Uploading Custom tracks into GBrowse on EuPathDB sites Last modified March 19, 2014

Creating custom tracks in GBrowse allows you to view data that is not integrated into EuPathDB. Use this feature to visualize your data in the context of other data on EuPathDB sites (gene models, RNA-sequencing, proteomics, etc.). Your data will remain private and is not retained by EuPathDB. This document offers suggestions for uploading your Bigwig, BED and BAM files such as from RNA Sequencing data.

General suggestions and comments:
- Many ‘omic files are quite large ranging from 100’s of MB to 100’s of GB. Even if you have a “fast” internet connection, your file uploads may take many, many minutes. If you have a “slow” internet connection, you should consider reformatting your files to a more compressed type, e.g. BigWig rather than BAM or use the URL option discussed below.
- If you encounter an upload failure, reset GBrowse using the Reset link and then try again.
- The option to upload ‘From a URL’ bypasses the browser upload limitations, resulting in shorter upload times and fewer errors. This is the best option for files > 2 GB, but it requires that you to make your file available on a public webserver at your location. You might need to ask a system administrator or webmaster for help with this.
- As mentioned above, EuPathDB will not incorporate your custom tracks into our database, thus custom tracks will disappear with new database releases or site maintenance affecting custom tracks.
BIGWIG FILES:

BigWig files ‘From a file’ – BigWig files can be uploaded and displayed in GBrowse using the ‘from a file’ option.

BigWig files ‘From a URL’ – BigWig files can be uploaded using the ‘from a URL’ option.

- Place the .bw file on a public site. Enter the URL for the .bw file in the text box provided and then click import. GBrowse will upload and display the .bw file. This option bypasses the browser upload limitations.

BAM FILES:

There are three ways to upload your BAM file to GBrowse:
1. ‘From a file’ option - directly from your computer
2. ‘From a URL’ option – place your file on a public server then upload using the file’s URL location
3. Convert your BAM to BigWig and upload the .bw file ‘From a file’
Details for these upload options are provided below.

1. BAM files ‘From a file’ – Using this option, files up to 1.8 GB can be uploaded, sorted and displayed.
a. BAM files larger than 1.8 GB will often fail to upload using the ‘from a file’ option. These files need to be pre-sorted, indexed and then loaded from a public URL. (see 2. BAM files ‘From a URL’ below).

b. Large BAM files can also be reduced in size by converting to BigWig format and uploaded using the ‘from a file’ option. For instructions see the section 3 below (Convert BAM to BigWig format in Galaxy and upload .bw ‘From a file’).

2. BAM files ‘From a URL’ BAM files can be displayed in GBrowse from a public URL.
   a. Place the .bam file and its corresponding .bai file (created by indexing against EuPathDB’s current version of the genome) in the same folder on a public site.
   b. Enter the URL for the .bam file in the text box provided. GBrowse will display the .bam file when ready. This option avoids time consuming sorting and processing because GBrowse uses the .bai file for that information.

3. Convert BAM to BigWig format in Galaxy and upload .bw ‘From a file’ Converting from BAM to BigWig format reduces your file size by approximately 85%. This option is recommended if the BAM file is larger than 2GB, if your internet connection is “slow” or if you are unable to post your BAM file to a public server.
   a. You must have a Galaxy account to perform this conversion. Go to https://usegalaxy.org/ and sign up for a free account.
   b. Upload the BAM file to your Galaxy account. If the file is larger than 2GB you will want to upload using an FTP client.
      i. Galaxy help document about Uploading data to your Galaxy account: https://wiki.galaxyproject.org/FTPUupload
      Note: the Galaxy community server address has recently changed. Do not use the address, main.g2.bx.psu.edu, shown in the tutorial. Instead use usegalaxy.org.
   c. Use the EuPathDB download system to retrieve the genome (.fasta format) used to create your BAM file. Be sure to retrieve the ‘correct’ version of the genome which contains genomic sequence IDs matching those in your BAM file.
      i. From the grey menu bar of a EuPathDB site, hover over Downloads and choose Data Files.
      ii. Navigate to the correct genome version in the Data Files directory. Data is organized by EuPathDB release number and then by organism, strain, and file type.
         For example: -> Current Release
->T. gondii ME49
->fasta
->data
->T. gondii ME49_Ge
ome.fasta

iii. Click on the file name to begin the download.

d. Designate your genome.fasta file as a ‘Custom Build’ in Galaxy
   i. Upload the genome.fasta file to your Galaxy account. Instructions for uploading files to Galaxy are in Step 3.a above.
   ii. After uploading the genome.fasta file, navigate the following path in Galaxy:
       ->User (last tab in menu bar)
       ->Custom Build
       -> Add a Custom Build.

iii. Complete Galaxy’s ‘Add a Custom Build’ form by naming your custom build, choose your genome.fasta file and click submit.

e. Convert the BAM file to BigWig format using the “Convert BAM to BigWig” option in the Edit Attributes, Convert Format tab:
   i. Set the Database Build of your BAM file to the custom build that you just added. Go to the Galaxy History containing your BAM file and click on the pencil icon next to your file name.

ii. Use the Attributes tab to set the Database/Build to the custom build. Click Save
iii. Use the “Convert BAM to BigWig” tool to create a new file with the bigwig format. Under the Convert Format tab choose Convert BAM to BigWig and click Convert.

f. Upload your new .bw file to GBrowse.
   i. Download the .bw to your computer by clicking on the file’s eye icon in the Galaxy History panel.

ii. Use the Add Custom Tracks ‘From a file’ in the GBrowse.

BED FILES:

BED files ‘from a file’ – A 1.2 GB BED file up to 2GB can be uploaded and displayed in GBrowse using the ‘from a file’ option. Improperly formatted BED files (columns out of order) will return the error: Invalid BED file: blockSizes != blockStarts at /var/www/w1.tritrypdb.org/cgi-bin/Bio/Graphics/Browser2/DataLoader/bed.pm line 27

BED files ‘from a URL’ – BED files can be displayed in GBrowse from a URL. Place the .bed file on a public server. Enter the URL for the .bed file in the text box provided and then click import.