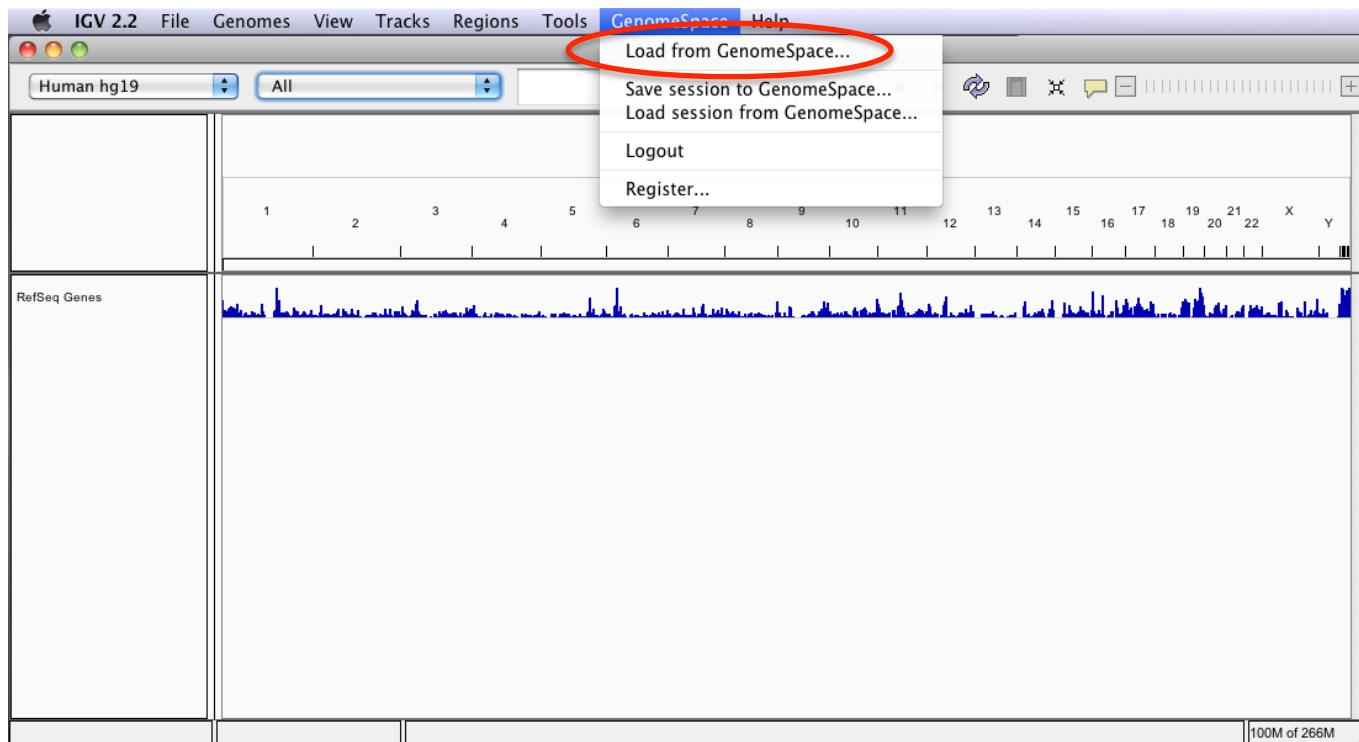
# Saving to GenomeSpace

Click “Save to GenomeSpace”  
from the context menu and then  
select a target directory

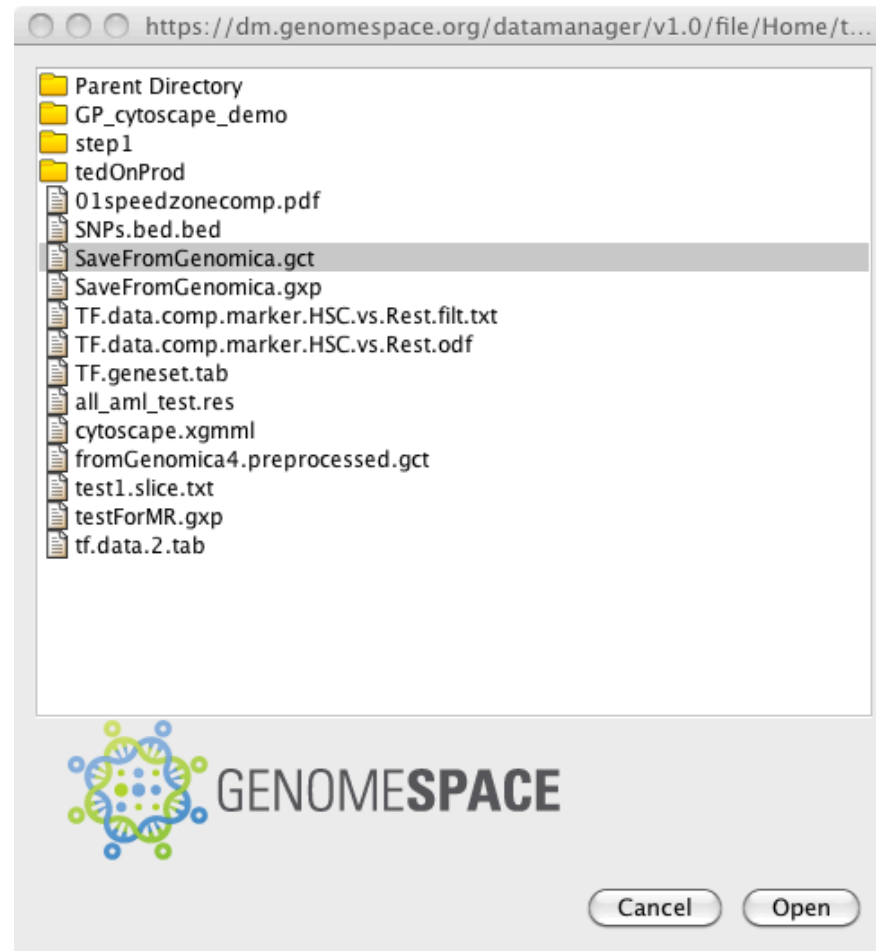


# Send to IGV

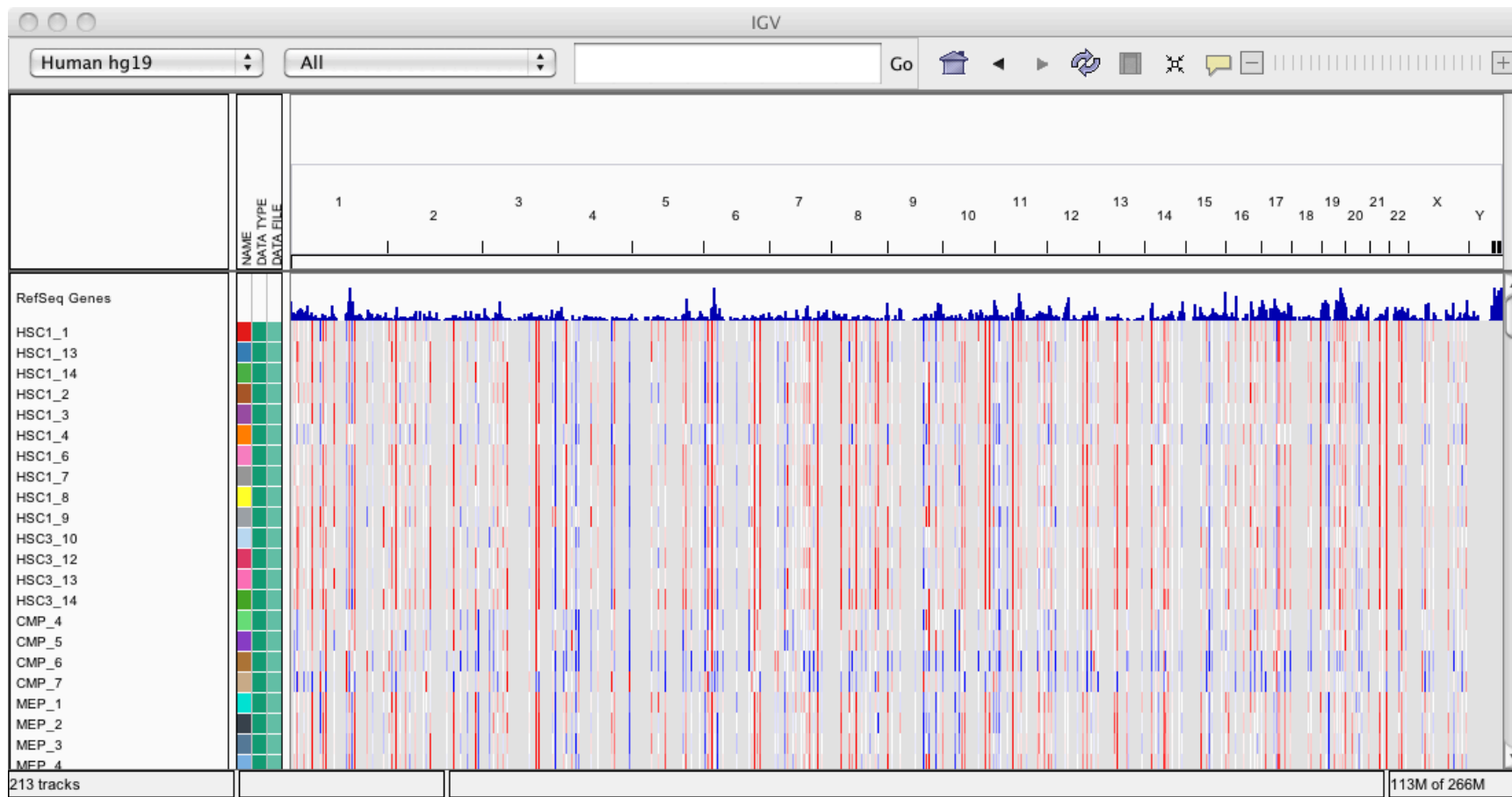
- In the GenomeSpace interface, launch IGV
  - Open the 'GenomeSpace' menu and 'Load from GenomeSpace'



# Select your file (from GenePattern)



# Visualize in IGV



# GenomeSpace UI

A detailed tour of the GenomeSpace  
User Interface

# Agenda

- File Management
- File operations
- Sharing with others
- Organizing your tools

# File Management

- Move a file or directory
- Copy ...
- Deleting ...
- Creating subdirectories
- Recent uploads



# File Operations

- Previewing a file
- Extracting rows and/or columns
- Format conversion

# File Preview

☐ 01speedzonecomp.pdf  
☐ SNPs.bed.bed  
☒ SaveFromGenomica.gct  
☐ SaveFromGenomica.gct  
☐ TF.data.1.filt.txt  
☐ TF.data.2.filt.txt  
☐ TF.data.3.filt.txt  
☐ all\_aml\_test.res  
☐ cytoscape.xgmml  
☐ fromGenomica4.preprocessed.gct  
☐ test1.slice.txt  
☐ testForMR.gxp  
☐ tf.data.2.tab

Preview

Extract rows / cols

Convert

Download

View file link

Sharing

Rename

Move

Delete

Welcome to GenomeSpace
 GenePattern
 +

<https://gsui.genomespace.org/jsui/gsui.html#>
cloud storage dropbox

GENOMESPACE BETA
 ted

Launch View Manage Recipes Help

cytoscape

ArrayExpress

Cistrome

Galaxy

GenePattern

Genomica

geWorkbench

IGV

Ins

File Preview

First 5000 bytes of SaveFromGenomica.gct  

```

#1.2
523 211
Name Description HSC1_1 HSC1_13HSC1_14HSC1_2 HSC1_3 HSC1_4 HSC1_6 HSC1_7 HSC1_8 HSC1_9 HSC3_10
NR2E3 10002 -0.47180805 -0.48530805 -0.14470805 -0.37860805 -0.029208057 -0.082408056 -0.33230805 -0.007408
ZBTB33 10009 -0.11712891 -0.08752891 0.23667109 0.1260711 -0.2489289 -0.05982891 0.03197109 0.1110710
          
```

Close

all\_aml\_test.res

cytoscape.xgmml

fromGenomica4.preprocessed.gct

test1.slice.txt

testForMR.gxp

tf.data.2.tab

ted

ted

ted

ted

ted

ted

1.83 Mb

1 Kb

1.31 Mb

0 bytes

40 Kb

6 Kb

6 minutes ago

6 minutes ago

6 minutes ago

6 minutes ago

6 minutes ago

6 minutes ago

# Extracting Rows and/or Columns

The screenshot shows the GenomeSpace web interface. The browser address bar displays <https://gsui.genomespace.org/jsui/gsui.html#>. The page header includes the "GENOMESPACE BETA" logo and a user profile "ted". A navigation bar contains links for File, Launch, View, Manage, Recipes, and Help. A dropdown menu is open under the "File" tab, listing actions: Upload, Create Subdirectory, Preview, **Extract rows and columns** (highlighted), Convert, Download, View Link URL, Sharing, Rename, Move, and Delete. The main content area shows a file browser for the "workingDir" of user "ted". A table lists files and folders with columns for Filename, Owner, Size, and Last Modified. The file "SaveFromGenomica.gct" is selected and highlighted in yellow.

Filename	Owner	Size	Last Modified
Up to: Home > ted > <b>workingDir</b>			
<input type="checkbox"/> GP_cytoscape_demo	ted		
<input type="checkbox"/> step1	ted		
<input type="checkbox"/> tedOnProd	ted		
<input type="checkbox"/> 01speedzonecomp.pdf	ted	394 Kb	7 minutes ago
<input type="checkbox"/> SNPs.bed.bed	ted	127 Kb	7 minutes ago
<input checked="" type="checkbox"/> <b>SaveFromGenomica.gct</b>	ted	1.17 Mb	7 minutes ago
<input type="checkbox"/> SaveFromGenomica.gxp	ted	1.21 Mb	7 minutes ago
<input type="checkbox"/> TF.data.comp.marker.HSC.vs.Rest.filt.txt	ted	2 Kb	6 minutes ago
<input type="checkbox"/> TF.data.comp.marker.HSC.vs.Rest.odf	ted	126 Kb	6 minutes ago
<input type="checkbox"/> TF.geneset.tab	ted	75 Kb	6 minutes ago
<input type="checkbox"/> all_aml_test.res	ted	1.83 Mb	6 minutes ago

©2012 The Broad Institute of MIT and Harvard | Funded by the National Human Genome Research Institute (NHGRI)

# Extracting rows and/or columns

- Check the columns you want to include
- Provide a first (and optionally last) row index to include
- Edit the file name and 'Save'

The screenshot shows the Genomespace web interface. A dialog box titled 'Extract rows and columns: SaveFromGenomica.gct' is open. It contains a table of columns with checkboxes for selection. The 'Save as' field is set to '/Home/ted/workingDir/SaveFromGenomica.slice.gct'. The 'Start at row' is 4, and the 'End at row' is empty. The 'delimiter' is set to 'Tab'. Below the column selection table, the first 10 lines of the data are displayed. At the bottom of the dialog are 'Save' and 'Close' buttons.

Click on column checkbox to select the column(s) you want to save.

Save as: /Home/ted/workingDir/SaveFromGenomica.slice.gct Start at row: 4 End at row: delimiter: Tab

[Toggle all columns](#)

	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1 #1.2													
2 523													
3 Name Description													
4 NR2E3	10002	-0.47180805	-0.48530805	-0.14470805	-0.37860805	-0.029208057	-0.082408056	-0.33230805	-0.007408057	-0.14750805	-0.24940805	-0.47610804	-0.49470805
5 ZBTB33	10009	-0.11712891	0.23667109	0.1260711	-0.2489289	-0.05982891	0.03197109	0.11107109	0.06697109	-0.3841289	-0.05062891	0.25927	
6 THRAP5	10025	-0.47442418	0.14357583	-0.113424174	-0.53862417	-0.38982418	-0.19352417	-0.21322417	0.006975829	-0.34362417	-0.46462417	-0.09522417	-0.21882
7 NR1H3	10062	1.8970824	1.2402824	1.0456824	1.8085824	1.3698825	1.4439825	1.4667825	1.5027825	1.4988824	1.7011825	1.1578825	1.71958
8 PREB	10113	-0.2005038	0.5561962	0.04779621	-1.4375038	0.21009621	-0.5052038	0.15909621	-0.30600378	0.10009621	-0.44520378	-0.22130379	0.03669
9 ZNF263	10127	-0.12586398	0.30643603	0.09873602	0.10793602	0.36153603	-0.17866398	0.40893602	-0.558464	-0.03476398	0.801736	-0.001463981	0.47363

Showing up to the first 10 lines of SaveFromGenomica.gct

Save Close

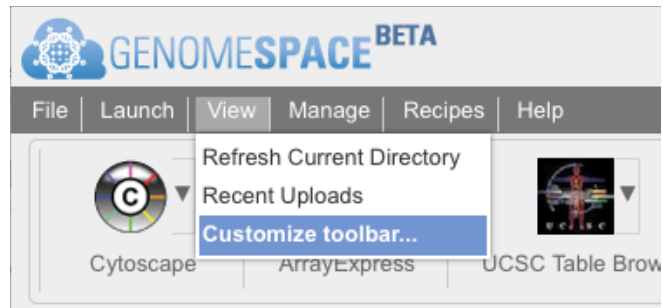
all\_aml\_test.res ted 1.83 Mb 6 minutes ago

©2012 The Broad Institute of MIT and Harvard | Funded by the National Human Genome Research Institute (NHGRI)

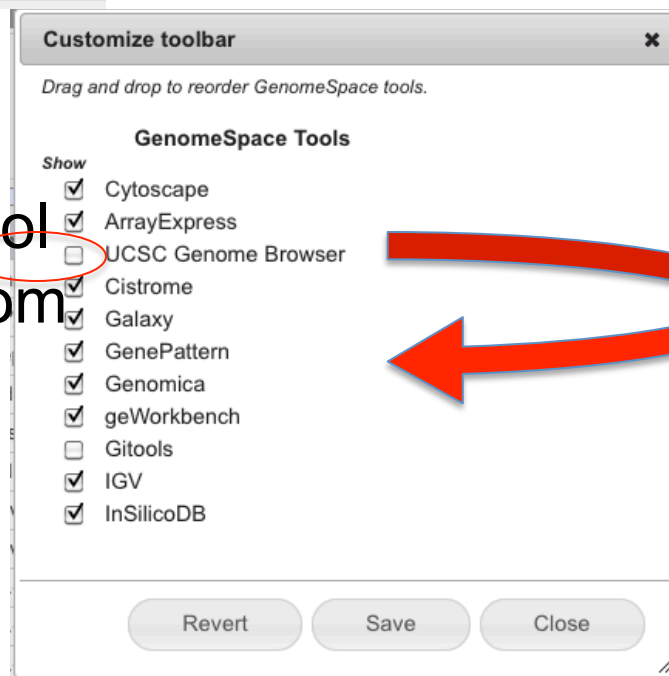
# Sharing with others

- Sharing files with
  - Individuals, groups
    - Creating groups for sharing
- Sharing links
  - With other GenomeSpace users
  - To people without GenomeSpace accounts

# Organizing tools



Uncheck the tool  
To remove it from  
The toolbar



Drag and drop  
tools in the list  
reorder them