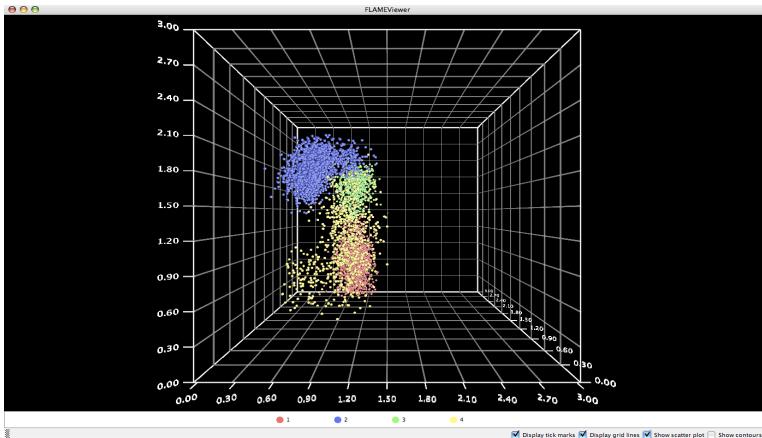
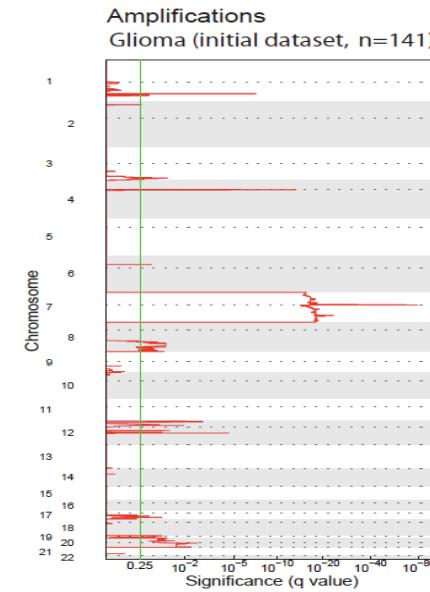


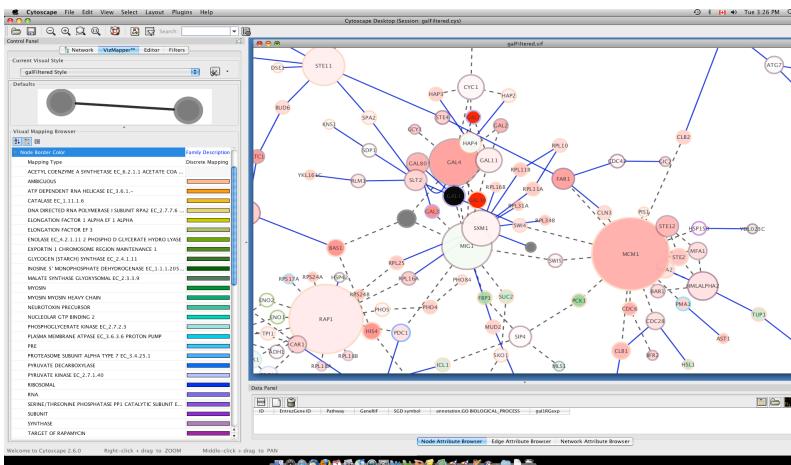
Other GenePattern Features



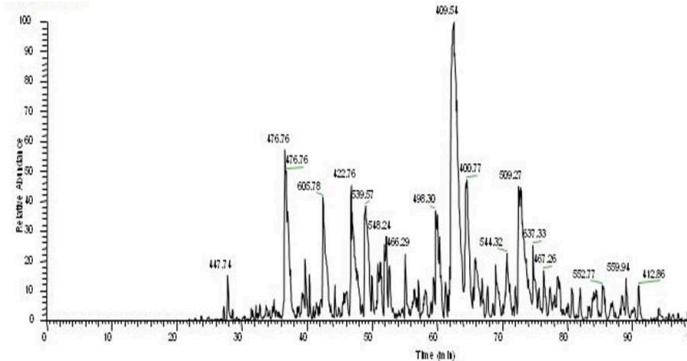
Flow Cytometry



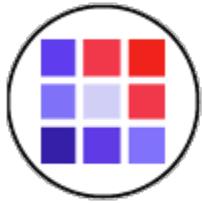
Sequence Variation Analysis



Network Analysis



Proteomics



Publish Your Notebooks

- Publish your own notebooks on the GenePattern Notebook Repository.
- Share notebooks with others.

The screenshot shows the GenePattern Notebook interface. On the left, there's a sidebar with tabs for 'Files', 'Running', and 'Public'. Under 'Files', there's a list of notebooks: 'Classification and Pre...', 'Differential Expression.ipynb', 'GenePattern Files in P...', 'Hierarchical Clustering.ipynb' (which is checked), 'K-means Clustering.ipynb', 'NMF Clustering.ipynb', 'Run an Analysis.ipynb', and 'Samples and Features.ipynb'. The main area has a title bar 'GenePattern Notebook' and a sub-header 'Publish Notebook to Repository'. A message box explains that publishing makes a copy available to anyone and requires manual update if changes are made. Below it, there are fields for 'Notebook Name' (set to 'Hierarchical Clustering'), 'Description' (set to 'How to perform hierarchical clustering analysis in GenePattern.'), 'Authors' (set to 'GenePattern Team'), and 'Quality' (set to 'Release'). At the bottom right of the dialog are 'Cancel' and 'Publish' buttons.



GenePattern Python Library

- Control a GenePattern server via Python
- Automatic integration with GenePattern cell data

```
import gp

# Create a GenePattern server proxy instance
gpserver = gp.GPServer('http://localhost:8080/gp', 'myusername', 'mypassword')

# Obtain GPTask by module name
module = gp.GPTask(gpserver, "PreprocessDataset")

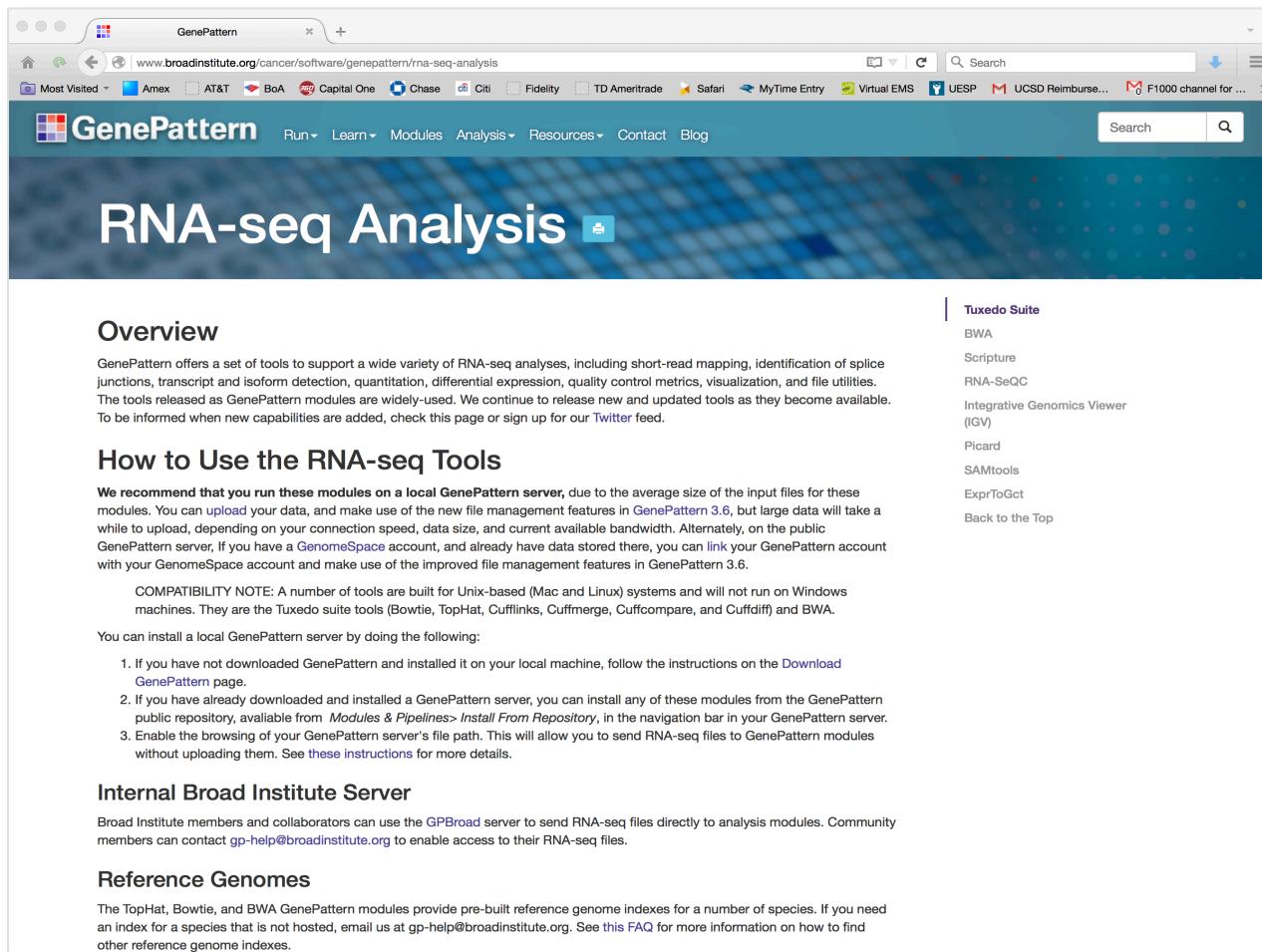
# Load module parameter data
module.param_load()

# Create a job specification
job_spec = module.make_job_spec()

# Upload a file to the server
uploaded_file = gpserver.upload_file("file_name", "/path/to/the/file/file_name")
job_spec.set_parameter("input.filename", uploaded_file.get_url())

# Submit the job to the GenePattern server
job = gpserver.run_job(job_spec)
```

RNA-Seq in GenePattern



The screenshot shows a web browser displaying the GenePattern RNA-seq Analysis page. The URL in the address bar is www.broadinstitute.org/cancer/software/genepattern/rna-seq-analysis. The page has a blue header with the GenePattern logo and navigation links for Run, Learn, Modules, Analysis, Resources, Contact, and Blog. A search bar is also present. The main content area features a large "RNA-seq Analysis" heading with a small icon. Below it is an "Overview" section containing text about the tools available for RNA-seq analysis. There is a "How to Use the RNA-seq Tools" section with compatibility notes and instructions for installing a local server. The "Internal Broad Institute Server" section provides information for members. The "Reference Genomes" section discusses pre-built indexes for various species. On the right side, there is a sidebar titled "Tuxedo Suite" listing several tools: BWA, Scripture, RNA-SeQC, Integrative Genomics Viewer (IGV), Picard, SAMtools, and ExprToGct. A link to "Back to the Top" is also provided.

RNA-seq Analysis

Overview

GenePattern offers a set of tools to support a wide variety of RNA-seq analyses, including short-read mapping, identification of splice junctions, transcript and isoform detection, quantitation, differential expression, quality control metrics, visualization, and file utilities. The tools released as GenePattern modules are widely-used. We continue to release new and updated tools as they become available. To be informed when new capabilities are added, check this page or sign up for our Twitter feed.

How to Use the RNA-seq Tools

We recommend that you run these modules on a local GenePattern server, due to the average size of the input files for these modules. You can upload your data, and make use of the new file management features in GenePattern 3.6, but large data will take a while to upload, depending on your connection speed, data size, and current available bandwidth. Alternately, on the public GenePattern server, if you have a [GenomeSpace](#) account, and already have data stored there, you can [link](#) your GenePattern account with your GenomeSpace account and make use of the improved file management features in GenePattern 3.6.

COMPATIBILITY NOTE: A number of tools are built for Unix-based (Mac and Linux) systems and will not run on Windows machines. They are the Tuxedo suite tools (Bowtie, TopHat, Cufflinks, Cuffmerge, Cuffcompare, and Cuffdiff) and BWA.

You can install a local GenePattern server by doing the following:

1. If you have not downloaded GenePattern and installed it on your local machine, follow the instructions on the [Download GenePattern](#) page.
2. If you have already downloaded and installed a GenePattern server, you can install any of these modules from the GenePattern public repository, available from [Modules & Pipelines: Install From Repository](#), in the navigation bar in your GenePattern server.
3. Enable the browsing of your GenePattern server's file path. This will allow you to send RNA-seq files to GenePattern modules without uploading them. See [these instructions](#) for more details.

Internal Broad Institute Server

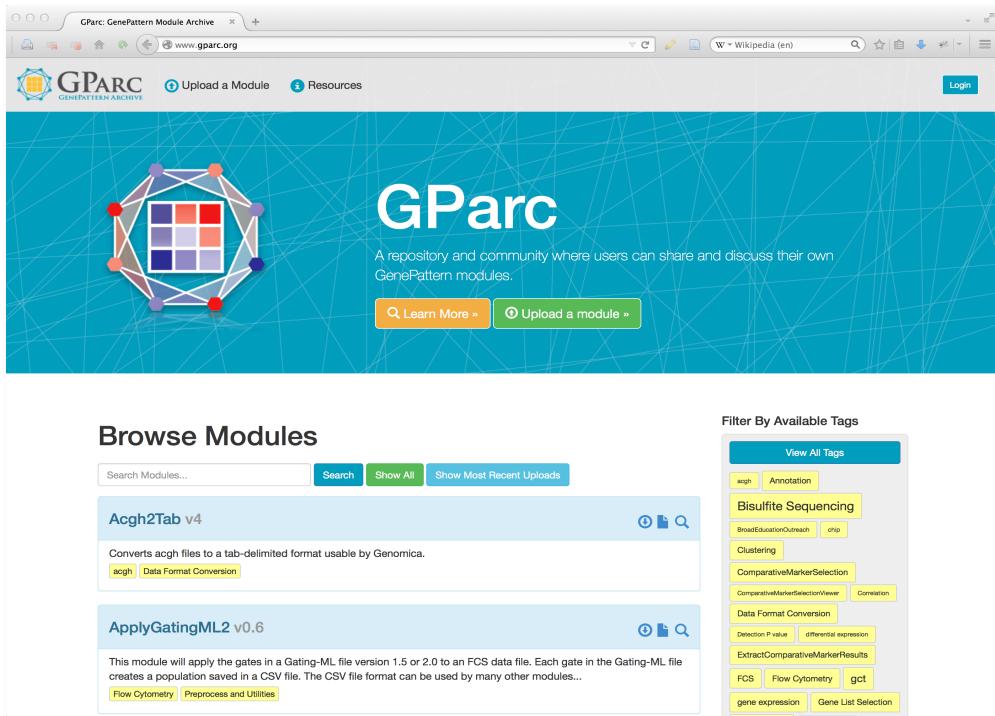
Broad Institute members and collaborators can use the [GPBroad](#) server to send RNA-seq files directly to analysis modules. Community members can contact gp-help@broadinstitute.org to enable access to their RNA-seq files.

Reference Genomes

The TopHat, Bowtie, and BWA GenePattern modules provide pre-built reference genome indexes for a number of species. If you need an index for a species that is not hosted, email us at gp-help@broadinstitute.org. See [this FAQ](#) for more information on how to find other reference genome indexes.

www.genepattern.org/rna-seq-analysis

GParc: the GenePattern Archive



- Find and download modules contributed by the worldwide GenePattern community
- Share your own modules easily
- Tag and comment on modules
- Get module updates

www.gparc.org

Resources

GenePattern

genepattern.org

Public GenePattern server

genepattern.broadinstitute.org

GenePattern Archive (GParc)

gparc.org

GenePattern Notebook

www.genepattern-notebook.org

Gene Set Enrichment Analysis (GSEA)

broadinstitute.org/gsea

GenomeSpace

genomespace.org

Integrative Genomics Viewer (IGV)

www.igv.org

Keep in touch!

Feature requests, bug reports, and
general help use our online forum

<https://groups.google.com/a/broadinstitute.org/forum/#!forum/gp-forum>

Mailing list to receive GenePattern news.
Sign up at www.genepattern.org/gp_mail.html

Follow us on Twitter: @GenePattern



Our Team

Peter Carr – Cambridge, MA

David Eby - Japan

Barbara Hill – Cambridge, MA

Arthur Liberzon – Cambridge, MA

Ted Liefeld – San Diego, CA

Clarence Mah – San Diego, CA

Michael Reich – San Diego, CA

Jim Robinson – San Diego, CA

Thorin Tabor – San Diego, CA

Edwin Juarez – San Diego, CA

Pablo Tamayo – San Diego, CA

Helga Thorvaldsdottir – Cambridge, MA

Douglass Turner – Cambridge, MA

PI

Jill P. Mesirov – San Diego, CA

