

# **GenePattern 3.1**

## **Web Client Guide**

## **Software Copyright**

### **The Broad Institute SOFTWARE COPYRIGHT NOTICE AGREEMENT**

This software and its documentation are copyright 2003-2008 by the Broad Institute/Massachusetts Institute of Technology. All rights are reserved.

This software is supplied without any warranty or guaranteed support whatsoever. Neither the Broad Institute nor MIT can be responsible for its use, misuse, or functionality.

# Web Client Guide

[Download pdf](#) 

GenePattern provides access to a broad array of computational methods used to analyze genomic data. Its extendable architecture makes it easy for computational biologists to add analysis and visualization modules, which ensures that GenePattern users have access to new computational methods on a regular basis.

If you are new to GenePattern, begin with the basics:

- The [Concepts Guide](#) provides a brief introduction to GenePattern: its primary objects (modules, pipelines, suites) and its client-server architecture. All other GenePattern documentation assumes that you are familiar with these concepts.
- The [Tutorial](#) provides a hands-on introduction to the GenePattern Web Client. Many users find that this is all they need to begin using GenePattern.

If you are familiar with GenePattern and are looking for more information about the Web Client, this is the right guide:

<a href="#">Getting Started</a>	Describes how to start and exit from the GenePattern Web Client. It provides an overview of the user interface, including the navigation bar and pop-up menus.
<a href="#">Running Modules and Pipelines</a>	Describes how to run GenePattern analyses and how to check job status.
<a href="#">Working with Analysis Results</a>	Describes how to display, save, and delete analysis results.
<a href="#">Working with Modules</a>	Describes how to install, create, edit, and delete modules.
<a href="#">Working with Pipelines</a>	Describes how to install, create, edit, and delete pipelines.
<a href="#">Working with Suites</a>	Describes how to install, create, edit, and delete suites.
<a href="#">Managing Modules, Pipelines, and Suites</a>	Provides detailed instructions for installing and deleting modules, pipelines, and suites. The previous sections summarize these actions and provide links to the detailed instructions provided in this section.
<a href="#">Managing the GenePattern Server</a>	Provides information for the GenePattern server administrator.

## Getting Started

This section describes how to start and exit from the GenePattern Web Client. It provides an overview of the user interface, including the navigation bar and pop-up menus.

- [Selecting a GenePattern Server](#)
- [Starting the GenePattern Web Client](#)
- [Exploring the User Interface](#)
- [Exiting from GenePattern](#)
- [Getting Help](#)

## Selecting a GenePattern Server

The server is the GenePattern engine. It hosts the GenePattern modules, pipelines, and suite; runs analyses; and stores analysis results. The GenePattern Web Client is the interface that you use to interact with the server. Before starting the GenePattern Web Client, you must know which GenePattern server you plan to use:

- **Public** GenePattern server: The Broad Institute hosts a publicly available GenePattern server at <http://genepattern.broad.mit.edu/gp/>.

- **Networked** GenePattern server: Many institutions have a GenePattern server installed in a central location, which can be accessed by researchers at that institution. To use a networked server, you must know its URL. Depending on how the GenePattern server was installed, you may also require a username and password.
- **Local** GenePattern server: You can install GenePattern and run the GenePattern server on your desktop computer. By default, the URL for a local GenePattern server is <http://localhost:8080/gp/>. To use a local server, install GenePattern and start the server as described below.

## Installing GenePattern

To install GenePattern, follow the instructions provided on the [Download GenePattern](#) page.

The same installation process is used by both the person installing a local GenePattern server for personal use and the person installing a networked GenePattern server for use by an institution. The difference is in the server configuration.

- If you are installing a local GenePattern server, use the default server settings.
- If you are installing a networked GenePattern, consider reconfiguring the GenePattern server as discussed in [Managing the GenePattern Server](#).

## Starting the GenePattern Server

Start the GenePattern server before starting the Web Client. The server takes several seconds to start; a client cannot connect to it while it is starting.

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



On Windows, while the server is starting, the cursor displays as an hourglass. The server is ready when the cursor returns to normal.

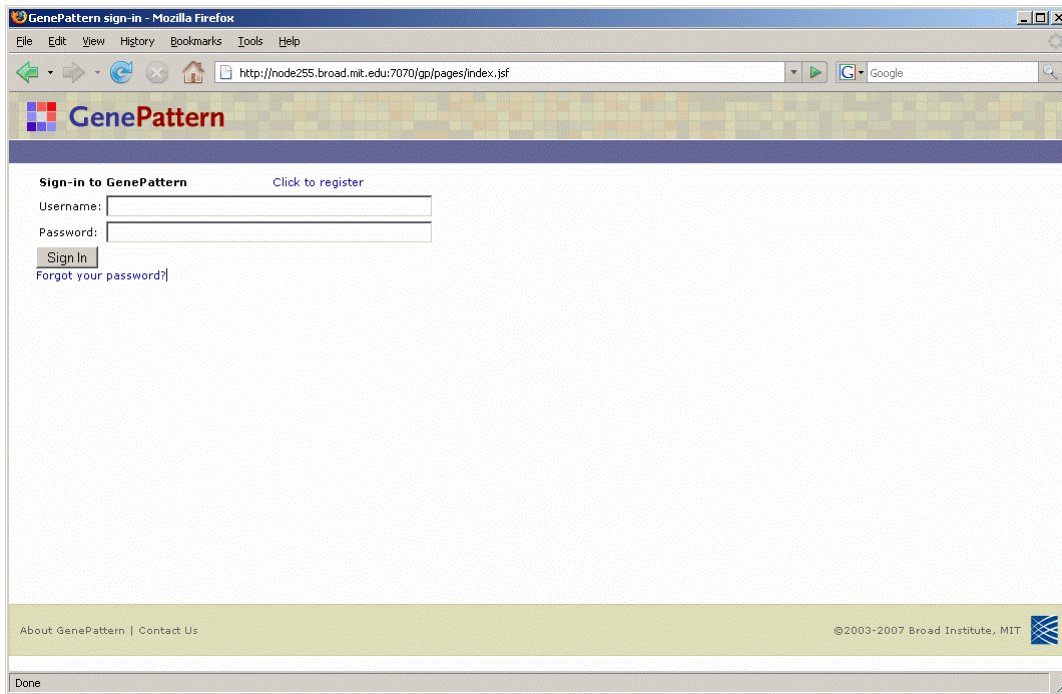
On Mac OS X, while the server is starting, the server icon bounces in the Dock. The server is ready when the icon stops bouncing.

## Starting the GenePattern Web Client

The GenePattern Web Client is a web-based application. To start the Web Client:

1. Open a web browser, such as Mozilla Firefox, Internet Explorer or Safari.
2. Enter the URL of the GenePattern server that you want to use. For example:
  - to use the public GenePattern server hosted by the Broad Institute, enter <http://genepattern.broad.mit.edu/gp/>.
  - to use the local GenePattern server that you installed and started on your desktop computer, enter <http://localhost:8080/gp/>.

The Web Client login page appears:



3. Enter your user name and password. If you do not have a GenePattern account, select *Click to register*.

Whether a GenePattern server requires passwords depends on how it is configured. The public GenePattern server requires passwords. By default, a local GenePattern server does not.

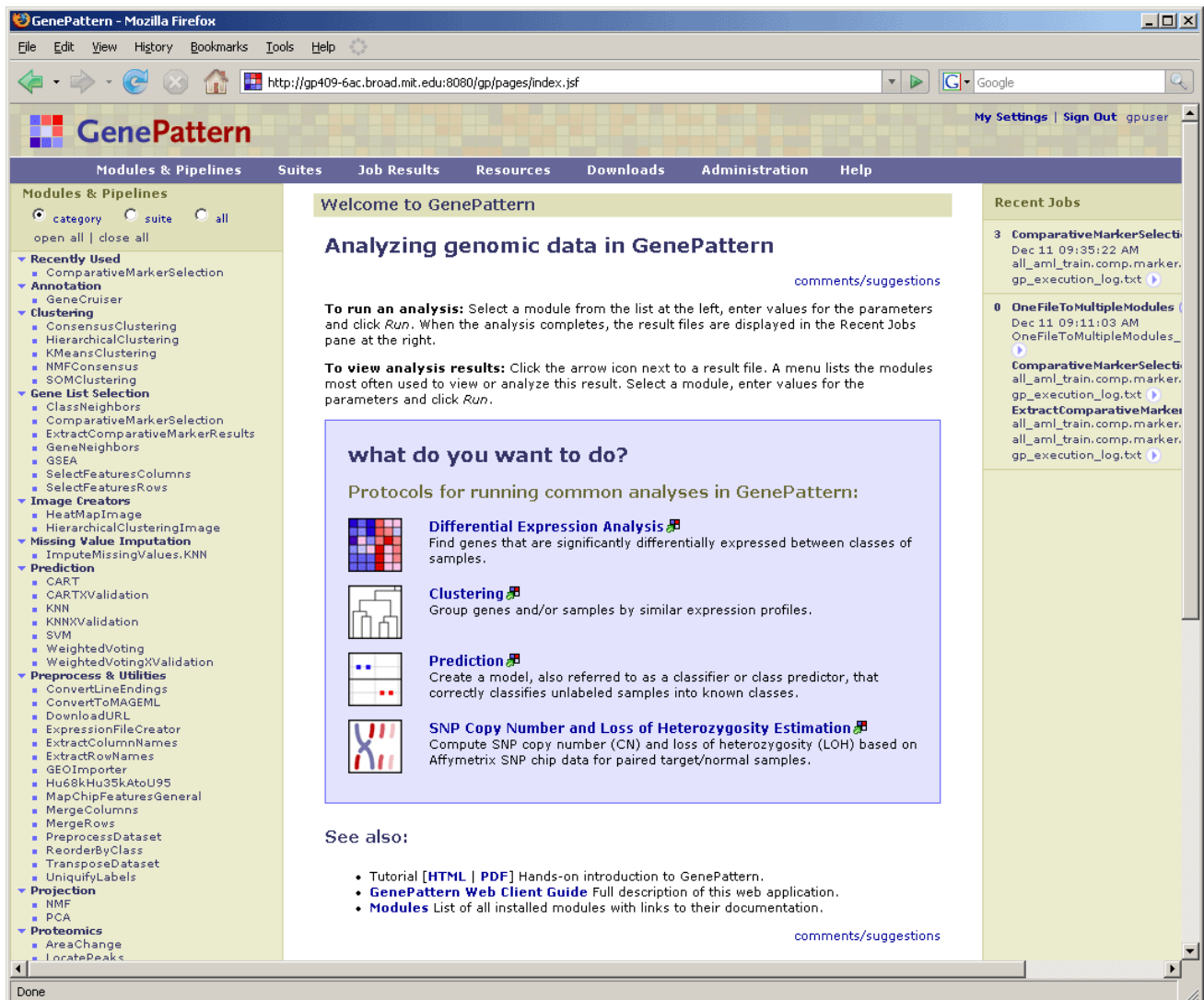
4. Click *Sign In*. The Web Client start page appears.

**Note:** If the Web Client cannot connect to the server, the browser displays a message such as “Unable to connect” or “Cannot display the webpage.”

- If you are using the public GenePattern server or a networked GenePattern server, wait a minute and try again.
- If you are using a local GenePattern server, [start the server](#) and then start the Web Client.

## Exploring the User Interface

When you start the Web Client, the Web Client displays the start page. To return to this page, click the *Modules & Pipelines* item in the navigation bar or the GenePattern icon in the title bar.



- The title bar includes:
  - GenePattern icon: Click the icon to return to this page.
  - *My Settings*: Click this link to modify your GenePattern account settings (see [My Settings](#)).
  - *Sign Out*: Click this link to exit from the GenePattern Web Client.
  - Your username: The GenePattern user name of the person currently logged into the Web Client.
- The navigation bar provides access to Web Client pages and operations (see [Navigation Bar](#)).
- The Modules & Pipelines pane lists the analysis modules and pipelines that you can run. Modules and pipelines not installed from the Broad repository are shown in red. Use the radio buttons to organize the modules and pipelines as you prefer:

- *Category* (default): Organizes analysis modules by functional category; pipelines are in the pipeline category. Each module is assigned to a category when it is created. To change the category, you must edit the module definition.
- *Suite*: Organizes modules and pipelines by suite. Suites are arbitrary collections of modules and pipelines. You can install suites from the Broad repository or create your own.

For more information about a suite, click *Suites>Manage* to display all available suites. Click a suite name to display its definition. Click the icon next to a suite name to display a menu of commands for working with the suite (see [Suites Menu](#)).

- *All*: Organizes modules and pipelines alphabetically by name.

- The center pane is the main display pane for the Web Client. The Web Client uses this pane to display information and to prompt you for input. Initially, the center pane contains a welcome page, which provides instructions for running common analysis protocols. To redisplay the welcome page, click the GenePattern icon in the title bar.
- The Recent Jobs pane lists the most recent analyses that you have run and the results files generated by those analyses, as well as any currently running analyses. Click the icon next to a job or file name to display a menu of commands for working with the job or file (see [Job Menu](#) and [File Menu](#)). Click [My Settings](#) to modify the History setting, which controls that number of analyses displayed in the Recent Jobs pane.

**Note:** If the URL of the Web Client is visible in your web browser, it points to the GenePattern server that you are using. The Web Client remains connected to the server throughout the session. The modules, pipelines and suites displayed in the Web Client are those installed on the server. When you run a module or pipeline, the Web Client runs the analysis on the server and stores the analysis results files on the server.

## Navigation Bar

The navigation bar, which appears on most Web Client pages, provides access to Web Client pages and operations. Click a link in the table to go to the section of this guide that describes that operation.

<b>Modules &amp; Pipelines</b>	Display the Web Client start page.
<a href="#">New Pipeline</a>	Create a pipeline.
<a href="#">New Module</a>	Create a module.
<a href="#">Install from repository</a>	Install a module or pipeline from the Broad repository.
<a href="#">Install from zip</a>	Install a module or pipeline from a zip file.
<a href="#">Manage</a>	Display installed modules or pipelines; delete modules or pipelines.
<b>Suites</b>	Display the Manage Suites page.
<a href="#">New</a>	Create a suite.
<a href="#">Install from repository</a>	Install a suite from the Broad repository.
<a href="#">Install from zip</a>	Install a suite from a zip file.
<a href="#">Manage</a>	Display installed suites; delete suites.
<b>Job Results</b>	Display the Results Summary page.
<a href="#">Results Summary</a>	Display jobs run on the server; delete jobs.
<b>Resources</b>	Display an overview of the resources.
Mailing List	Display the form you use to join a low-traffic GenePattern mailing list.
Report Bugs	Display the form you use to contact the GenePattern team to report bugs, provide feedback, or ask questions.
User Forum	Display the Yahoo GenePatternUserGroup page, which you can use to contact other GenePattern users.
Contact Us	Display a form, which you can use to send questions and comments to the GenePattern team.
<b>Downloads</b>	Display an overview of the available downloads.
Install Desktop Client	Download and install the Desktop Client software.
Programming Libraries	Download and install GenePattern libraries for use with Java, MATLAB, or R.
Public Datasets	Download sample datasets for use with GenePattern.
<b>Administration</b>	Display the Server Settings page.

<a href="#">Server settings</a>	Modify settings that affect the GenePattern server.
<b>Help</b>	Display the Web Client start page.
Tutorial	Display the <a href="#">Tutorial</a> , which provides a hands-on tour of the Web Client and Desktop Client.
Concepts Guide	Display the <a href="#">Concepts Guide</a> , which provides a brief introduction to GenePattern: its primary objects (modules, pipelines, suites) and its client-server architecture.
Desktop Client Guide	Display the <a href="#">Desktop Client Guide</a> , which describes the Desktop Client and how to use it.
Web Client Guide	Display this guide, which describes the Web Client and how to use it.
Programmers Guide	Display the <a href="#">Programmers Guide</a> , which provides guidelines for writing modules and instructions for accessing GenePattern from the Java, MATLAB, and R programming environments.
Module Documentation	Display a list of the modules and pipelines installed on your server, with brief descriptions and links to the module/pipeline documentation.
File Formats	Display the <a href="#">File Formats Guide</a> , which describes all file formats and provides instructions for creating input files.
Release Notes	Display the <a href="#">Release Notes</a> , which describes new features and known issues in this release.
FAQ	Display the GenePattern list of <a href="#">Frequently Asked Questions</a> .
About GenePattern	Display the release date and build number of the GenePattern server software.

## Job Menu

When the Web Client displays an analysis job and its results, click the icon next to the job name to display a menu of commands for working with that job. For more information, see [Working with Analysis Results](#).

Download	Download a zip file containing all analysis results files for this job.
Reload	Display the analysis and its parameters in the center pane, with the parameters set to the values used for this analysis job.
Delete	Delete the analysis job and its analysis results files from the GenePattern server.
Info	Display the parameter values and the analysis results files for this job.
View Java Code View MATLAB Code View R Code	Display the command line that you would use to run this job in the Java, MATLAB, or R programming environments. These commands are useful for programmers who want to access GenePattern from one of these programming environments or from their own applications.

## File Menu

When the Web Client displays an analysis job and its results, click the icon next to the file name to display a menu of commands for working with that file. For more information, see [Working with Analysis Results](#).

Delete	Delete the file from the GenePattern server.
Save	Download the file from the GenePattern server.
Create Pipeline	Create a GenePattern pipeline that includes the modules and parameters necessary to reproduce this results file.
<i>List of modules</i>	List modules that commonly use this type of file as an input parameter. Select an analysis to display its parameters in the center pane, with this results file specified as the first input parameter.



## Suites Menu

When the Web Client displays a suite, click the icon next to the suite name to display a menu of commands for working with that suite. For more information, see [Working with Suites](#).

Edit	Available only for suites that you have created. Display the Edit GenePattern Suite page, which you can use to modify your suite.
Delete	Delete the suite from the GenePattern server.
Export excluding dependents	Create a zip file that contains the definition of the suite, but not the modules or pipelines in the suite. The zip file can be used to install the suite on another GenePattern server (see <a href="#">Exporting and Installing Suites Using Zip Files</a> ). Installing the suite from this zip file will not install any modules or pipelines in the suite; they must already be installed on the GenePattern server or be installed separately.
Export including dependents	Create a zip file that contains the definition of the suite, as well as the modules and/or pipelines in the suite. The zip file can be used to install the suite on another GenePattern server (see <a href="#">Exporting and Installing Suites Using Zip Files</a> ). Installing the suite from this zip file will also install the modules and pipelines in the suite (unless they are already installed on the GenePattern server).

## My Settings

Use *My Settings* to change your GenePattern account information:

1. Click *My Settings* (in the title bar of the Web Client). The Web Client displays the Account Information page.
2. From the left pane, select the information that you want to modify. The Web Client displays the related page.
3. Modify and save your account information.
4. Optionally, return to step 2 to make additional changes.

GenePattern Account Information - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://gp409-6ac.broad.mit.edu:8080/gp/pages/accountInfo.jsf

Google

GenePattern

My Settings | Sign Out gpuser

Modules & Pipelines Suites Job Results Resources Downloads Administration Help

Account Information

- Change Email
- Change Password
- History
- Visualizer Memory

Current Email No email set

New Email

Confirm New Email

Save

About GenePattern | Contact Us

©2003-2008 Broad Institute, MIT

Done

- **Change Email:** Change the email address for your GenePattern account on this server. The GenePattern server uses your email address for the following:

- When you run a module or pipeline, the Web Client sends the job to the GenePattern client and displays the Job Status page. On that page, you can click the *email notification* check box to have the Web Client send you email when the job completes.
- If you forget your password and request a new password, the Web Client sends your new password to this email address.
- **Change Password:** Change the password for your GenePattern account on this server.  
**Note:** By default, GenePattern servers are installed without password protection. For information about enabling password protection, see [Securing the Server](#).
- **History:** Use this option to specify the number of recent analyses listed in the Recent Jobs pane on the Web Client start page.
- **Visualizer Memory:** Specify the Java virtual machine configuration parameters (such as VM memory settings) to be used when running visualization modules. By default, this option is used to specify the amount of memory to allocate when running visualization modules (-Xmx512M).

## Exiting from GenePattern

To exit from the GenePattern Web Client, click *Sign Out* in the top right corner of the title bar.

To shutdown a GenePattern server:

1. Exit from the Web and Desktop clients.
  - To exit from the GenePattern Web Client, click *Sign Out* in the top right corner of the title bar.
  - To exit from the GenePattern Desktop Client, if it is running, select *File>Quit*.
2. Shutdown the server by double-clicking the Stop GenePattern Server icon, shown below. **Alternative:** If you are using Windows, you can simply close the console window.



StopGenePatternServer

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.

If a GenePattern client remains running after you shutdown the server, the client displays error messages as it attempts to contact the server.

## Getting Help

The [GenePattern web site](#) provides an overview of GenePattern and its analysis modules, as well as links to the GenePattern software and documentation. The documentation is your primary source for help with GenePattern:

- The [Concepts Guide](#) provides a brief introduction to GenePattern: its primary objects (modules, pipelines, suites) and its client-server architecture. All other GenePattern documentation assumes that you are familiar with these concepts.
- The [Tutorial](#) provides a hands-on tour of the Web Client and Desktop Client.
- The [Desktop Client Guide](#) describes the Desktop Client and how to use it.
- The [Web Client Guide](#), this guide, describes the Web Client and how to use it.
- The [Programmers Guide](#) provides guidelines for writing modules and instructions for accessing GenePattern from the Java, MATLAB, and R programming environments.
- The [Modules](#) page lists the modules and pipelines in the Broad repository, with links to their documentation.
- The [File Formats Guide](#) describes all file formats and provides instructions for creating input files.
- The [Release Notes](#) describe new features and known issues in this release.
- [Frequently Asked Questions](#) answers common questions about GenePattern.

## Running Modules and Pipelines

An analysis module runs a single analysis. A pipeline runs a series of analysis modules. If you are unfamiliar with GenePattern modules and pipelines, see the [Concepts Guide](#).

- [Running a Module or Pipeline](#)
- [Visualizer Modules](#)
- [Entering Parameters](#)
- [Checking Analysis Results](#)
- [Rerunning an Analysis](#)

## Running a Module or Pipeline

You can run a module or pipeline in two ways:

- From the Modules & Pipelines pane of the Web Client start page.
- From the File menu of an analysis results file. Using this method chains the selected module to the module that created the analysis results file, which is useful for creating pipelines (see [Creating Pipelines](#)).

To run a module or pipeline:

1. Select the module or pipeline to run in one of two ways:
  - In the Modules & Pipelines pane, select the module or pipeline to run. The parameters appear in the center pane.
  - From the Recent Jobs pane (or the Results Summary page), click the icon next to an analysis results file and select the module to run. The parameters appear in the center pane. The first input file parameter is set to the analysis results file.
2. Enter values for the parameter fields (see [Entering Parameters](#)).
3. Click *Run*. The Web Client sends the job to the GenePattern server and displays the Job Status page. When the GenePattern server completes the job, the Web Client displays the results on the Job Status page (see [Checking Analysis Results](#)). How long a job takes to complete depends on the size of your dataset and the analysis that you are running.
4. Click *Return to Modules & Pipelines Start* to return to the Web Client start page. The Recent Jobs pane shows the job that you just ran.

Modules run from a pipeline work exactly the same as those run directly from the Web Client. Analysis results files generated by a pipeline are exactly the same as those generated by modules run directly from the Web Client.

## Visualizer Modules

GenePattern analysis and visualization modules require Java 1.5.

- Analysis modules are run on the GenePattern server, which must have Java 1.5 installed.
- Visualizer modules are run on the GenePattern client; therefore, the computer you are using to run the Web Client must also have Java 1.5 installed.

When you run a visualizer module, the Web Client sends the job to the GenePattern server and displays the Job Status page. When the job completes, the Web Client displays the results on the Job Status page and opens the visualizer. If you move off the Job Status page before the Web Client opens the visualizer, the visualizer is not displayed.

Pipelines can include analysis and visualization modules. Visualizer modules in a pipeline are displayed just as if you ran them directly from the Web Client. Pipelines can also include nested pipelines; however, visualizer modules in a nested pipeline are not displayed.

## Entering Parameters

When you select a module or pipeline, the Web Client displays its parameters:

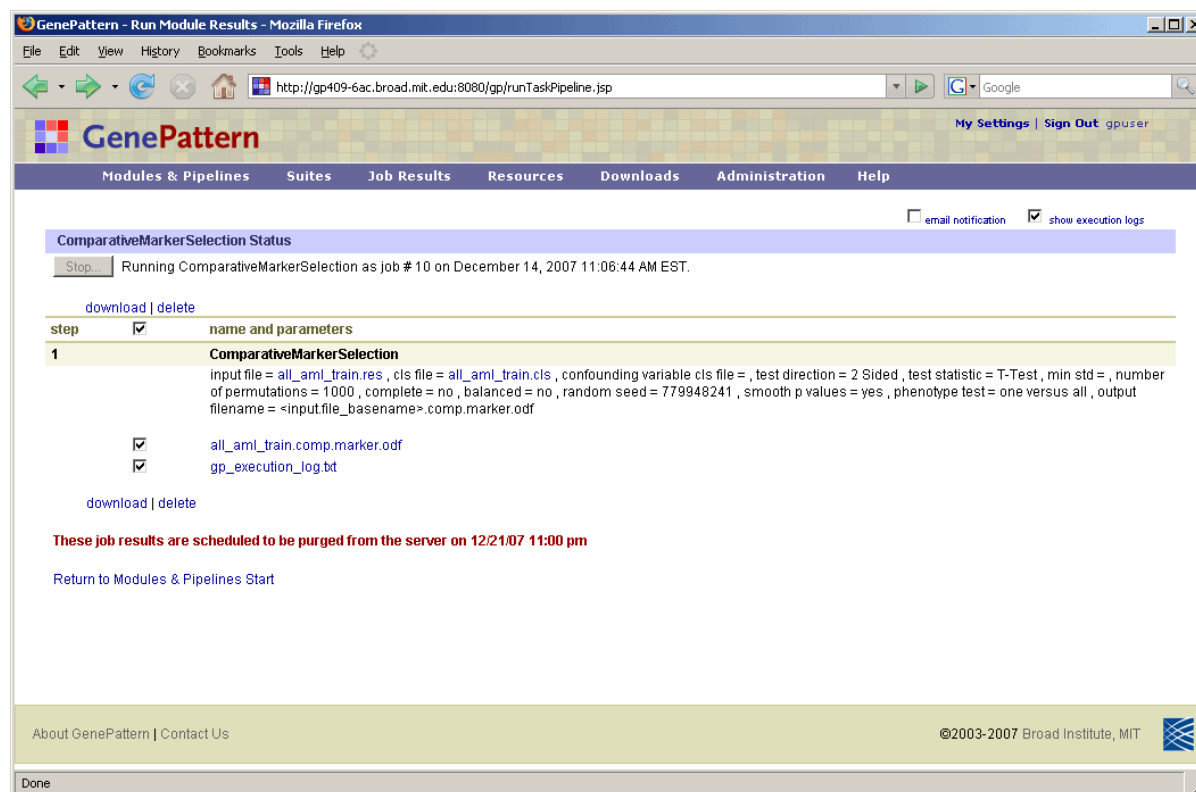
The screenshot shows the GenePattern web interface in a Mozilla Firefox browser. The main content area displays the 'ComparativeMarkerSelection' module form, version 3.10. The form has a 'Show parameter descriptions' checkbox at the top right. Below it, there are several parameter fields, each with a description and a 'Browse...' button for file selection. The parameters are: 'input file\*' (required), 'cls file\*' (required), 'confounding variable cls file' (optional), 'test direction\*' (required, dropdown menu), 'test statistic\*' (required, dropdown menu), 'min std' (optional), 'number of permutations\*' (required, text input), and 'complete\*' (required, dropdown menu). At the top of the form are 'Run' and 'Reset' buttons, and links for 'properties', 'export', and 'help'. On the left side, there is a sidebar with 'Modules & Pipelines' and 'Recently Used' sections. On the right side, there is a 'Recent Jobs' section showing a list of previous analyses.

The parameter form is similar for all modules and pipelines:

- The form lists each parameter followed by an entry box for the parameter value. Most modules have parameters; most pipelines do not.
- Below each parameter is a brief description of that parameter. Use the *Show parameter descriptions* check box at the top of the form to hide and show the descriptions.
- Required parameters are marked by an asterisk (\*). Other parameters are optional.
- When a parameter requires a file name, you can enter it in many ways:
  - Click *Browse* (next to the file parameter box) to display the file selection window. Navigate to the desired file, select it, and click *Open*. Leave the *Upload File* radio button selected.
  - From the *Recent Jobs* pane, click the icon next to an analysis results file of the desired type and select *Send To parameter-name*. Leave the *Specify URL* radio button selected. This has the benefit of chaining this analysis to the previous analysis, which can be useful for creating pipelines (see [Creating Pipelines](#)).
- At the top and bottom of the form, are the following options:
  - *Run*: Starts the analysis.
  - *Reset*: Resets all parameters to their default values.
  - *Properties*: Displays the module or pipeline definition form. The module definition form lists the program that implements the analysis. The pipeline definition form lists the modules run by the pipeline.
  - *Export*: Creates a zip file that contains the module or pipeline definition. The zip file can be used to install this module or pipeline on another GenePattern server.
  - *Edit*: Available only for modules that you have created. Displays the Update Module form, which allows you to update your module.
  - *Help*: Displays the module or pipeline documentation.
- At the bottom of the form, the *View Code* field displays the code (Java, MATLAB, or R) used to run the module with the parameters that you have entered. This can be useful for programmers writing batch procedures or new modules.

## Checking Analysis Results

When you run a module or pipeline, the Web Client sends the analysis job to the GenePattern server and displays the job status page. When the job completes, the Web Client updates the status page to list the analysis results files. Analysis results are stored on the GenePattern server for a period of time (by default, one week) and then deleted. The text shown in a **bold red** font tells you when the analysis results will be deleted.



From the status page, while a job is running, you can:

- Select the email notification check box to have the GenePattern server send you email when the job is finished. If you are running a job that may take an extended period of time, select this check box and continue with other work. You can run other analyses or exit from the Web Client. The GenePattern server continues to work on the job and sends you email when it is finished. When you receive the email, start the Web Client and review the analysis results.
- Click *Stop* to cancel the job.

After the job is finished, the Web Client displays the analysis results files and you can:

- Click an analysis results file to display the file in a text viewer.
- Click *Download* to copy the selected files from the GenePattern server to your own directories. To select a file, select the check box next to it.
- Click *Delete* to delete the selected files from the GenePattern server. To select a file, select the check box next to it.
- Select/clear the *show execution logs* check box to show/hide the execution log files.

Click *Return to Modules & Pipelines Start* to return to the Web Client start page. The job and its results files are displayed in the Recent Jobs pane of the Web Client start page and on the Results Summary page. The file commands on the status page are also available in the Recent Jobs pane (see [Exploring the User Interface](#)) and on the *Job Results>Results Summary* page (see [Job Results Page](#)).

## Rerunning an Analysis

To rerun an analysis:

1. Display your analysis jobs in one of two ways:
  - Click *Modules & Pipelines* to display the Web Client start page, where your most recent jobs are listed in the Recent Jobs pane.

- Click *Job Results>Results Summary* to display the Job Results page, which lists all of your analysis jobs.
2. Click the icon next to the job that you want to rerun and select *Reload*. The parameters appear in the center pane set to the values that were used for this job.
  3. Optionally, modify the parameter values.
  4. Click *Run*.

## Working with Analysis Results

When you run a module or pipeline, GenePattern runs the analysis job on the GenePattern server. Analysis results are stored on the GenePattern server for a period of time (by default, one week) and then deleted. If you are unfamiliar with how GenePattern runs modules and pipelines, see the [Concepts Guide](#).

The following table summarizes the different ways you can work with analysis results:

Displaying analysis results	<p>To display analysis results:</p> <ol style="list-style-type: none"> <li>1. Display your analysis jobs in one of two ways: <ul style="list-style-type: none"> <li>▪ Click <i>Modules &amp; Pipelines</i> to display the Web Client start page, where your most recent jobs are listed in the Recent Jobs pane.</li> <li>▪ Click <i>Job Results&gt;Results Summary</i> to display all of your analysis jobs.</li> </ul> <p>For each job, the Web Client lists all files stored on the server for that job.</p> </li> <li>2. Display a file: <ul style="list-style-type: none"> <li>▪ Click the file name to display the file in a text viewer.</li> <li>▪ Click the icon next to the file and select a module to run that module against the file. Modules with “Viewer” in their name are designed to display analysis results files.</li> </ul> </li> </ol> <p>For more information, see <a href="#">Displaying Analysis Results</a>.</p>
Saving analysis results	<p>To save results persistently (beyond the period of time they are stored on the server), you must download the files to a more permanent location. To save results:</p> <ol style="list-style-type: none"> <li>1. Display your analysis jobs, as shown above.</li> <li>2. Save one or more of the files for a job: <ul style="list-style-type: none"> <li>▪ To save a single file, click the icon next to a file and select <i>Save</i>. GenePattern saves the file to the location that you select.</li> <li>▪ To save all of the files, click the icon next a job and select <i>Download</i>. GenePattern creates a zip that contains all of the results files.</li> </ul> </li> </ol>
Deleting analysis results	<p>If you no longer need your analysis results, you can delete the files from the server. To delete the files:</p> <ol style="list-style-type: none"> <li>1. Display your analysis jobs, as shown above.</li> <li>2. Delete one or more of the files for a job: <ul style="list-style-type: none"> <li>▪ To delete a single file, click the icon next to a file and select <i>Delete</i>.</li> <li>▪ To delete all of the files, click the icon next to a job and select <i>Delete</i>.</li> </ul> </li> </ol>

## Displaying Analysis Results

When you run a module or pipeline, the files generated by the module/pipeline are stored on the GenePattern server. The module author determines the content and format of the generated files; however, by convention, each module generates the following files:

- **Analysis results files** are typically formatted text files that contain the results of the analysis. Most analysis results files are intended to be used as input to subsequent analyses. Although these files can be viewed in a text viewer, the amount of information in the files may make them difficult to read.

If an analysis module generates an analysis results file that requires examination and interpretation, there is generally a corresponding visualization module that you can use to display the results. Visualization modules have “Viewer” in the title. For example:

- Results from `ComparativeMarkerSelection` are viewed using the `ComparativeMarkerSelectionViewer`
- Results from modules in the Prediction category are viewed using the `PredictionResultsViewer`
- Results from `SOMClustering` are viewed using the `SOMClusteringViewer`
- Gene expression datasets are often viewed using the `HeatMapView`
- **Execution log files** are text files that describe how the analysis was run. A module might generate one or more of the following log files:
  - `gp_task_execution_log.txt`: contains the parameter values used to run the analysis, which is useful for reproducing analysis results.
  - `stdout.txt`: contains "standard output" messages; that is, comments generated as the analysis module runs.
  - `stderr.txt`: contains "standard error" messages; that is, information about errors (if any) that occurred during the analysis.

Log files can be viewed using a text viewer.

To display analysis results:

1. Display your analysis jobs in one of two ways:
  - Click *Modules & Pipelines* to display the Web Client start page, where your most recent jobs are listed in the Recent Jobs pane (see [Exploring the User Interface](#)).
  - Click *Job Results>Results Summary* to display all of your analysis jobs (see [Job Results Page](#)).

For each job, the Web Client lists all files stored on the server for that job.

2. To display an analysis results files in a visualization module, click the icon next to the file and select the module to use. `GenePattern` displays the analysis parameters in the center pane. The first input file parameter is set to this analysis results file.
3. To display an analysis log file in a text viewer, click the file name.

## Job Results Page

When you run a module or pipeline, the analysis job is run on the `GenePattern` server and the analysis results are stored on the server. The Recent Jobs pane of the Web Client start page lists your most recent analysis jobs and their analysis results (see [Exploring the User Interface](#)). The Job Results page lists all your analysis jobs and their analysis results.

To display the Job Results page, click *Job Results>Results Summary*:

GenePattern | Job Results - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://gp409-6ac.broad.mit.edu:8080/gp/pages/jobResults.jsf

Google

My Settings | Sign Out gpuser

Modules & Pipelines Suites Job Results Resources Downloads Administration Help

Job Results

☐ show everyone's jobs ☒ show execution logs

sort jobs by: job id number sort files by: name Sort

job	module/file	size	submitted	completed	status
11	ExtractComparativeMarkerResults		Dec 14 11:09:59 AM	Dec 14 11:10:02 AM	finished
	gp_execution_log.txt	1.0 KB		Dec 14 11:10:02 AM	
	all_aml_train.comp.marker.filt.txt	1.0 KB		Dec 14 11:10:02 AM	
	all_aml_train.comp.marker.filt.res	20.0 KB		Dec 14 11:10:02 AM	
10	ComparativeMarkerSelection		Dec 14 11:06:44 AM	Dec 14 11:07:06 AM	finished
	gp_execution_log.txt	1.0 KB		Dec 14 11:07:07 AM	
	all_aml_train.comp.marker.odf	1.8 MB		Dec 14 11:07:06 AM	
3	ComparativeMarkerSelection		Dec 11 09:35:05 AM	Dec 11 09:35:22 AM	finished
	gp_execution_log.txt	1.0 KB		Dec 11 09:35:23 AM	
	all_aml_train.comp.marker.odf	664.0 KB		Dec 11 09:35:22 AM	

delete

1

About GenePattern | Contact Us

©2003-2008 Broad Institute, MIT

Done

From this page, you can

- View the jobs run on your GenePattern server. To view only your own jobs, clear the *show everyone's modules* check box. To view all jobs, select the check box.
- View the results files for each job:
  - Most modules generate both analysis results files and execution log files. To display only the analysis results files, clear the *show execution logs* check box. To display all generated files, select the check box.
  - Click *open all* or *close all* to expand all jobs (displaying both jobs and their generated files) or collapse all jobs (displaying only the jobs and hiding their generated files).
- Delete jobs and/or results files from your GenePattern server. Deleting a job deletes all of its results files.
- Sort the jobs and/or files:

sort jobs by: job id number sort files by: name Sort

- Use the drop-down lists to select the job and file attributes to use for the sort.
- Use the up and down arrow next to each drop-down list to select the sort order (ascending or descending).
- Click *Sort*.

For each job, the Web Client displays similar information:

- Job number.
- Name of the module or pipeline that was run. Click the icon following the job name to display the [job menu](#).
- List of generated files stored on the GenePattern server. For each file:
  - File name. Click the icon following the file name to display the [file menu](#).
  - Size of the file.
  - Date and time the file was created.

Click the icon to the left of the job name to hide/show the generated files.

- Date and time the job was submitted.
- Date and time the job was completed.
- Job status: running or finished.



## Working with Modules

Analysis and visualization modules are at the heart of GenePattern. **Analysis modules** provide computational methods and tools for gene expression analysis, proteomics data analysis, SNP analysis, and data preprocessing and conversion. **Visualization modules** display your data and analysis results graphically. If you are unfamiliar with GenePattern modules and pipelines, see the [Concepts Guide](#).

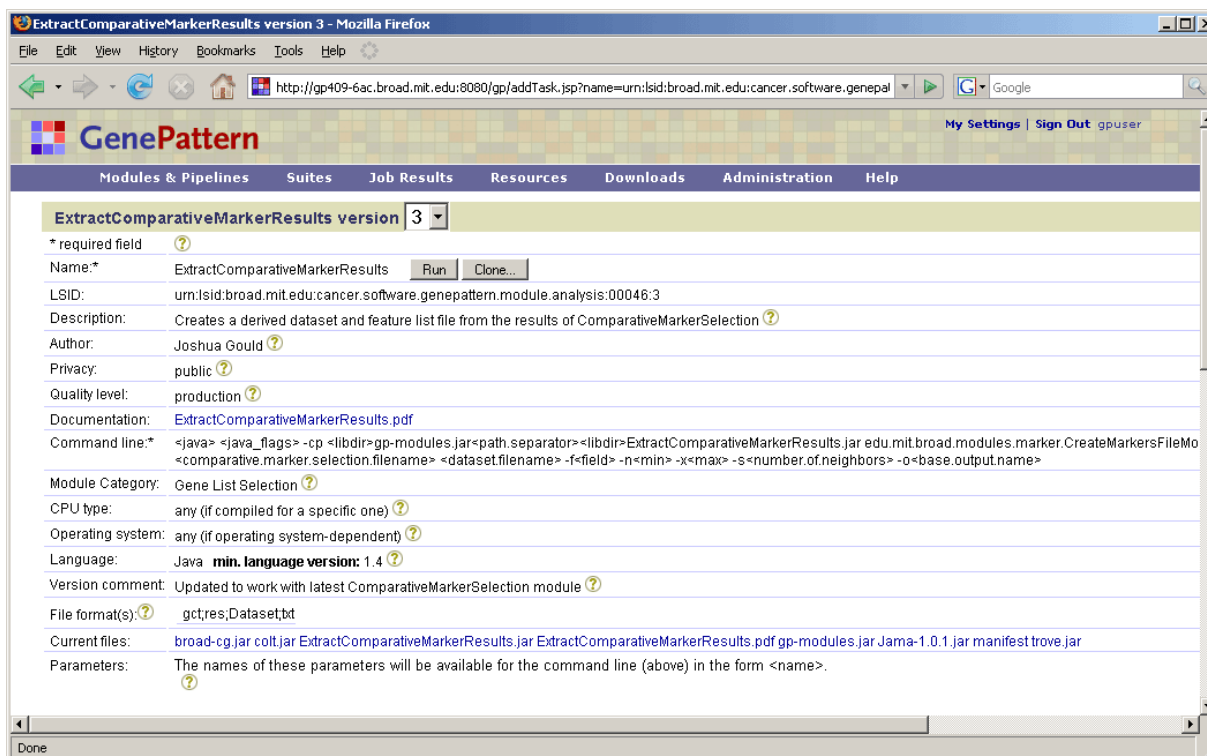
The following table summarizes the different ways you can work with GenePattern modules.

Run a module	Select a module, enter its parameters and click <i>Run</i> . For more information, see <a href="#">Running Modules and Pipelines</a> .
Display module definitions	<p>A module's definition includes the author, the command line used to invoke module, and the programs used to execute module.</p> <p>To display a module's definition, click <i>Modules &amp; Pipelines</i> and select the module. When the Web Client displays the module parameters, click <i>Properties</i>. The Web Client displays the <a href="#">module definition page</a>.</p>
Share modules with other users	<p>Zip files provide a convenient way to share modules with other GenePattern users:</p> <ul style="list-style-type: none"><li>• To export a module to a zip file, click <i>Modules &amp; Pipelines</i> and select the module to export. When the Web Client displays the module parameters, click <i>Export</i>.</li><li>• To install a module from a zip file, click <i>Modules &amp; Pipelines&gt;Install from zip</i>.</li></ul> <p>For more information, see <a href="#">Exporting and Installing Modules &amp; Pipelines Using Zip Files</a>.</p>
Install modules from the repository	The Broad Institute maintains a repository of modules, pipelines, and suites. To install modules from the Broad repository, click <i>Modules &amp; Pipelines&gt;Install from Repository</i> . For more information, see <a href="#">Installing Modules &amp; Pipelines from the Repository</a> .
Create modules	An analysis module invokes a program that executes the desired function. To create a module, you must write the program that implements the analysis and then create the GenePattern module that invokes that program. For more information, see <a href="#">Creating Modules</a> .
Edit modules	You can edit a module that you have created or copy a public module and edit your copy of the public module. For more information, see <a href="#">Editing Modules</a> .
Delete modules	To delete a module from your GenePattern server, click <i>Modules &amp; Pipelines&gt;Manage</i> . For more information, see <a href="#">Managing Modules &amp; Pipelines</a> .

## Module Definition Page

To display the definition of a module:

1. Click *Modules & Pipelines* to display the Web Client start page.
2. Select the module to display.
3. When the Web Client displays the module parameters, click *Properties*. The Web Client displays the module definition page, as shown below.



The module definition page is a read-only version of the page used to create and edit the module. The field descriptions below are presented in two parts: a brief description of the field, which is generally sufficient if you are viewing the module definition, and additional details, which are necessary if you are creating or editing the module definition.

- **Name.** Name of the module.

The name of the module is used to identify the module in the user interface of the GenePattern clients. The name should be a short but descriptive, without spaces or punctuation, and may include both upper- and lower-case characters.

Example: ConsensusClustering

- **LSID.** The Life Science Identifier (LSID) used to uniquely identify a GenePattern module.

You cannot create or edit LSIDs. The GenePattern server automatically assigns an LSID to each version of a module. If you are unfamiliar with GenePattern versioning, see the [Concepts Guide](#).

- **Description.** Brief description of the module.

GenePattern displays the description, sometimes in abridged form, in forms and drop-down lists in the clients and in generated code when creating scripts from pipelines. The description should be a sentence or short paragraph that documents succinctly what your module does and why someone would want to use it.

Example from ConsensusClustering: Resampling-based clustering method

- **Author.** The author's name and affiliation (company or academic institution).

GenePattern displays the author name as part of the module definition. This is a comment-only field. If you share this module, the author field allows other users to credit the author when citing the module and to contact the author with questions, suggestions, or enhancement ideas.

Example from ConsensusClustering: Stefano Monti, Broad Institute

- **Privacy.** Modules may be marked as public or private:

- Public modules may be accessed by anyone using the GenePattern server.
- Private modules may be accessed only by the person who installed or created the module (or by an administrator).

When you create a module, by default, it is marked private. When you are ready for others to use your module, change the privacy to public.

- **Quality level.** One of three terms that indicates the author's confidence in the robustness of the module: development, preproduction, and production.

When you create a module, by default, its quality level is development. Although these terms have no strict definitions, they are useful for setting user expectations. If you share this module, set the quality level appropriately.

- **Command line.** Command line used to launch the module. Values enclosed in angle brackets are replaced by specific values before the command executes.

When you create/update a module, the command line is critical and must be platform-independent. You define the command line as a combination of fixed and dynamic text. GenePattern resolves the dynamic text to build the command line that executes the module. For more information, see [Defining the Module Command Line](#).

- **Module category.** Category name assigned to this module, which is used to organize modules and pipelines. Pipelines are always assigned to the category name *pipeline*.

When you create/update a module, you can choose an existing category name or create a new category name. If your module fits into an existing category, such as Preprocess & Utilities, select that category from the drop-down list; otherwise, click the *New* button to add a new category. The Web Client creates the drop-down list of categories dynamically based on the categories of the modules installed on your GenePattern server. If you delete the last module in a given category, that category is removed from the drop-down list.

- **CPU type.** Indicates the type of CPU required to run the module, or *any* if the module runs on any type of CPU.

When you create/update a module, if your code is compiled for a specific platform (Intel, Alpha, PowerPC, etc.), select that platform from the drop-down list. GenePattern enforces CPU requirements when it runs the module.

- **Operating system.** Indicates the operating system required to run the module, or *any* if the module runs on any operating system.

When you create/update a module, if your code requires a specific operation system (Windows, Linux, MacOS, etc.), select that operating system from the drop-down list. GenePattern enforces operating system requirements when it runs the module.

- **Language.** Indicates the programming language used to implement the module.

- **min. language version.** Indicates the version of the programming language used to implement the module.

GenePattern does not enforce programming language requirements. However, including the language version information in the module definition gives prospective users a hint concerning system requirements.

- **Version comment.** Describes changes made to the module in this version.

When you update a module, briefly describe the changes that you have made. When GenePattern clients display a drop-down list of versions, the comments for each version are visible in the drop-down list.

- **File format(s).** Lists the file formats used by the module.

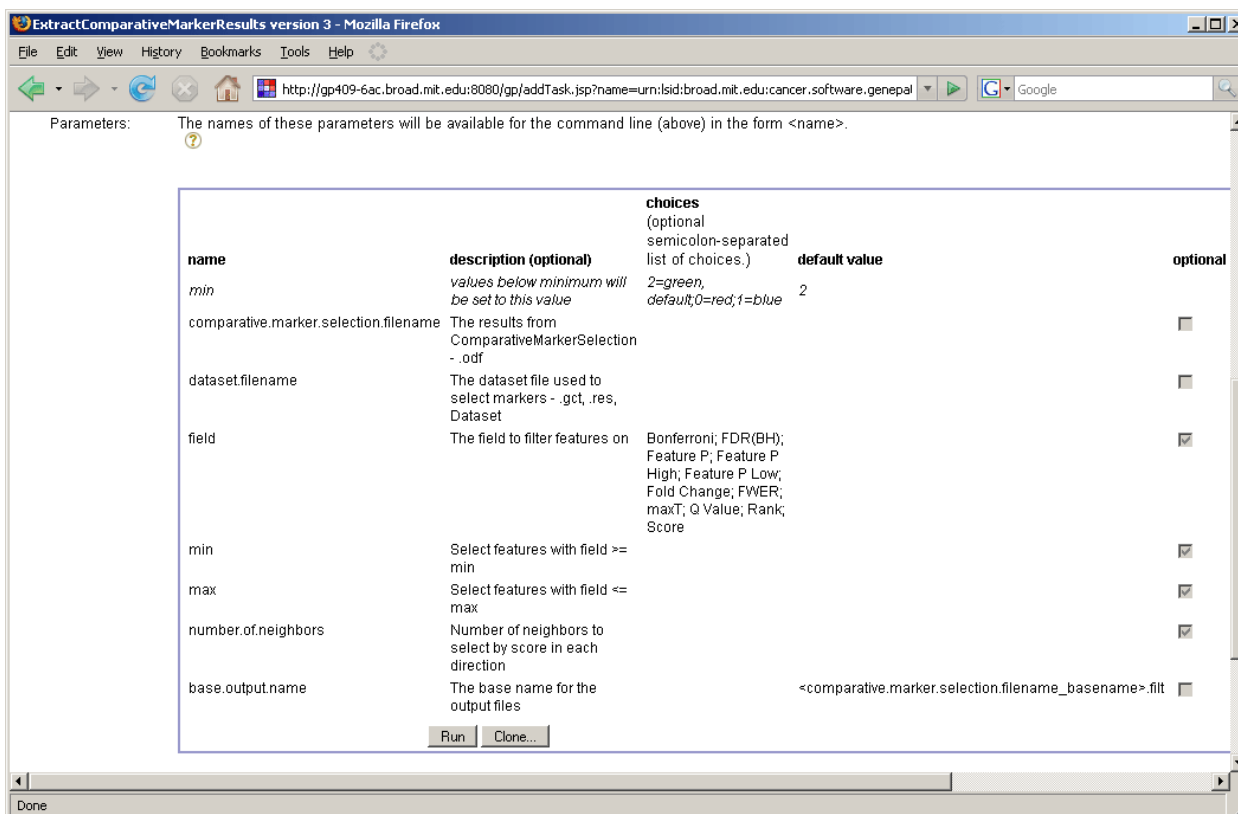
When you create/update a module, use this field to select all file formats used by the module. To select multiple file formats from the list, use ctrl-click or shift-click. If the module uses a file format not included in the list, click the *New* button to add that file format.

- **Current files:** Lists the support files packaged with the module.

When you create/update a module, you must specify all files used in your module, including scripts, libraries, property files, DLLs, executable programs, documentation, and so on. For more information, see [Adding Module Support Files](#).

- **Parameters:** Lists the module parameters.

When you create/update a module, you must define each parameter in the module command line. For more information, see [Defining the Module Parameters](#).



## Defining the Module Command Line

Defining the command line that launches your module is a critical piece of defining your GenePattern module. The text of the command line must be variable to address different parameter values, CPU platforms, and operating systems. To ensure that your module runs under different conditions, the command line that you enter on the module definition page for will be a combination of fixed and variable text. You specify the variable text in the form of substitution variables enclosed in angle brackets. GenePattern replaces the substitution variables with their assigned values before invoking the command.

Use substitution variables to reference:

- **Module parameters.** To [define module parameters](#), use the *Parameters* section of the module definition form. To reference a module parameter in the command line, enter the parameter name enclosed in angle brackets; for example, <filename>. Each parameter defined in the *Parameters* section must be mentioned in the command line, unless you have specified that it is an optional parameter. If you provide a default value for a parameter, the default value is used when a user invokes the module and fails to specify a value for the parameter.
- **Server configuration properties.** To set server configuration properties or create new server configuration properties, select *Administration>Server Settings* and edit the Custom settings, as described in [Modifying Server Settings](#). For a complete list of server configuration properties, see the GenePattern `/resources/genepattern.properties` file. To reference a configuration property in the command line, enter the property name enclosed in angle brackets; for example, <libdir>.
- **Java system properties.** For more information about these properties, see [Java system properties](#). To reference a Java system property in the command line, enter the property name enclosed in angle brackets.
- **Programming environment path names.** GenePattern provides three server configuration properties, <java> <perl> <R>, that resolve to the full path names for Java, Perl, and R. These are the three languages used to build the modules currently provided in the Broad repository. If your program is run by another interpreter, such as Python or LISP, you can use an absolute pathname (for example, `/usr/bin/python`) or create a configuration property for the interpreter pathname (for example, `python=/usr/bin/python`). If your program is an executable file, you do not need an interpreter, so do not need such a property. To create a server configuration property, select *Administration>Server Settings* and edit the Custom settings, as described in [Modifying Server Settings](#).

The following table lists the substitution variables for the most commonly used server configuration properties:

<java>	path to Java; usually the one running the GenePattern server
--------	--

<perl>	path to Perl, installed with GenePattern server on Windows, otherwise the one already installed on your system
<R>	path to a program that runs R and takes as input a script of R commands. R is installed with the GenePattern server on Windows and MacOS
<java_flags>	Java virtual machine configuration parameters (such as VM memory settings) from the <a href="#">Server Settings page</a> (Java Flag settings)
<libdir>	directory where the module's support files are stored
<job_id>	job number
<name>	name of the module being run
<filename_basename>	for each input file parameter, the filename without the file extension or directory  <b>Note:</b> In the property name, <i>filename</i> is the name of the input file parameter. For example, if you have an input file parameter named input.filename, the substitution property name is <input.filename_basename>. The next two properties are similar.
<filename_extension>	for each input file parameter, the extension without the filename or directory
<filename_file>	for each input file parameter, the input filename without the directory
<path.separator>	Java classpath delimiters (: or ;), useful for specifying a classpath for Java-based modules
<file.separator>	/ or \ for directory delimiter
<line.separator>	newline, carriage return, or both for line endings
<user.dir>	current directory where the job is executing
<user.home>	user's home directory
<parameter_name>	value of the named parameter; for example, if you have a parameter named arg1, use the substitution property <arg1> to include the value of that parameter on the command line

The following example uses substitution variables for two configuration properties, <java> and <libdir>, and one module parameter, <arg1>:

```
<java> -cp <libdir>mymodule.jar com.foo.MyModule <arg1>
```

To execute the module, GenePattern locates the Java runtime and asks to execute the MyModule class using code from the module support file mymodule.jar. The value of the arg1 parameter is passed as an argument to MyModule.

**Note:** if your module is designed to accept a standard input stream and/or write to a standard output stream, you can use redirection syntax when describing the command line. To redirect a file to the input stream, enter the text \< followed by the input file parameter. To redirect the standard output or standard error streams to a named file, enter the text \> or \>& followed by the name of the output file. In the following example, the LogTransform module reads its input from the standard input stream and writes its output to the standard output stream:

```
<perl> <libdir>log_transform.pl \< <input.filename> \> <output.file>
```

## Defining the Module Parameters

When you create/update a module, you must define each module parameter. GenePattern uses the definition that you supply to prompt users for input when they run your module. Use the Parameters section of the Create/Update module form to enter each parameter. Every parameter entered in this section must appear in the command line, unless you mark the parameter as optional. The following example shows the Parameters section for the Create/Update module form for ExtractComparativeMarkerResults:

Following are descriptions of each field. The first line of the Parameters section provides an example parameter.

- Enter the list of choices as a semi-colon delimited list. For each choice, specify the setting accepted by the program, an equal sign, and the text to be displayed. For example, enter this list of choices:

to create a drop-down list that contains the following choices:

Choices	Associated command-line component
Hierarchical clustering	hierarchical
Self-organizing map	SOM

Non-negative Matrix Factorization	NMF
nothing specified	
pi	3.14159265

- **Default value.** Specify a default value to be supplied on the module's command line in cases where the user does not supply a value when invoking the module. This is not the same as the module's own internal defaults. Instead, this allows the GenePattern module declaration author to create a default, even when none exists internally within the module.

Default values for parameters that have a choice list must be either blank or one of the values from the choice list. Any other setting results in an error message. If no default for a choice list is provided, the first entry on the list will be the default.

The default value may use substitution variables, just like the rest of the command line. So a valid default for an output file might be <input.filename\_basename>.foo, meaning that the output file will have the same stem as the input.filename parameter, but will have a .foo extension.

- **Optional.** Select the check box to indicate that the parameter is optional. Optional parameters do not have to be specified on the command line. When a user fails to enter a value for an optional parameter, nothing is added to the command line for that parameter.
- **Prefix when specified.** If an optional parameter requires a text prefix on the command line, use this field to specify the prefix. For example, you might need to write "-F *filename*" to pass in a filename; however, if the filename is optional, you do not necessarily want to specify "-F ". To solve this problem, set the prefix to "-F " (note that the ending space is included). If the optional filename is provided, then the prefix is also added.
- **Type.** Declaration of the type of an input parameter allows the GenePattern client to present the value to the user in the appropriate format. Parameter type choices are: text, integer, floating point, and input file. (If you specify an input file parameter, when a user runs the module and specifies the input file, GenePattern uploads the entire file to the server, not just the file name.)
- **File Format.** When you select a parameter type of input file, a drop-down list of file formats appears in the file format column. Select the valid file format(s) for this parameter. To select multiple file formats, use CTRL-click. If your module requires an input file format not included in the list, scroll back to the **File format(s)** field and click *New* to add that format to the list.

## Adding Module Support Files

When you create/update a module, you must specify all files used in the module, including scripts, libraries, property files, DLLs, executable programs, documentation, and so on. All files are copied to the GenePattern server and may be referenced in the command line field using the syntax: <libdir>filename. You can specify as many files as needed, provided you have the space available on your GenePattern server.

### Adding and Removing Support Files

To specify a file, use the Support files field:

1. Click the Browse button next to an entry field to select a file. The *Support files* field provides five (5) entry fields, so you can enter at most five files at a time.
2. Click *Save*, near the bottom of the form. GenePattern saves all changes you have made to the module since you last saved, adds the files to the module and copies them to the server, creates a new version of the module, and displays the status message:

Installation of your *ModuleName* module (version *x*) is complete.  
Run *ModuleName*

3. To continue editing the module, click your *ModuleName* in the first sentence. In the Create/Update Module form, the files that you added are now listed in the **Current files** field.
4. Repeat this procedure to add additional files.

To remove files that you have already added, use the **Current files** field:

1. Select the file from the *Current files* list.
2. Click *Delete*. Deleting a file from the Current files list deletes all changes since the last time you saved the module. The Web Client displays a warning message reminding you of this fact.



3. Click OK to delete the file (and any other changes made since the last time you saved).

### Including Documentation

Public modules should always include documentation. Although GenePattern requires no special format, the GenePattern team suggests PDF files for maximum portability. The module documentation should include instructions for using the module, a detailed description of each input parameter, a detailed description of each output file (both its format and content), and explain the algorithm or reference the paper, journal, or book that explains it.

The documentation that you provide with your module is automatically available to GenePattern users. As a GenePattern user, when you select a module in GenePattern client, the client displays a form that includes the module parameters and a *Help* button. When you click the *Help* button, GenePattern examines the list of support files for the module and displays the first file that has a standard documentation extension. If you have provided a documentation file with your module, that file is displayed as the help file; if you have not provided a documentation file, GenePattern displays a message indicating the no information is available. (By default, the standard documentation file extensions are .doc, .pdf, .txt, .html, and .rdf. You can modify this list of extensions by editing the files.doc property in the GenePattern `/resources/genepattern.properties` file.)

## Creating Modules

Creating a GenePattern module is a two-step process:

1. Write a program that executes the desired function. You can write the program in the language of your choice; for example, you can use a compiled language, such as C, to create an executable or use a scripting language, such as Perl, to create a script that is run by an interpreter. The [GenePattern Programmer's Guide](#) provides guidelines for writing programs that will be run as GenePattern modules.
2. Use the GenePattern Web Client to create a GenePattern module that invokes the program that you have written. It takes just a few minutes to enter the necessary information. Once you have done so, you can run the module and share it with others. You can decide which parameters from the algorithm to expose to the user and can replace command line parameter names that are hard to remember with names that are self-explanatory. You can also create drop-down list choices for parameters to reduce the possibility of invoking the module with incorrect values.

To create a module that invokes the program that you have written (or otherwise obtained):

1. Click *Modules & Pipelines>New Module*. The Web Client displays the Create Module form:



The screenshot shows the 'add GenePattern module - Mozilla Firefox' window. The address bar contains 'http://gp409-6ac.broad.mit.edu:8080/gp/addTask.jsp'. The GenePattern logo is at the top left, and 'My Settings | Sign Out gpuser' is at the top right. A navigation bar includes 'Modules & Pipelines', 'Suites', 'Job Results', 'Resources', 'Downloads', 'Administration', and 'Help'. The main section is titled 'Create Module'. It contains several form fields: 'Name:\*' (required, no spaces), 'LSID:', 'Description:', 'Author:' (with a hint '(name, affiliation)'), 'Privacy:' (set to 'private'), 'Quality level:' (set to 'development'), 'Command line:\*' (a large text area), 'Module Category:' (with a 'New...' button), 'CPU type:' (set to 'any'), 'Operating system:' (set to 'any'), 'Language:' (set to 'any') with a 'min. language version:' field, 'Version comment:', 'File format(s):' (with a list box containing 'Actual Peaks', 'Actual Peaks to EM-matched Peaks', and 'adj', and a 'New...' button), and 'Support files:' (with a text description and a 'Browse...' button). The status bar at the bottom says 'Done'.

2. Enter values for the fields. For descriptions of the fields, click Help or see the [Module Definition Page](#).
3. Click Save to create the module. GenePattern checks the following:
  - Every parameter not marked as optional is included in the command line.
  - Every parameter in the command line parameter is a parameter, environment variable, or system property.
  - Module parameter names are valid.

If no errors are found, GenePattern copies the support files to the server and makes the module available to the GenePattern clients.

**Note:** By adding a module, a user can execute arbitrary code on the GenePattern server. Because arbitrary code may include malicious code, take precautions to protect your server: for example, employ virus scanner software and restrict access to appropriately privileged (non-root) users. For more information about securing your server, see [Securing the Server](#).

## Tutorial: Creating a Module

Following is a brief tutorial that creates a module named `log_transform`. The program invoked by this module is a Perl script, `gp_tutorial_files/log_transform/log_transform.pl`, which log-transforms all positive values in a data set and sets all negative or zero values to zero. This Perl script is part of the GenePattern tutorial data set, which is downloaded during a full installation of GenePattern and is also available on the [GenePattern web site](#).

To create the `log_transform` module:

1. Click *Modules & Pipelines* > *New Module*. The Web Client displays the Create Module form. (The module definition form, with the tutorial data filled in, appears at the end of the procedure.)
2. Enter the following information in the first few fields:
  - **Name:** LogTransform
  - **Description:** Log transform a data (gct) file.

- **Author:** Your name and affiliation.
  - **Privacy:** Select private. This means that only you (or an administrator) can view and run the module. Public allows all users connected to this server to view and run the module.
  - **Quality level:** Select preproduction. This indicates that you have finished development, but are not yet ready for production.
3. Enter the following text in the *Command line* field (if you are working online, cut and paste this text into the field):
- ```
<perl> <libdir>log_transform.pl -F <input.filename> -o <output.file>
```
- Typically, you enter the command line as a combination of fixed text and variables defined by GenePattern. This allows the command line to be independent of the operating environment and allows different values to be specified at different invocations of the command. This command line uses the following variables:
- `<perl>` represents the full path to the Perl installation used by GenePattern.
  - `<libdir>` represents the full path to the directory that contains the files for this module, including the program file.
  - The Perl script, `log_transform.pl`, expects two parameters, an input file and an output file name: `-F <input.filename> -o <output.file>`. When your program has parameters, you include them in the command line and also define them in the *Parameters* field, as described below.
4. Enter the following information in the next few fields:
- **Module category:** Preprocess & Utilities
  - **CPU type:** Any
  - **Operating system:** Any
  - **Language:** Perl
  - **Version comment:** Initial version
  - **File format(s):** Select gct from the list of formats.
5. Use the *Support files* field to upload your program, any other files needed to execute the module, and documentation for the module (a file with a text extension, such as .pdf, .doc, or .txt):
1. Click the *Browse* button at the end of the first line in the *Support files* field. The Web Client displays the File Upload window.
  2. Navigate to the `gp_tutorial_files/log_transform` directory, select the file `log_transform.pl`, and click *Open*. This is the script that implements the module.
  3. Click the *Browse* button at the end of the second line in the *Support files* field. The Web Client displays the File Upload window.
  4. Select the file `LogTransform.pdf` and click *Open*. This is the documentation for the module. When a GenePattern user displays your module and clicks the *Help* button, GenePattern displays the support file that has a text extension (such as .pdf, .doc, or .txt).
6. Use the *Parameters* field to describe your two program parameters: `input.filename` and `output.file`. The parameter names and descriptions that GenePattern displays when a user runs your module are the parameter names and descriptions that you provide here.

In the first row of the *Parameters* field, enter the following information for `input.filename`:

- **name:** `input.filename`
- **description:** The dataset to be transformed (gct format).
- **type:** choose input file
- **file format:** choose gct

In the second row, enter the following information for `output.file`:

- **name:** `output.file`
- **description:** The name of the new transformed file.
- **type:** choose text (this is a name that the user enters)

- 
- add GenePattern module - Mozilla Firefox
- File Edit View History Bookmarks Tools Help
- http://gp409-6ac.broad.mit.edu:8080/gp/addTask.jsp
- Google
- GenePattern** My Settings | Sign Out gpuser
- Modules & Pipelines Suites Job Results Resources Downloads Administration Help
- ### Create Module
- \* required field ?
- Name:\*  \* (required, no spaces) ?
- LSID:
- Description:  ?
- Author:  (name, affiliation) ?
- Privacy:  ?
- Quality level:  ?
- Command line:\*  ?
- Module Category:  New... ?
- CPU type:  (if compiled for a specific one) ?
- Operating system:  (if operating system-dependent) ?
- Language:  min. language version:  ?
- Version comment:  ?
- File format(s): ?  New...
- Support files: The actual program plus any required libraries will be accessible to your command line as <libdir>filename ?
- |                                                                       |           |
|-----------------------------------------------------------------------|-----------|
| C:\gp_tutorial_files\gp_tutorial_files\log_transform\log_transform.pl | Browse... |
| C:\gp_tutorial_files\gp_tutorial_files\log_transform\LogTransform.pdf | Browse... |
|                                                                       | Browse... |
|                                                                       | Browse... |
|                                                                       | Browse... |
- Done

## Editing Modules

1. Click *Modules & Pipelines* to display the Web Client start page.

2. Display a module definition form in one of the following ways:
  - Select a module that you created. When the Web Client displays the module parameters, click *Edit*.
  - Select a public module. When the Web Client displays the module parameters, click *Properties*. On the module definition form, click *Clone* to create a copy of the module. You created the copy, so you can edit it.
3. Edit the module definition form. Click *Help* for descriptions of the fields.
4. Click *Save* to create a new version of the module.

## Working with Pipelines

A GenePattern pipeline defines a sequential series of modules to be run. Modules run from a pipeline work exactly the same as those run directly from the Web Client. If you are unfamiliar with GenePattern pipelines, see the [Concepts Guide](#).

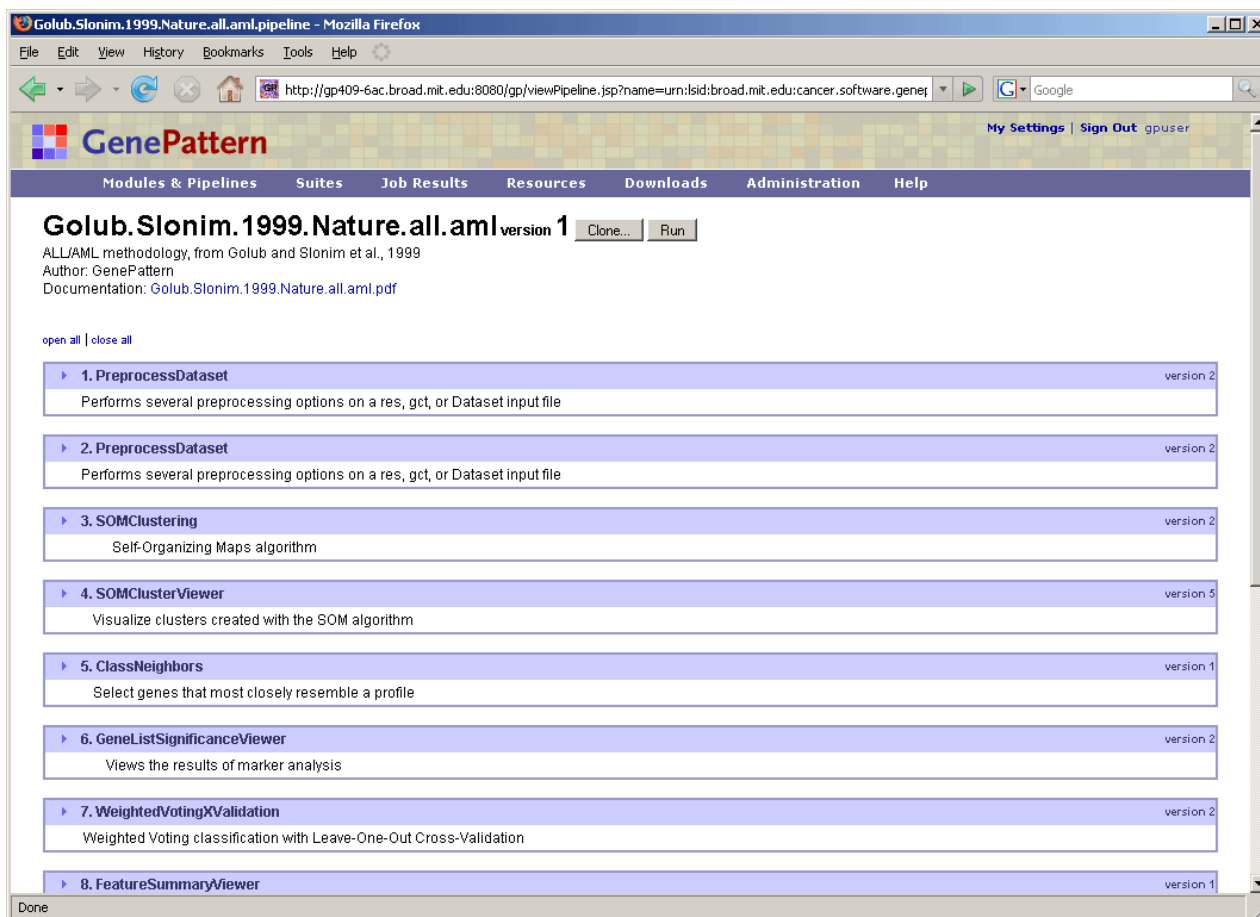
The following table summarizes the different ways you can work with GenePattern pipelines.

|                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|---------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Run a pipeline                        | Select a pipeline, enter its parameters and click <i>Run</i> . For more information, see <a href="#">Running Modules and Pipelines</a> .                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| Display pipeline definitions          | A pipeline's definition lists the pipeline's author, the modules to be run and their parameters. To display the pipeline definition, click <i>Modules &amp; Pipelines</i> and select the pipeline. When the Web Client displays the pipeline parameters, click <i>Properties</i> . The Web Client displays the <a href="#">pipeline definition page</a> .                                                                                                                                                                                                                       |
| Share pipelines with other users      | <p>Zip files provide a convenient way to share pipelines with other GenePattern users.</p> <ul style="list-style-type: none"> <li>• To export a pipeline to a zip file, click <i>Modules &amp; Pipelines</i> and select the pipeline to export. When the pipeline parameters (if any) appear in the center pane, click <i>Export</i>.</li> <li>• To install a pipeline from a zip file: click <i>Modules &amp; Pipelines&gt;Install from zip</i>.</li> </ul> <p>For more information, see <a href="#">Exporting and Installing Modules &amp; Pipelines Using Zip Files</a>.</p> |
| Install pipelines from the repository | The Broad Institute maintains a repository of modules, pipelines, and suites. To install pipelines from the Broad repository, click <i>Modules &amp; Pipelines&gt;Install from Repository</i> . For more information, see <a href="#">Installing Modules &amp; Pipelines from the Repository</a> .                                                                                                                                                                                                                                                                              |
| Create pipelines                      | You can create an empty pipeline and add modules to it, or you can start with an analysis results file and have GenePattern create a pipeline that recreates that analysis results file. For more information, see <a href="#">Creating Pipelines</a> .                                                                                                                                                                                                                                                                                                                         |
| Edit pipelines                        | You can edit a pipeline that you have created or clone a public pipeline and edit your copy of the public pipeline. For more information, see <a href="#">Editing Pipelines</a> .                                                                                                                                                                                                                                                                                                                                                                                               |
| Delete pipelines                      | To delete a pipeline from your GenePattern server, click <i>Modules &amp; Pipelines&gt;Manage</i> . For more information, see <a href="#">Managing Modules &amp; Pipelines</a> .                                                                                                                                                                                                                                                                                                                                                                                                |

## Pipeline Definition Page

To display the definition of a pipeline:

1. Click *Modules & Pipelines* to display the Web Client start page.
2. Select the pipeline to display.
3. When the Web Client displays the pipeline parameters, click *Properties*. The Web Client displays the pipeline definition page, as shown below.



On the pipeline definition page, you can:

- Click the icon next to a module to show/hide the module parameters.
- Click *open all* and *close all* to show and hide all module parameters.
- Click *Clone* to create (and edit) your own copy of this pipeline.
- Click *Run* to run the pipeline.
- Click *Edit* to edit the pipeline. The *Edit* button, not shown here, is only available for pipelines that you have created.

## Creating Pipelines

To create a pipeline:

1. Display a pipeline definition form in one of the following ways:
  - To create an empty pipeline, click *Modules & Pipelines>New Pipeline*. The Web Client displays an empty pipeline definition form, as shown in the first figure below.
  - To create a pipeline based on an analysis results file, click the icon next to the analysis results file and select *Create Pipeline* from the pop-up menu. The Web Client displays a populated pipeline definition form, as shown in the second figure below.

GenePattern adds modules to the pipeline based on the following logic: add the module that created the results file; check the module's input file parameters; if the input file for the module was the output file of a previous module, add the previous module; check that module's input file parameters; continue to walk back through the chain of modules, adding modules to the pipeline, until you reach the initial input file.

  - To create a copy of an existing pipeline, click *Modules & Pipelines* and select the pipeline. When the Web Client displays the pipeline parameters, click *Properties*. On the pipeline definition form, click *Clone* to create a copy of the pipeline. The Web Client displays a populated pipeline definition form, as shown in the second figure below.
2. Complete the pipeline definition by editing the pipeline definition form, as described in [Editing Pipelines](#).

3. Click Save to create the pipeline.

When you create an empty pipeline, the Web Client displays an empty pipeline definition form:

The screenshot shows the 'new pipeline - Mozilla Firefox' window. The address bar shows the URL: <http://gp409-6ac.broad.mit.edu:8080/gp/pipelineDesigner.jsp>. The GenePattern logo is at the top left, and navigation links 'My Settings | Sign Out gpuser' are at the top right. The main menu includes 'Modules & Pipelines', 'Suites', 'Job Results', 'Resources', 'Downloads', 'Administration', and 'Help'. The 'Pipeline Designer' section contains the following form fields:

- Pipeline name\*:
- Description:
- Author:  name, affiliation
- Privacy:
- Version comment:
- Documentation:  - LSID:
- Module Category & Name:
-

At the bottom, there are links 'About GenePattern | Contact Us' and a copyright notice '©2003-2007 Broad Institute, MIT'.

When you create a pipeline from a results file or by cloning, the Web Client displays a populated pipeline definition form:

The screenshot shows the 'Job11 - Mozilla Firefox' window. The address bar shows the URL: <http://gp409-6ac.broad.mit.edu:8080/gp/pipelineDesigner.jsp?name=urn%3AIsid%3A8080.gpuser.Gp409-6ac>. The GenePattern logo is at the top left, and navigation links 'My Settings | Sign Out gpuser' are at the top right. The main menu includes 'Modules & Pipelines', 'Suites', 'Job Results', 'Resources', 'Downloads', 'Administration', and 'Help'. The 'Pipeline Designer - Job11 version 1' section contains the following form fields:

- Pipeline name\*:
- Description:
- Author:  name, affiliation
- Privacy:
- Version comment:
- Documentation:  - LSID: urn:lsid:8080.hkuehn.Gp409-6ac.broad.mit.edu:genePatternmodules:4:1

Below the form, there is a section for '1. ComparativeMarkerSelection: version 3.9 (latest)'. It includes a description: 'Computes significance values for features using several metrics, including FDR(BH), Q-Value, maxT, FWER, Feature-Specific P-Value, and Bonferroni.' and a link to the documentation: [ComparativeMarkerSelection.pdf CompMarkSelAppNote.pdf](#). The table below lists the parameters:

| prompt when run          | parameter name                 | value                                                                                                                                                                                                                    | description                                         |
|--------------------------|--------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------|
| <input type="checkbox"/> | input file:                    | <input type="text" value="&lt;GenePatternURL&gt;getfile.jsp?task=&lt;LSID&gt;&amp;file=all_train.res"/> <input <="" td="" type="button" value="Browse..."/> <td>The input file - .res, .gct, .odf type=Dataset</td>      | The input file - .res, .gct, .odf type=Dataset      |
| <input type="checkbox"/> | cls file:                      | <input type="text" value="&lt;GenePatternURL&gt;getfile.jsp?task=&lt;LSID&gt;&amp;file=all_train.cls"/> <input <="" td="" type="button" value="Browse..."/> <td>The class file - .cls</td>                               | The class file - .cls                               |
| <input type="checkbox"/> | confounding variable cls file: | <input type="text" value="&lt;GenePatternURL&gt;getfile.jsp?task=&lt;LSID&gt;&amp;file=all_train.cls"/> <input <="" td="" type="button" value="Browse..."/> <td>The class file containing the confounding variable.</td> | The class file containing the confounding variable. |
| <input type="checkbox"/> | test direction:                | <input type="text" value="2 Sided"/>                                                                                                                                                                                     | The test to perform                                 |

## Editing Pipelines

To edit a pipeline:

1. Click *Modules & Pipelines* to display the Web Client start page.
2. Select the pipeline that you want to edit. The Web Client displays the pipeline parameters form.
3. Display a pipeline definition form in one of the following ways:

- Click *Edit* (at the top of the form). The *Edit* option appears only if you created the pipeline and the LSID authority of the pipeline matches the LSID authority of your GenePattern server.
- Click *Properties*. When the Web Client displays the pipeline definition form, click *Clone* (at the top of the form) to create a copy of the pipeline. You create a copy of the pipeline and GenePattern assigns it the LSID authority of your GenePattern server; therefore, you can edit that copy of the pipeline.

4. Edit the pipeline definition, as described below.

5. Click *Save* to create a new version of the pipeline.

The remaining topics in this section describe how to edit the pipeline definition:

- [Editing the Pipeline Description](#)
- [Adding Modules](#)
- [Removing Modules](#)
- [Reordering Modules](#)

## Editing the Pipeline Description

The fields at the top of the pipeline definition form describe the pipeline:

The screenshot shows the GenePattern Pipeline Designer interface in a Mozilla Firefox browser. The page title is 'copyOfGolub.Slonim.1999.Nature.all.aml - Mozilla Firefox'. The address bar shows the URL: 'http://localhost:8080/gp/pipelineDesigner.jsp?name=urn:lsid:8080.gpuser.Gp409-6ac.broad.mit.edu:genepatternmodules:5:1'. The GenePattern logo is at the top left, and navigation links like 'Modules & Pipelines', 'Suites', 'Job Results', 'Resources', 'Downloads', 'Administration', and 'Help' are at the top. The main content area is titled 'Pipeline Designer - copyOfGolub.Slonim.1999.Nature.all.aml version 1'. It contains several form fields: 'Pipeline name\*' (copyOfGolub.Slonim.1999.Nature.all.aml), 'Description' (ALL/AML methodology, from Golub and Slonim et al., 1999), 'Author' (GenePattern), 'Privacy' (private), 'Version comment' (1.0), and 'Documentation' (Golub Slonim.1999.Nature.all.aml.pdf). Below these fields is a table for '1. PreprocessDataset:' with columns for 'prompt when run', 'parameter name', 'value', and 'description'. The table lists parameters like 'input filename', 'output file', 'output file format', 'filter flag', and 'preprocessing flag' with their respective values and descriptions.

- Pipeline name:** The name of the pipeline. When naming pipelines, note the following:
  - Pipeline names can include alphanumeric characters, periods (.), and underscores (\_).
  - Pipeline names must not include spaces or special characters such as: exclamation points (!), at signs (@), pound signs (#), dollar signs (\$), percent signs (%), carets (^), ampersands (&), and asterisks (\*).
  - For cross-platform compatibility, avoid the following names: con, prn, aux, nul, com1, com2, com3, com4, lpt1, lpt2, and lpt3. Machines running Windows cannot accept files with these names, regardless of the file extension.
  - GenePattern does not prevent you from using the same name for multiple pipelines; however, using unique names is strongly recommended.
- Version:** To the right of the name field is a drop-down list of versions. By default, you are editing the most recent version of the pipeline. To edit a different version, select that version from the drop-down list.

Editing a pipeline creates a new version of the pipeline; it does not change the existing version of the pipeline. If you are unfamiliar with GenePattern versioning, see the [Concepts Guide](#).

- Description:** A brief description of the pipeline, which is displayed when a user runs the pipeline or views the pipeline definition.

- **Author:** The author's name and affiliation (company or academic institution). This is a comment-only field. If you share this pipeline, the author field allows other users to credit the author and to contact the author with questions, suggestions, or enhancement ideas.
- **Privacy:** Select Private (default) or Public. A private pipeline can be accessed only by the person who created or installed the pipeline (or by an administrator); a public pipeline can be seen and run by all users.
- **Version comment:** A brief description of this version. When GenePattern clients display a drop-down list of versions, the comments for each version are visible in the drop-down list.
- **Documentation:** The full path name of the file that contains the pipeline documentation. Documentation is strongly encouraged for public pipelines. Click *Browse* to select the (previously created) pipeline documentation file.
- **LSID:** The Life Science Identifiers (LSIDs) for this pipeline. You cannot create or edit LSIDs. The GenePattern server automatically assigns an LSID to each version of a pipeline. If you are unfamiliar with GenePattern versioning, see the [Concepts Guide](#).

## Adding Modules

Each module in a pipeline has its own definition; for example:

To add a module:

1. Save changes that you have made to the pipeline.

When you add a new module in the middle of a pipeline, file name parameters for subsequent modules may be lost if they have not yet been saved. Saving the pipeline avoids this potential problem.

2. Select the *Add Another Module* button. The Web Client starts a new module definition **after** the current module:

**Note:** The only module in a new empty pipeline definition is the start of the first module.

3. Select a category of module from the *Category* list. The Web Client displays a drop-down list of modules in that category:

4. Select a module from the drop-down list of modules. The Web Client displays the definition form for that module:



1. SOMClustering:
version 2

Self-Organizing Maps algorithm  
documentation: [SOMClustering.pdf](#)

| prompt when run          | parameter name    | value                                                             | description                                                                                                                                                                                                                                                                         |
|--------------------------|-------------------|-------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/> | dataset filename: | <input type="text"/> <span style="float: right;">Browse...</span> | Dataset (res, gct, or odf dataset)                                                                                                                                                                                                                                                  |
| <input type="checkbox"/> | output stub:      | <input type="text" value="cdataset.filename_base"/>               | The base output file name - .odf                                                                                                                                                                                                                                                    |
| <input type="checkbox"/> | cluster range:    | <input type="text" value="2:3"/>                                  | Range of values can be entered and the program will automatically run the algorithm for each number of clusters in the range.                                                                                                                                                       |
| <input type="checkbox"/> | seed range:       | <input type="text" value="42"/>                                   | The seed for the random number generator, is exposed to allow the user to recreate a given session at a later time, (as opposed to allowing the program to generate a random initial seed which could result in different outcomes even if all the other parameters are identical). |
| <input type="checkbox"/> | iterations:       | <input type="text" value="50000"/>                                | How many times the algorithm should try to refine the clusters. Initially, this value can be set low for faster exploration, but should be set high (e.g., 50,000) for good convergence.                                                                                            |
| <input type="checkbox"/> | cluster by:       | <input type="text" value="rows"/>                                 | Whether to cluster by rows or columns                                                                                                                                                                                                                                               |
| <input type="checkbox"/> | som rows:         | <input type="text" value="0"/>                                    | Setting this and som.cols to a non zero value will override cluster.range and the computation will be for the specified geometry                                                                                                                                                    |
| <input type="checkbox"/> | som cols:         | <input type="text" value="0"/>                                    | Setting this and som.rows to a non zero value will override cluster.range and the computation will be for the specified geometry                                                                                                                                                    |
| <input type="checkbox"/> | initialization:   | <input type="text" value="Random_Vectors"/>                       | How to select initial random centroids                                                                                                                                                                                                                                              |
| <input type="checkbox"/> | neighborhood:     | <input type="text" value="Bubble"/>                               | Neighborhood function determining how centroids near the target centroid are updated                                                                                                                                                                                                |
| <input type="checkbox"/> | alpha initial:    | <input type="text" value="0.1"/>                                  | Initial learning weight for centroid updates                                                                                                                                                                                                                                        |
| <input type="checkbox"/> | alpha final:      | <input type="text" value="0.005"/>                                | Final learning weight for centroid updates                                                                                                                                                                                                                                          |
| <input type="checkbox"/> | sigma initial:    | <input type="text" value="5.0"/>                                  | Initial sigma to determine update neighborhood size around the target centroid                                                                                                                                                                                                      |
| <input type="checkbox"/> | sigma final:      | <input type="text" value="0.5"/>                                  | Final sigma to determine update neighborhood size around the target centroid                                                                                                                                                                                                        |

Add Another Module
Delete SOMClustering

5. At the top of the module definition form is the module name, followed by a drop-down list of versions. By default, you are using the most recent version of the module. To have the pipeline run a different version of the module, select that version from the drop-down list. A version is included in the drop-down list only if that version of the module is installed on the GenePattern server.
6. Enter the parameter values for the module. For most parameters, you enter a value, select a value from a drop-down list, or use the default value supplied by GenePattern. For input file parameters, select a file or use an output file from a previous module:
  - To select a file, click the *Browse* button to the right of the parameter field. The Web Client displays the Upload File window. Select a file and click *Open*. The selected file becomes the value for that parameter field.
  - To have the input file be an output file generated by a previous module, choose a previous module from the *use output from* drop-down list. If the selected module generates multiple output files, the Web Client displays a drop-down list of the generated output files. Select the output file that you want to use.

To have GenePattern prompt the user for a parameter value when the pipeline is run:

1. Select the *Prompt when run* box to the left of the parameter field. The Web Client removes the parameter value from the pipeline definition.

1. ComparativeMarkerSelection:
version 4

Computes significance values for features using several metrics, including FDR(BH), Q-Value, maxT, FWER, Feature-Specific P-Value, and Bonferroni.  
documentation: [ComparativeMarkerSelection.pdf](#) [CompMarkSelAppNote.pdf](#)

| prompt when run                     | parameter name                 | value                                                             | description                                         |
|-------------------------------------|--------------------------------|-------------------------------------------------------------------|-----------------------------------------------------|
| <input checked="" type="checkbox"/> | input file:                    | <a href="#">set prompt when run display settings...</a>           | The input file - .res, gct, .odf type=Dataset       |
| <input checked="" type="checkbox"/> | cls file:                      | <a href="#">set prompt when run display settings...</a>           | The class file - .cls                               |
| <input type="checkbox"/>            | confounding variable cls file: | <input type="text"/> <span style="float: right;">Browse...</span> | The class file containing the confounding variable. |

2. Optionally, click *set prompt when run display settings* to modify the default text of the prompt.

The pipeline definition form has no mechanism for adding a module **before** an existing module; therefore, to add a module at the **beginning** of the pipeline: add the new\_first module below the old\_first module, delete the old\_first module, and (if necessary) recreate the old\_first module below the new\_first module.

## Removing Modules

To remove a module, click the delete button for that module definition:

4. SOMClusterViewer:
version 5

Visualize clusters created with the SOM algorithm  
documentation: [SOMClusterViewer.pdf](#)

| prompt when run          | parameter name        | value                                                         | description             |
|--------------------------|-----------------------|---------------------------------------------------------------|-------------------------|
| <input type="checkbox"/> | som cluster filename: | <input type="text"/> <input type="button" value="Browse..."/> | SOM Properties ODF file |
|                          | or use output from    | 3 SOMClustering                                               | 1st output              |

## Reordering Modules

The pipeline definition form has no mechanism for reordering module. You must delete the module that you want to move and recreate it in its new position.

## Passing Parameters to a Pipeline

When you add a module to a pipeline, you specify the parameter values for that module. Optionally, you can have GenePattern prompt the user for one or more parameter values when the pipeline is run. As described in [Adding Modules](#), you select the *Prompt when run* box to the left of a parameter and, optionally, click *set prompt when run display settings* to modify the description of the parameter. GenePattern prompts the user for these parameter values each time the pipeline is run.

Occasionally, a pipeline requires that the same input file be specified for multiple parameters. For example, consider the following scenario:

1. Run ComparativeMarkerSelection. Prompt the user for the *input filename* parameter (an expression dataset file) and the *cls filename* parameter (a class file).
2. Run ExtractComparativeMarkerSelection. Set the *comparative marker selection filename* parameter to the output file from ComparativeMarkerSelection and prompt the user for the *dataset filename* parameter (the same expression dataset file used for ComparativeMarkerSelection).

Ideally, you want to prompt the user for an expression dataset file and use that file for both the *input filename* (ComparativeMarkerSelection) and *dataset filename* (ExtractComparativeMarkerSelection) parameters. To do that, add the ConvertLineEndings module to the start of your pipeline:

1. Run ConvertLineEndings. Prompt the user for the *input filename* parameter. ConvertLineEndings takes an input file and creates an output file that is similar (it converts the line endings in the file to those used by perl on the host operating system).
2. Run ComparativeMarkerSelection. Set the *input filename* parameter to the output file from ConvertLineEndings and prompt the user for the *cls filename* parameter (a class file).
3. Run ExtractComparativeMarkerSelection. Set the *comparative marker selection filename* parameter to the output file from ComparativeMarkerSelection and set the *dataset filename* to the output file from ConvertLineEndings.

## Working with Suites

Use suites to group modules and pipelines into packages that have related functionality; for example, you might create a suite that contains the modules that you most commonly use to analyze new data files. Suites help you to organize and work with modules and pipelines. If you are unfamiliar with GenePattern suites, see the [Concepts Guide](#).

The following table summarizes the different ways you can work with GenePattern suites.

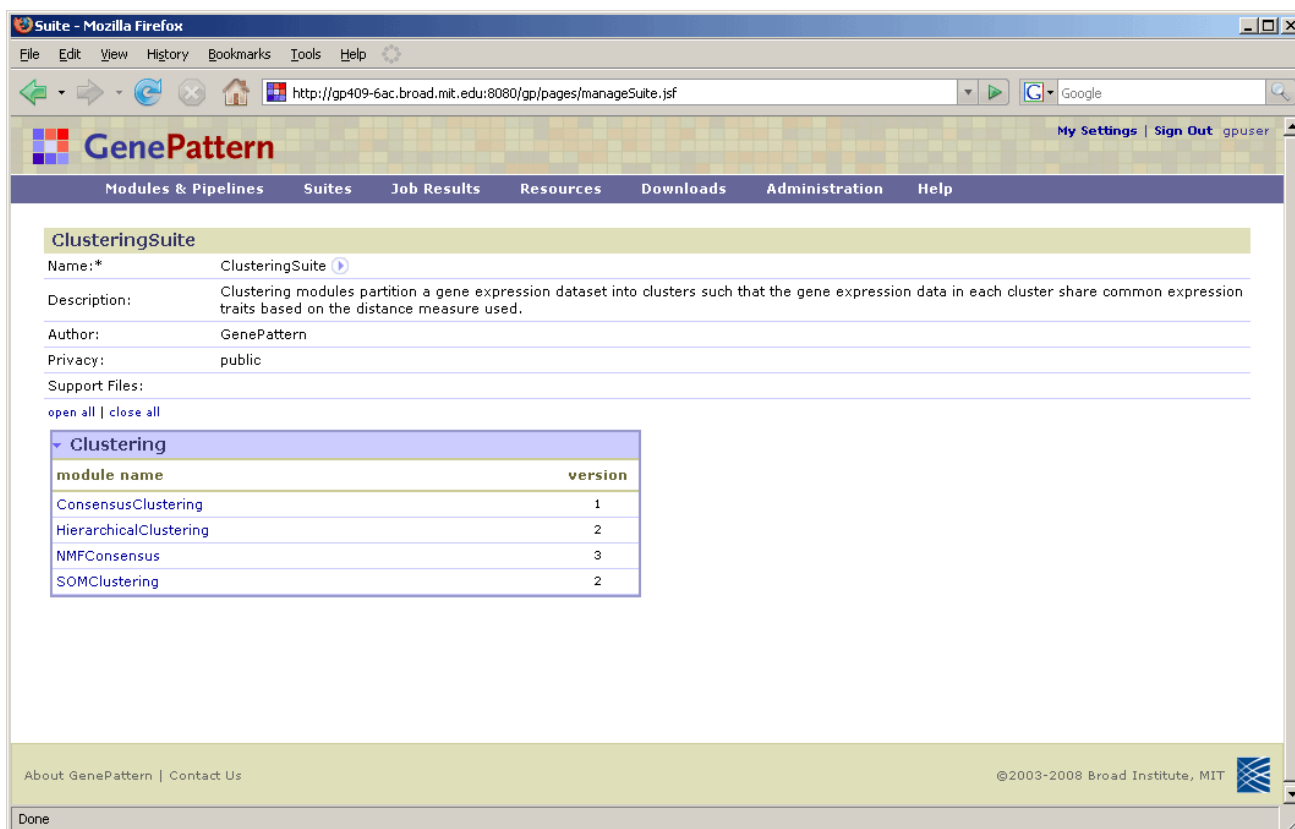
|                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Display suite definitions          | <p>To display the suite definition:</p> <ol style="list-style-type: none"> <li>1. Click <i>Suites&gt;Manage</i>.</li> <li>2. Click the name of the suite that you are interested in.</li> </ol>                                                                                                                                                                                                                                                                                                                                                      |
| Share suites with other users      | <p>Zip files provide a convenient way to share suites with other GenePattern users.</p> <ul style="list-style-type: none"> <li>• To export a suite to a zip file, click <i>Suites&gt;Manage</i>, click the icon next to a suite to display the suite menu, and select <i>Export Including dependents</i> or <i>Export excluding dependents</i>.</li> <li>• To install a suite from a zip file: click <i>Suites&gt;Install from zip</i>.</li> </ul> <p>For more information, see <a href="#">Exporting and Installing Suites Using Zip Files</a>.</p> |
| Install suites from the repository | <p>The Broad Institute maintains a repository of modules, pipelines, and suites. To install suites from the Broad repository, click <i>Suites&gt;Install from Repository</i>. For</p>                                                                                                                                                                                                                                                                                                                                                                |

|               |                                                                                                                                                                      |
|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|               | more information, see <a href="#">Installing Suites from the Repository</a> .                                                                                        |
| Create suites | To create a suite, click <i>Suites&gt;New</i> . For more information, see <a href="#">Creating Suites</a> .                                                          |
| Edit suites   | You can edit a suite that you have created or copy a public suite and edit your copy of the public suite. For more information, see <a href="#">Editing Suites</a> . |
| Delete suites | To delete a suite from your GenePattern server, click <i>Suites&gt;Manage</i> . For more information, see <a href="#">Managing Suites</a> .                          |

## Suite Definition Page

To display the definition of a suite:

1. Click *Suites>Manage*.
2. Click the name of the suite that you want to display. The Web Client displays the suite definition page, as shown below.



- **Name:** The name of the suite.
- **Description:** A brief description of the suite.
- **Author:** The author's name and affiliation (company or academic institution). This is a comment-only field.
- **Privacy:** Select Private (default) or Public. A private suite may be accessed only by the person who created or installed the suite (or by an administrator); a public suite can be accessed by all users.
- **Support files:** Any files included with the suite. Public suites generally include documentation.
- **Modules & Pipelines:** The remainder of the definition form lists all of the modules and pipelines that are on your GenePattern server by category. The ClusteringSuite in this figure contains modules from only one category, the Clustering category.

From the suite definition page, you can:

- Click the icon next to the suite name to display the [suite menu](#).
- Click *open all* and *close all* to show and hide the modules/pipelines in all categories.

- Click the icon next to a category to show/hide the modules/pipelines in that category.

## Creating Suites

To create a suite:

1. Click *Suites>New*. The Web Client displays an empty suite definition form.
2. Define the suite by entering values for each field (see [Suite Definition Page](#)). To add modules and pipelines to the suite:
  - Select the check box next to a module/pipeline to add that module/pipeline to the suite.
  - Select the check box next to a category name to add all modules and pipelines in that category to the suite.
3. Click Save. GenePattern creates the suite.

The screenshot shows the 'Create GenePattern Suite' form. The 'Name' field is empty. The 'Description' field is empty. The 'Author' field is 'gpuser'. The 'Privacy' dropdown is 'all users'. There are three 'Browse...' buttons for 'Support Files'. The 'Annotation' section is expanded, showing 'GeneCruiser' selected. The 'Projection' section is expanded, showing 'NMF' and 'PCA' selected. The 'Clustering' section is expanded, showing 'ConsensusClustering', 'HierarchicalClustering', and 'KMeansClustering' selected. The 'Proteomics' section is expanded, showing 'AreaChange' and 'LocatePeaks' selected.

## Editing Suites

To view or edit a suite:

1. Click *Suites>Manage*. The Web Client displays the Manage Suites page, which lists the suites on your GenePattern server.
2. To view the suite definition, click the suite name.
3. To edit the suite definition:
  1. Click the icon next to the suite name and select *Edit*. This option is available only if you created the suite.
  2. Edit the suite definition.
  3. Click Save to update the suite definition.

## Managing Modules, Pipelines, and Suites

An analysis module runs a single analysis. A pipeline runs a series of analysis modules. Suites group modules and pipelines into packages that have related functionality, which helps you to organize and work with modules and pipelines. If you are unfamiliar with GenePattern modules, pipelines and suites, see the [Concepts Guide](#).

- [Installing Modules & Pipelines from the Repository](#)

- [Exporting and Installing Modules & Pipelines Using Zip Files](#)
- [Managing Modules & Pipelines](#)
- [Installing Suites from the Repository](#)
- [Exporting and Installing Suites Using Zip Files](#)
- [Managing Suites](#)

## Installing Modules & Pipelines from the Repository

The Broad Institute maintains a repository of modules and pipelines that are freely available to the public. To install these modules and pipelines on your GenePattern server:

1. Click *Modules & Pipelines>Install from Repository*. The Web Client displays the Install Modules & Pipelines from Repository page, as shown below.
2. Select the modules and pipelines to install.

The search fields at the top of the page allow you to select any version of a module that is not installed on your server or more recent versions of any module already installed on your server. To install an older version of a module already installed on your server, delete the current version(s) installed on your server. You can then use this page to install any version of the module (which is no longer installed on your server).

3. Click *Install Checked*.

The screenshot shows the GenePattern web interface in a Mozilla Firefox browser. The page title is 'GenePattern - Install Modules & Pipelines From Repository'. The URL is 'http://gp409-6ac.broad.mit.edu:8080/gp/pages/taskCatalog.jsf'. The page has a navigation bar with links: Modules & Pipelines, Suites, Job Results, Resources, Downloads, Administration, and Help. The main content area is titled 'Install Modules & Pipelines From Repository'. It contains search checkboxes: 'Search for new modules to install' (checked), 'Search for updates of the currently installed modules' (checked), and 'Search for up to date modules' (unchecked). There are also checkboxes for 'Operating System': 'Platform Independent' (checked) and 'Windows' (checked). A 'Search' button is present. Below the search options, it says 'Found 20 modules' and 'Install Checked'. A table lists the found modules:

| Name (Version)                                                         | Module Type | Details                                                                                                                                      |
|------------------------------------------------------------------------|-------------|----------------------------------------------------------------------------------------------------------------------------------------------|
| <input checked="" type="checkbox"/> CompareSpectra<br>1                | Proteomics  | author: D.R. Mani, Broad Institute<br>size: 2.8 MB<br>requirements: R 2.0.1, any OS                                                          |
| <input checked="" type="checkbox"/> HierarchicalClustering.MATLAB<br>0 | Clustering  | author: Gad Getz, gp-help@broad.mit.edu<br>size: 175.0 KB<br>required patches: MCRInstaller4.0 (81.4 MB)<br>requirements: MATLAB, Windows OS |

Use the top section of the form to find the modules to install. To update the list of modules/pipelines, select the modules/pipelines to search for and click *Search*:

- *Search for new modules to install*: Displays modules and pipelines that are in the Broad repository and not on your server.
- *Search for updates of the currently installed modules*: Displays modules and pipelines where the Broad repository contains a more recent version than the version installed on your server.
- *Search for up to date modules*: Displays modules and pipelines where the Broad repository contains versions not installed on your server.
- *Operating system*: Filters the search results to display only modules and pipelines that run the selected operating system platform(s).

For each module and pipeline, the Web Client displays similar information:

- In the *Name (Version)* column:
  - Module/pipeline name.

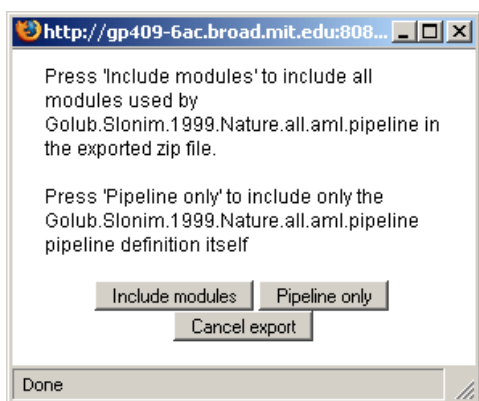
- Drop-down list of available versions. By default, the most recent version is selected.
- Brief description of the module/pipeline.
- *documentation* link, which displays the module/pipeline documentation.
- *download zip* link, which downloads a zip file of the module/pipeline. Downloading the zip file allows you to examine the source files before installing them on your server. You can then install the module/pipeline from the zip file, as described in [Exporting and Installing Suites Using Zip Files](#).
- In the *Module type* column, the category label assigned to the module/pipeline.
- In the *Details* column, the author, size, and operating system requirements for the module/pipeline.

## Exporting and Installing Modules & Pipelines Using Zip Files

Zip files provide a convenient means of sharing your modules and pipelines with other GenePattern users. You can export a module or pipeline to a zip file. The zip file can then be used to install the module or pipeline on another GenePattern server.

To export a module or pipeline to a zip file:

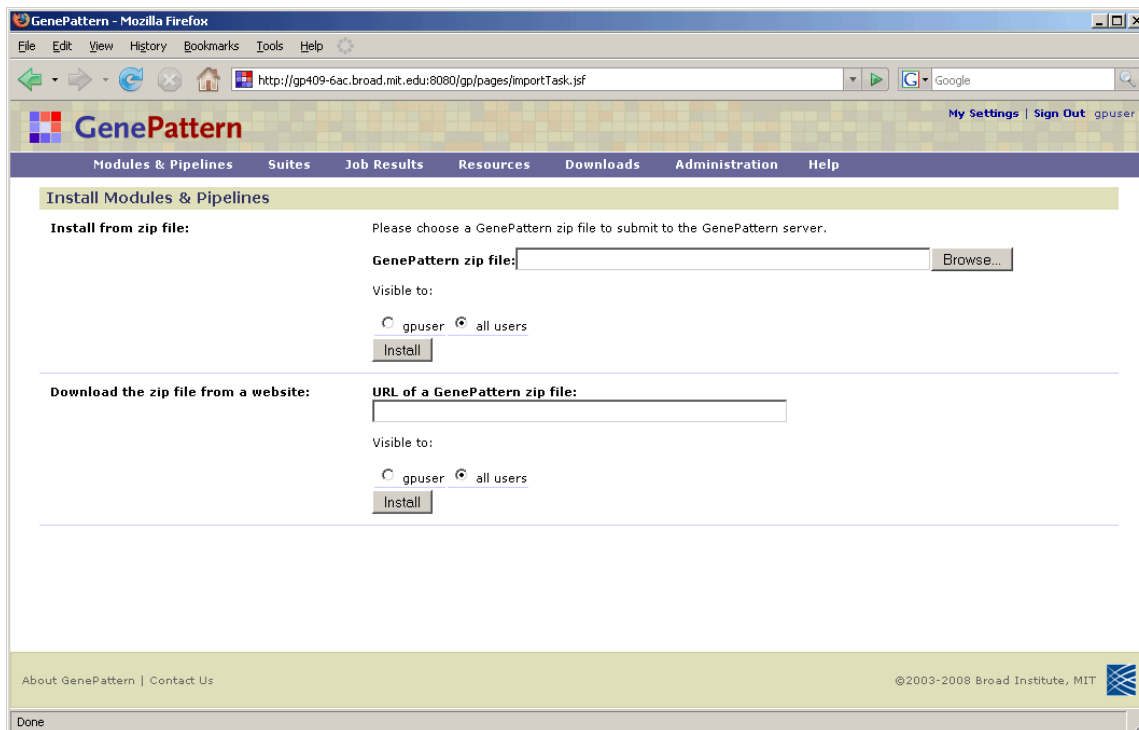
1. Click *Modules & Pipelines*. The Web Client displays the start page.
2. In the Modules & Pipelines pane, select the module or pipeline to export. The analysis parameters (if any) appear in the center pane.
3. Click *Export*.
4. If you are exporting a pipeline, the Web Client displays the following dialog:



- Click *Include modules* to create a zip file that includes all modules in the pipeline.
  - Click *Pipeline only* to create a zip file that contains only the pipeline definition. (If you install a pipeline without installing the modules that it uses, when you run the pipeline, GenePattern lists the missing modules and prompts you to edit the pipeline or install the modules.)
5. The Web Client creates a zip file that contains the module/pipeline and prompts you to save the file to your local drive.

To install a module or pipeline from a zip file:

1. Click *Modules & Pipelines>Install from zip*. The Web Client displays the Install Modules & Pipelines page, as shown below.
2. Select the file to install in one of two ways:
  - To select a zip file from a directory, click *Browse*.
  - To select a zip file from a web site, enter the URL for the zip file.
3. Use the *Visible to* radio buttons to select a private or public installation:
  - Your user name: Choose this option for a private installation. Only you (or an administrator) can view and run the module/pipeline.
  - All users (default): Choose this option for public installation. Any user connected to this server can view or run the module/pipeline.
4. Click *Install*.

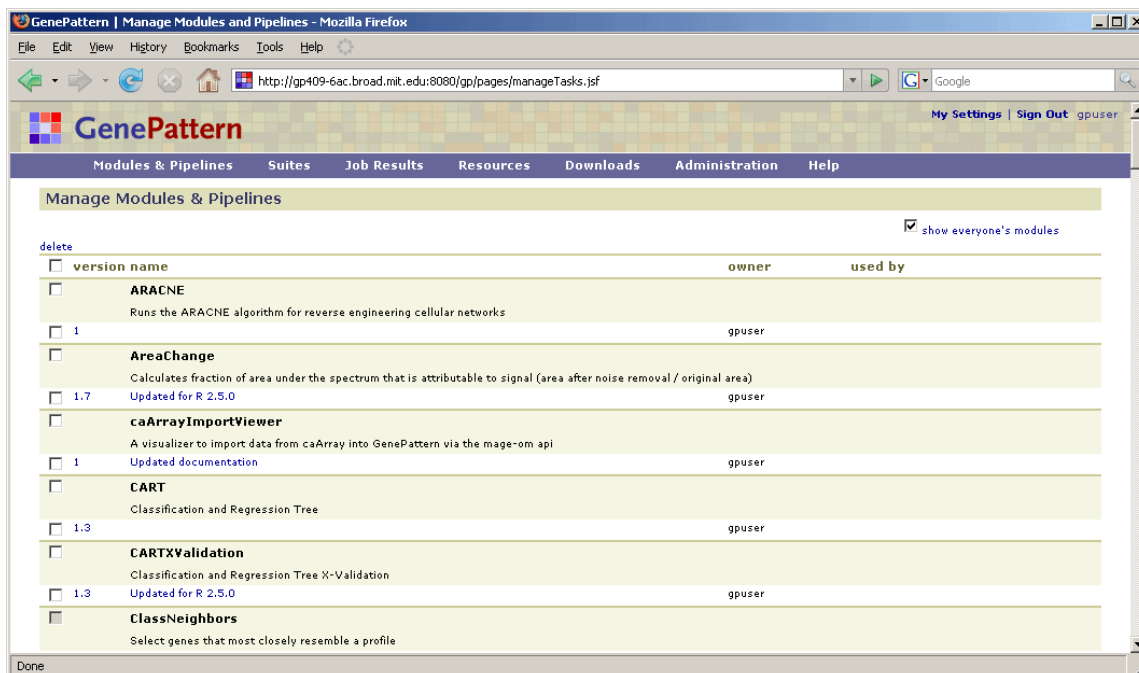


## Managing Modules & Pipelines

Click *Modules & Pipelines* > *Manage* to display the Manage Modules & Pipelines page. From this page, you can

- View the modules/pipelines installed on your GenePattern server. To display only your own modules/pipelines, clear the *show everyone's modules* check box. To display all modules/pipelines, select the check box. Private modules owned by other users are not visible unless you are an administrator.
- Delete modules/pipelines from your GenePattern server.

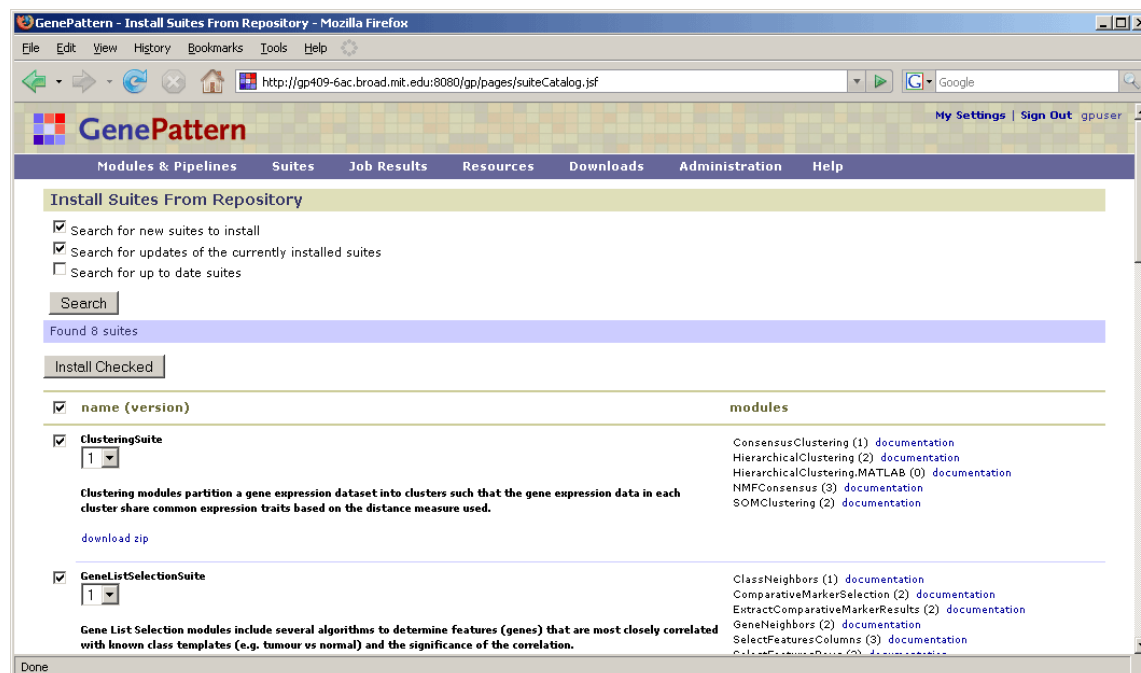
**Note:** If the check box next to the module/pipeline is grayed out, you cannot delete it. Typically, this occurs when the module/pipeline is used by another module, pipeline, or suite. To delete the module/pipeline, you must first delete the module, pipeline, or suite that uses it or remove it from that module, pipeline, or suite.



## Installing Suites from the Repository

The Broad Institute maintains a repository of suites that are freely available to the public. To install these suites on your GenePattern server:

1. Click *Suites>Install from Repository*. The Web Client displays the Install Suites from Repository page, as shown below.
2. Select the suites to install.  
Installing a suite installs its modules, if they are not yet installed on the server.
3. Click *Install Checked*.



Use the top section of the form to find the suites to install. To update the list of suites, select the suites to search for and click *Search*:

- *Search for new suites to install*: Displays suites that are in the Broad repository and not on your server.
- *Search for updates of the currently installed suites*: Displays suites where the Broad repository contains a more recent version than the version installed on your server.
- *Search for up to date suites*: Displays suites where the Broad repository contains versions not installed on your server.

For each suite, the Web Client displays similar information:

- In the *Name (Version)* column:
  - Suite name. Click the icon next to the name to display a menu of commands for working with the suite (see [Suites Menu](#)).
  - Drop-down list of available versions. By default, the most recent version is selected.
  - One-line description of the suite.
  - *download zip* link, which downloads a zip file of the suite. Downloading the zip file allows you to examine the source files before installing them on your server. You can then install the suite from the zip file, as described in [Exporting and Installing Suites Using Zip Files](#).
- In the Modules column, a list of the modules and pipelines in the suite.

## Exporting and Installing Suites Using Zip Files

Zip files provide a convenient means of sharing your suites with other GenePattern users. You can export a suite to a zip file. The zip file can then be used to install the suite on another GenePattern server.

To export a suite to a zip file:



1. Click *Suites>Manage*.
2. Click the icon following the name of the suite that you want to export. The Web Client displays the suite menu.
3. Click one of the following:
  - *Export excluding dependents*: Creates a zip file that contains the definition of the suite, but not the modules or pipelines in the suite. Installing the suite from this zip file will not install any modules or pipelines in the suite; they must already be installed on the GenePattern server or be installed separately.
  - *Export Including dependents*: Creates a zip file that contains the definition of the suite, as well as the modules and/or pipelines in the suite. Installing the suite from this zip file will also install the modules and pipelines in the suite (unless they are already installed on the GenePattern server).

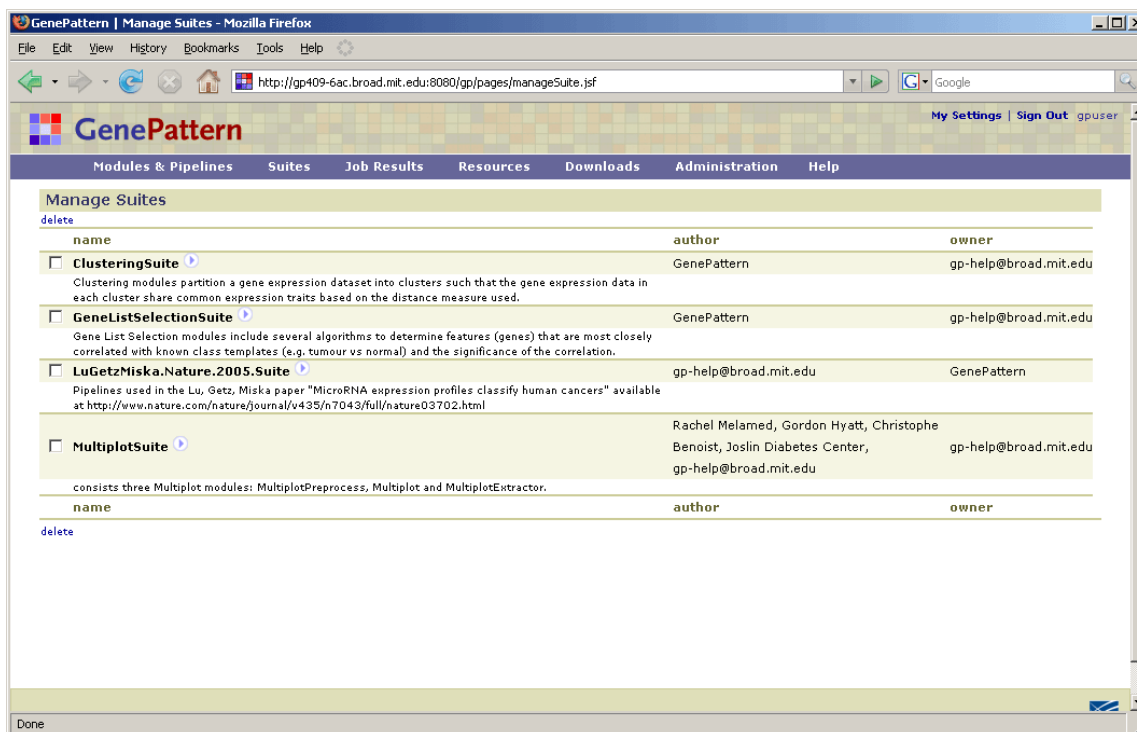
To install a suite from a zip file:

1. Click *Suites>Install from zip*.
2. Select the file to install. You can identify the zip file using a file specification or a URL.
3. Use the *Visible to* radio buttons to select a private or public installation:
  - Your user name: Choose this option for a private installation. Only you (or an administrator) can view and run the module/pipeline.
  - All users (default): Choose this option for public installation. Any user connected to this server can view or run the module/pipeline.
4. Click *Install*.

## Managing Suites

Click *Suites>Manage* to display the Manage Suites page. From this page, you can

- View the suites installed on your GenePattern server.
- Delete suites from your GenePattern server. If the check box next to the module/pipeline is grayed out, you cannot delete it. Deleting a suite does not delete its modules.
- Click the icon next to a suite name to display a menu of commands for working with the suite (see [Suites Menu](#)).



## Managing the GenePattern Server

GenePattern can be run standalone on a small machine or separated into its client and server components to take advantage of a more powerful compute server. When you installed your GenePattern server, you set the basic server configuration options for your site. This section describes how to modify GenePattern server configuration options and how to manage other aspects of the GenePattern server environment:

- [Modifying Server Settings](#)
- [Using a Queuing System](#)
- [Securing the Server](#)
- [Changing the GenePattern Database \(HSQL to Oracle\)](#)

## Modifying Server Settings

To modify the configuration of your GenePattern server, use the Server Settings page:

1. Click *Administration>Server Settings* to display the Server Settings page.
2. From the Server Settings pane, select the server setting that you want to modify. The Web Client displays a page of related server configuration options.
3. Modify and save the server configuration options.
4. Optionally, return to step 2 to change additional settings.

The following table summarizes the server settings. For more detail, click a link in the table.

|                                     |                                                                                                          |
|-------------------------------------|----------------------------------------------------------------------------------------------------------|
| <a href="#">Access</a>              | Specify which clients have access to the server.                                                         |
| <a href="#">Advanced</a>            | Specify software source directories and other low-level configuration options.                           |
| <a href="#">Command Line Prefix</a> | Specify commands and qualifiers to be prepended to the command line used to invoke a module or pipeline. |
| <a href="#">Custom</a>              | Create new server configuration options.                                                                 |
| <a href="#">Database</a>            | Specify configuration options for the GenePattern database.                                              |
| <a href="#">File Purge</a>          | Specify how long files remain on the server before being deleted.                                        |

|                                       |                                                                                                                                                             |
|---------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <a href="#">GenePattern Log</a>       | Display the log file for the GenePattern server.                                                                                                            |
| <a href="#">Programming Languages</a> | Specify the root directories for the programming languages used by GenePattern and the Java flags to be added to Java command lines executed by the server. |
| <a href="#">Proxy</a>                 | If your organization has a web proxy between the GenePattern server and the internet, specify the proxy information required to access the internet.        |
| <a href="#">Repositories</a>          | Specify the URL used to access the module repository and the suite repository.                                                                              |
| <a href="#">Shut Down Server</a>      | Shutdown the GenePattern server.                                                                                                                            |
| <a href="#">Web Server Log</a>        | Display the log file for the web server used by the GenePattern server.                                                                                     |

## Access

Use the Access page to define which GenePattern clients have access to the GenePattern server. The localhost (127.0.0.1) computer cannot be denied access to the locally installed GenePattern server. This prevents you from inadvertently denying yourself access to the server.

Using the Access page to control which computers have access to the GenePattern server is the simplest way to secure your server. You can also control access to your server based on user authentication and user permissions, as described in [Securing the Server](#). The Access page filters are applied before any user-specific authentication or permissions are checked. If your computer cannot access the server, you cannot access the server regardless of your username/password or permissions.

- Click *Standalone* to allow only local clients to connect to the server; that is, you can access this GenePattern server only from the computer that it is running on.
- Click *Any Computer* (default) to allow any client to connect to the server.
- Click *These Domains* to allow only clients from specific domains to connect to the server. Enter a comma-separated list of domains or IP addresses in the text box, for example:  
`broad.mit.edu,dfci.harvard.edu,mit.edu`.

GenePattern scans all incoming connection attempts. If they match in whole or in part any domain name or IP address in this list, the server allows access; otherwise, the server redirects the connection to a page indicating that the server does not allow access.

Click *Save* to save your changes. Click *Restore* to return to the value set at installation.

## Advanced

The Advanced page contains directory specifications for the GenePattern source files and other low-level configuration options. You rarely need to modify these options.

| Advanced Configurations                                                    |                                             |
|----------------------------------------------------------------------------|---------------------------------------------|
| Patch Qualifiers:                                                          | build.tag.os.arch.os.name.os.version.java.v |
| Location for Patches:                                                      | C:/Program Files/GenePatternServer/patche   |
| Location for Ant:                                                          | <java> -cp <tomcatCommonLib>/tools.jar -ja  |
| Directory for Resources:                                                   | ./resources                                 |
| Directory for Modules:                                                     | ./taskLib                                   |
| Directory for TomcatCommonLib:                                             | C:/Program Files/GenePatternServer/Tomc     |
| Directory for Webapp:                                                      | C:/Program Files/GenePatternServer/Tomc     |
| Directory for GenePattern Log:                                             | logs/genepattern.log                        |
| Classpath for Pipelines:                                                   | <webappDir><file.separator>WEB-INF<file.s   |
| Main Class for Pipelines:                                                  | org.genepattern.server.webapp.RunPipeline   |
| VM args for Pipelines:                                                     | -DGenePatternURL=<GenePatternURL>           |
| Decorator for Pipelines:                                                   | org.genepattern.server.webapp.RunPipeline   |
| Installed Patch LSIDs:                                                     | urn:lsid:broad.mit.edu:cancer.software.gene |
| Desktop Client Install URL:                                                | http://www.broad.mit.edu/cancer/software/ge |
| Number of Threads:                                                         | 20                                          |
| <input type="button" value="Save"/> <input type="button" value="Restore"/> |                                             |

Click **Save** to save your changes. Click **Restore** to return to the values set at installation.

## Command Line Prefix

Use the Command Line Prefix page to define commands and qualifiers to be prepended to the command line used to invoke a module or pipeline. For example, use this page to prepend commands and qualifiers that execute modules and pipelines on a cluster farm, as described in [Using a Queuing System](#).

| Command Line Prefix                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                                              |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|--------|---------------------|----------------------------|---------|----------------------------------|--|--------------------|--|---------------------------------|--|-------------|--|---------------------|--|
| Default Command Prefix:                                                                                                                                                                                                                                                                                                                                                                                                                                         | <input type="text"/> <input type="button" value="save default"/>                             |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
| The default command prefix is used for all modules that do not have an explicit prefix setting below.                                                                                                                                                                                                                                                                                                                                                           |                                                                                              |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
| Add New Prefix:                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                                                              |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
| name:                                                                                                                                                                                                                                                                                                                                                                                                                                                           | <input type="text"/> content: <input type="text"/> <input type="button" value="add prefix"/> |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
| <table border="1"> <thead> <tr> <th>name</th> <th>content</th> </tr> </thead> <tbody> <tr> <td>default</td> <td></td> </tr> </tbody> </table>                                                                                                                                                                                                                                                                                                                   |                                                                                              | name   | content             | default                    |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
| name                                                                                                                                                                                                                                                                                                                                                                                                                                                            | content                                                                                      |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
| default                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                              |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
| <table border="1"> <thead> <tr> <th>module</th> <th>command prefix name</th> </tr> </thead> <tbody> <tr> <td>ComparativeMarkerSelection</td> <td>default</td> </tr> <tr> <td>ComparativeMarkerSelectionViewer</td> <td></td> </tr> <tr> <td>ConvertLineEndings</td> <td></td> </tr> <tr> <td>ExtractComparativeMarkerResults</td> <td></td> </tr> <tr> <td>HeatMapView</td> <td></td> </tr> <tr> <td>MyComparativeMarker</td> <td></td> </tr> </tbody> </table> |                                                                                              | module | command prefix name | ComparativeMarkerSelection | default | ComparativeMarkerSelectionViewer |  | ConvertLineEndings |  | ExtractComparativeMarkerResults |  | HeatMapView |  | MyComparativeMarker |  |
| module                                                                                                                                                                                                                                                                                                                                                                                                                                                          | command prefix name                                                                          |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
| ComparativeMarkerSelection                                                                                                                                                                                                                                                                                                                                                                                                                                      | default                                                                                      |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
| ComparativeMarkerSelectionViewer                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                                              |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
| ConvertLineEndings                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                              |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
| ExtractComparativeMarkerResults                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                                                              |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
| HeatMapView                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                                              |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
| MyComparativeMarker                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                                              |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |
| For Modules ...                                                                                                                                                                                                                                                                                                                                                                                                                                                 | ... Use Prefix ... <input type="button" value="add mapping"/>                                |        |                     |                            |         |                                  |  |                    |  |                                 |  |             |  |                     |  |

To prepend text to all (or most) command lines executed by the GenePattern server:

1. Enter the desired commands and qualifiers in the *Default Command Prefix* field.
2. Click **save default**. GenePattern displays the updated content of the *default* prefix. The *name/content* table in the middle of the form lists the *default* prefix and its content. The previous illustration shows the *default* prefix with no content.

When GenePattern executes a module or pipeline, it constructs the appropriate command line, prepends the *default* prefix to that command line, and then executes the command line.

To prepend text only to command lines that invoke specific modules or pipelines:

1. In the *Add New Prefix* field, enter a name for the prefix and the commands and qualifiers to prepend to the command line.

2. Click *add prefix*. GenePattern creates the new prefix, updates its content, and adds the prefix to the *name/content* table in the middle of the form.
3. At the bottom of the form, select one or more module(s)/pipeline(s), select your new prefix, and click *add mapping*. GenePattern adds the prefix information to the *module/command prefix name* table.

When GenePattern executes a module or pipeline listed in the *module/command prefix name* table, it constructs the appropriate command line, prepends the specified prefix to that command line, and then executes the command line. (When GenePattern executes a module or pipeline *not* listed in that table, it constructs the appropriate command line, prepends the *default* prefix to that command line, and then executes the command line.)

## Custom

Use the Custom page to define your own configuration options.

When you create a module, the custom configuration options are available as substitution variables in the module command line. For example, if you define a custom property "foo", you can use <foo> in the command line to pass the value of the custom configuration option to your module. In the Broad repository, for example, the LandmarkMatching and PeakMatch modules use the custom configuration option `pepperPrefix`.

1. In the *name* field, enter a name for the configuration option.
2. In the *content* field, enter a value for the configuration option.
3. Click *add setting*. GenePattern adds the option to the table at the bottom of the form.

## Database

Use the Database page to set configuration options for the GenePattern database. The following figure shows the HSQL options. You rarely need to change these options.

Click *Save* to save your changes. Click *Restore* to return to the value set at installation.

## File Purge

Use the File Purge page to specify when analysis results files are deleted from the server:

File Purge Settings

Purge Jobs After: 7 days
Purge Time: 23:00
Save Restore

File purge settings define how often and, at what time of day, intermediate result files created on the server are deleted (purged). To never delete files set the purge frequency to -1. Note that this could result in large amounts of hard drive space being used over time.

- Use *Purge Jobs After* to specify the number of days the server keeps the analysis results files. To prevent the server from automatically deleting the files, set this value to -1.
- Use *Purge Time* to specify what time of day (24-hour format) the server deletes the files.

Click *Save* to save your changes. Click *Restore* to return to the values set at installation.

## GenePattern Log

Use the GenePattern Log page to view warnings and messages generated by the GenePattern server. (Use the [Web Server Log](#) page to view messages generated by the web server that GenePattern uses.)

GenePattern Log

Gene Pattern log file from node258.broad.mit.edu on Fri Nov 16 15:28:01 EST 2007

```

2007-11-06 16:44:39.834 ERROR [Thread-43] (GenePatternAnalysisTask.java:3505) - Cannot inst
2007-11-06 16:44:39.843 ERROR [Thread-43] (TaskCatalogBean.java:317) - org.genepattern.serv
2007-11-06 16:59:53.576 ERROR [http-8080-Processor23] (HibernateSessionRequestFilter.java:4
2007-11-07 12:06:31.234 ERROR [Thread-130] (GenePatternAnalysisTask.java:3505) - Cannot inst
2007-11-07 12:06:31.236 ERROR [Thread-130] (TaskCatalogBean.java:317) - org.genepattern.serv
2007-11-07 12:07:28.784 ERROR [http-8080-Processor12] (HibernateSessionRequestFilter.java:4

```

## Programming Languages

The Programming Languages page contains two sections.

- Use Programming Language Configurations to specify the root directories for the programming languages used by GenePattern:

Programming Language Configurations

Java: C:/Program Files/GenePatternServer/jre/bin
Perl: C:/Program Files/GenePatternServer/perl/bin
R 2.5 Home: C:/Program Files/GenePatternServer/R25/b
R 2.0.1 Home: C:/Program Files/GenePatternServer/R/bin/
Save Restore

When you install GenePattern, you install the programming languages used by GenePattern. If you have alternate programming language installations that you prefer to use, use this page to point to those installations.

- Use Programming Language Options to add Java virtual machine configuration parameters (such as VM memory settings) for the execution of Java-based modules:

Programming Language Options

Java VMOptions: -Xmx512M
Java Visualizer VMOptions: -Xmx512M
R Options: -no-save -quiet -slave -no-restore
Save Restore

Java VMOptions are passed to the VM of most Java language modules. You can use these options to increase the amount of memory allotted to a Java language module if you are experiencing OutOfMemory errors.

For details of java options run  
java -X  
at a command line.

R Options are passed to the R interpreter of most R language modules. You can use these options to increase the amount of memory allotted to an R language module if you are experiencing vector allocation errors.

For details of R options run  
R --help  
at a command line.

- Use *Java VMOptions* to add flags to Java command lines executed on the GenePattern server. Changes to these options take effect when you stop and restart the server.

- Use *Java Visualizer VMOptions* to add flags to Java command lines executed on the GenePattern client (the visualizer modules). Changes to these options take effect the next time you run a visualizer module.
- Use *R Options* to add flags to Java command lines executed in the R programming environment. Changes to these options take effect when you stop and restart the server.

Click **Save** to save your changes. Click **Restore** to return to the value set at installation.

## Proxy

If your server is behind a firewall, use the Proxy page to set the HTTP and FTP Proxy information. Without the proxy information, the server cannot download modules, pipelines, or suites from the repository maintained by the Broad Institute. If you do not know the proxy information, contact your systems administrator.

Proxy Settings

Proxy Host:   
Proxy Port:   
Proxy Username:   
Proxy Password:

Proxy settings are needed only to connect to the Module Repository, and only if your organization has a web proxy between you and the internet. Username and password are required only if your proxy requires authentication. The username and password values will not be saved to the config file for security reasons and need to be re-entered after a server restart.

Click **Save** to save your changes. Click **Restore** to return to the values set at installation.

## Repositories

Use the Repositories page to identify the location of the repository to be accessed by the GenePattern server when you install modules and pipelines or suites from the repository. By default, it points to the module repository maintained by the Broad Institute. If you would like to implement and maintain a module repository at your site, contact the GenePattern help desk ([gp-help@broad.mit.edu](mailto:gp-help@broad.mit.edu)).

Module Repositories

Previous Module Repository URL:   
Module Repository URL:

Suite Repositories

Previous Suite Repository URL:   
Suite Repository URL:

Click **Save** to save your changes. Click **Restore** to return to the values set at installation. Click **Remove** to delete the selected URL from the list.

## Shut Down Server

You can shutdown the GenePattern server by clicking the link on this page. For easier ways of shutting down the server, see [Exiting from GenePattern](#).

Shut Down Server

Warning: The GenePattern server can NOT be restarted from this page.

[Click here to shut down the GenePattern server.](#)

## Web Server Log

Use the Web Server Log page to view messages generated by the web server that GenePattern uses. (Use the [GenePattern Log](#) page to view warnings and messages generated by the GenePattern server.)

| WebServer Log                                                                               |  |
|---------------------------------------------------------------------------------------------|--|
| Web Server log file from node258.broad.mit.edu on Fri Nov 16 15:28:01 EST 2007              |  |
| 2007-11-06 16:44:39.834 ERROR [Thread-43] (GenePatternAnalysisTask.java:3505) - Cannot inst |  |
| 2007-11-06 16:44:39.843 ERROR [Thread-43] (TaskCatalogBean.java:317) - org.genepattern.serv |  |
| 2007-11-06 16:59:53.576 ERROR [http-8080-Processor23] (HibernateSessionRequestFilter.java:8 |  |
| 2007-11-06 16:59:53.582 ERROR [http-8080-Processor23] (StandardWrapperValve.java:253) - Ser |  |

## Using a Queuing System

This section describes how you can integrate the GenePattern server with your queuing system. If you need help integrating the GenePattern server into your environment in other ways, contact the GenePattern help desk ([gp-help@broad.mit.edu](mailto:gp-help@broad.mit.edu)).

Queuing systems such as the Load Sharing Facility (LSF) and the Sun Grid Engine (SGE) allow computational resources to be used effectively. If you have such a queuing system, you typically want the GenePattern server to use it. To configure your GenePattern server to execute jobs using LSF or SGE:

- Add the GenePatternURL property to the GenePattern configuration file, `GenePatternServer/resources/genepattern.properties`, specifying the URL of your server. For example:

```
GenePatternURL=http\://myserver.company.com:8080/gp/
```

When you run a pipeline, the GenePattern server uses this URL to construct the links to the output files.

By default, the GenePatternURL property is not set. When you run a pipeline, the GenePattern server derives the URL at run time based on the current IP address of the host server. This is ideal for a user running on a laptop, where the IP address may change at startup. However, if you are using a queuing system, the derived URL is incorrect: it is based on the IP address of the queuing system server rather than the GenePattern server.

- Use the Command Line Prefix options of the Server Settings page to have the GenePattern server add the required options to the command line each time it executes a module.

**Note:** When using a queuing system, short jobs will take slightly longer because they must be dispatched to the queuing system. However, on a heavily used server, using a queuing system to manage job executions should improve performance overall, especially for compute-intensive and long-running jobs.

For example, if you are using LSF, modify the Command Line Prefix options as follows:

1. Click *Administration>Server Settings* to display the Server Settings page.
2. Select Command Line Prefix from the Server Settings pane. The Web Client displays the Command Line Prefix page.
3. Enter the following text in the *Default Command Prefix* field:

```
bsub -K -o lsf_log.txt
```

- The `-K` flag instructs the `bsub` command to wait for the job to complete before returning.
  - The `-o` flag specifies the file to which the job writes standard output and standard error messages.
4. Optionally, set the environment variables `BSUB_QUIET` and `BSUB_QUIET2` to prevent `bsub` from printing common job messages to standard out:
    - Setting `BSUB_QUIET` prevents `bsub` from printing the messages `<<Job is submitted to default queue <normal>>>` and `<<Waiting for dispatch>>`.
    - Setting `BSUB_QUIET2` prevents `bsub` from printing the message `<<Job is finished>>`.

Another alternative is to create a script that sets the environment variables and then executes the job using LSF or SGE. The command prefix would then execute the script. For example:

1. Create the shell script to set the variables and execute the job using LSF. The script executes in the `jobResults` directory for the job; for example, for job 3248, the script executes in the `GenePattern/Tomcat/webapps/gp/jobResults/3248/` directory. The following script sets the environment variables, submits the job to the LSF queue, waits for the job to complete, saves stdout to a new file, `stdout.txt`, and saves stderr to a new file, `stderr.txt`. By convention, GenePattern considers a job to fail if there is any output to stderr.

```
#!/bin/bash
#
# Submit the job to LSF
# Save lsf out and err files into the jobResults directory.
```



```
#
lsf_err=.lsf.err;
cmd_out=cmd.out;
BSUB_QUIET=
BSUB_QUIET2=
export BSUB_QUIET
export BSUB_QUIET2
# submit the job and wait (-K) for the job to complete
bsub -q genepattern -K -o .lsf_%J.out -e $lsf_err "$@$@ > $cmd_out"
# sleep to allow for NFS delay
sleep 2;
# If there is stdout from the job, pipe to stdout of this script,
then delete the output file
if [ -e $cmd_out ]
then
    cat $cmd_out >&1;
    rm $cmd_out;
fi
# If there is stderr from the job, pipe to stderr of this script then
delete stderr file
if [ -e $lsf_err ]
then
    cat $lsf_err >&2;
    rm $lsf_err;
fi
```

2. Click *Administration>Server Settings* to display the Server Settings page.
3. Select Command Line Prefix from the Server Settings pane. The Web Client displays the Command Line Prefix page.
4. Enter the following text in the *Default Command Prefix* field:  
`/fully/qualified/path/to/lsf_default.sh`
5. The script shown here saves the lsf log file into the job results directory. In the Web Client, the log files are displayed with the other job result files. If you do not want the log files displayed in the Web Client, edit the `/resources/genepattern.properties` file and set the following property:

```
jobs.FilenameFilter=.lsf*
```

## Securing the Server

Securing your GenePattern server helps you to control who has access to which operations. Since GenePattern is primarily a web application (including SOAP interfaces) running on a web server, general approaches for securing your web servers are applicable to your GenePattern server. In addition, GenePattern provides several security features that can easily be used by non-technical users to control access to the server.

This section describes several ways to secure your GenePattern server:

- [Access Filtering](#)
- [Password Protection](#)
- [User Accounts](#)
- [User Authentication](#)
- [User Permissions](#)
- [Secure Sockets Layer \(SSL\) Support](#)

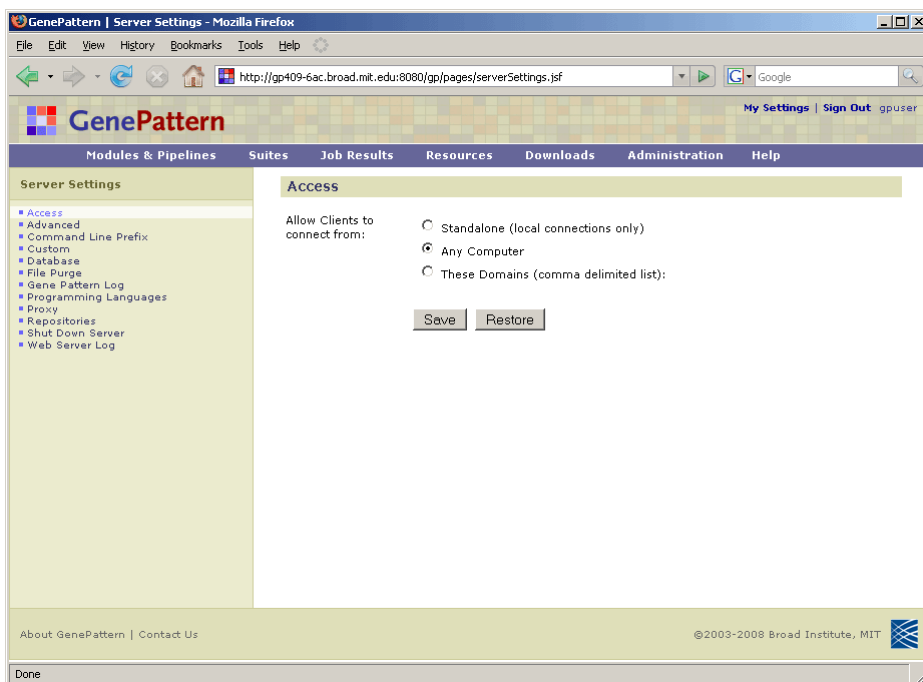
### Access Filtering

Use the Access page to define which GenePattern clients have access to the GenePattern server. The localhost (127.0.0.1) computer cannot be denied access to the locally installed GenePattern server. This prevents you from inadvertently denying yourself access to the server.

Access filtering prevents users from connecting to your GenePattern server unless they come from a known computer. This is the simplest way to secure your GenePattern server because you control access filtering from the Server Administration page of the Web Client. The Access page filters are applied before any user-specific authentication or permissions are checked. If your computer cannot access the server, you cannot access the server regardless of your username/password or permissions.

To use access filtering (as described in [Modifying Server Settings](#)):

1. Click *Administration*>*Server Settings*.
2. Use the Access page to determine which clients have access to your GenePattern server:



- Click *Standalone* to allow only local clients to connect to the server; that is, you can access this GenePattern server only from the computer that it is running on.
- Click *Any Computer* (default) to allow any client to connect to the server.
- Click *These Domains* to allow only clients from specific domains to connect to the server. Enter a comma-separated list of domains or IP addresses in the text box, for example:  
`broad.mit.edu,dfci.harvard.edu,mit.edu`.

GenePattern scans all incoming connection attempts. If they match in whole or in part any domain name or IP address in this list, the server allows access; otherwise, the server redirects the connection to a page indicating that the server does not allow access.

## Password Protection

By default, your GenePattern server requires only a user name to authenticate a GenePattern user. You can easily add password protection by modifying the GenePattern server properties.

To add password protection to your server, modify the GenePattern server properties:

1. Edit the GenePattern configuration file,  
`GenePatternServer/resources/genepattern.properties`.
2. Set the `requirePassword` property to `true`: `requirePassword=true`.
3. Save the `genepattern.properties` file.
4. Restart the GenePattern server.

When you add password protection to your server:

- Existing users are assigned a blank password. The first time a user signs in with a blank password, the GenePattern server requires the user to set the password.

- New users are required to register before using the GenePattern server. By registering, the user creates a GenePattern account with an associated username, password, and email address.
- The sign-in screen prompts you for a username and password. If you forget your password, click the *Forgot your password?* link and GenePattern emails you a temporary password.

Assigning passwords to existing user accounts prevents anyone from inadvertently or intentionally logging into and taking control of another user's account. After adding password protection to your server, set passwords for existing users as follows:

1. Sign into GenePattern using the username of the existing user. If you do not know the usernames of the existing users, the GenePattern help desk ([gp-help@broad.mit.edu](mailto:gp-help@broad.mit.edu)) can help you retrieve the names from your server.
2. When GenePattern prompts you to set a password, select a password for that user.
3. After setting the password, GenePattern displays the Change Email page ([My Settings](#)). If the email address has not been set, set it. This is the address GenePattern uses to send the user a new password if necessary.
4. Sign out and repeat the process for the next user.
5. After setting passwords for all users, let them know that passwords have been set. You do not need to send the users their passwords. Simply ask users to sign into GenePattern and click the *Forgot your password?* link to have GenePattern send a temporary password.

## User Accounts

By default, anyone can create a new GenePattern account. You can easily prevent this by modifying the GenePattern server properties.

To restrict account creation on your server, modify the GenePattern server properties:

1. Edit the GenePattern configuration file,  
GenePatternServer/resources/genepattern.properties.
2. Set the create.account.allowed property to false: create.account.allowed=false.
3. Save the genepattern.properties file.
4. Restart the GenePattern server.

Existing users can sign into GenePattern as usual. New users cannot sign into GenePattern. A new user account must be manually created by editing the GenePattern database. After the user account is added to the database, the user can sign into GenePattern. For more information, contact the GenePattern help desk ([gp-help@broad.mit.edu](mailto:gp-help@broad.mit.edu)).

## User Authentication

The GenePattern server can be configured to provide user name and password protection. Alternative user authentication can easily be added by any capable Java programmer. The remainder of this section is written for such a programmer.

GenePattern user authentication is performed by a servlet filter installed in front of the GenePattern web application in its web.xml file. To implement alternative authentication, implement a new servlet filter that performs the desired authentication and a new login page that requests additional information from the user. The new servlet filter must, after authenticating a user, set a session attribute in the `java.servlet.http.HttpSession` object called `userID`. All pages in the GenePattern server and all database interactions will use this as the user's user name.

To see the source code for the default `AuthenticationFilter.java` and the `web.xml` files, click the following links: [AuthenticationFilter.java](#), [web.xml](#).

Once you have written and compiled the new `ServletFilter`:

1. Place the jar file containing the new `ServletFilter` into the following directory:  
\*/GenePatternServer/Tomcat/webapps/gp/WEB-INF/lib
2. Modify the GenePattern server's web.xml document.  
\*/GenePatternServer/Tomcat/webapps/gp/WEB-INF/web.xml

**Note:** It is important to maintain the existing order of the servlet filters in the web.xml document as they are used in the order they are defined in the document. The Authentication filter must come before the Authorization filter for the Authorization filter to work.

3. Change the definition of the AuthenticationFilter to use the class that you have provided.
4. Add any necessary configuration elements that it requires.
5. Restart the GenePattern server for the changes to take effect.

**Note:** If you look at the code for the default Authentication Filter ([AuthenticationFilter.java](#)), you will see that it allows requests through that have a parameter called `jsp_precompile` that have come from the `localhost`. If you do not allow these requests through unauthenticated, you will see a series of errors when you start the GenePattern server as it attempts to precompile the JSP pages. These are **not** fatal errors, but they slow down server response for users the first time that pages are accessed following a server restart.

## User Permissions

The GenePattern server has a built-in authorization mechanism based on two configuration files in the `GenePatternServer/resources` directory:

- `userGroups.xml` defines user groups
- `permissionMap.xml` defines which user groups have which permissions

User authorization is based on the user's identity as provided at GenePattern login (for more information, see [User Authentication](#)). Once a user has been authenticated, the GenePattern server uses these two configuration files to determine the user's permissions. Keep file permissions applied to these files to ensure that they are not modified unintentionally.

The default configuration files allow all users to perform all actions: [userGroups.xml](#), [permissionMap.xml](#). One simple way to limit access to your server is to edit the `userGroups.xml` file and name your system administrators; for example: [userGroups.xml \(modified\)](#).

### userGroups.xml

The `userGroups.xml` file defines groups of users to whom permissions may be assigned.

Any number of groups may be created as long as their names are unique. If the user named `*` is included in a group, as shown in the default settings, then any and all users are considered to be a part of that group. A user may belong to any number of groups.

### permissionMap.xml

The `permissionMap.xml` file maps permission names to user groups; that is, you use this file to state that members of a particular group have a particular permission.

- **Groups:** Any number of groups may be assigned a permission by adding additional `<group name="thename" />` elements to a permission. The presence of a group named `*` means that any and all groups have that permission, which means any and all users have that permission.
- **Permissions:** Permissions are predefined by GenePattern and should not be added or removed. The following table describes the predefined permissions, which are listed in the default `permissionMap.xml` file and are the only permissions recognized by GenePattern.

**Note:** No explicit permission is required to run public modules/pipelines, or private modules/pipelines that you have created. No explicit permission is required to edit or delete your own modules, pipelines, suites, or jobs.

|                              |                                                                                                                                                                                                                                                                                                                                     |
|------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>createModule</b>          | Permits creation of a module. Creation refers to any action that adds a module to the server, including create, install from repository, install from zip, and clone.                                                                                                                                                               |
| <b>createPrivatePipeline</b> | Permits creation of a private pipeline (a pipeline visible only to its creator). Creation refers to any action that adds a private pipeline to the server, including create, install from repository, install from zip, and clone.<br><br><b>Note:</b> To install the modules in a pipeline, you must have createModule permission. |
| <b>createPrivateSuite</b>    | Permits creation of a private suite (a suite visible only to its creator). Creation refers to any action that adds a private suite to the server, including create, install from repository, install from zip, and clone.<br><br><b>Note:</b> To install the modules in a suite, you must have createModule permission.             |

|                             |                                                                                                                                                                                                                                                                                      |
|-----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>createPublicPipeline</b> | Permits creation of a public pipeline. Creation refers to any action that adds a public pipeline to the server, including create, install from repository, install from zip, and clone.<br><b>Note:</b> To install the modules in a pipeline, you must have createModule permission. |
| <b>createPublicSuite</b>    | Permits creation of a public suite. Creation refers to any action that adds a public suite to the server, including create, install from repository, install from zip, and clone.<br><b>Note:</b> To install the modules in a suite, you must have createModule permission.          |
| <b>adminJobs</b>            | Permits viewing and deleting jobs and associated files owned by other users.                                                                                                                                                                                                         |
| <b>adminModules</b>         | Permits viewing and deleting private modules owned by other users. Permits deleting public modules.<br><b>Note:</b> No explicit permission is required to view public modules.                                                                                                       |
| <b>adminPipelines</b>       | Permits viewing and deleting private pipelines owned by other users. Permits deleting public pipelines.<br><b>Note:</b> No explicit permission is required to view public pipelines.                                                                                                 |
| <b>adminSuites</b>          | Permits viewing and deleting private suites owned by other users. Permits deleting public suites.<br><b>Note:</b> No explicit permission is required to view public suites.                                                                                                          |
| <b>adminServer</b>          | Permits access to <i>Administration&gt;Server Settings</i> and all actions on the Server Settings page, including modifying server settings and shutting down the server.                                                                                                            |

## User Authorization

GenePattern authorization, like GenePattern [user authentication](#), is performed by a servlet filter installed in front of the GenePattern web application in its web.xml file. To implement alternative authorization, you need to implement a new servlet filter that performs the desired authorization. It is important to maintain the existing order of the servlet filters in the web.xml document as they are used in the order they are defined in the document. The Authentication filter must come before the Authorization filter for the Authorization filter to work. To see the source code for the default `AuthorizationFilter.java` and the `web.xml` files, click the following links: [AuthorizationFilter.java](#), [web.xml](#).

## Secure Sockets Layer (SSL) Support

This section describes how you can modify the GenePattern web application to run on a web server that is configured to use the HTTPS protocol, where essentially the regular http requests are routed through a secure sockets layer (SSL) making them much harder for hackers to access. If you have installed your GenePattern server onto a web server other than the default Tomcat instance it is distributed with, configure your web server according to its instructions and then follow Step 2 below.

**Note:** When running under SSL, programming language clients and the GenePattern Desktop Client may not be able to connect to your GenePattern server.

### Step 1. Configure Tomcat for SSL support

Follow the instructions available at <http://tomcat.apache.org/tomcat-5.5-doc/ssl-howto.html> to configure the Tomcat instance for using SSL. In doing so, you will modify the Tomcat configuration file, which is located in the `GenePatternServer/Tomcat/conf` directory.

### Step 2. Configure GenePattern for SSL

Once the Tomcat (or other web server) has been configured for SSL, modify the GenePattern configuration file, `GenePatternServer/resources/genepattern.properties`, to ensure that its properties are in synch with the web server:

1. Add a new key, `java.net.ssl.trustStore=<path to keystore>`.

This should point to the keystore you created when configuring Tomcat (above) or some keystore that GenePattern can use to establish SSL connections.

2. Modify the value for the key `GENEPATTERN_PORT` to use the https port you selected when configuring Tomcat (above).

3. Modify the value for the key `GenePatternURL` to use the `https` protocol and the `https` port you selected, for example:

`http://localhost:8080/gp` becomes `https://localhost:8443/gp`

Save the `genepattern.properties` file and restart your server. Any bookmarked links to your GenePattern server must be updated to the new protocol and port.

## Changing the GenePattern Database (HSQL to Oracle)

The GenePattern database has been implemented in both HSQL and Oracle. The GenePattern installation builds the HSQL database and sets the GenePattern server properties to reference that database. To use the Oracle implementation instead, build the Oracle database and modify the GenePattern properties to reference that database, as described in the following procedure:

1. Build the Oracle database using the `.sql` file provided with the GenePattern installation:  
`/GenePatternServer/resources/analysis_oracle-30.sql`
2. Edit the `GenePatternServer/resources/genepattern.properties` file:
  1. Locate the Database Parameters section.
  2. Comment out the HSQL statements.
  3. Uncomment the Oracle statements.
  4. Modify the Oracle URL, username, and password for your site. If necessary, modify the Oracle scheme name.
3. Start the GenePattern server.

## Documentation Update History

| Version | Release date  | Comments                                       |
|---------|---------------|------------------------------------------------|
| 3.1.1   | July 2008     | Updated <a href="#">Using a Queuing System</a> |
| 3.1     | December 2007 | GenePattern 3.1 Release                        |
| 3.0     | April 2007    | GenePattern 3.0 Release                        |