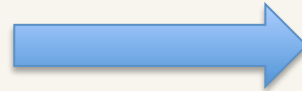


Getting data into Galaxy

Tools

- [BEDTools](#)
- [Get Data](#)
- [ENCODE Tools](#)
- [Lift-Over](#)
- [Text Manipulation](#)
- [Filter and Sort](#)
- [Join, Subtract and Group](#)
- [Convert Formats](#)
- [Extract Features](#)



Tools

- [BEDTools](#)
- [Get Data](#)
 - [BioMart Central server](#)
 - [Upload File](#) from your computer
 - [UCSC Main](#) table browser
 - [UCSC Archaea](#) table browser
 - [BX main](#) browser

Getting data into Galaxy

Galaxy / UIUC

Analyze Data Workflow Shared Data Visualization Help User

Using 57.2 Gb

Tools

- BEDTools
- Get Data
 - BioMart Central server
 - Upload File from your computer**
 - UCSC Main table browser
 - UCSC Archaea table browser
 - BX main browser
- 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
- Statistics
- Graph/Display Data
- Regional Variation
- Multiple regression
- Multivariate Analysis
- Multiple Alignments
- Metagenomic analyses
- FASTA manipulation
- NCBI BLAST+

Upload File (version 1.1.3)

File Format:
Auto-detect

Which format? See help below

File:
Choose File no file selected

TIP: Due to browser limitations, uploading files larger than 2GB is guaranteed to fail. To upload large files, use the URL method (below) or FTP (if enabled by the site administrator).

URL/Text:

Here you may specify a list of URLs (one per line) or paste the contents of a file.

Files uploaded via FTP:

File	Size	Date
Your FTP upload directory contains no files.		

This Galaxy server allows you to upload files via FTP. To upload some files, log in to the FTP server at biocluster.igb.illinois.edu using your Galaxy credentials (email address and password).

Convert spaces to tabs:
 Yes
Use this option if you are entering intervals by hand.

Genome:
----- Additional Species Are Below -----

Execute

History

Unnamed history 0 bytes

Your history is empty. Click 'Get Data' on the left pane to start

- Click “Choose File”, and proceed to pick your file from your computer.
- Click Execute to start the upload
- **2Gb file size limit**

Getting data into Galaxy

Galaxy / UIUC

Analyze Data Workflow Shared Data Visualization Help User Using 57.2 Gb

Tools

- BEDTools
- Get Data
 - BioMart Central server
 - Upload File from your computer**
 - UCSC Main table browser
 - UCSC Archaea table browser
 - BX main browser
- 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
- Statistics
- Graph/Display Data
- Regional Variation
- Multiple regression
- Multivariate Analysis
- Multiple Alignments
- Metagenomic analyses
- FASTA manipulation
- NCBI BLAST+

Upload File (version 1.1.3)

File Format:
Auto-detect
Which format? See help below

File:
Choose File no file selected
TIP: Due to browser limitations, uploading files larger than 2GB is guaranteed to fail. To upload large files, use the URL method (below) or FTP (if enabled by the site administrator).

URL/Text:

Here you may specify a list of URLs (one per line) or paste the contents of a file.

Files uploaded via FTP:

File	Size	Date
Your FTP upload directory contains no files.		

This Galaxy server allows you to upload files via FTP. To upload some files, log in to the FTP server at biocluster.igb.illinois.edu using your Galaxy credentials (email address and password).

Convert spaces to tabs:
 Yes
Use this option if you are entering intervals by hand.

Genome:
----- Additional Species Are Below -----

Execute

History

Unnamed history 0 bytes

Your history is empty. Click 'Get Data' on the left pane to start

- This is generally not advisable, since usually files are very large; you can paste small files in.

Getting data into Galaxy

The screenshot shows the Galaxy / UIUC interface. The top navigation bar includes 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The right side shows 'Using 57.2 Gb' and a 'History' panel with 'Unnamed history' (0 bytes) and a message: 'Your history is empty. Click 'Get Data' on the left pane to start'.

The main content area is titled 'Upload File (version 1.1.3)'. It features several sections:

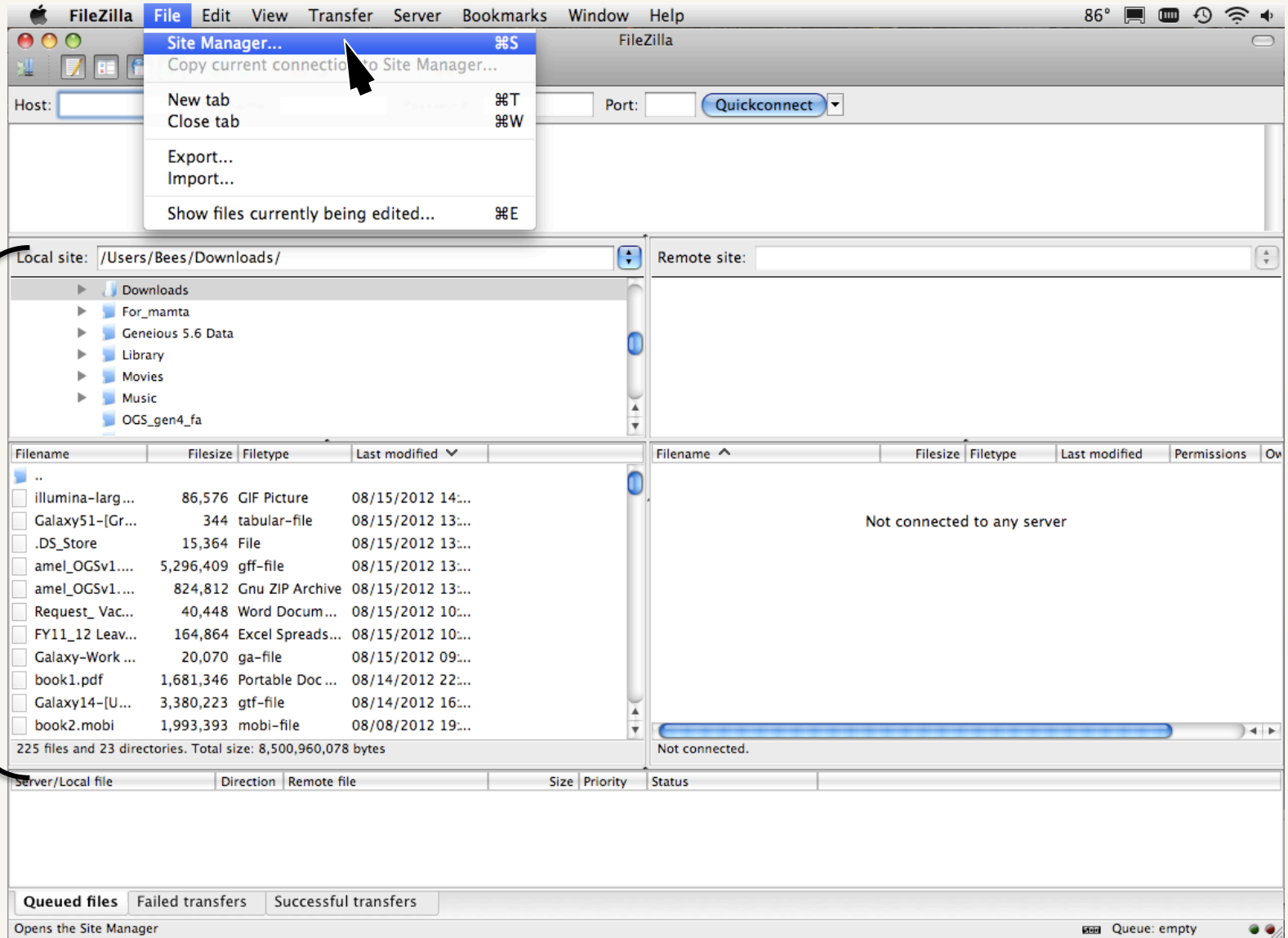
- File Format:** A dropdown menu set to 'Auto-detect' with the text 'Which format? See help below'.
- File:** A 'Choose File' button followed by 'no file selected'. A tip below reads: 'TIP: Due to browser limitations, uploading files larger than 2GB is guaranteed to fail. To upload large files, use the URL method (below) or FTP (if enabled by the site administrator)'.
- URL/Text:** A large text input field with the instruction: 'Here you may specify a list of URLs (one per line) or paste the contents of a file.'
- Files uploaded via FTP:** A table with columns 'File', 'Size', and 'Date'. Below the table, it states: 'Your FTP upload directory contains no files.' A note below explains: 'This Galaxy server allows you to upload files via FTP. To upload some files, log in to the FTP server at biocluster.igb.illinois.edu using your Galaxy credentials (email address and password).'
- Convert spaces to tabs:** A checkbox labeled 'Yes' with the instruction: 'Use this option if you are entering intervals by hand.'
- Genome:** A dropdown menu currently showing '----- Additional Species Are Below -----'.

At the bottom of the main content area is an 'Execute' button. The left sidebar contains a 'Tools' menu with categories like 'BEDTools', 'Get 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', 'Statistics', 'Graph/Display Data', 'Regional Variation', 'Multiple regression', 'Multivariate Analysis', 'Multiple Alignments', 'Metagenomic analyses', 'FASTA manipulation', and 'NCBI BLAST+'. The 'Upload File from your computer' option under 'Get Data' is circled in black.

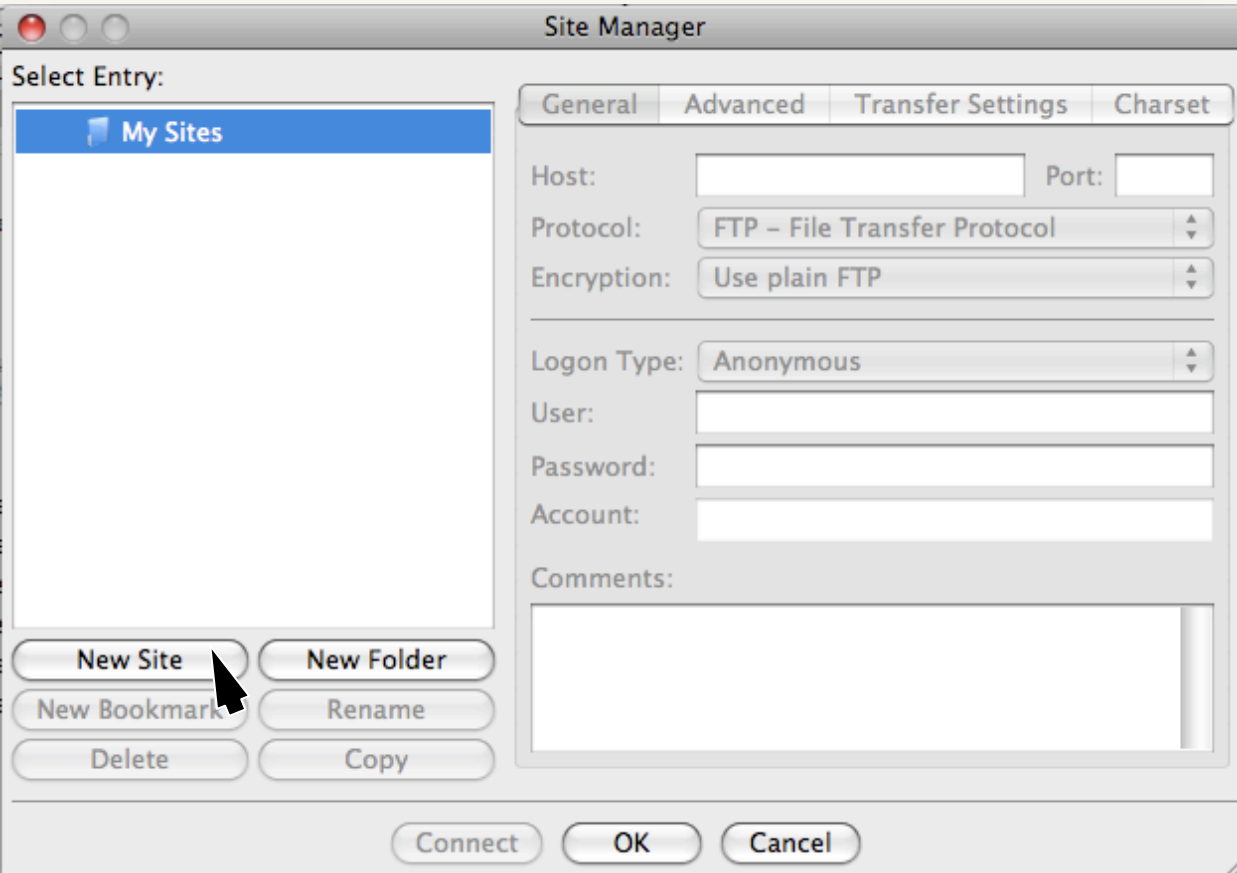
- You can upload files (SFTP) using an FTP client like Filezilla, Cyberduck, Fugu

Getting data into Galaxy

Directory structure of your current computer

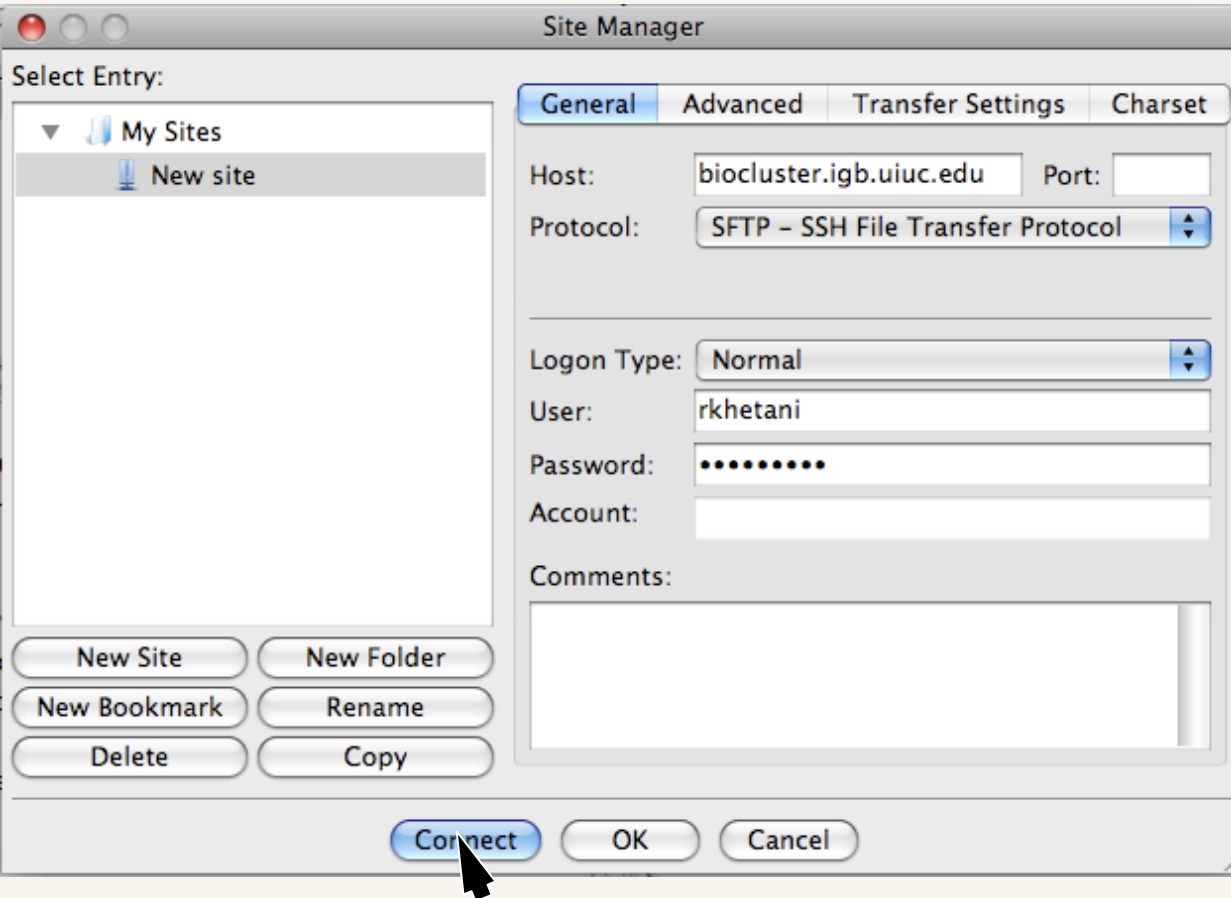


Getting data into Galaxy



- Add a new site

Getting data into Galaxy



- In the “Host:” field type biocluster.igb.uiuc.edu
- Leave “Port:” empty
- Make sure you select SFTP from the “Protocol:” pull-down menu
- Change “Logon Type:” to Normal
- Enter your galaxy username and password
- Leave “Account:” empty
- Click Connect

Getting data into Galaxy

The screenshot shows the FileZilla interface with a split window view. The top bar indicates the connection to 'New site - sftp://rkhetani@biocluster.igb.uiuc.edu - FileZilla'. The 'Host', 'Username', 'Password', and 'Port' fields are visible, along with a 'Quickconnect' button. The command history shows 'ls' and 'mtime ".ssh"', with a response of '1343749064'. The local site is '/Users/Bees/Downloads/' and the remote site is '/home/n-z/rkhetani'. The local site view shows a tree structure with folders like 'Downloads', 'For_mamta', 'Geneious 5.6 Data', 'Library', 'Movies', 'Music', and 'OGS_gen4_fa'. The remote site view shows a tree structure with folders like 'home', 'n-z', and 'rkhetani'. The 'rkhetani' folder is expanded, showing a list of files and directories. A red arrow points to the 'galaxy-upload' directory in the remote site list. The file list table is as follows:

Filename	Filesize	Filetype	Last modified	Permissions
..				
illumina-larg...	86,576	GIF Picture	08/15/2012 14:...	
Galaxy51-(Gr...	344	tabular-file	08/15/2012 13:...	
.DS_Store	15,364	File	08/15/2012 13:...	
ameI_OGSv1....	5,296,409	gff-file	08/15/2012 13:...	
ameI_OGSv1....	824,812	Gnu ZIP Archive	08/15/2012 13:...	
Request_ Vac...	40,448	Word Docum...	08/15/2012 10:...	
FY11_12 Leav...	164,864	Excel Spread...	08/15/2012 10:...	
Galaxy-Work ...	20,070	ga-file	08/15/2012 09:...	
book1.pdf	1,681,346	Portable Doc ...	08/14/2012 22:...	
Galaxy14-(U...	3,380,223	gtf-file	08/14/2012 16:...	
book2.mobi	1,993,393	mobi-file	08/08/2012 19:...	

225 files and 23 directories. Total size: 8,500,960,078 b

Server/Local file | Direction | Remote file

Queued files | Failed transfers | Successful t

- You will see that both sides of the split windows will fill up
- Click on the “galaxy-upload” directory
- Drag and drop files from your folder into the galaxy-upload folder

Getting data into Galaxy

Galaxy / UIUC

Analyze Data Workflow Shared Data Visualization Help User

Using 57.2 Gb

Tools

BEDTools

Get Data

- BioMart Central server
- Upload File from your computer**
- UCSC Main table browser
- UCSC Archaea table browser
- BX main browser

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

Statistics

Graph/Display Data

Regional Variation

Multiple regression

Multivariate Analysis

Multiple Alignments

Metagenomic analyses

FASTA manipulation

NCBI BLAST+

Upload File (version 1.1.3)

File Format: Auto-detect

Which format? See help below

File: Choose File no file selected

TIP: Due to browser limitations, uploading files larger than 2GB is guaranteed to fail. To upload large files, use the URL method (below) or FTP (if enabled by the site administrator).

URL/Text:

Here you may specify a list of URLs (one per line) or paste the contents of a file.

Files uploaded via FTP:

File	Size	Date
<input type="checkbox"/> Your file...		

This Galaxy server allows you to upload files via FTP. To upload some files, log in to the FTP server at biocluster.igb.illinois.edu using your Galaxy credentials (email address and password).

Convert spaces to tabs: Yes

Use this option if you are entering intervals by hand.

Genome: ----- Additional Species Are Below -----

Execute

History

Unnamed history 0 bytes

Your history is empty. Click 'Get Data' on the left pane to start

- If you click on “Upload File” again, the file you transferred will appear in the area that was previously empty.
- This is the best way to get large files into galaxy