GenePattern
Tutorial

Software Copyright

The Broad Institute
SOFTWARE COPYRIGHT NOTICE AGREEMENT

Table of Contents

Introduction

GenePattern provides access to many of the computational methods used to analyze genomic data. Using GenePattern, you can easily run analyses, visualize results, and build pipelines to recreate analytic processes that use a combination of tools. Its extendable architecture makes it easy for computational biologists to add analysis and visualization modules, which ensures that you have access to new computational methods on a regular basis.

GenePattern has three components:

● the server, which is the GenePattern engine

● the Java Client, which is the primary GenePattern interface for most users

● the Web Client, which is primarily used for administrative tasks, such as adding and deleting analysis modules

This tutorial introduces you to all three components, focusing primarily on the GenePattern Java Client. The tutorial, which takes approximately 2 hours to complete, walks you through the following steps:

● Start the GenePattern server

● Start the Java Client

  ▪ Explore the Java Client interface

  ▪ Run analysis tools

  ▪ Review analysis results

  ▪ Save and delete analysis results files

  ▪ Create a pipeline that runs a sequence of analyses

● Start the Web Client

  ▪ Explore the Web Client interface

  ▪ Run analysis tools and pipelines

  ▪ Delete the pipeline that you created in the Java Client

● Stop the GenePattern Java Client, the GenePattern Web Client, and the GenePattern server

All of the information you need to successfully complete this tutorial is contained in the tutorial. For users who like additional discussion along the way, the tutorial includes pointers to the GenePattern Java Client Guide and the GenePattern Web Client Guide. Feel free to follow these links or to ignore them, depending on your learning style.

Installing GenePattern

This tutorial assumes that you have installed the GenePattern server, the analysis and visualization modules, the Java Client, the Web Client, and the sample datasets. If you have not installed all of these components, please go back and follow the installation instructions. If your GenePattern server is running, but you are not sure which components were installed, you can revisit the installation summary:

1. Click here or type http://localhost:8080/gp/installFrame.jsp into your web browser. The GenePattern login page appears.

2. Enter your user name, usually your e-mail address, and click *sign in*. The installation summary appears.

Starting the GenePattern Server

You must start the GenePattern server before you can start either the GenePattern Java Client or the GenePattern Web Client. The server takes several seconds to start; clients cannot connect to it while it is starting.

To start the GenePattern server, double-click the Start GenePattern Server icon, shown below. By default, the GenePattern installation places this icon on your desktop.

StartGenePatt
ernServer

On Windows, while the server is starting, a console window displays messages. The server is ready when the console displays the following message:

```
***********************************************************************
GenePattern server version 2.0.0 build 2.0.0.358 built Thu, Dec 15 2005 is ready
***********************************************************************
```

On Mac OS X, while the server is starting, the server icon bounces in the Dock. The server is ready when the icon stops bouncing. If you want to read the messages, see the Console.

Using the GenePattern Java Client

This part of the tutorial introduces you to the GenePattern Java Client, which is the primary interface for GenePattern. In this part, you learn how to:

● Start the Java Client.

● Run analysis tools.

● Review analysis results.

● Save and delete analysis results files.

● Create a pipeline that runs a sequence of analyses.

Starting the Java Client

To start the GenePattern Java Client, double-click the GenePattern Java Client icon, shown below. By default, the GenePattern installation places this icon on your desktop.

GenePattern
Client

A message informs you that the Java Client is retrieving modules and jobs from the GenePattern server. After a few seconds, the Java Client displays its welcome message:
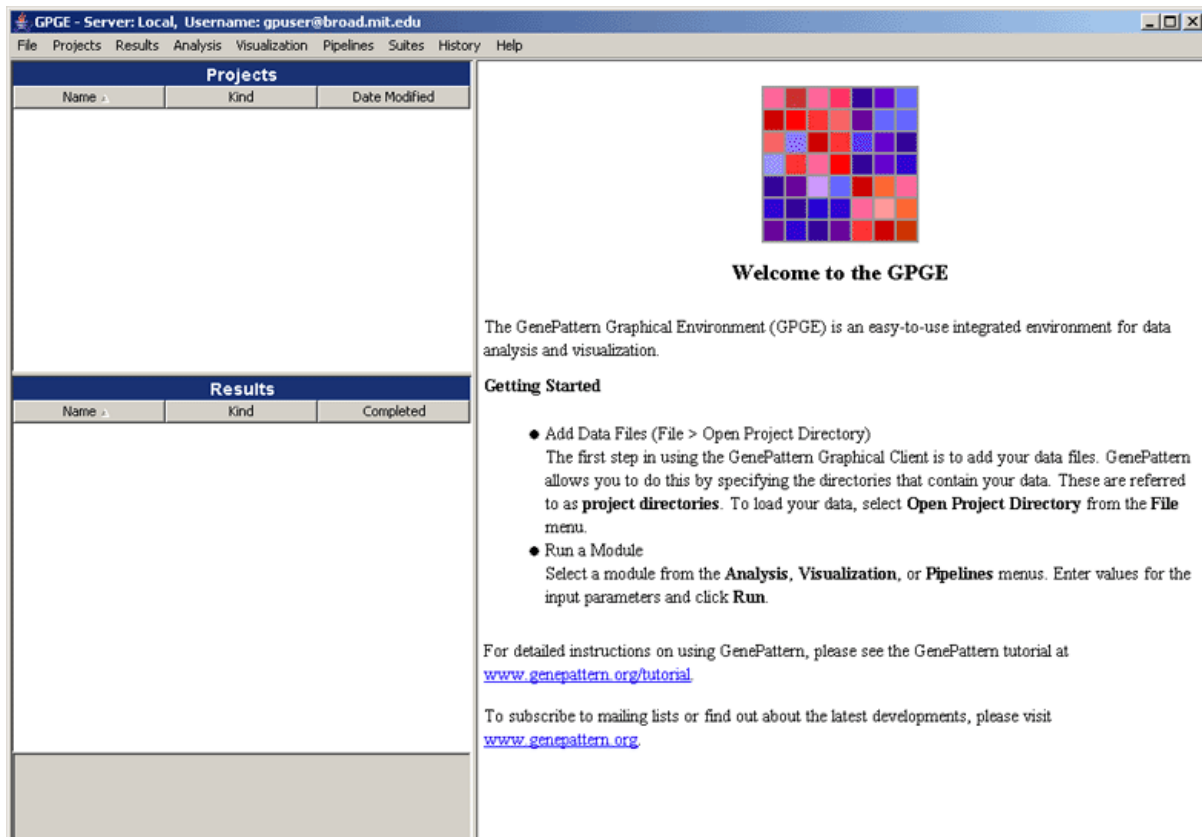
Notice the following elements of the Java Client window:

5

- The title bar shows you the GenePattern server and user name that you are using. The Java Client, which remains connected to the server throughout the session, runs all of your analyses on this server. Typically, your GenePattern user name is your e-mail address.

- The Projects panel lists your project directories; that is, the directories holding the data files that you are analyzing. When you first start the Java Client, this panel is empty.

- The Results panel lists each analysis that you have run and its generated analysis results files. When you first start the Java Client, this panel is empty.

- The panel on the right, which displays the welcome message, is the main window. The Java Client uses this window to display information and to prompt you for input.

- The menu bar lists the Java Client commands.

  For more information: see Menu Bar in the *GenePattern Java Client Guide*.

Running Analysis and Visualization Tools

Now that you have started the GenePattern server and GenePattern Java Client, you are ready to run your first analysis. In this section, you will learn how to:

- Add a project directory to the Projects panel.

- Run an analysis against the data files in that directory.

- Track the progress of the analysis as it runs.

- Review the analysis results when the analysis completes.

When you installed GenePattern, you downloaded a number of sample datasets, including the `gp_tutorial_files/all_aml` directory. In this first analysis, you will run the Comparative Marker Selection analysis against the `all_aml_train` (res and cls) files in the `gp_tutorial_files/all_aml` directory. This analysis determines which genes in the dataset (res) file are most closely correlated with phenotypes in the class (cls) file.

Adding a Project Directory

In GenePattern, the directory that holds the files that you want to analyze is called a project directory. You do not have to add your project directories to the GenePattern Java Client, but doing so makes it easier to work with the files in those directories.

In this exercise, you are analyzing the files in the `gp_tutorial_files/all_aml` directory. To add this project directory to the GenePattern Java Client:

1. Select *File>Open Project Directory*. GenePattern displays the Choose a Project Directory window:
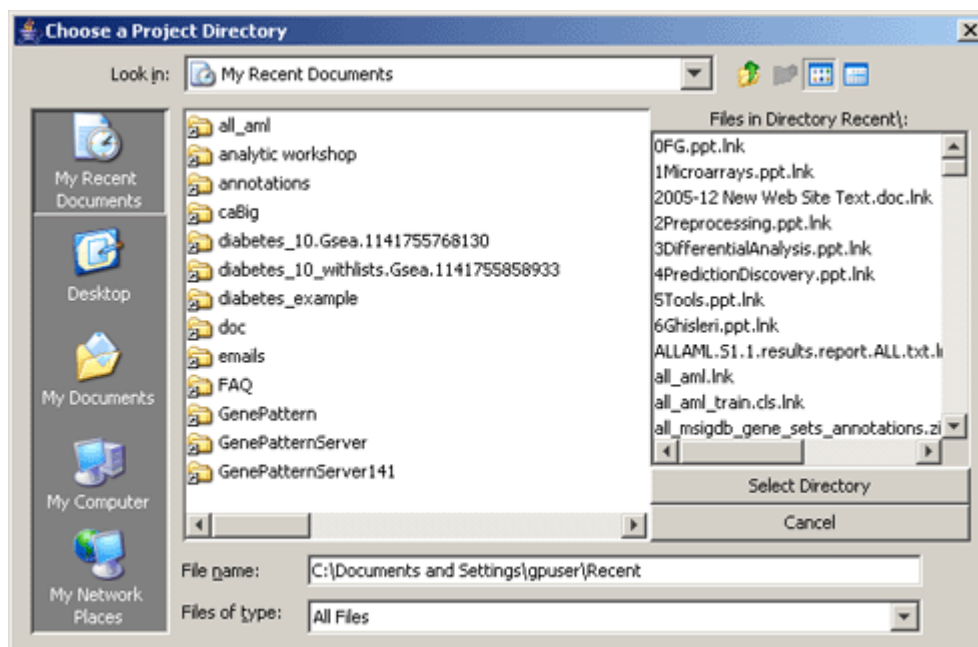
2. In the Choose a Project Directory window, navigate to the `gp_tutorial_files/all_aml` directory and click *Select Directory*. GenePattern adds that directory to the Projects pane:
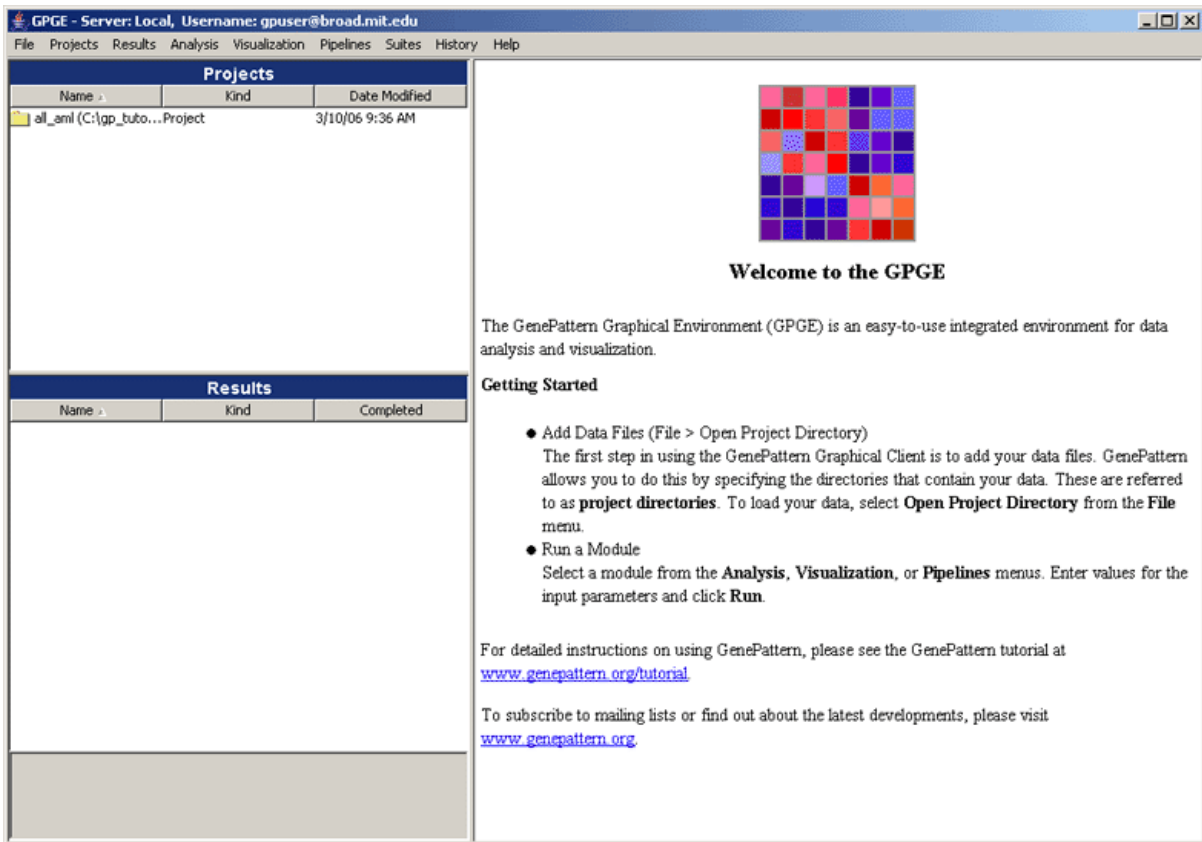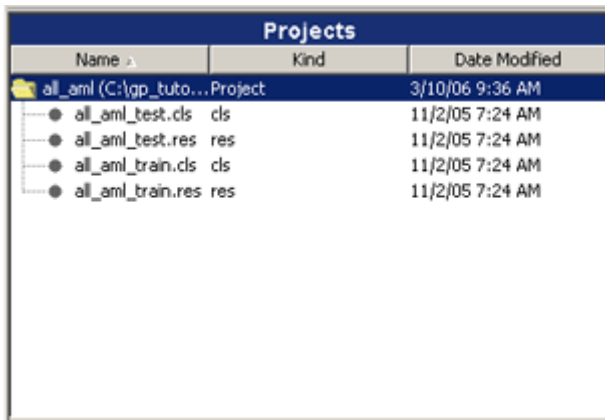

3. In the Projects pane, double-click the directory name to see the files in the directory:

For more information: see [Projects Pane](#) in the *GenePattern Java Client Guide*.

Running the Analysis

Now, you are ready to run the Comparative Marker Selection analysis against the `all_aml_train` (res and cls) files in your project directory. To do so:

1. Select *Analysis>Gene List Selection>ComparativeMarkerSelection*. The analysis and its parameters appear in main window:
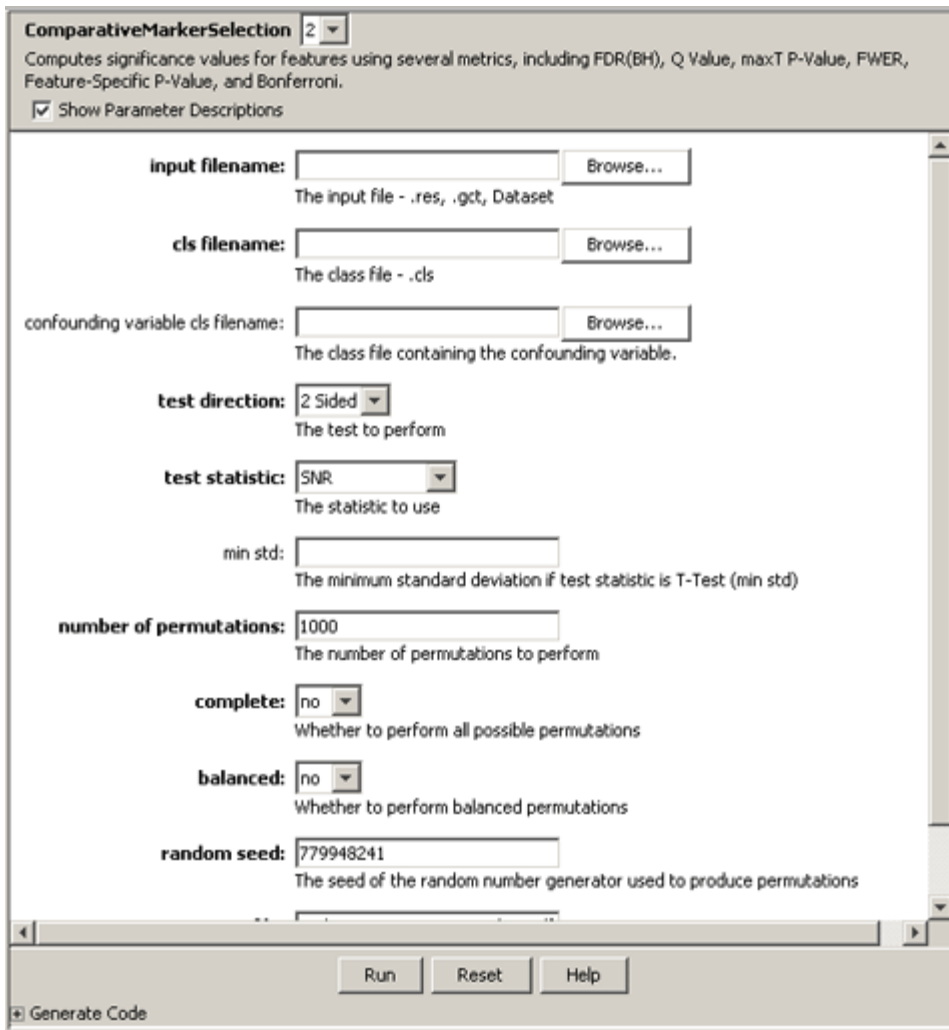
2. For the *input filename* parameter, select the `all_aml_train.res` file:

   1. Right-click the `all_aml_train.res` file in the Projects pane. The following context menu appears:
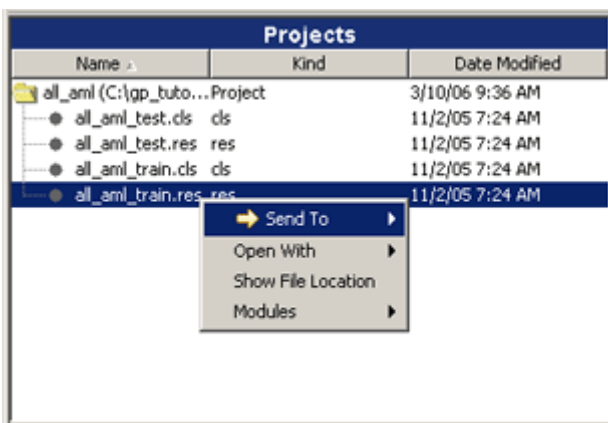
2. Select *Send to>input filename*. GenePattern sends the file to the *input filename* parameter.

Most commands are available from either the menu bar or a context menu. In this step, you used a context menu to set the *input filename* parameter. In the next step, you use the menu bar to set the *cls filename* parameter.

3. For the *cls filename* parameter, select the `all_aml_train.cls` file:

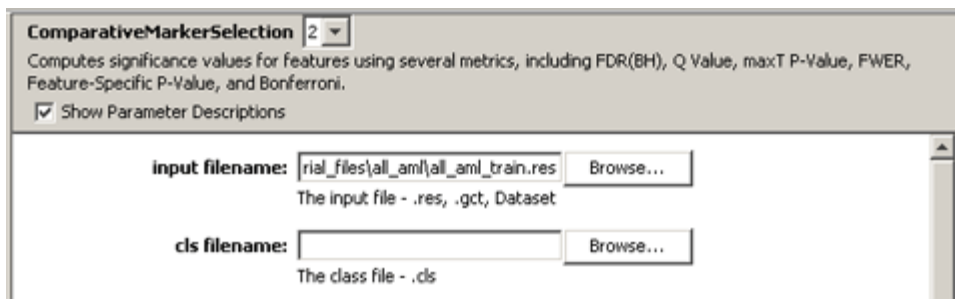   1. Click the `all_aml_train.cls` file in the Projects pane to select it.

    2. Select *Projects>Send to>cls filename*. GenePattern sends the file to the *cls filename* parameter:
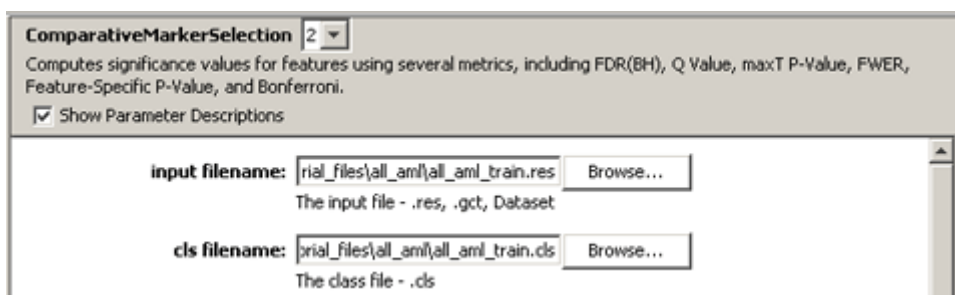
Leave the remaining parameters set to their default values.

4. Optionally, click *Help* (at the bottom of the main window) to view the documentation for this analysis. Each analysis module has its own documentation.

5. Click *Run* to start the analysis. The analysis appears in the Results panel with a status of Processing:
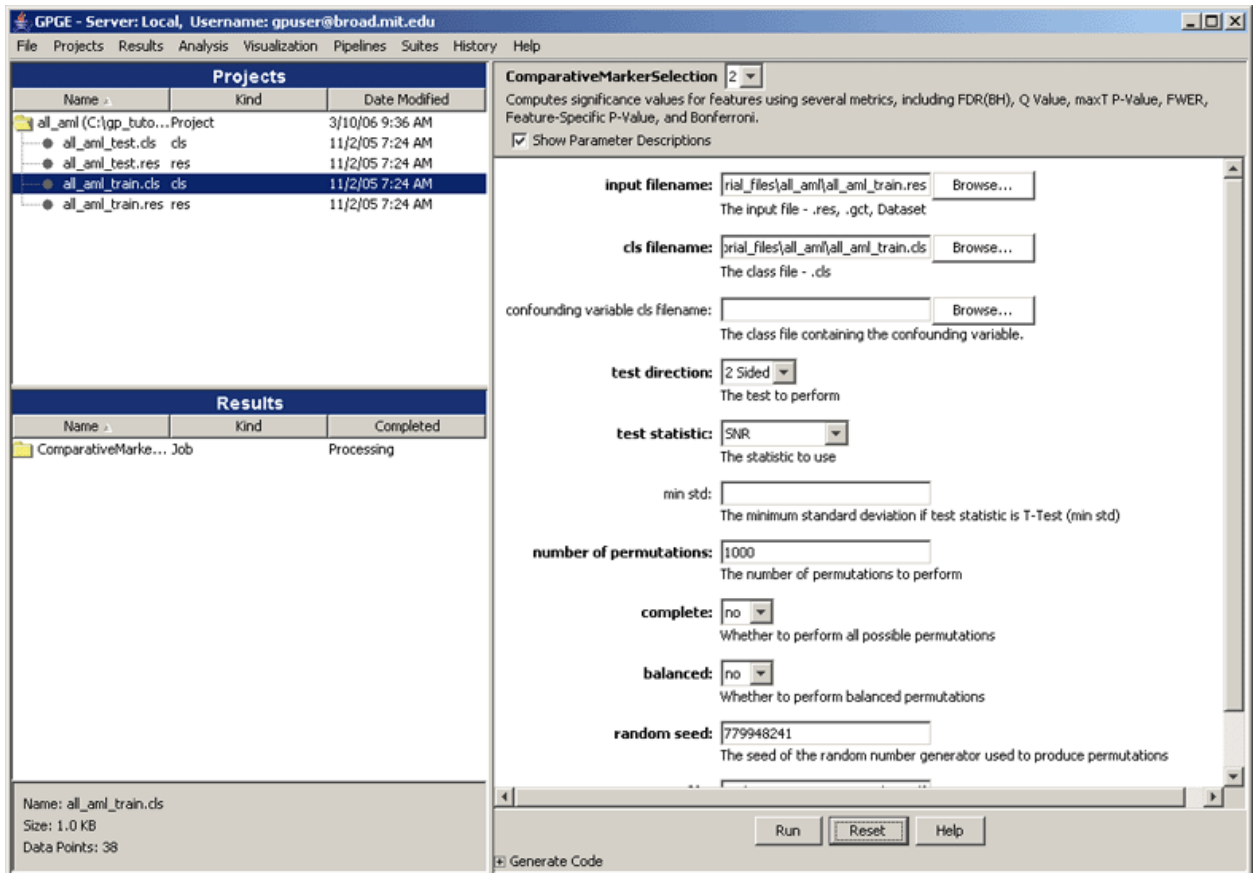
For more information: see [Running Analysis and Visualization Modules](#) in the *GenePattern Java Client Guide*.

Checking Job Status

When you click *Run* to start an analysis, the GenePattern Java Client sends your analysis request to the GenePattern server. The GenePattern server runs your analysis. When the analysis completes, the GenePattern server notifies the GenePattern Java Client. The GenePattern Java Client then notifies you in two ways:

● In the Results pane, the Java Client updates the Completed column to show the data and time the analysis finished:

● By default, the Java Client also adds each finished analysis to the Recently Completed Jobs window. If this window is not open, the Java Client opens it:
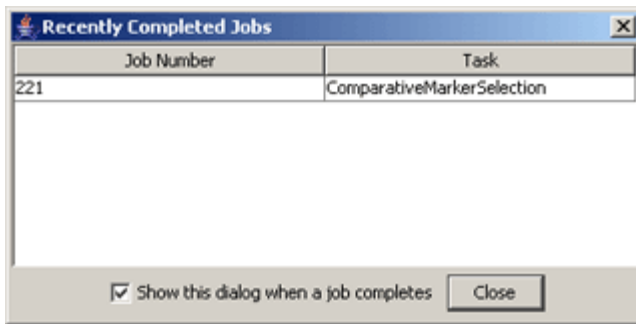
You can use *File>Alert On Job Completion* to control whether or not the Java Client uses the Recently Completed Jobs window. If the menu item has a check mark, GenePattern updates the window, opening it if necessary; otherwise, it does not. To toggle the check mark, select the menu item.

Click *Close* to close the Recently Completed Jobs window.

Reviewing the Results

When your analysis completes, the analysis results files appear under the analysis name in the Results pane. Double-click the analysis name to see the results files:
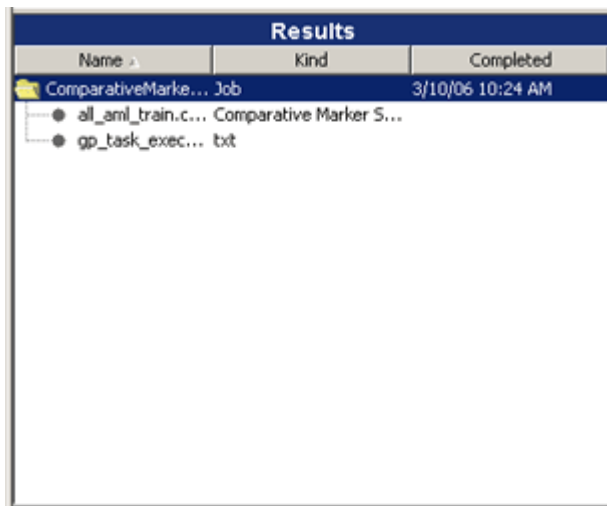

For more information: see Results Pane in the *GenePattern Java Client Guide*.

An analysis generates multiple analysis results files, including the results of the analysis, an execution log file, and sometimes additional message or error logs. You can view analysis results in many ways:

- *Results>Open With>Text Viewer* displays the selected file in a text viewer. This is generally the best way to view log files.

- *Results>Open With>Default Application* displays the selected file in the default application (if any) associated with this file type. This can be useful when you have used your operating system to assign default applications to particular file types.

- *Results>Modules* lists the modules in the Visualization menu that can be used to view the selected file. You select the module to use and run it as you did the analysis. This is generally the best way to view analysis results since the Visualization modules are designed for this purpose.

In this exercise, you use the text viewer to view the execution log and then use the Comparative Marker Selection Viewer module to view the results of the Comparative Marker Selection analysis:

1. Right-click the execution log file, `gp_task_execution_log.txt`, and select *Open With>Text Viewer*.
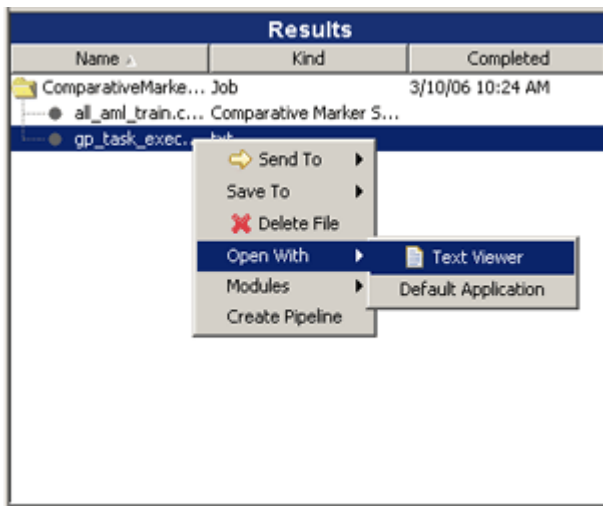
GenePattern displays the execution log in a text viewer:

2. Close the text viewer by clicking the Close (X) icon in the upper right corner.

3. Right-click the analysis results file, `all_aml_train.comp.marker.odf`, and select *Modules>ComparativeMarkerSelectionViewer*. The visualization module and its parameters appear in main window. Notice that the file you selected has been sent to the *input filename* parameter:

12

ComparativeMarkerSelectionViewer 2 ▼
View the results from ComparativeMarkerSelection
☑ Show Parameter Descriptions

input filename: 1, all_aml_train.comp.marker.odf   Browse...
The input file

Run   Reset   Help

⊞ Generate Code
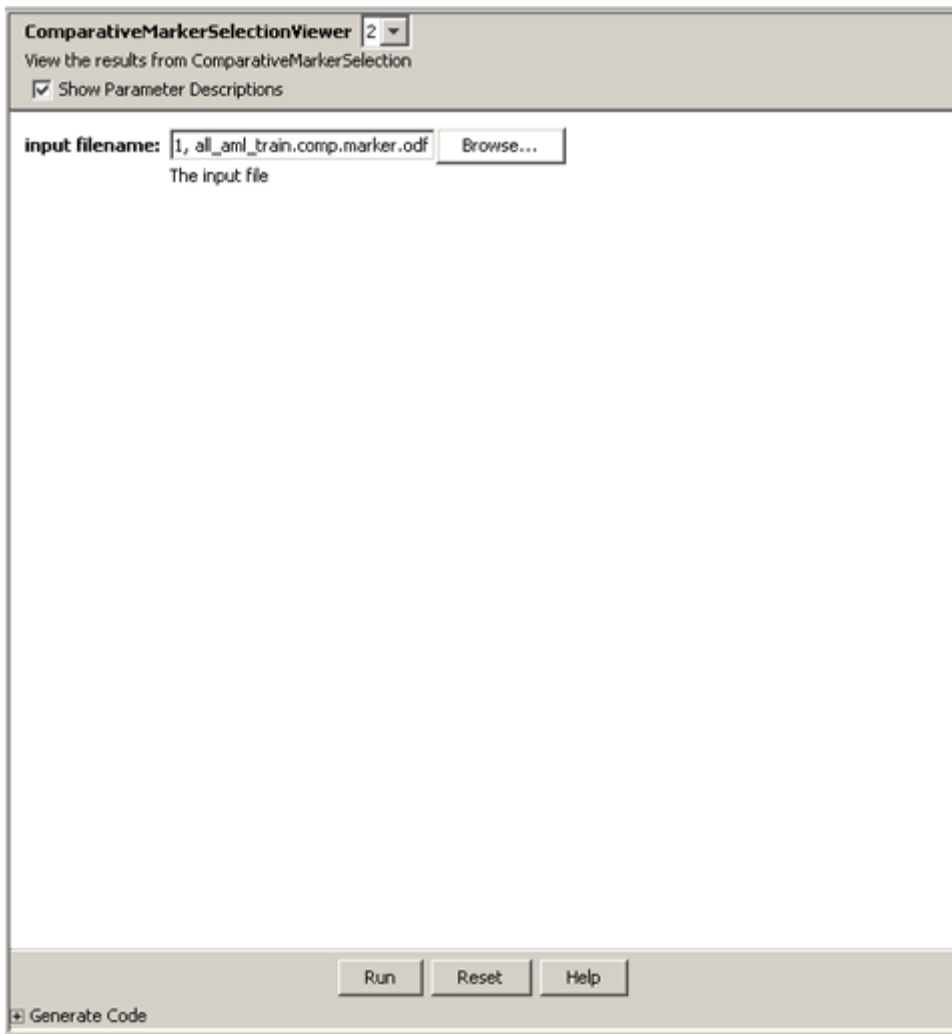
4. Click *Run* to run the visualization module. After a few moments, the Comparative Marker Selection Viewer window appears:

Comparative Marker Selection - all_aml_train.comp.marker.odf

File Edit View GeneCruiser Window

Descending Rank vs. Score

■ Upregulated in ALL (3829) ■ Upregulated in AML (3300)

Row	Rank	Upregulated In	Feature	Score	Feature P	FDR(BH)	Q Value	Bonferroni	FWER	maxT
1	1	AML	M55150_at	-1.469	0	0	0	0	0.001	0.001
2	2	AML	U50136_rna1_at	-1.421	0	0	0	0	0.001	0.001
3	3	AML	X95735_at	-1.406	0	0	0	0	0.001	0.001
4	4	ALL	U22376_cds2_s_at	1.34	0	0	0	0	0.001	0.001
5	5	AML	M16038_at	-1.203	0	0	0	0	0.002	0.002
6	6	AML	M23197_at	-1.196	0	0	0	0	0.002	0.002
7	7	AML	M84526_at	-1.191	0	0	0	0	0.002	0.002
8	8	AML	Y12670_at	-1.185	0	0	0	0	0.002	0.002
9	9	AML	U82759_at	-1.164	0	0	0	0	0.002	0.002
10	10	AML	D49950_at	-1.14	0	0	0	0	0.002	0.002
11	11	ALL	X59417_at	1.125	0	0	0	0	0.002	0.002
12	12	AML	M27891_at	-1.109	0	0	0	0	0.003	0.003
13	13	AML	X17042_at	-1.106	0	0	0	0	0.004	0.004
14	14	ALL	U05259_rna1_at	1.103	0	0	0	0	0.004	0.004
15	15	AML	U12471_cds1_at	-1.07	0	0	0	0	0.006	0.006
16	16	AML	U46751_at	-1.064	0	0	0	0	0.007	0.007
17	17	ALL	M92287_at	1.044	0	0	0	0	0.01	0.01
18	18	AML	Y00787_s_at	-1.041	0	0	0	0	0.01	0.01
19	19	AML	L08246_at	-1.035	0	0	0	0	0.012	0.012
20	20	ALL	X74262_at	1.028	0	0	0	0	0.014	0.014
21	21	AML	M80254_at	-1.028	0	0	0	0	0.014	0.014
22	22	ALL	L13278_at	1.024	0	0	0	0	0.015	0.015
23	23	ALL	M31211_s_at	1.024	0	0	0	0	0.015	0.015
24	24	AML	M62762_at	-1.023	0	0	0	0	0.015	0.015
25	25	ALL	U09087_s_at	1.014	0	0	0	0	0.017	0.017
26	26	AML	M81933_at	-1.008	0	0	0	0	0.017	0.017
27	27	AM...	M96326_rna1_at	-1.005	0	0	0	0	0.018	0.018

Note: If the top portion of the viewer is blank, click *Window>Descending Rank vs. Score* to refresh the viewer.

5.   In the Viewer window, use the Window menu to view other feature-specific plots.

6.   To view the documentation for the Comparative Marker Selection Viewer, return to the Java Client window and click *Help*. Like analysis modules, each visualization module has its own documentation.

7.   To close the Viewer window, return to the Viewer window and click the Close (X) icon in the upper right corner.

In the Java Client window, you might notice that the Comparative Marker Selection Viewer module is not listed in the Results pane. This is because the viewer runs on the Java Client, not the GenePattern server.

Running Additional Analyses

You have successfully completed your first analysis. You ran the Comparative Marker Selection analysis against the all_aml_train.res and all_aml_train.cls files to determine which genes in the dataset (res) file are most closely correlated with phenotypes in the class (cls) file. The results of that analysis are in the analysis results file, all_aml_train.comp.marker.odf.

Now, you will run the Extract Comparative Marker Results analysis against the analysis results file, all_aml_train.comp.marker.odf. The Extract Comparative Marker Results analysis allows you to create a new dataset based on the results of the Comparative Marker Selection analysis.

Running the Analysis

To run the Extract Comparative Marker Results analysis:

1.   Right-click the analysis results file, all_aml_train.comp.marker.odf, and select *Modules>ExtractComparativeMarkerResults*. The analysis and its parameters appear in the main window. Notice that the file you selected has been sent to the *comparative marker selection filename* parameter:

ExtractComparativeMarkerResults 2 ▾
Creates a dataset and feature list file from the results of ComparativeMarkerSelection
☑ Show Parameter Descriptions

comparative marker selection filename: | 1, all_aml_train.comp.marker.odf | Browse...
The results from ComparativeMarkerSelection - .odf

dataset filename: | | Browse...
The dataset file that the markers were selected from - .gct, .res, Dataset

field: | Rank ▾ |
The field to filter features on

min: | |
Select features with field >= min

max: | |
Select features with field <= max

number of neighbors: | |
Number of neighbors to select by score in each direction

base output name: | election.filename_basename>.filt |
The base name for the output files

◄ | ► |

Run | Reset | Help

⊞ Generate Code

2. For the *dataset filename* parameter, select the `all_aml_train.res` file:

   1. Right-click the `all_aml_train.res` file in the Projects pane.

   2. Select *Send to>dataset filename*. GenePattern sends the file to the *dataset filename* parameter.

3. For the *field* parameter, leave the default value of Rank.

4. For the *max* field, enter 100.

   By entering these values, you are creating a dataset based on the top 100 features (genes) in the analysis results file, `all_aml_train.comp.marker.odf`.

5. Click *Run* to start the analysis. The analysis appears in the Results panel with a status of Processing.

   When the analysis completes, the Java Client updates the Completed column to show the data and time the analysis finished and (optionally) adds the finished analysis to the Recently Completed Jobs window.

6. Click *Close* to close the Recently Completed Jobs window.

For more information: see <u>Running Analysis and Visualization Modules</u> in the *GenePattern Java Client Guide*.

Reviewing the Results

When the Extract Comparative Marker Results analysis completes, the analysis results files appear under the analysis name in the Results pane. Double-click the analysis name to see the results files:

Now, you will use the Heat Map Viewer to review the analysis results:

1. In the Results pane, find the analysis results file, `all_aml_train.comp.marker.filt.res`. (To widen the Name column, click-and-drag the line between the Name and Kind column headings.)

2. Right-click the analysis results file and select *Modules>HeatMapViewer*. The visualization module and its parameters appear in main window. Notice that the selected file has been sent to the *input filename* parameter.

3. Click *Run* to run the visualization module. After a few moments, the Heat Map Viewer window appears. The viewer displays dataset expression values in a heat map format: the largest values are displayed as the reddest (hot), the smallest values as the bluest (cool), and the intermediate values as shades of blue and red:
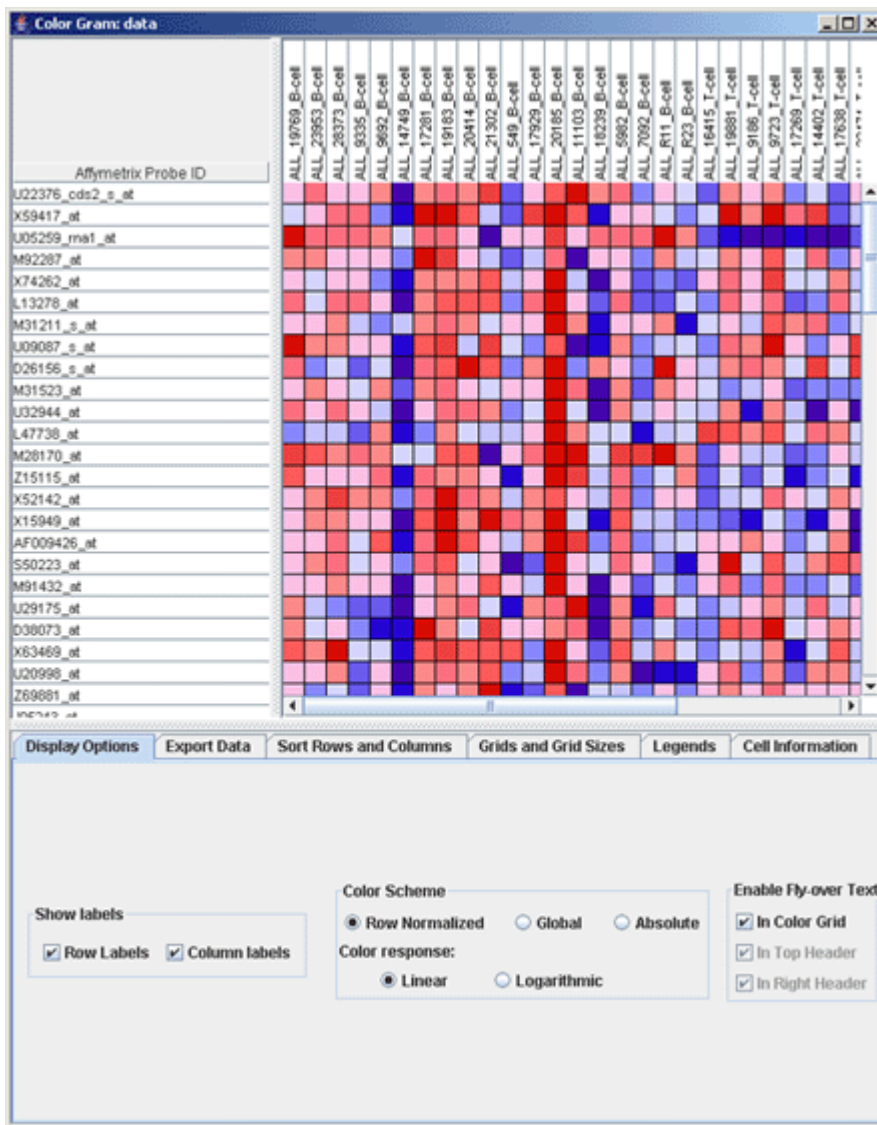
4. In the Viewer window, adjust the grid size for a better view:

   1. Click the Grids and Grid Sizes tab.

   2. Click-and-drag the *Adjust the size of the cells* slider 4 or 5 points to the left. The Viewer displays the grid using smaller size cells.

   You can see by visual inspection which genes are closely correlated with which phenotype.

5. Optionally, to view the documentation for the Heat Map Viewer, return to the Java Client window and click *Help*.

6. To close the Viewer window, return to the Viewer window and click the Close (X) icon in the upper right corner.

You have successfully run an Extract Comparative Marker Results analysis and reviewed the results using the Heat Map Viewer.

Using Pipelines

As you have seen, the GenePattern Java Client makes it easy to run individual analyses and to review analysis results. GenePattern pipelines take this one step further: they make it easy to run multiple analyses. You can define a pipeline to run multiple analyses against a single dataset or to run a sequence of analyses, where the output from one analysis becomes the input for a subsequent analysis.

In this tutorial, you have so far run two analyses: Comparative Marker Selection and Extract Comparative Marker Results. The analysis results file from your first analysis became the input file for your second analysis. By running these two analyses, you produced a new dataset that contains the 100 genes in your dataset

(`all_aml_train.res`) that are most closely correlated with phenotypes in your class file
(`all_aml_train.cls`).

In this section, you will

- Create a pipeline that duplicates your analysis results.

- Edit the pipeline so that it operates on any set of data (res and cls) files.

**For more information**: see <u>Working with Pipelines</u> in the *GenePattern Java Client Guide*.

Creating a Pipeline Based on Previous Results

You can create a pipeline in one of two ways:

- You can create an empty pipeline and then add to it the analysis modules that you want the pipeline to run.

- You can create a pipeline based on an analysis results file and have the Java Client add to the pipeline the analysis modules used to create that results file. After creating the pipeline, the Java Client uses the following steps to determine which modules to add to the pipeline:

  1. It adds the module that created the results file.

  2. It checks the module's input file parameter.

  3. If the input file for the module was the output file of a previous module, the Java Client adds the previous module and returns to step 2; otherwise, there are no more modules to be added to the pipeline.

In this exercise, you will create a pipeline based on the analysis results file from the Extract Comparative Marker Results analysis:

1. Right-click the analysis results file, `all_aml_train.comp.marker.filt.res`, and select *Create Pipeline*. The Java Client prompts you for a pipeline name:
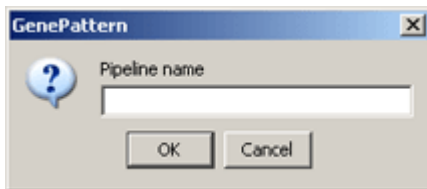
2. Enter a name for the pipeline, my.comparative.marker, and click *OK*. The Java Client displays the new pipeline and its definition:
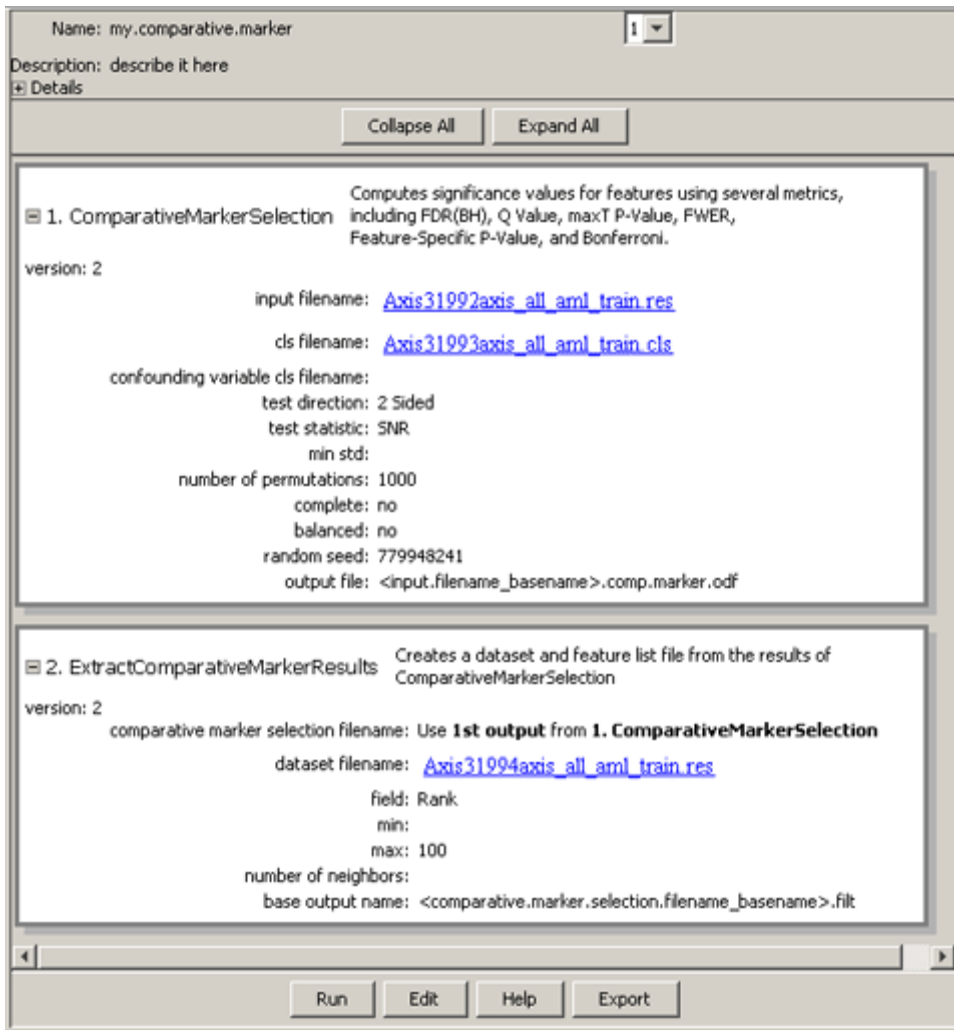
The Java Client created the pipeline and then added to it the module that created the results file, Extract Comparative Marker Results. The input file for the Extract Comparative Marker Results module was the output file from Comparative Marker Selection, so the Java Client added the Comparative Marker Selection module to the pipeline.

Adding Tasks to a Pipeline

The pipeline contains the two analysis modules used to create the analysis results file: Comparative Marker Selection and Extract Comparative Marker Results. In your original analysis, after creating the analysis results file, you used the Heat Map Viewer to review the results.

To add the Heat Map Viewer module to your pipeline:

1. Click *Edit* (at the bottom of the window). The Java Client opens the pipeline definition for editing:
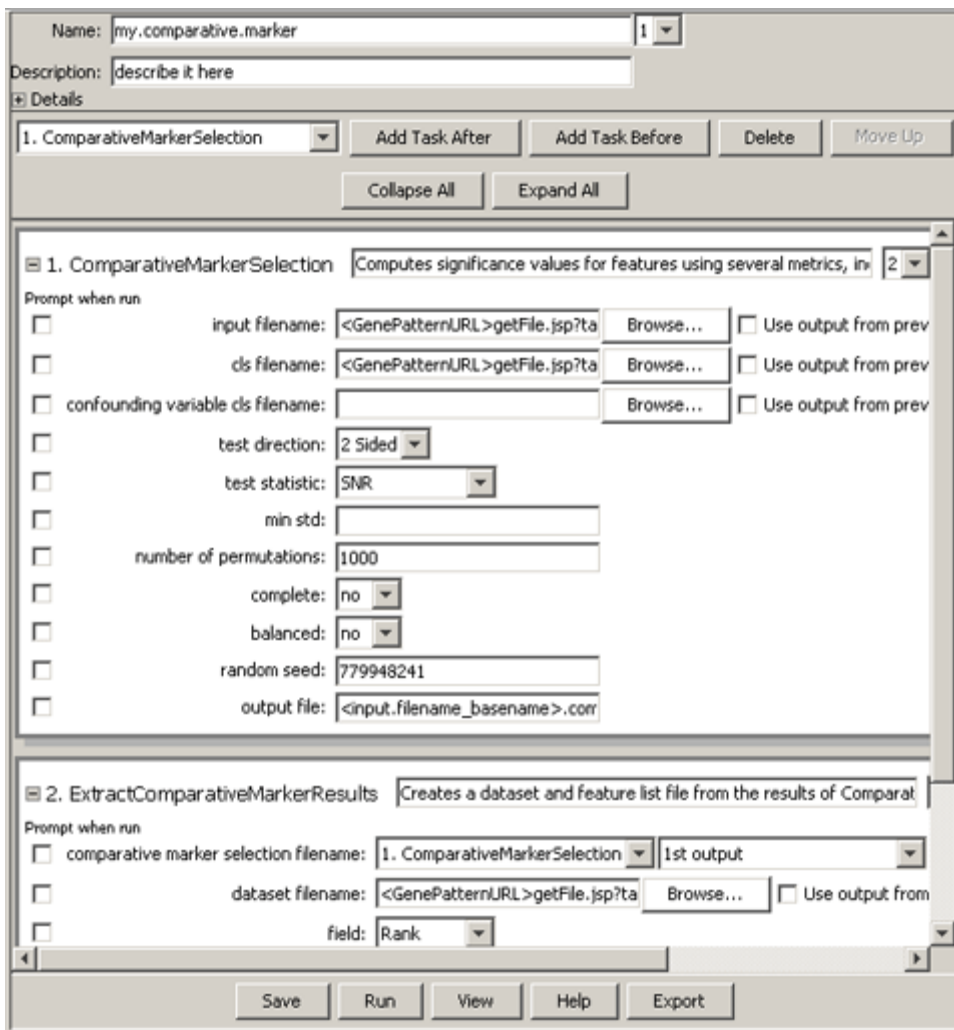
2. Click on the second module in the pipeline, ExtractComparativeMarkerResults, to select it.

3. Click *Add Task After* (near the top of the window). The Java Client displays the Add Task window:
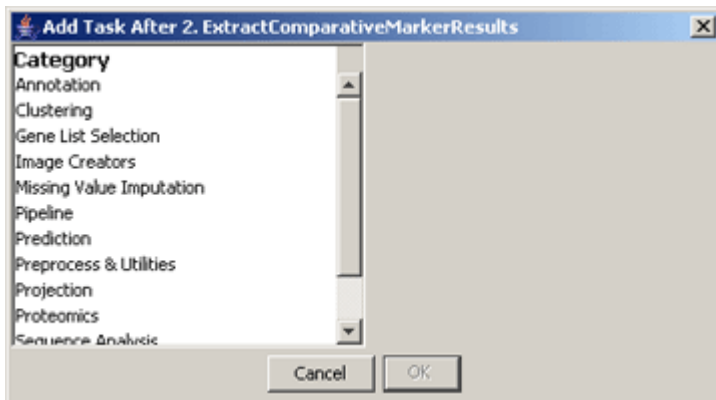

4. Select Visualizer from the Category list. (You may need to scroll the list.) The Java Client populates the Task list with all of the modules in the Visualization menu:
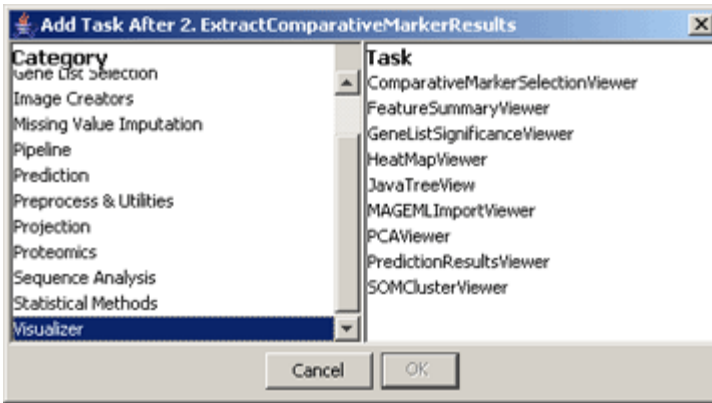
5. Select HeatMapViewer from the Task list and click *OK*. The Java Client adds the HeatMapViewer module after the ExtractComparativeMarkerResults module in the pipeline definition window.

6. Make the value of the *filename* parameter for the HeatMapViewer the analysis results file from the ExtractComparativeMarkerResults module:

   1. Select the *Use output from previous task* check box. The Java Client adds the Choose Task drop-down list to the *filename* parameter field.

   2. Select ExtractComparativeMarkerResults from the Choose Task drop-down list. The ExtractComparativeMarkerResults module generates multiple analysis results files; therefore, the Java Client adds the Dataset drop-down list to the *filename* parameter field.

   3. Select res from the Dataset drop-down list.
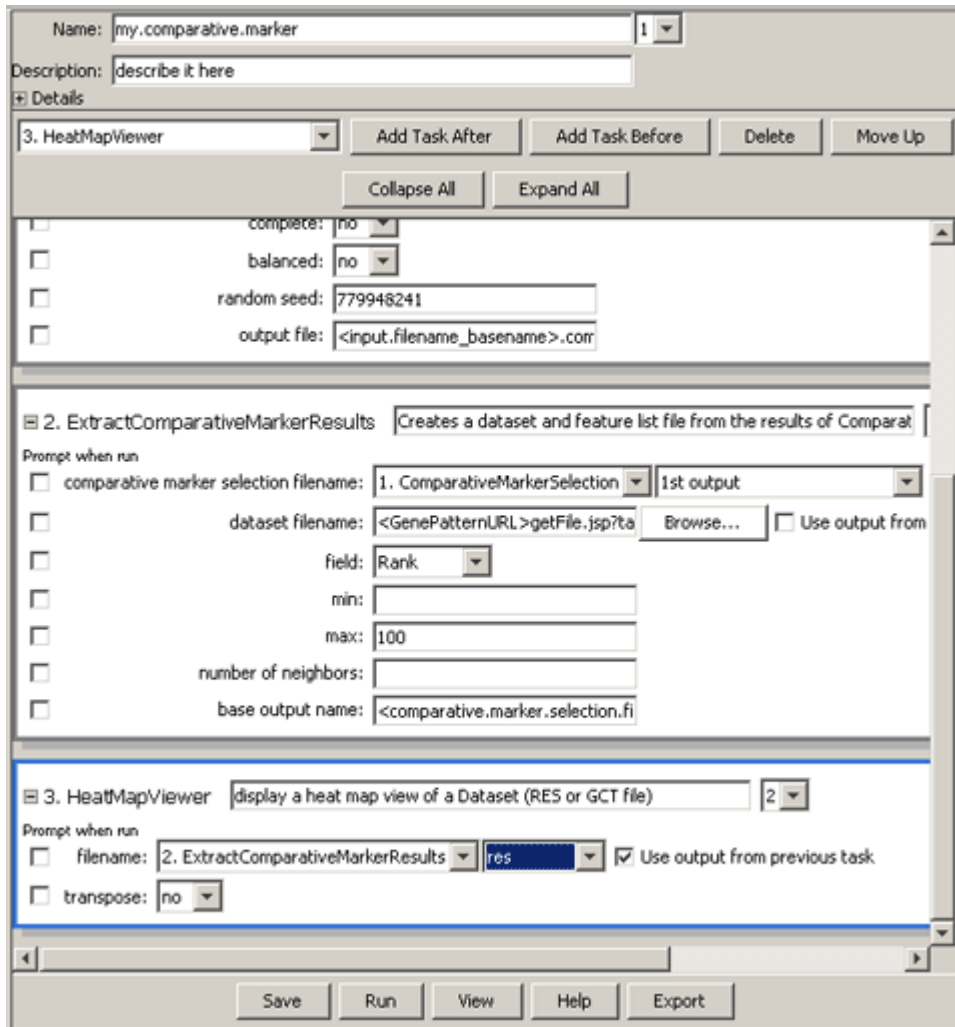
7. Click *Save* (at the bottom of the window) to save your changes. The Java Client saves the pipeline and displays the following confirmation message:
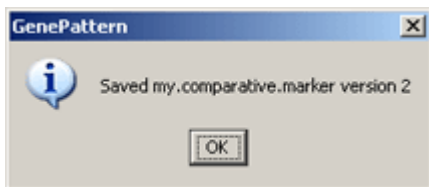
8. Click *OK* to close the confirmation message window.

Editing a pipeline always creates a new version of this pipeline. This ensures that you can always reproduce any analysis that you have run, even if you subsequently change the pipeline that you originally used to run the analysis.

For more information: see Understanding Version Numbers in the *GenePattern Java Client Guide*.

Running a Pipeline

To run the pipeline:

1. Select *Pipelines>my.comparative.marker*. (Notice that the pipeline names are color-coded; the pipelines you create are displayed in pink.) The Java Client displays your pipeline in the main window:
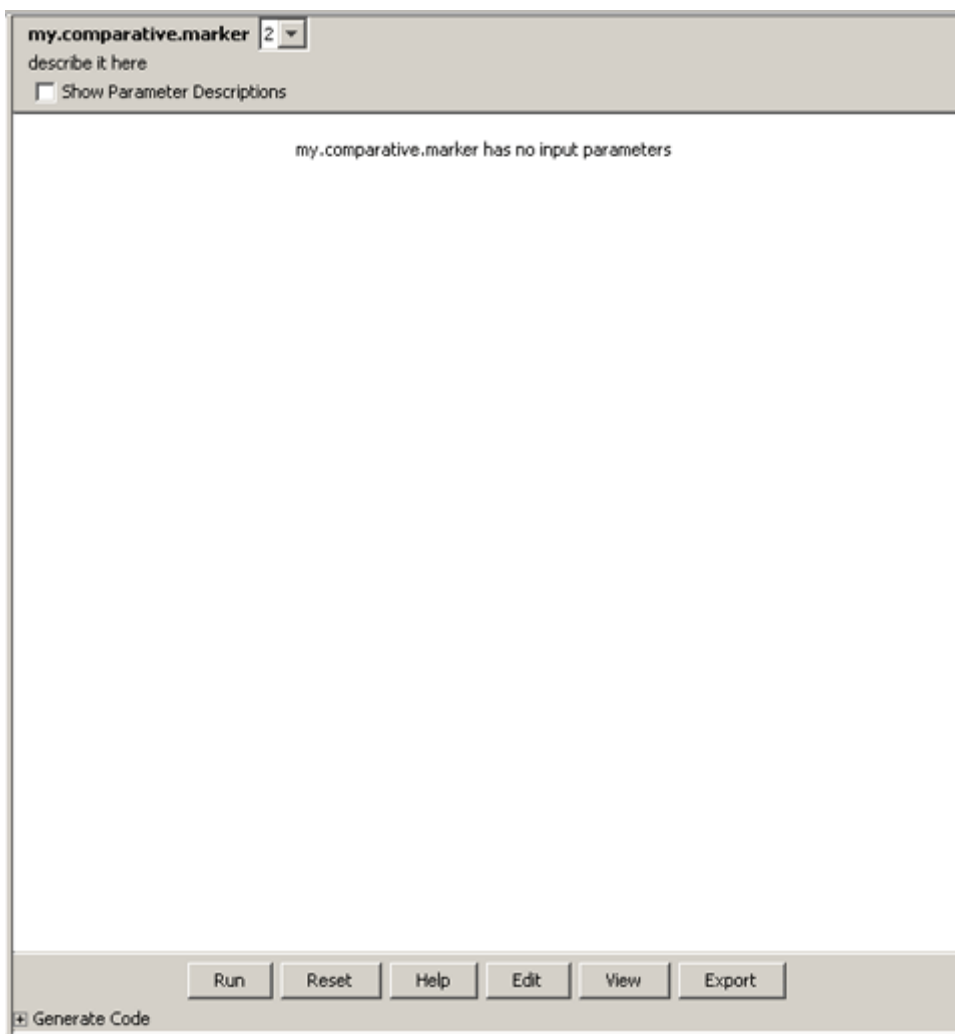
2. Click *Run* to run the pipeline. The pipeline name appears in the Results window with a status of processing. After a short time, the Heat Map Viewer appears and the pipeline completes.

When the pipeline completes, the Java Client updates the Completed column to show the data and time the analysis finished and (optionally) adds the finished pipeline to the Recently Completed Jobs window.

3. Close the Heat Map Viewer and the Recently Completed Jobs window.

4. Double-click the pipeline name in the Results pane to see the analysis results files. Notice that all of the results files for all of the analyses appear under the pipeline name.

   Analysis and visualization modules run from a pipeline work exactly the same as those run directly from the Java Client. If you run a Viewer from a pipeline, it gets displayed just as if you ran it directly from the Java Client. Analysis results files generated by a pipeline are exactly the same as those generated by modules run directly from the Java Client.

Adding Parameters to a Pipeline

You have created a pipeline that duplicates your original analysis: it runs the Comparative Marker Selection analysis on the `all_aml_train` data (res and cls) files, uses the analysis results as input to the Extract Comparative Marker Results analysis, and then displays the analysis results using the Heat Map Viewer.

Now, you want to edit the pipeline. Rather than having a pipeline that analyzes the `all_aml_train` data (res and cls) files, you want to create a pipeline that can analyze any set of data (res and cls) files. When you run the pipeline, you want the Java Client to prompt you for the data (res and cls) files to be analyzed.

To edit the pipeline:

1. Select *Pipelines>my.comparative.marker*. The Java Client displays your pipeline in the main window.

2. Click *Edit*. The Java Client opens the pipeline definition for editing.

   The data (res and cls) files to be analyzed are named in the *input filename* and *cls filename* parameters of the Comparative Marker Selection analysis. These are the parameter values that you want the Java Client to prompt you for when run the pipeline.

3. Click on the *Prompt when run* check box next to the *input filename* parameter of the Comparative Marker Selection analysis. The Java Client removes the parameter value from the pipeline definition.

4. Click on the *Prompt when run* check box next to the *cls filename* parameter of the Comparative Marker Selection analysis. The Java Client removes the parameter value from the pipeline definition.

5. Click *Save* to save your changes. The Java Client saves the pipeline and displays a confirmation message.

6. Click *OK* to close the confirmation message window.

Now, you are ready to run your edited pipeline. Your project directory contains two sets of data files: `all_aml_train` (res and cls) and `all_aml_test` (res and cls). When the Java Client prompts you for the *input filename* and *cls filename* parameters, you will enter the `all_aml_test` (res and cls) files.

To run your edited pipeline:

1. Select *Pipelines>my.comparative.marker*. The Java Client displays your pipeline and its parameters in the main window:
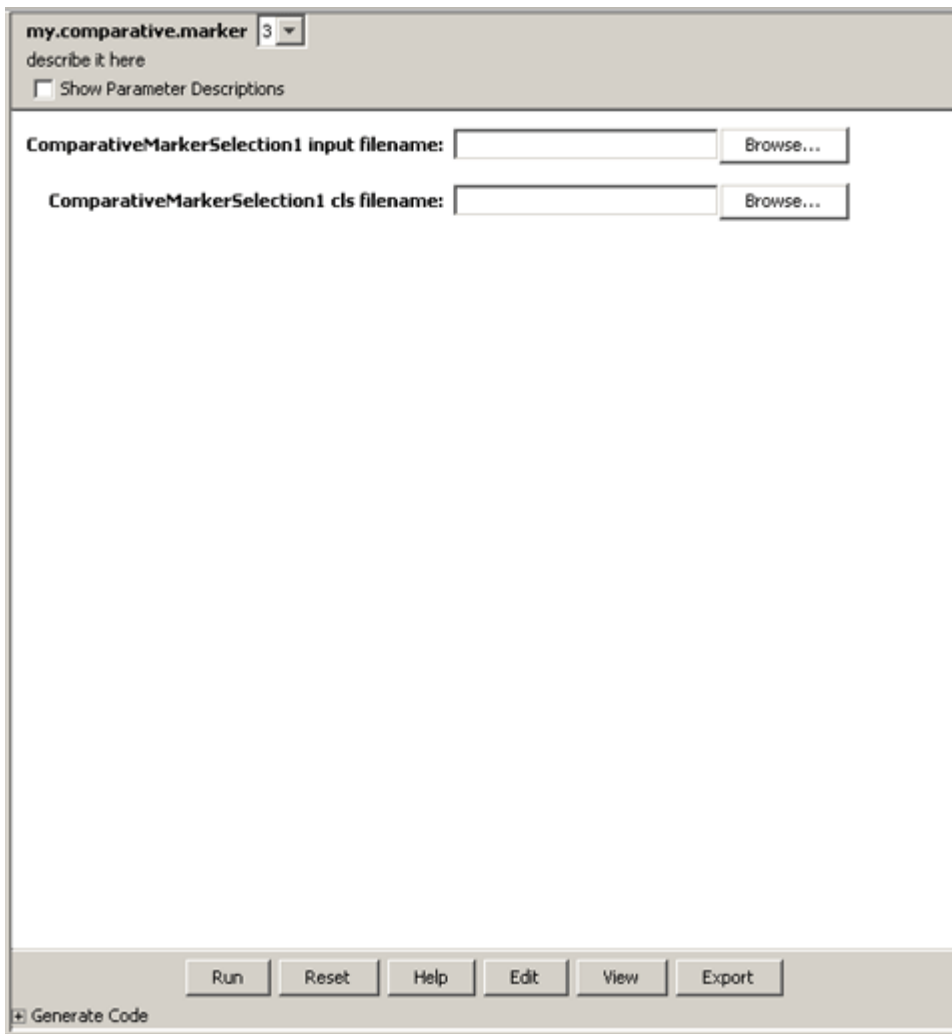
2. Drag-and-drop the `all_aml_test.res` file from the Projects pane to the *ComparativeMarkerSelection1 input filename* parameter.

3. Drag-and-drop the `all_aml_test.cls` file from the Projects pane to the *ComparativeMarkerSelection1 cls filename* parameter.

4. Click *Run* to run the pipeline. The pipeline name appears in the Results window with a status of processing. After a short time, the Heat Map Viewer appears and the pipeline completes.

5. Close the Heat Map Viewer and the Recently Completed Jobs window.

Saving and Deleting Analysis Results

As described earlier in the tutorial, when you run an analysis or a pipeline, the Java Client sends your request to the GenePattern server. The GenePattern server runs the analysis and notifies the Java Client when the analysis or pipeline completes. The analysis results files that appear in the Results pane of the Java Client window are actually stored on the GenePattern server. Server storage is temporary and analysis results files are deleted after they have been on the server for a certain length of time (by default, one week).

To save your analysis results files, you must copy each file from the server to a more permanent location. If you do not need your analysis results, you can delete them from the server at any time, which also removes them from the Results pane.

To save an analysis results file:

1. In the Results pane, click the file that you want to save. For example, click the analysis results of the ExtractComparativeMarkerResults analysis, `all_aml_train.comp.marker.filt.res`:

2. Select *Results>Save To*. GenePattern displays the *Save To* submenu, which lists your project directories and an *Other* item.

3. Select your project directory. GenePattern copies the file to the selected directory and then refreshes the Projects pane to include the new file:
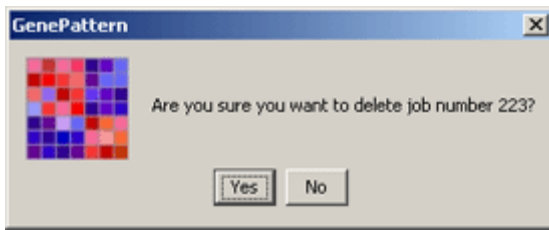

You can delete an analysis results file in much the same way that you save one: click the file that you want to delete and select *Results>Delete File*. Or, you can delete an analysis and all of its results files. To delete an analysis and its results files:

1. In the Results pane, click the analysis that you want to delete. For example, click the ExtractComparativeMarkerResults analysis:

2. Select *Results>Delete Job.* A confirmation window appears:

25

3. Select *Yes* to delete the job and its analysis results files. The GenePattern Client deletes the analysis and all its results files from the GenePattern server and the Results pane. Notice that the file you saved is not affected; it remains in your project directory.

For more information: see Working with Analysis Results in the *GenePattern Java Client Guide*.

Next Steps

Click *Help>Getting Started* to redisplay the welcome message in the Java Client window.

This part of the tutorial introduced you to the GenePattern Java Client. Leave the Java Client open as you begin the next part, which introduces you to the GenePattern Web Client.

Using the GenePattern Web Client

This part of the tutorial introduces you to the GenePattern Web Client, which is primarily used for administrative tasks, such as adding and deleting analysis and visualization modules. In this part, you learn how to:

●   Start the Web Client.

●   Run analysis tools and pipelines.

●   Delete pipelines.

Starting the GenePattern Web Client

To start the GenePattern Web Client:

1. Double-click the GenePattern Web Client icon, shown below. By default, the GenePattern installation places this icon on your desktop.

GenePatternH
ome.html

The GenePattern login page appears:
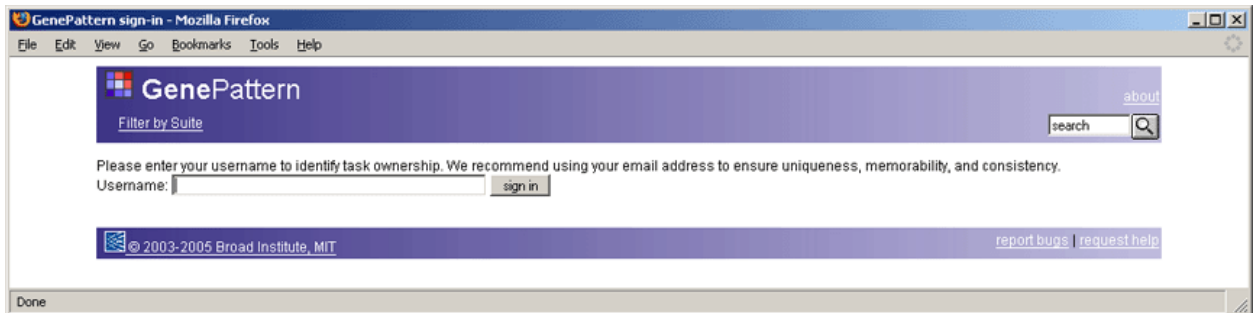
2. Enter your user name, usually your e-mail address, and click *sign in*. The Web Client home page appears, as shown below.

Note: Most GenePattern operations are available from both the Java Client and the Web Client. A few operations are available only from the Web Client. The following figure shows these operations circled in red.

26

Following are brief descriptions of the elements on the Web Client home page:

● The pipeline and task drop-down lists in the title bar provide quick access to pipelines and tasks (analysis and visualization modules).

● The Recent Jobs column lists the analyses that you have run.

● The Tasks panel provides access to the analysis and visualization modules.

● The Pipelines panel provides access to pipelines.

● The Tasks and Pipelines section of the third column provides additional links for working with modules and pipelines.

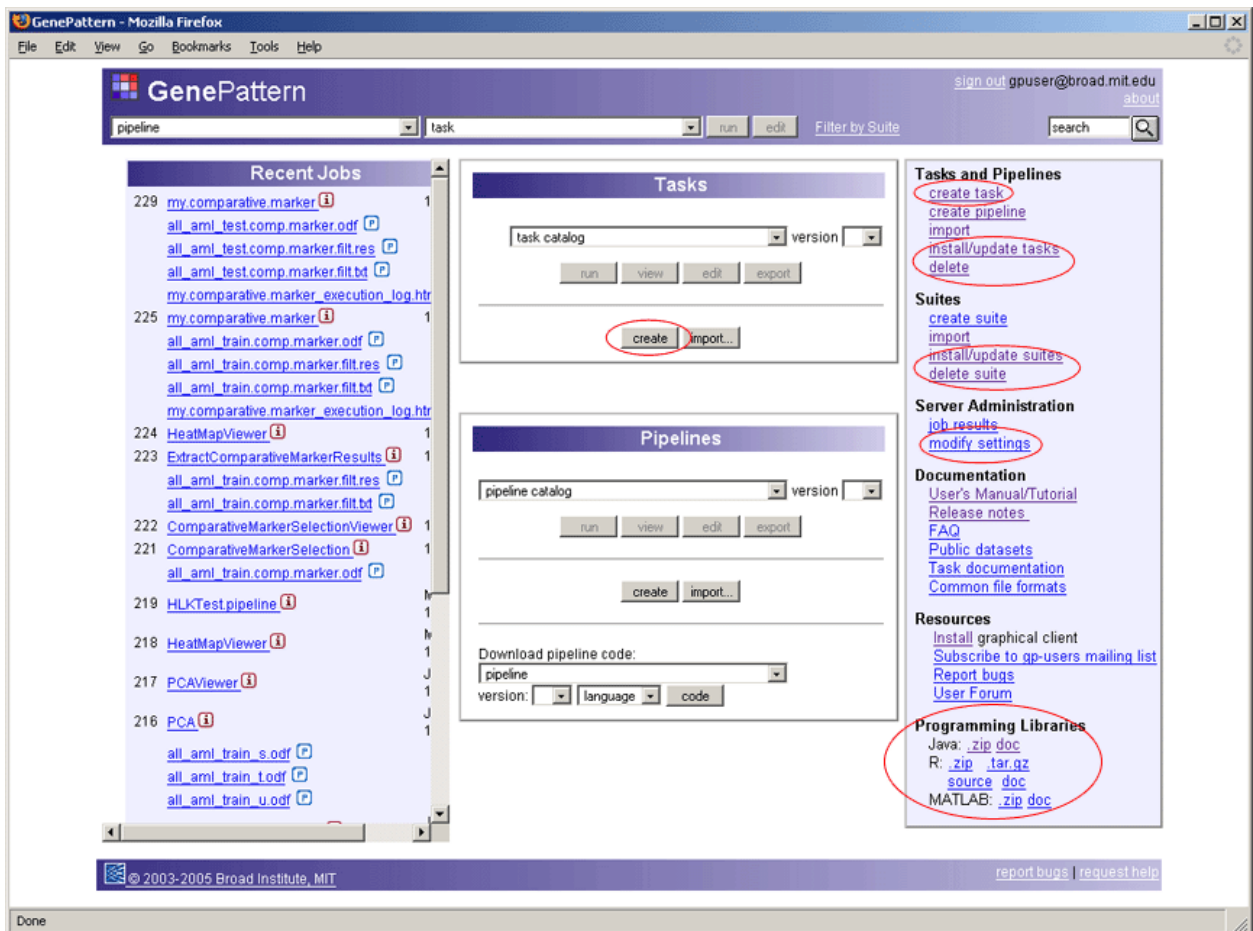● The Suites section provides access to suites, which allow you to group modules. The introduction to the Java Client in this tutorial did not discuss suites, but they are available from both the Java Client and the Web Client.

   For more information: see Working with Suites in the *GenePattern Java Client Guide*.

● The Server Administration section allows you to save and delete analysis results (*job results*) and modify server configuration settings (*modify settings*).

● The Documentation section provides access to all of the GenePattern documentation, including the documentation for each module and pipeline (*Task documentation*).

● The Resources section provides access to useful links that are also available on the GenePattern web page.

● The Programming Libraries section is for programmers interested in integrating GenePattern into their programming environments.

   For more information: see the GenePattern Programmer's Guide.

As you begin working with the Web Client, click the GenePattern icon in the title bar at any time to return to the Web Client home page.

Running an Analysis

Typically, you run analyses from the Java Client; however, you can also run them from the Web Client. In this section, you run the Comparative Marker Selection analysis and then use the Comparative Marker Selection Viewer to review the analysis results.

To run the Comparative Marker Selection analysis:

1. Select ComparativeMarkerSelection from the task drop-down list in the title bar.

2. Click *Run* (to the right of the drop-down list). The Web Client displays the module and its parameters:
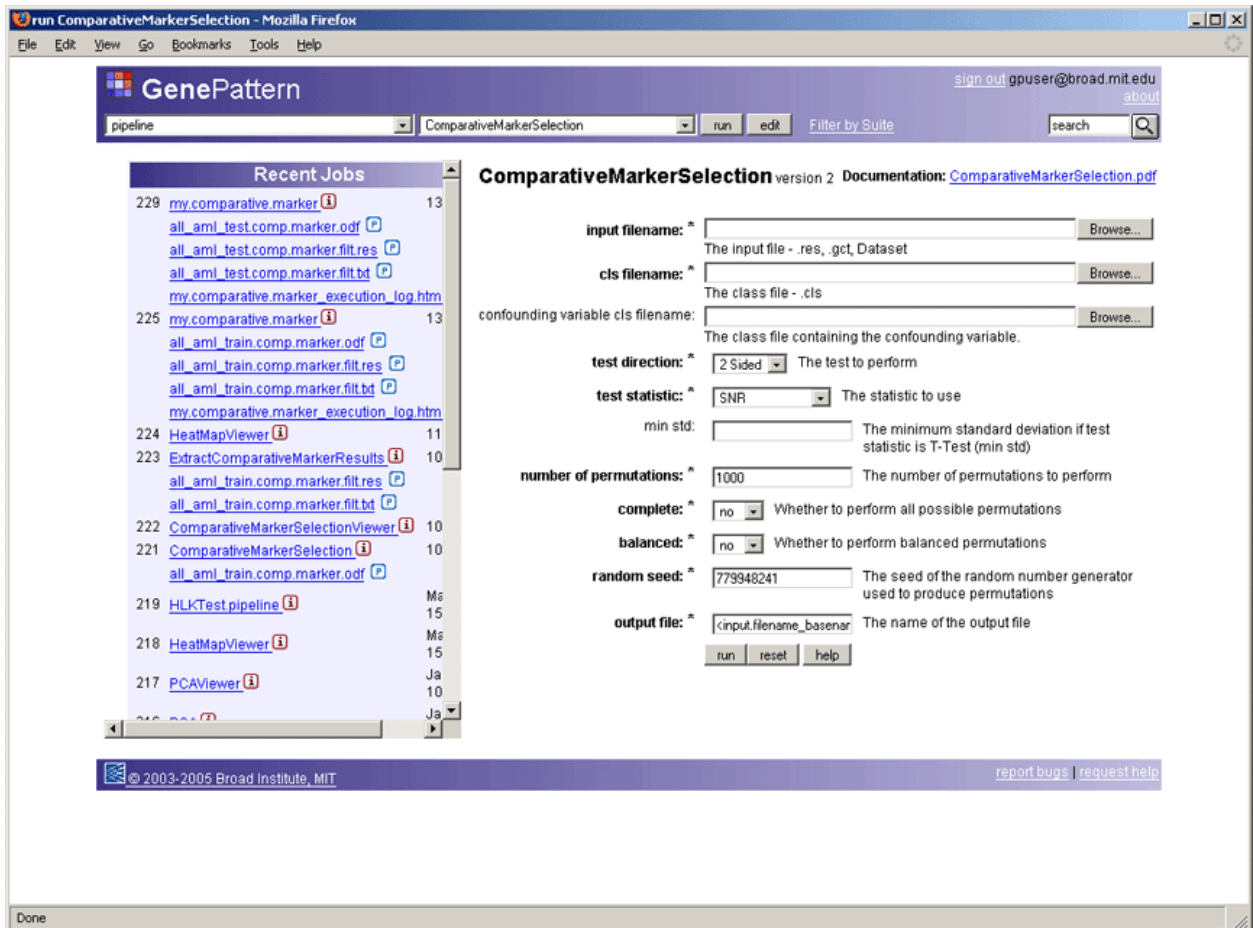
3. For the *input filename* parameter, select the `all_aml_train.res` file in the `gp_tutorial_files/all_aml` directory:

    1. Click the *Browse* button next to the right of the parameter field. The File Upload window appears:
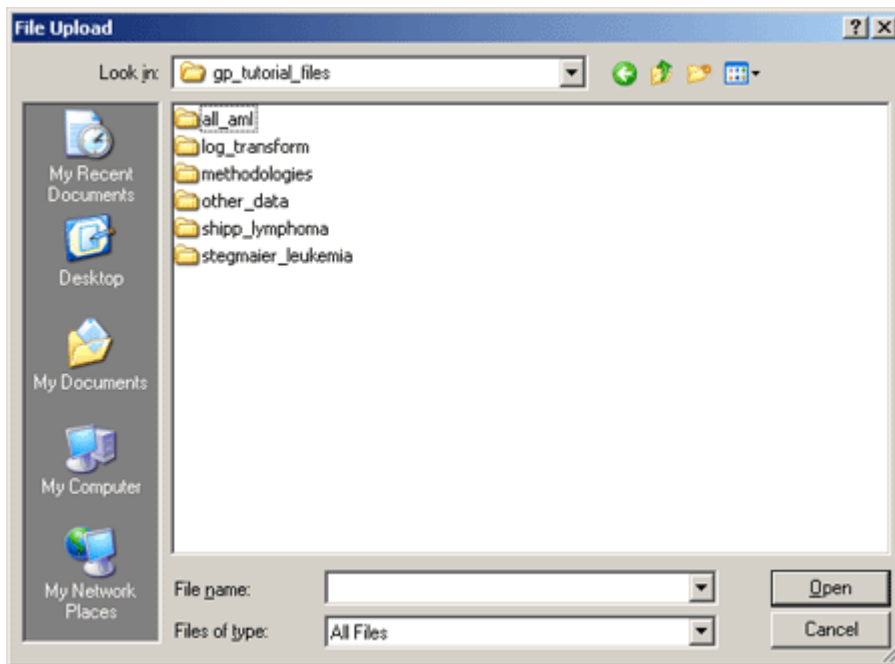
2. Navigate to the `gp_tutorial_files/all_aml` directory, select the `all_aml_train.res` file, and click *Open*. The file appears in the *input filename* parameter field.

4. For the *cls filename* parameter, select the `all_aml_train.cls` file in the `gp_tutorial_files/all_aml` directory. The file appears in the *cls filename* parameter field.

5. Click *Run* at the bottom of the form. The log window appears. After a few minutes, the analysis completes and the analysis results files appear in the log window:

The analysis results files listed here are stored on the GenePattern server. You can use the *download selected results* button to save the analysis results files and the *delete selected results* button to delete them. For now, leave the analysis results files on the server.

To view the analysis results in the Comparative Marker Selection Viewer:

1. Click the GenePattern icon in the title bar to display the Web Client home page.

   Notice that the Recent Jobs list shows the ComparativeMarkerSelection analysis and its analysis results file.
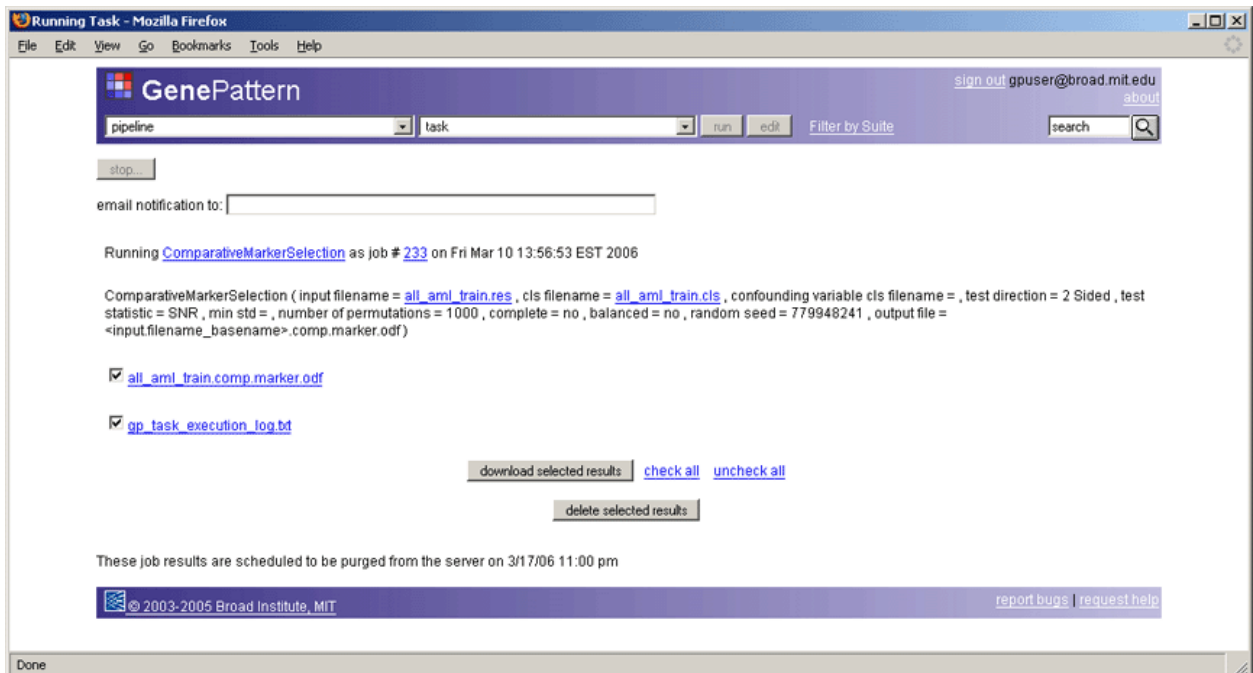
29

2. Select ComparativeMarkerSelectionViewer from the task drop-down list in the title bar.

3. Click *Run* to the right of the drop-down list. The Web Client displays the module and its parameters.

4. In the Recent Jobs list, right-click the analysis results file from the ComparativeMarkerSelection analysis and select *Copy Link Location*:
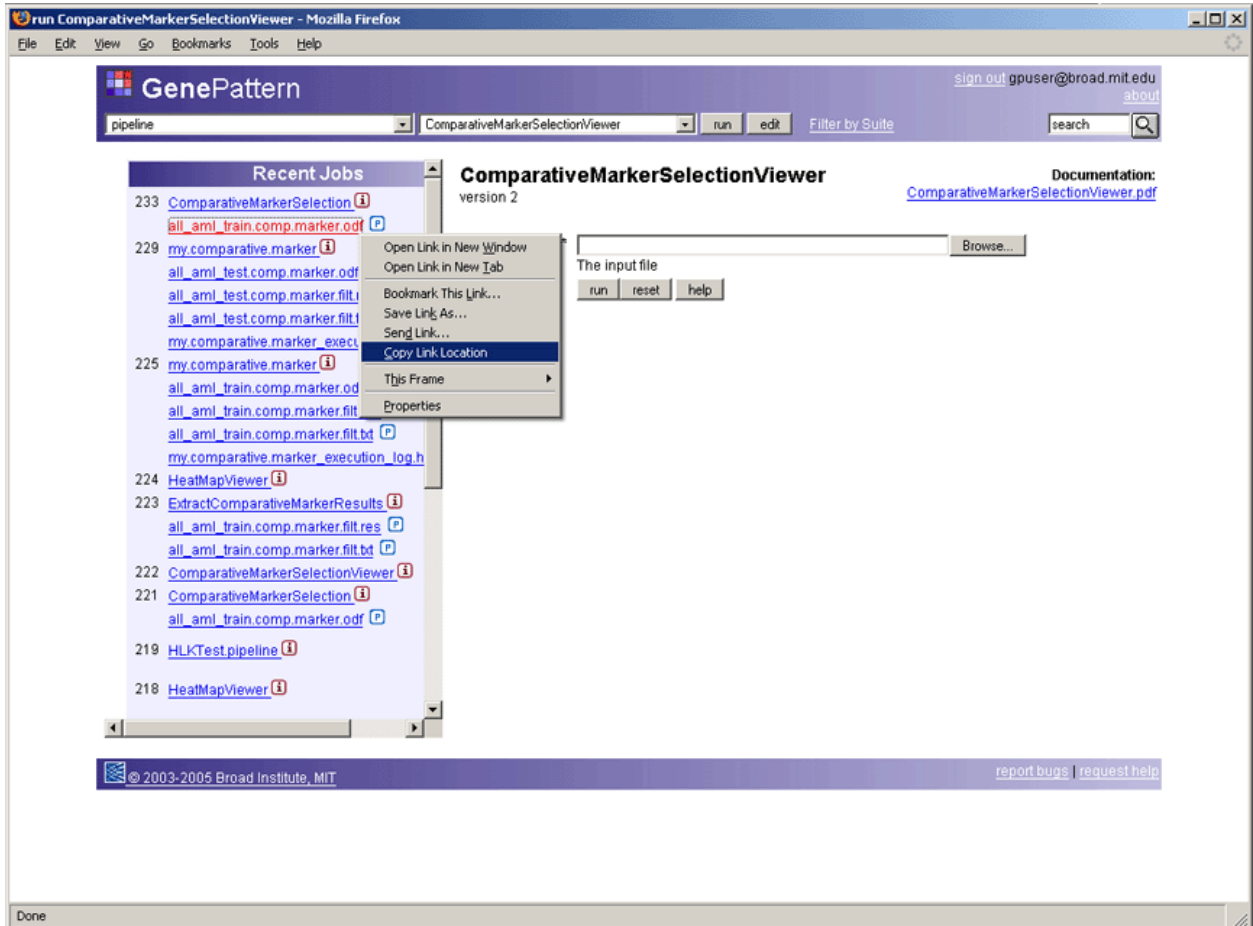
The Web Client copies the file name to the clipboard.

5. Paste the copied file name into the *input filename* parameter field:

   ▪ On Windows, click in the parameter field and press CTRL+v.

   ▪ On Mac OS X, click in the parameter field and press Apple+v.

6. Click *Run* (below the parameter). The log window appears. After a moment, the Comparative Marker Selection Viewer appears.

7. Close the Viewer window by clicking the Close (X) icon in the upper right corner.

8. Click the GenePattern icon in the Web Client title bar to display the Web Client home page.

Creating a Pipeline

The Web Client, like the Java Client, allows you to create an empty pipeline or create a pipeline from an analysis results file.

● To create an empty pipeline, click the Create button in the Pipelines panel in the middle column, as shown below.

● To create a pipeline from an analysis results file, click the P icon next to the file in the Recent Jobs list, as shown below.
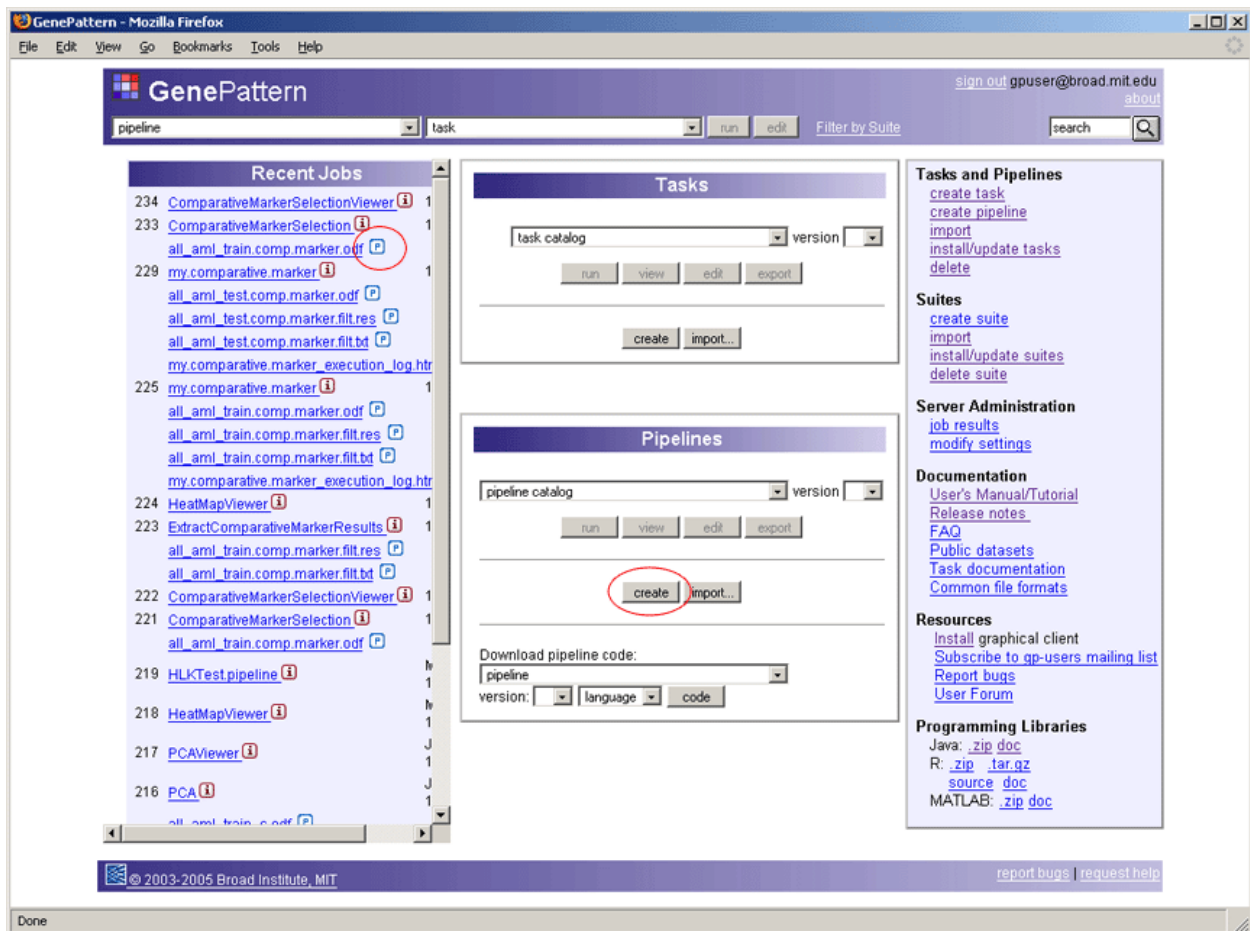
In general, it is easier to create pipelines from the Java Client rather than from the Web Client; therefore, the tutorial does not walk you through creating a pipeline in the Web Client.

For more information: see Managing Pipelines in the *GenePattern Web Client Guide*.

Deleting a Pipeline

You can create pipelines from either the Java Client or the Web Client; however, you can delete pipelines only from the Web Client.

In this exercise, you will delete the my.comparative.marker pipeline that you created from the Java Client:

1. Click the *delete* link under the Tasks and Pipelines heading in the third column. The Delete Tasks page appears:

2. Scroll down to your my.comparative.marker pipeline. Pipeline names are color-coded; the pipelines that you created are displayed in pink.

3. Select all versions of your pipeline. You can click the check box next to each version or click the *check all* link to the right of the version check boxes.

4. Scroll to the bottom or top of the Delete Tasks page and click *delete selected tasks*. A confirmation window appears:
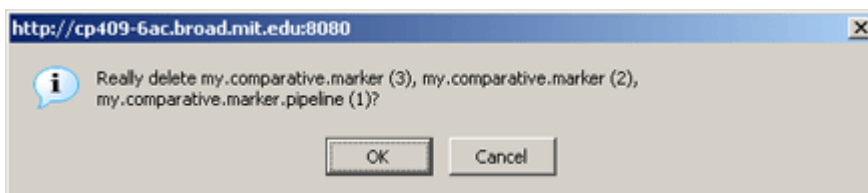
5. Click *OK* to delete the selected pipelines. The Web Client deletes the pipelines and redisplays the Delete Tasks page.

6. Scroll the list to confirm that you have deleted all versions of your my.comparative.marker pipeline.

7. Click the GenePattern icon in the Web Client title bar to display the Web Client home page.

Learning More about the Web Client

This tutorial provides only a brief introduction to the GenePattern Web Client. If you using GenePattern primarily to run its analysis and visualization modules or to create and run pipelines, you rarely need to use the GenePattern Web Client; you can use the GenePattern Java Client instead. If you are responsible for managing the GenePattern server or you are a programmer interested in integrating GenePattern into your programming environment, you are strongly encouraged to learn more about the Web Client.

For more information: see the GenePattern Web Client Guide.

Synchronizing the Clients and Server

You have now explored the GenePattern Web Client and the GenePattern Java Client. Both of these clients work with the GenePattern server. For example, when you run an analysis on the client, the client sends your request to the server, which runs the analysis. The analysis results files are stored on the server.

The GenePattern Web Client is implemented as a series of web pages. When you click a link in the Web Client, your web browser displays the appropriate page. As necessary, the web browser reads information from the GenePattern server to display the web page. Therefore, the information displayed in the Web Client reflects the current content of the GenePattern server.

The GenePattern Java Client is implemented as a software application. When you start the Java Client, it connects to the GenePattern server, reads information from the server, and builds the application interface. The Analysis, Visualization, Pipelines, and Suites menus and the contents of the Results pane all reflect information read from the GenePattern server at a particular point in time. Periodically, the Java Client retrieves updated information from the GenePattern server; for example, when you run an analysis, the Java Client updates the Results pane to include the new analysis results files. However, you may need to refresh the Java Client to ensure that the information it displays reflects the current content of the GenePattern server.

To refresh the Java Client:

1. Return to the Java Client window.

2. Examine the analyses in the Results pane. Notice that the ComparativeMarkerSelection analysis that you ran from the Java Client appears in the Results pane, but the ComparativeMarkerSelection analysis that you ran from the Web Client does not.

3. Click *File>Refresh>Jobs*. The Java Client connects to the GenePattern server, retrieves information about all analyses that you have run, and updates the Results pane.

4. Examine the analyses in the Results panel. Notice that both ComparativeMarkerSelection analyses now appear in the Results panel.

5. Examine the Pipelines menu. Notice that the menu still includes the my.comparative.marker pipeline, which you deleted from the Web Client.

6. Click *File>Refresh>Modules*. The Java Client connects to the GenePattern server, retrieves the list of modules (analysis modules, visualization modules, pipelines, and suites) from the server, and updates the Analysis, Visualization, Pipelines, and Suites menus.

7. Examine the Pipelines menu. Notice that the menu no longer includes the my.comparative.marker pipeline.

Shutting Down GenePattern

When you are finished working with GenePattern, you can shutdown both the clients and the server. Remember that the clients are connected to the server; therefore, you want to exit from the clients before you shutdown the server. If you shutdown the server first, the clients will display error messages as they attempt to contact the server.

To exit from GenePattern:

1. Exit from the Java and Web clients.

   ▪ To exit from the GenePattern Java Client, select *File>Quit*.

   ▪ To exit from the GenePattern Web Client, click the *sign out* link in the top right corner of the title bar and then close the web browser window.

2. Shutdown the server by double-clicking the Stop GenePattern Server icon, shown below. Alternative: if you are using Windows, you can close the console window.

StopGenePatt
ernServer

On Windows, when you shutdown GenePattern, the GenePattern console window closes. On Mac OS X, the GenePattern server icon disappears from the Dock. On Linux machines, the GenePattern server exits silently.

Learning More About GenePattern

Thank you for taking this time to learn about GenePattern!

As you continue to work with GenePattern, please visit the GenePattern web site, which provides an overview of the GenePattern analysis modules, as well as links to the GenePattern software and documentation.

We welcome your feedback. If you have suggestions, comments, or questions that are not answered by the documentation, contact the GenePattern help desk (gp-help@broad.mit.edu).