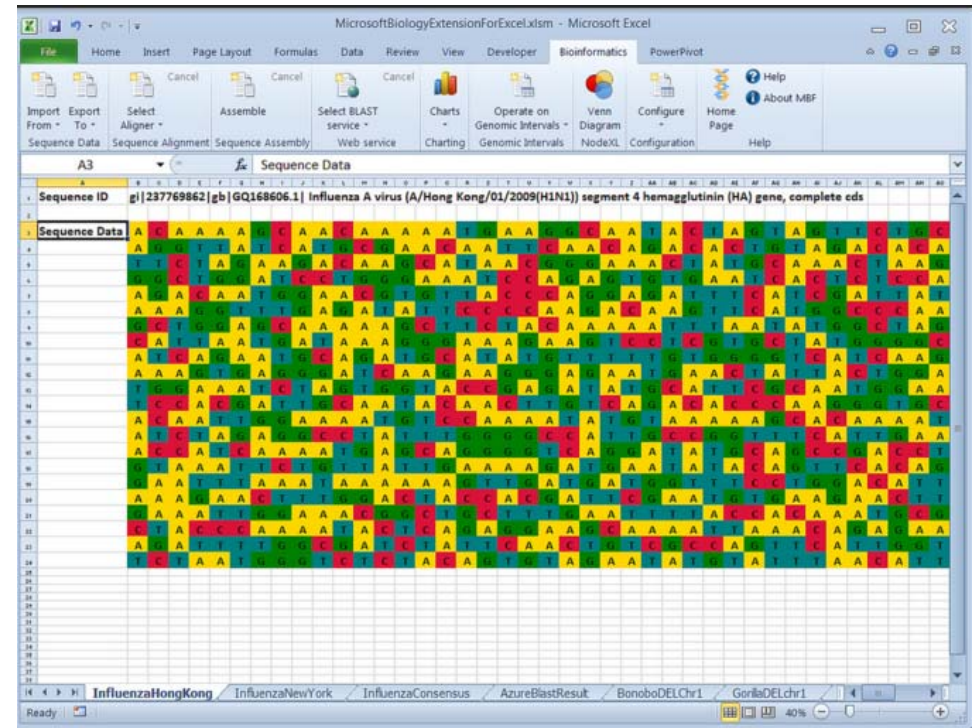


A screenshot of the Galaxy web interface. The top navigation bar includes links for "Analyze Data", "Workflow", "Shared Data", "Visualization", "Help", and "User". A left sidebar lists various tool categories like "Get Data", "Text Manipulation", and "NGS: Mapping". The main content area displays the text: "Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy [start here](\"#\") or consult our [help resources](\"#\")." Below this, the title "What is Galaxy?" is followed by the subtitle "A walkthrough of the Galaxy interface". At the bottom of the main area is the URL <https://galaxyproject.org>. On the right, a "History" panel shows a list of recent jobs, including "68: RawData", "71: FastQC on data", and "70: FastQC on data".

- ❖ What Galaxy is not?
- ❖ What is Galaxy?
- ❖ Why Galaxy makes life simpler?
- ❖ Galaxy's interface
- ❖ A mini practical

What Galaxy is not?

- Galaxy is not Excel.
- Galaxy is not a programming language.
- Galaxy is not Unix/Linux
- Galaxy does not replace Unix/Linux

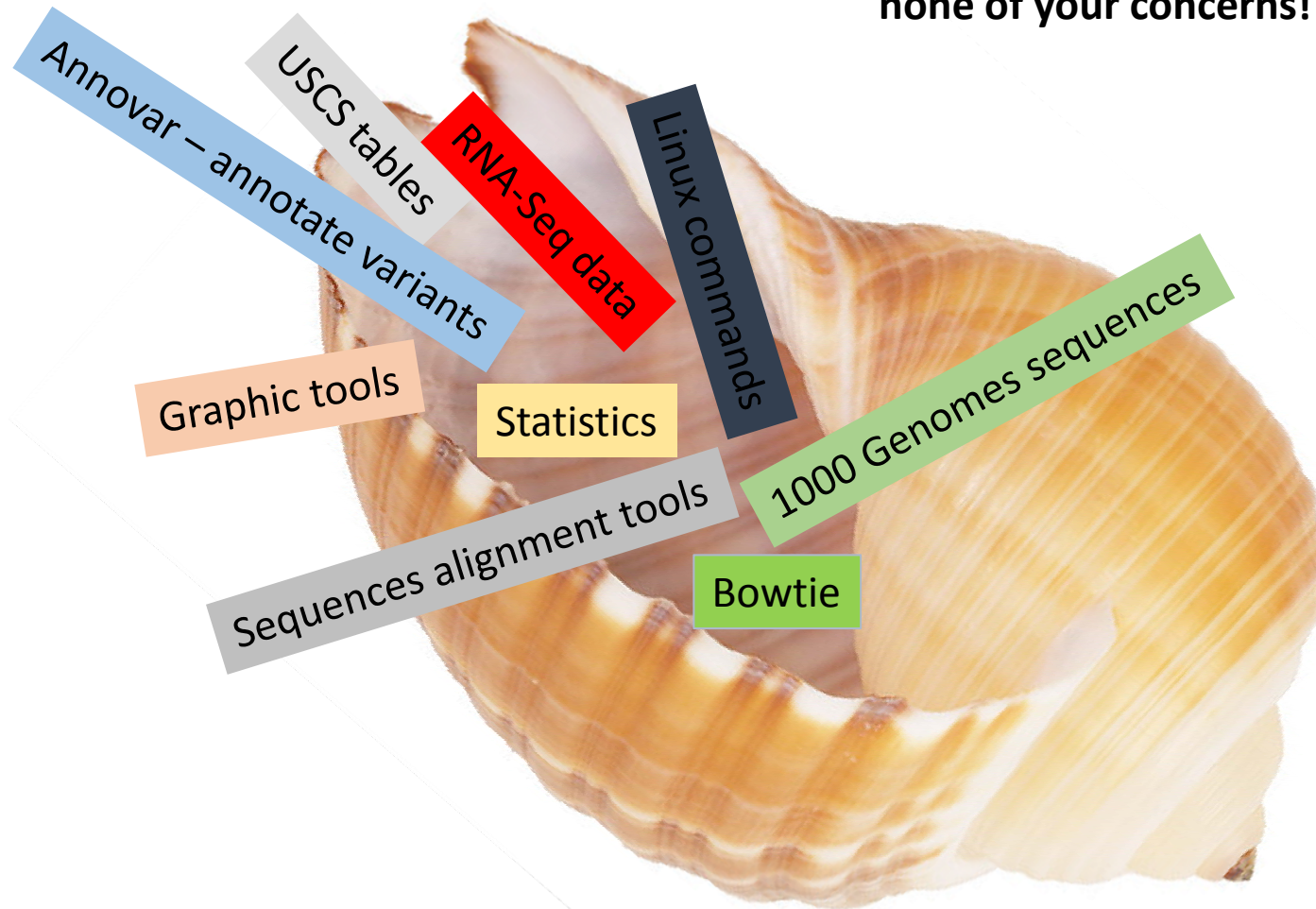
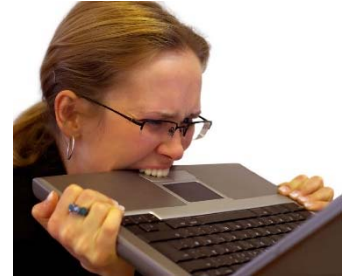


The Microsoft Biology Initiative includes several Microsoft biology tools that enable biology and bioinformatics researchers to be more productive in making scientific discoveries. One such tool, the Microsoft Research Biology Extension for Excel, displays the contents of a FASTA file containing an Influenza A virus sequence. By importing FASTA data into Excel, researchers are better able to visualize and analyze information.

https://www.flickr.com/photos/msr_redmond/5073031709

What is Galaxy?

Implementation, configuration,
version management, OS
compliance, updates, memory
management – **are (almost)
none of your concerns!**



What is Galaxy?

- Galaxy is a free, open, web-based platform for data intensive biomedical research.
- Galaxy has a simple interface to many popular bioinformatics tools and databases
- Galaxy generates reproducible *workflows* (pipelines)
- Galaxy can be used as a data hub

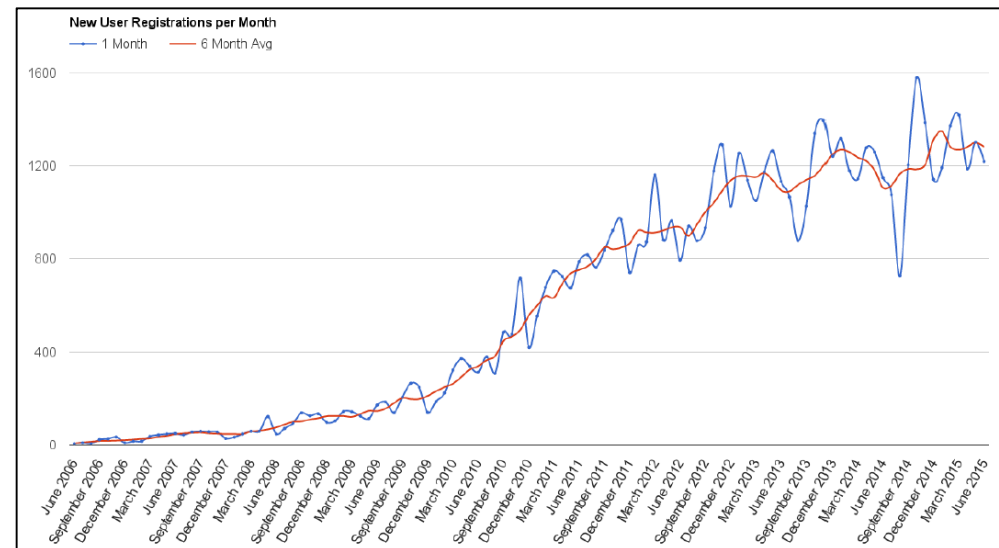
Who makes Galaxy?



+



=



Why Galaxy makes life simpler?

- Galaxy will teach you about command-line options of bioinformatics tools
- Galaxy has a simple ways to process text files
- Galaxy will teach you to build reproducible *workflows* (pipelines)
- Galaxy will help you handle large databases
- Galaxy can be customized
- Many online videos and documentation are available
- Learning Galaxy is as simple as falling off the bicycle. You don't forget it.

Some resources

The main Galaxy server:

<https://usegalaxy.org>

Galaxy wiki:

<https://wiki.galaxyproject.org>

Biostar forum:

<http://biostar.usegalaxy.org>

Learning how to use Galaxy:

<https://wiki.galaxyproject.org/Learn>

<https://wiki.galaxyproject.org/Teach/Resources>

<http://bioinformatics.sph.harvard.edu/ngs-workshops/courses/introduction-to-galaxy/>

Mailing lists:

List Link	Description	List Email Address	Archives (see Searching)
galaxy-announce	Announcements of interest to the Galaxy community. <i>Moderated and low volume.</i>		Nabble, Mailman
galaxy-dev	Local installation, configuration, and tool integration help, or to propose new features. <i>High volume.</i>	<galaxy-dev AT lists DOT galaxyproject DOT org>	Nabble, GMail, Mail- Archive.com, OSDir, Mailman
galaxy-proteomics	Proteomics related questions, announcements, and anything else of interest to the Galaxy proteomics community. <i>Low volume.</i>	<galaxy-proteomics AT lists DOT galaxyproject DOT org>	Nabble, Mailman
galaxy-training	If you have a question about teaching bioinformatics with Galaxy, or about teaching Galaxy itself, then this is a good place to find an answer. This is also the official mailing list of the Galaxy Training Network (GTN).	<galaxy-training AT lists DOT galaxyproject DOT org>	Nabble, Mailman
galaxy-france	Cette liste est destinée à l'information (et aux discussions) de la Communauté française Galaxy. (This list is for announcements to (and discussion within) the French Galaxy Community. Most list content is in French.) <i>Faible volume / Low volume.</i>	<galaxy-france AT lists DOT galaxyproject DOT org>	Nabble, Mailman
galaxy-uk	Mailing list for the Galaxy-UK community. This is not a support list.	<galaxy-uk AT groups DOT galaxyproject DOT org>	Google Groups, Nabble
galaxy-commits	Galaxy source control commit messages.		Mailman

Galaxy has four important interface areas

Main Galaxy menu bar



The screenshot displays the Galaxy web interface with three main sections highlighted by red arrows and labels:

- Tools and commands:** A vertical panel on the left containing a search bar and a list of tool categories such as Get Data, Send Data, Lift-Over, Text Manipulation, Datamash, Convert Formats, Filter and Sort, Join, Subtract and Group, Fetch Alignments/Sequences, NGS: QC and manipulation, NGS: Mapping, NGS: RNA Analysis, NGS: SAMtools, NGS: BamTools, NGS: Picard, NGS: VCF Manipulation, NGS: Peak Calling, NGS: Variant Analysis, NGS: RNA Structure, NGS: DuplexNovo, Operate on Genomic Intervals, Statistics, Graph/Display Data, CloudMap, Phenotype Association, BEDTools, Genome Diversity, EMBOSS, Regional Variation, FASTA manipulation, Multiple Alignments, Metagenomic Analysis, Multiple regression, Multivariate Analysis, and Motif Tools.
- Main content and results area:** The central workspace showing the output of a tool. It includes a header for "#FastQC 0.11.2 pass" and a table of sequence quality metrics. The table has columns for "Base", "Mean", and "Pass". The data shows various quality scores across different bases and positions.
- History:** A panel on the right showing a list of recent jobs. Each job entry includes a name (e.g., "78: Filter on data 7", "77: Filter on data 7", "74: Map with BWA-MEM on data 72 and data 70 (mapped reads in BAM format)", "73: Bowtie2 on data 72 and data 70 (sorted reads (sorted BAM)", "72: FastQC on data 68: RawData", "71: FastQC on data 68: Webpage", "70: FastQC on data 67: RawData", "69: FastQC on data 67: Webpage", "68: SampleX CTCA GA L001_R2_00 1.fastq", "67: SampleX CTCA GA L001_R1_00 1.fastq"), a size (e.g., "170.9 MB"), a format (e.g., "format: fastq, database: hg_g1k_v37"), and an "uploaded fastq file" button.

The menu bar

Analyze Data

Workflow

Main working
environment

Analyze Data

Workflow

Shared Data ▾

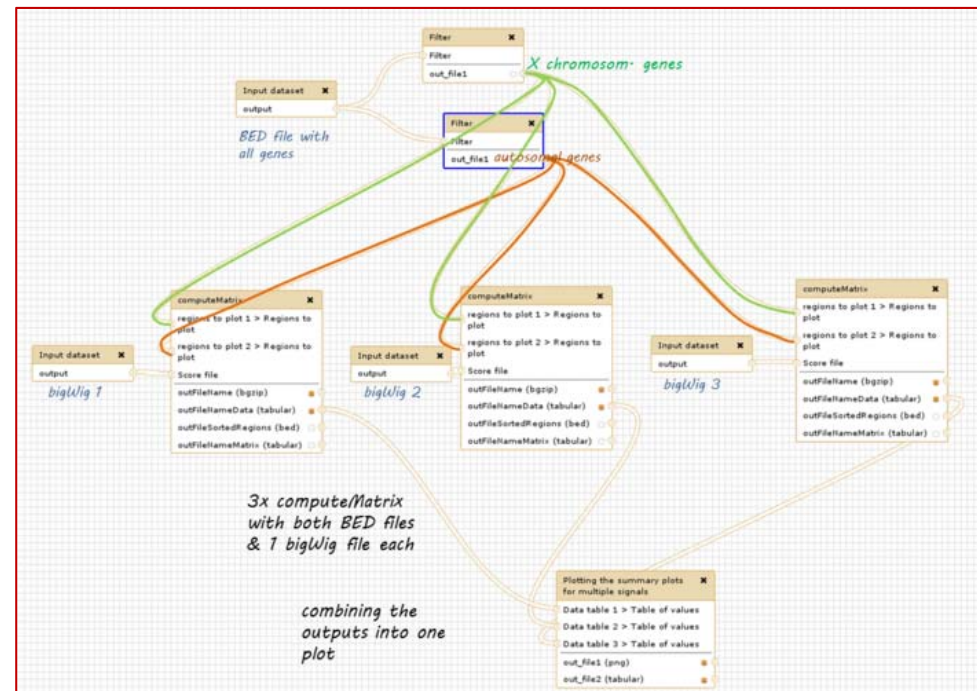
Visualization ▾

Help ▾

User ▾



Observe, edit, and create your
workflow



Analyze Data

Workflow

Shared Data ▾

Visualization ▾

Help ▾

User ▾



Access published data, histories, workflows
Share your data

Data Libraries deprecated

Data Libraries

Published Histories

Published Workflows

Published Visualizations

His

Analyze Data Workflow Shared Data Visualization Help User

New Track Browser
Saved Visualizations

Visualize genomic data

Analyze Data Workflow Shared Data Visualization Help User

Galaxy help

- Galaxy Biostar
- Ask a question
- Support
- Search
- Mailing Lists
- Videos
- Wiki
- How to Cite Galaxy
- Interactive Tours
- Terms and Conditions

th Percentile 90th Percentile
.0
.0
.0

Analyze Data Workflow Shared Data Visualization Help User

Login to record your history,
access saved history, and datasets

Logged in as eelhaik@gmail.com
Preferences
Custom Builds
Logout

Saved Histories
Saved Datasets
Saved Pages
API Keys

0th Percentile 90th Percentile
4.0

History menu

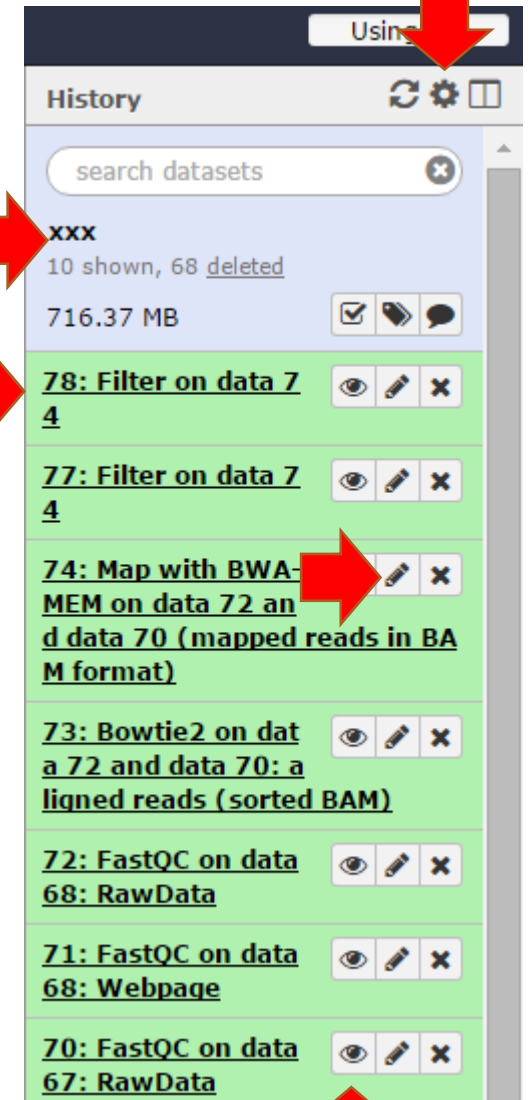
Give a meaningful name to your history
(you will create several 'histories')

Click on a file name to access various
details and see a (tiny) file preview

Rename a file

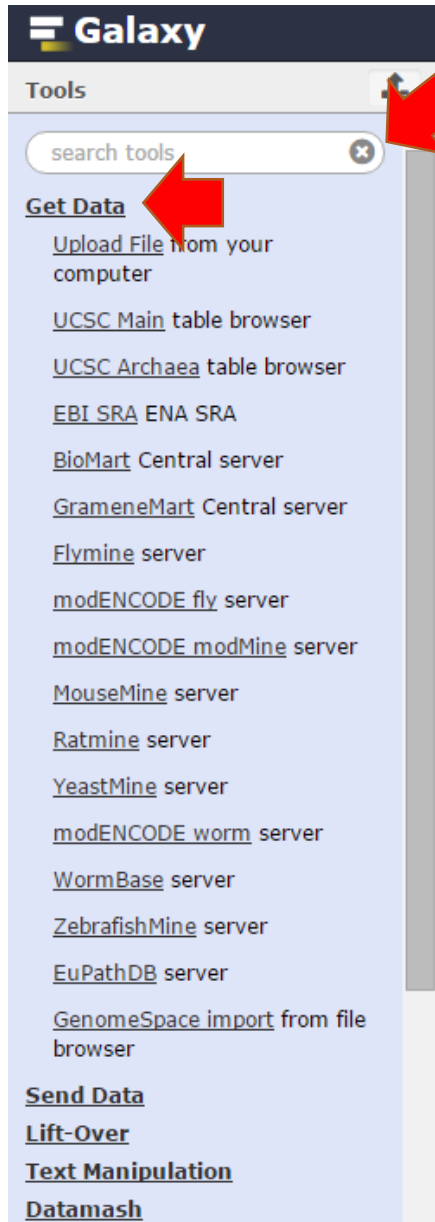
You can do other stuff while one job is running

More important commands



View the file/s
(enable scratch book)

Tools menu

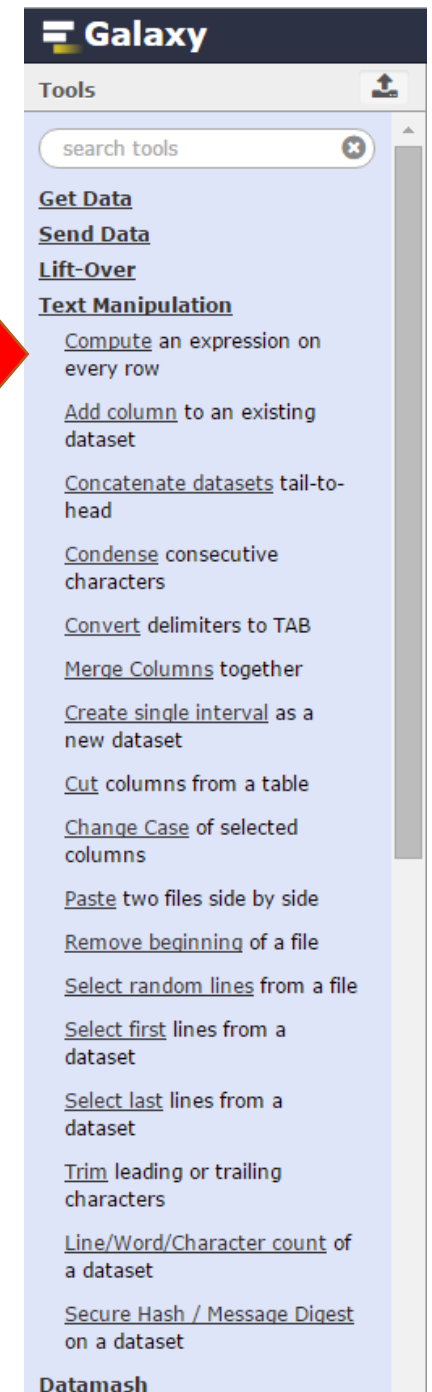


The screenshot shows the Galaxy web interface's 'Tools' menu. At the top is a search bar labeled 'search tools'. Below it, the menu is organized into sections: 'Get Data' (containing links like 'Upload File from your computer', 'UCSC Main table browser', 'UCSC Archaea table browser', 'EBI SRA ENA SRA', 'BioMart Central server', 'GrameneMart Central server', 'Flymine server', 'modENCODE fly server', 'modENCODE modMine server', 'MouseMine server', 'Ratmine server', 'YeastMine server', 'modENCODE worm server', 'WormBase server', 'ZebrafishMine server', 'EuPathDB server', and 'GenomeSpace import from file browser'), 'Send Data', 'Lift-Over', 'Text Manipulation', and 'Datamash'. Red arrows point from the text annotations to the search bar and the 'Get Data' section.

Often quicker to search for tool

Lots of options for getting data

Non-bioinformatics tools



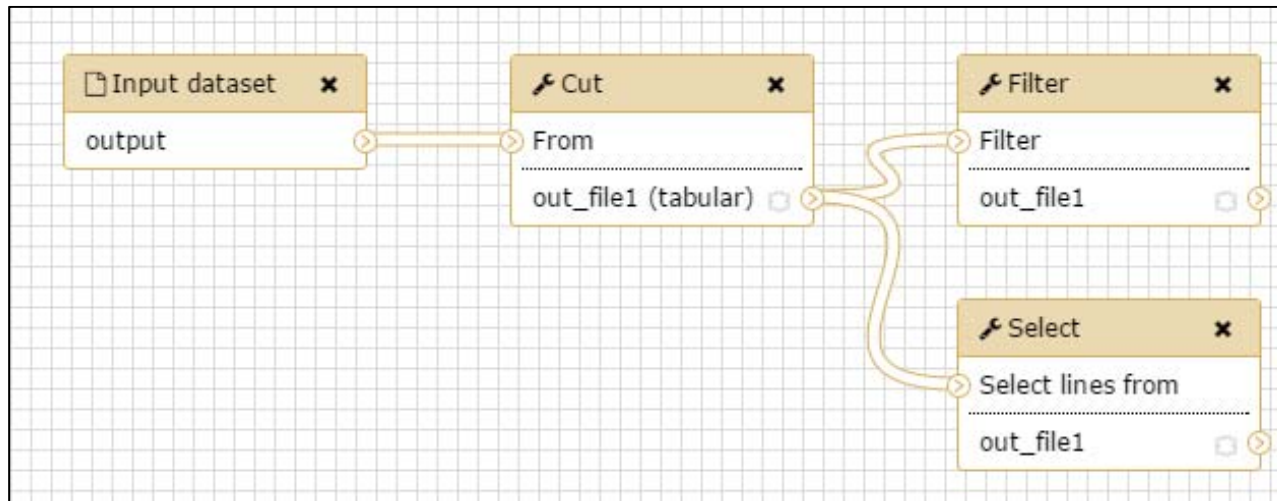
This screenshot shows a different view of the Galaxy 'Tools' menu, focusing on non-bioinformatics tools. It features a search bar at the top. The visible tools are categorized under 'Get Data', 'Send Data', 'Lift-Over', and 'Text Manipulation'. The 'Text Manipulation' section is expanded, showing a long list of tools such as 'Compute an expression on every row', 'Add column to an existing dataset', 'Concatenate datasets tail-to-head', 'Condense consecutive characters', 'Convert delimiters to TAB', 'Merge Columns together', 'Create single interval as a new dataset', 'Cut columns from a table', 'Change Case of selected columns', 'Paste two files side by side', 'Remove beginning of a file', 'Select random lines from a file', 'Select first lines from a dataset', 'Select last lines from a dataset', 'Trim leading or trailing characters', 'Line/Word/Character count of a dataset', and 'Secure Hash / Message Digest on a dataset'. The 'Datamash' tool is listed at the bottom. A red arrow points from the text annotation to the 'Text Manipulation' section.

Not all Galaxies are created equal

- The Galaxy interface continues to evolve
- Not all Galaxy installations will have the same set of tools
- Not all Galaxy installations will have the same **versions** of tools
- This has been a brief walkthrough!

Give it a try

Count the number of human SNPs with a “missense” functions in chr 21
Count the number of human SNPs with a “stop-loss” functions in chr 21



Count missense variants
4 shown, 9 [deleted](#)
56.24 MB

☒

13: Select on data 3

7: Filter on data 3

3: Cut on data 1

1: UCSC Main on Human: snp142Common (chr21:1-46709983)

Try also to:

Preview each file, View each file, View **all** files (with Scratchbook), Rename them
Try manipulating one of your files (view & rename file, download the file)
History (Rename, Export history (only if files are small), Create new history)