**Gramene Tutorial and Exercises: ASPB 2011 Workshop**

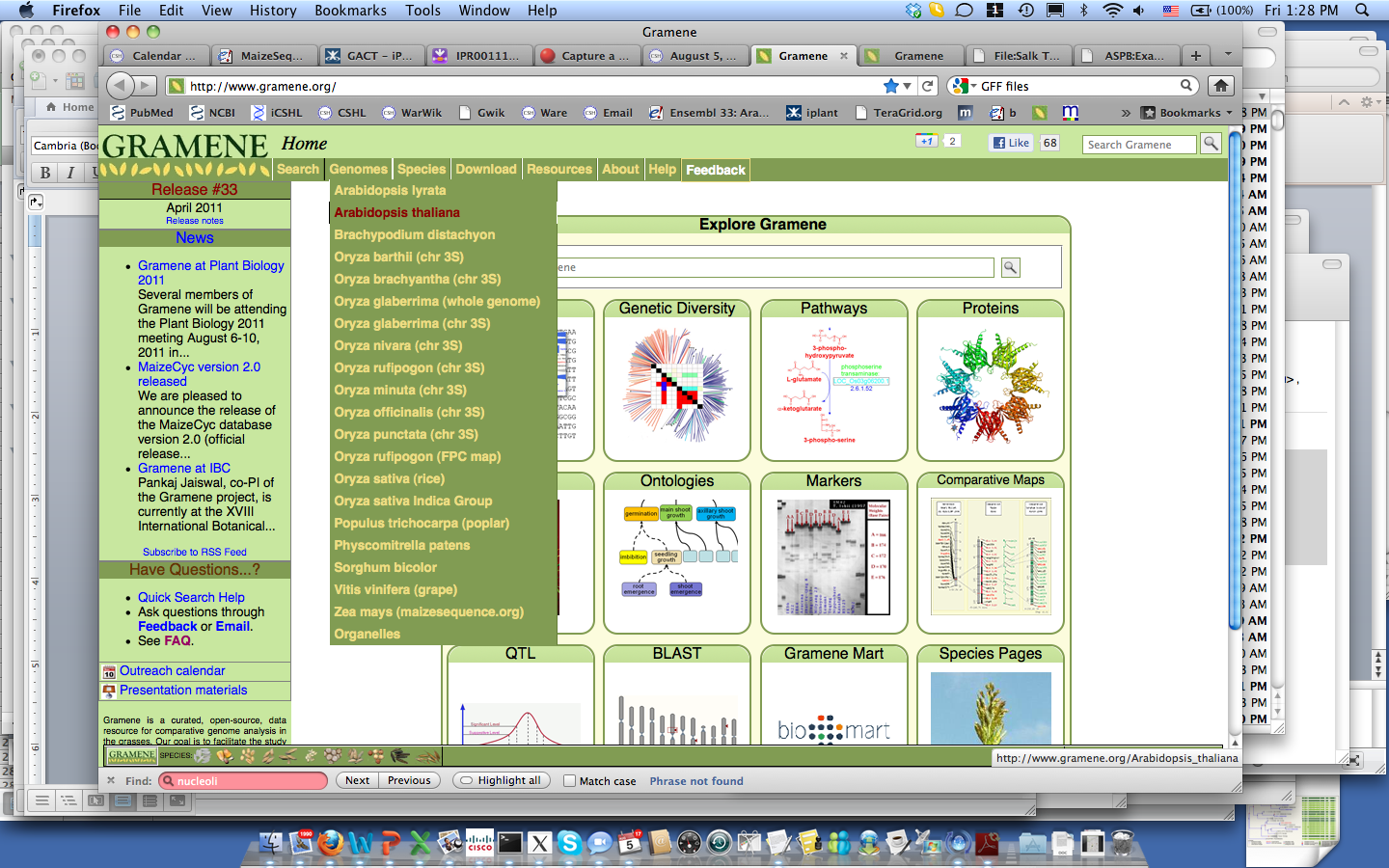
**Viewing Custom Data Tracks and 3rd Party Data in the Genome Browser:**

**Background:** Gramene provides several methods for loading your own genomic data or 3rd party data to be viewed as custom tracks within the genome browsers. Any data that is positionally mapped to a reference assembly, for example gene annotations, marker locations, or quantitative data, can be viewed privately in the browser alongside other tracks available in Gramene. Furthermore, once data is loaded, you can export high-resolution, publication quality images of your data.

This exercise teaches three ways to load custom or 3rd party data into the browser: 1) upload data from your own computer; 2) upload data on a remote server using an URL; 3) Attach a DAS track from a DAS registry. This exercise will also explain various ways to format your data and customize their appearance. Regardless of the methods employed it is crucial that the information to be displayed is mapped to the same version of the reference assembly as hosted in Gramene. Assembly version information is documented on the Description Page for each genome in Gramene (e.g. <http://www.gramene.org/Arabidopsis_thaliana/Info/Index>).

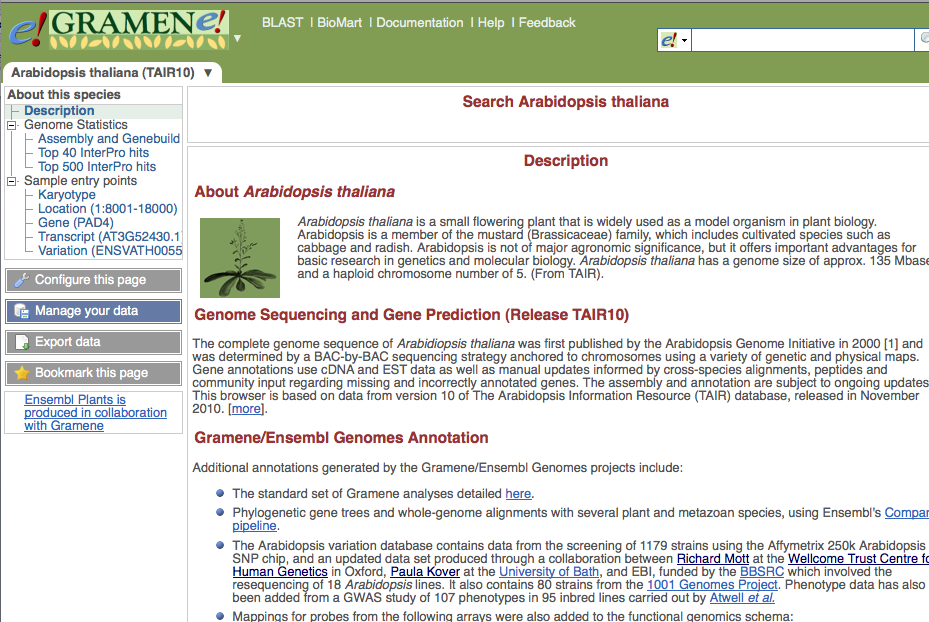
**Uploading data from your own computer:**

1. For this exercise you will first need some example data. Go to <http://outreach.gramene.org/gramene-outreach/images/f/f9/Salk_T-DNA_gff.txt> and save the file onto your computer before proceeding. This is an example of a GFF file[[1]](#footnote-1) giving coordinates of T-DNA insertions[[2]](#footnote-2) into the Arabidopsis Salk genome.
2. Go to Gramene ([www.gramene.org](http://www.gramene.org)) then navigate to the Arabidopsis genome page using the “Genomes” drop down menu.



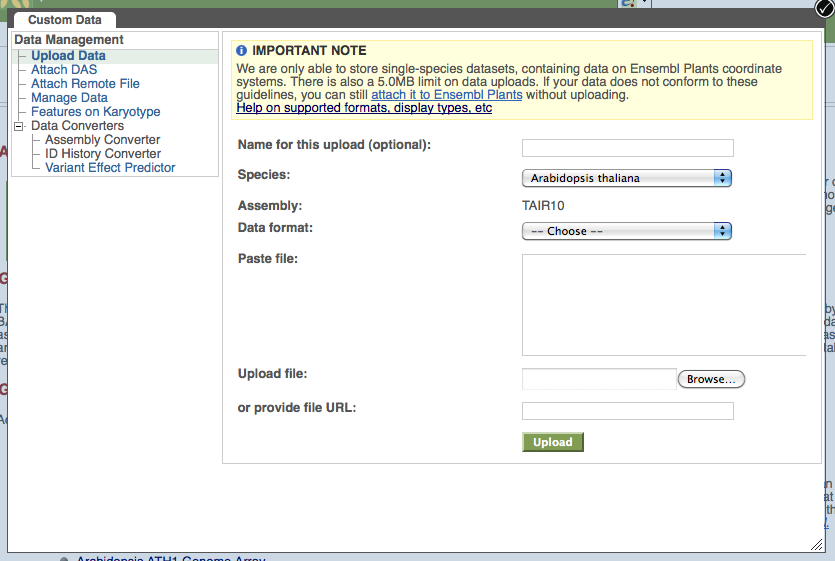
(www.gramene.org)

1. Now click on “Manage your data” on the left-hand panel

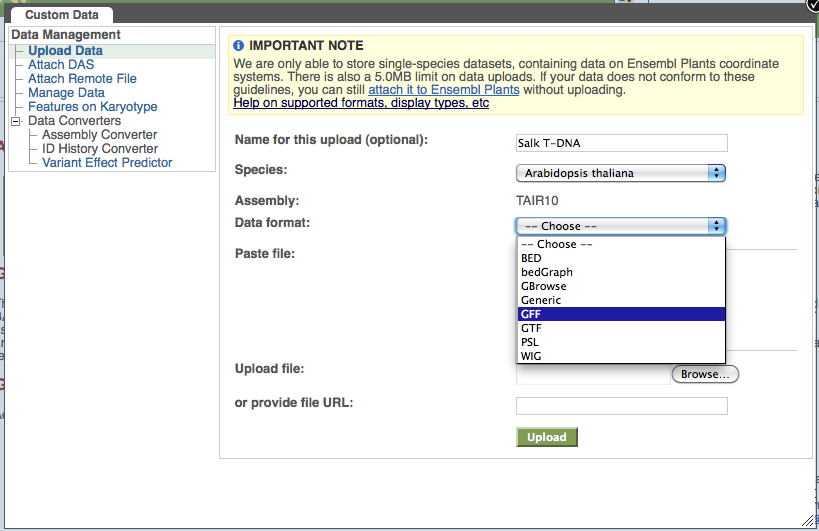


(http://www.gramene.org/Arabidopsis\_thaliana/Info/Index)

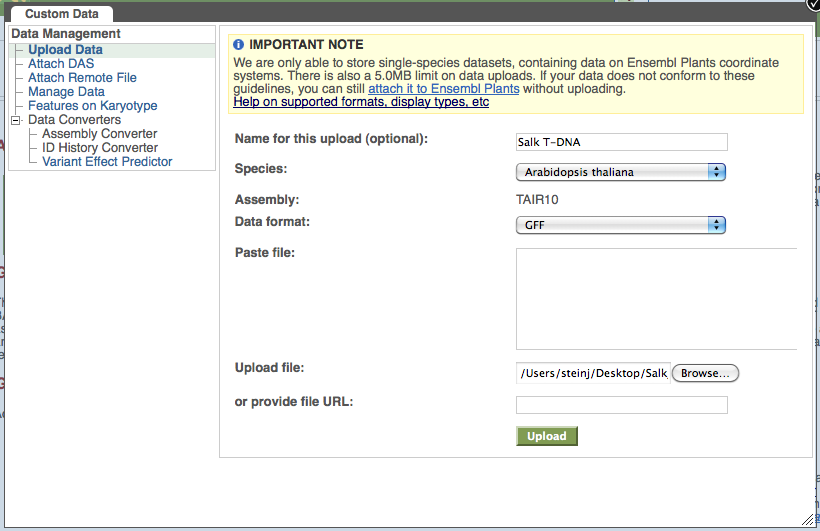
1. A pop-up menu will appear. Click on “Upload Data” on the left-hand menu panel to display the following screen:



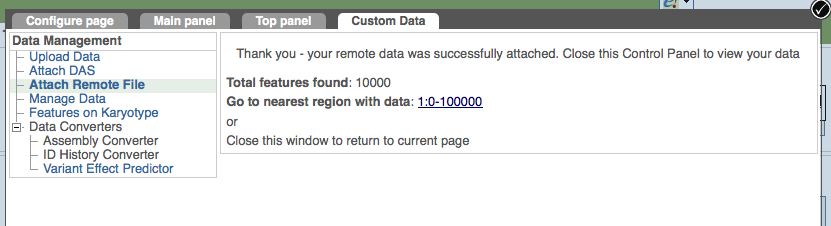
1. Select Arabidopsis thaliana from the species drop-down menu if not already showing. Choose “GFF” from the “Data format” drop-down menu. Also you can enter a name that will be displayed along side the track in the browser.



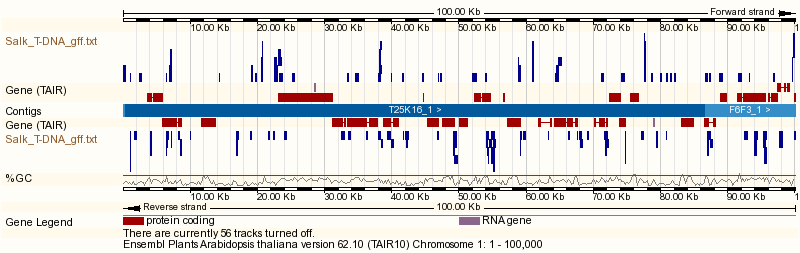
1. Use the “Browse…” option to upload “Salk\_T-DNA\_gff.txt” from your computer then click the green “Upload” button. (Note: you could also copy and paste your file into the text window).



1. If successful you will get a confirmation window with a link to a genomic location that contains the uploaded data:



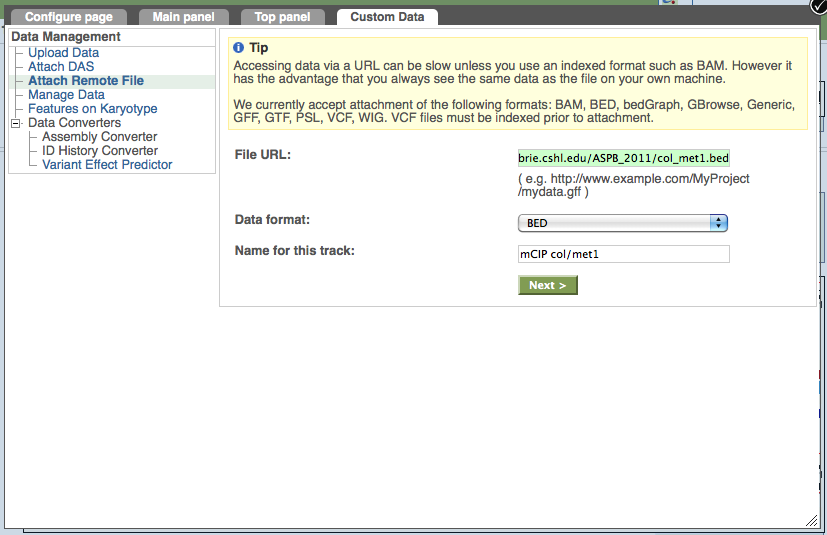
1. Click on the link and you should now see a browser view with your added track!



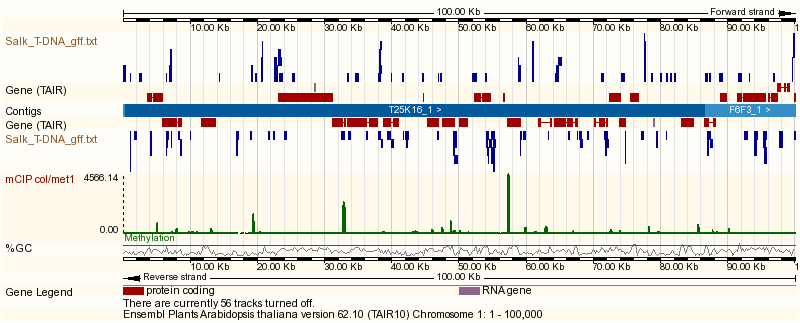
(http://www.gramene.org/Arabidopsis\_thaliana/Location/View?r=1:1-100000)

**Uploading data from a remote server using an URL**

1. Data can be stored on a remote server and accessed using an URL. In this exercise you will load quantitative DNA methylation data2 that is stored on a machine at Cold Spring Harbor Labs. The method is very similar to the preceding exercise.
2. As before click on “Manage your data” to bring up the pop-up menu. This time choose “Attach remote file”. In the “File URL:” box paste the following <http://brie.cshl.edu/ASPB_2011/col_met1.bed>.
3. The link leads to a BED file, another standard format for representing genomic data[[3]](#footnote-3). Thus you should choose “BED” in the Data format drop-down menu. Also give your new track a name (e.g. mCIP col BU/UB).



1. Click “Next” and once again you will get a confirmation page with a link to a browser location that displays the data. Click on this link. The track shows quantitative data displayed as a histogram-like graph with y-axis values ranging from the minimum to the maximum of the data.



1. If time allows you can follow the same steps to load additional data sets from this methylation study:

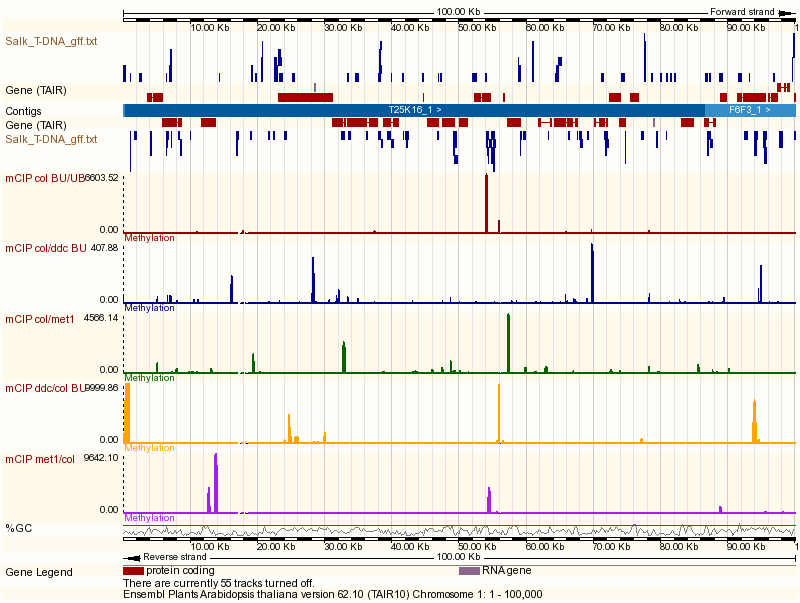
<http://brie.cshl.edu/ASPB_2011/met1_col.bed>

<http://brie.cshl.edu/ASPB_2011/col_BU_UB.bed>

<http://brie.cshl.edu/ASPB_2011/col_ddc.bed>

<http://brie.cshl.edu/ASPB_2011/ddc_col.bed>

Notice that each of the loaded tracks has a different color. Later in this exercise you will learn how to specify color and height of the graphs within the BED data file.

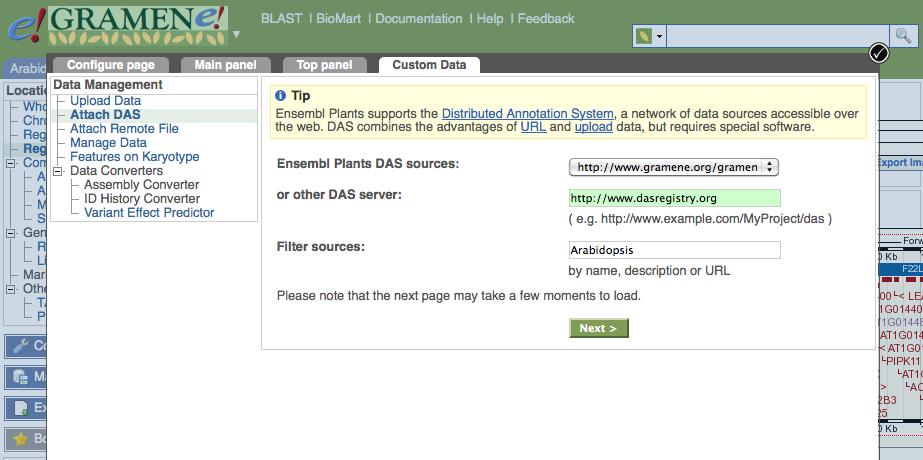


(http://www.gramene.org/Arabidopsis\_thaliana/Location/View?r=1:1-100000)

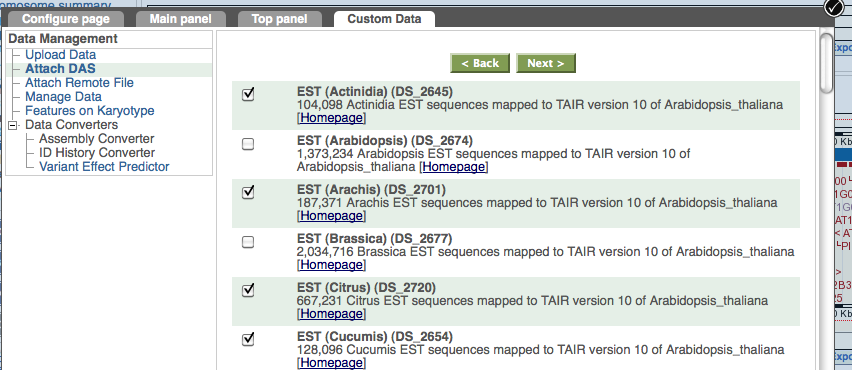
**Attach a DAS track from a DAS registry**

DAS stands for Distributed Annotation System. This is a standard protocol for sharing third party data over the internet via what is known as a DAS registry. Several existing DAS registries are available for accessing data for Arabidopsis and other reference genomes (mostly comprised of non-model EST alignments). A comprehensive resource is at www.dasregistry.org.

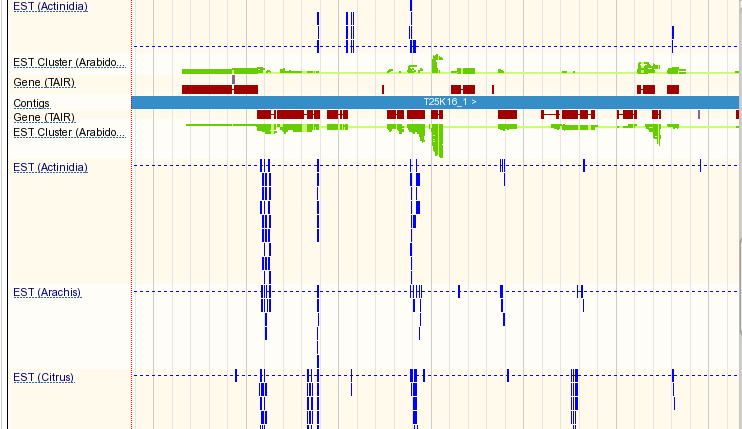
1. By this time you have no doubt noticed the “Attach DAS” option within the “Manage your data” pop-up menu. Click on this now.
2. From the drop-down menu you can select from several available DAS sources. Select the following option: “http://www.gramene.org/gramenedas/das”, enter “Arabidopsis” into the “Filter” window and click the green “Next” button. (Note that other DAS registries can be specified manually in the text window as shown).



1. A menu will appear listing available data sets. Select any number of these using the check boxes and proceed by clicking “Next”.



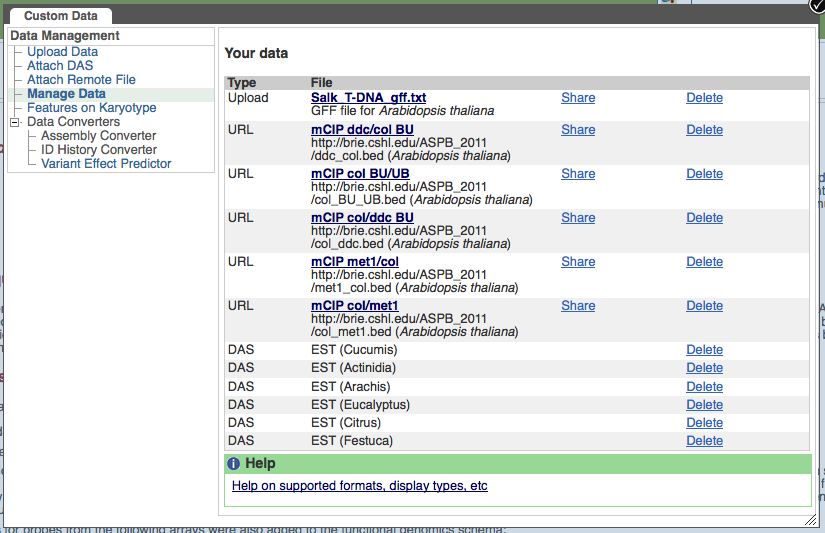
1. After confirmation of success you can go to the browser to view the added DAS tracks.



(Image of DAS tracks)

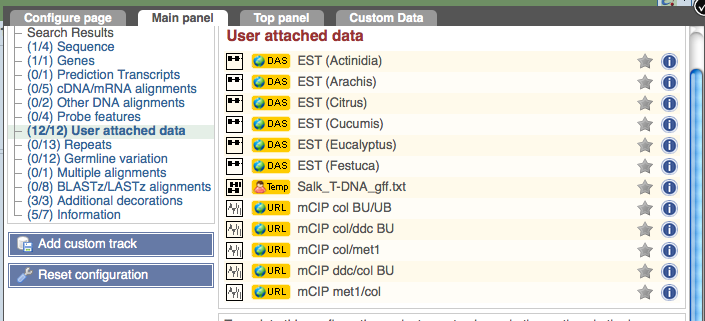
**Managing Custom Data**

Keeping track of your data, documentation of their sources, and the ability to delete them is achieved by selecting the “Manage Data” option in the pop-up menu that you have accessed in the previous exercises. Simply click “Delete” to get rid of any data you don’t want. Note that the “Share” link is not supported at this time.



**Turning Custom Tracks On and Off**

1. Choose to display your tracks as you would any other track in the browser. Click “Configure this page” from the left-hand panel on the browser and select the “Main panel” tab at the top of the pop-up menu.
2. You will find a “User attached data” option in the left-hand menu panel within the pop-up. Your data will be listed with check boxes to turn on and off these tracks.

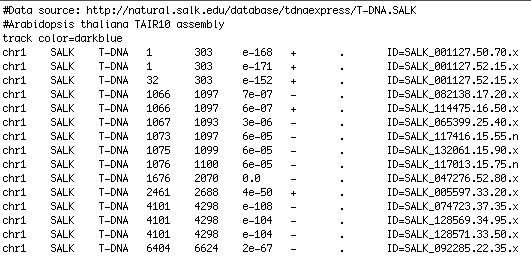


**Configuring data files to control appearance in the browser**

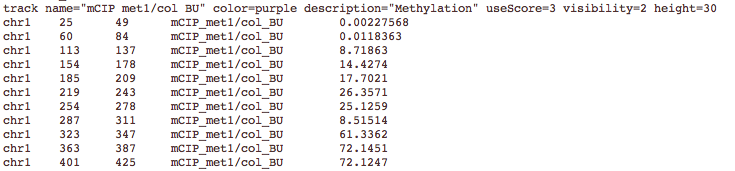
For GFF and BED type files you can control how the browser will render the color and other visual characteristics of the data by including a header line with the first word ‘track’ in the data file. Note that such files do not require this header, but if you don’t include one then all of your data will be a glaringly bright red by default and it won’t be possible to specify how quantitative data is displayed. After the word ‘track’ various parameters can be set using a “parameter=value” format. See examples below.

Color is specified using the color parameter, for example “color=orange”. Most common color names are accepted and their shade can be specified as light or dark as in the example GFF file below.

To specify rendering of a histogram plot you need to include “useScore=3” in the track header. Height of the y-axis is controlled using the height parameter as in “height=10” which would produce a plot that is shorter than the one shown in the example BED file shown below. Additional options and more thorough description can be found at http://www.gramene.org/info/website/upload/index.html#formats.



Example GFF file



Example BED file

1. GFF stands for Generic Feature Format. For more information see: <http://gmod.org/wiki/GFF> [↑](#footnote-ref-1)
2. T-DNA and methylome data derived and downloaded from http://signal.salk.edu/ [↑](#footnote-ref-2)
3. For details on the BED specification see <http://useast.ensembl.org/info/website/upload/bed.html> [↑](#footnote-ref-3)