



The PATRIC Command Line Interface

THE P₃ SCRIPTS FOR ACCESSING PATRIC DATA AND
SERVICES

Where to Find Help

In HELP -> TUTORIALS find the main Command-Line Interface tutorial.

PATRIC Documentation

Search docs

CONTENTS

Website Tutorials

- Genomics
- Transcriptomics
- Protein Tools
- Metabolomics
- Users
- The PATRIC Command Line Interface**
- User Guides (ARCHIVE)
- PATRIC User Guides
- Webinars
- News
- Systems Architecture

Docs » Website Tutorials [Edit on GitHub](#)

Website Tutorials

Genomics

- Assembling a genome at PATRIC
- Genome Annotation at PATRIC
- Exploring a private genome in PATRIC
- BLAST Service
- Similar Genome Finder
- Submitting a Variant analysis job at PATRIC
- Submitting a Trn-Seq job at PATRIC
- Building Phylogenetic trees in PATRIC
- Using the PATRIC Metagenomic Binning Service
- Analyzing Your Metagenome Bins

Transcriptomics

- Expression Import
- Submitting an RNA-Seq job at PATRIC
- Examining Transcriptomics Data

Protein Tools

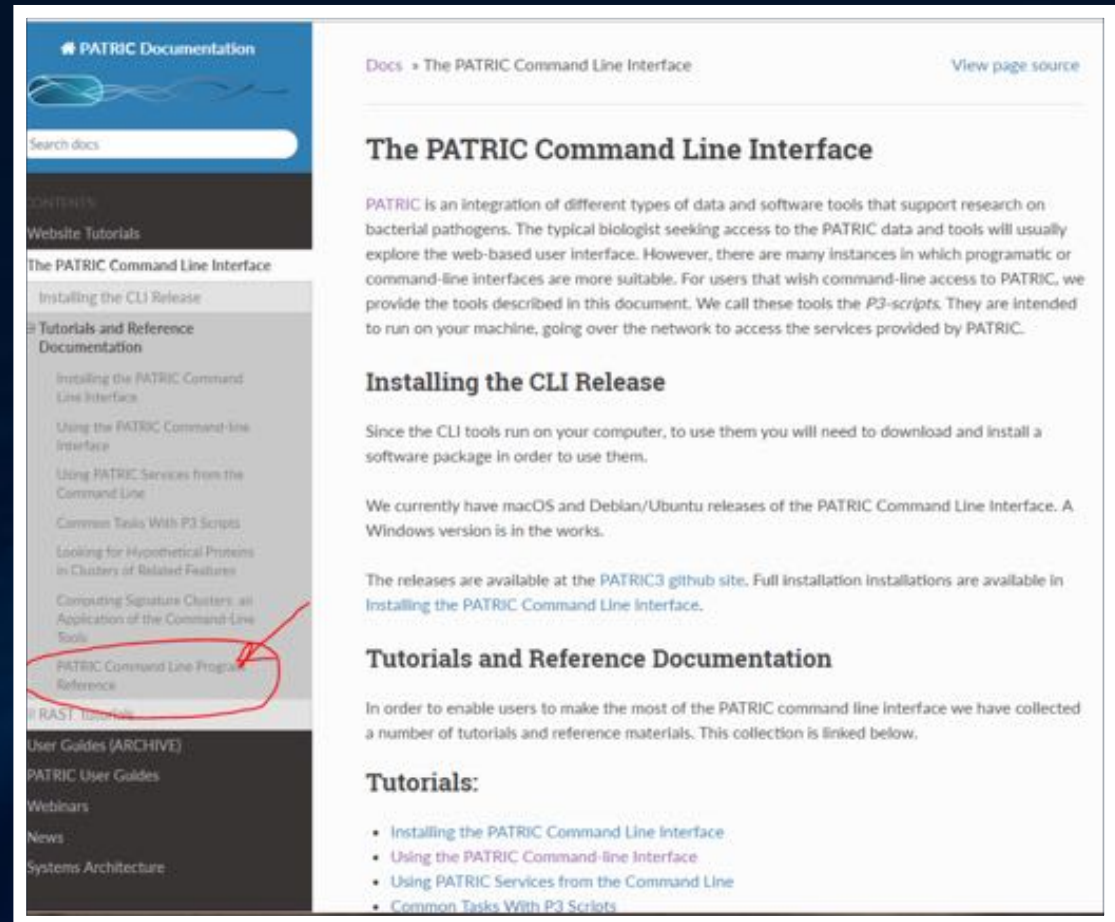
- Comparing Protein Families across Individual genomes and genome groups
- Proteome Comparison: Comparing annotated proteins across genomes.
- Formatting an Excel file to rapidly scan comparison genomes - Proteome Comparison

Metabolomics

- Comparing Pathways in PATRIC across Individual genomes or genome groups

Where to Find Help

Inside the main tutorial, find all commands listed in the Command Line Program Reference



The screenshot shows the PATRIC Documentation website. The left sidebar contains a navigation menu with the following items: PATRIC Documentation, Search docs, CONTENTS, Website Tutorials, The PATRIC Command Line Interface, Installing the CLI Release, Tutorials and Reference Documentation (highlighted with a red circle), PATRIC Command Line Program Reference (circled in red), RAST, User Guides (ARCHIVE), PATRIC User Guides, Webinars, News, and Systems Architecture. The main content area displays the page title 'The PATRIC Command Line Interface' and includes sections for 'Installing the CLI Release' and 'Tutorials and Reference Documentation'. A list of tutorials is provided at the bottom of the main content area.

PATRIC Documentation

Search docs

CONTENTS

Website Tutorials

The PATRIC Command Line Interface

Installing the CLI Release

Tutorials and Reference Documentation

Installing the PATRIC Command Line Interface

Using the PATRIC Command-line Interface

Using PATRIC Services from the Command Line

Common Tasks With P3 Scripts

Looking for Hypothetical Proteins in Clusters of Related Features

Computing Signature Clusters: an Application of the Command-Line Tools

PATRIC Command Line Program Reference

RAST

User Guides (ARCHIVE)

PATRIC User Guides

Webinars

News

Systems Architecture

Docs » The PATRIC Command Line Interface [View page source](#)

The PATRIC Command Line Interface

PATRIC is an integration of different types of data and software tools that support research on bacterial pathogens. The typical biologist seeking access to the PATRIC data and tools will usually explore the web-based user interface. However, there are many instances in which programmatic or command-line interfaces are more suitable. For users that wish command-line access to PATRIC, we provide the tools described in this document. We call these tools the *P3-scripts*. They are intended to run on your machine, going over the network to access the services provided by PATRIC.

Installing the CLI Release

Since the CLI tools run on your computer, to use them you will need to download and install a software package in order to use them.

We currently have macOS and Debian/Ubuntu releases of the PATRIC Command Line Interface. A Windows version is in the works.

The releases are available at the [PATRIC3 github site](#). Full installation instructions are available in [Installing the PATRIC Command Line Interface](#).

Tutorials and Reference Documentation

In order to enable users to make the most of the PATRIC command line interface we have collected a number of tutorials and reference materials. This collection is linked below.

Tutorials:

- [Installing the PATRIC Command Line Interface](#)
- [Using the PATRIC Command-line Interface](#)
- [Using PATRIC Services from the Command Line](#)
- [Common Tasks With P3 Scripts](#)

Where to Find Help

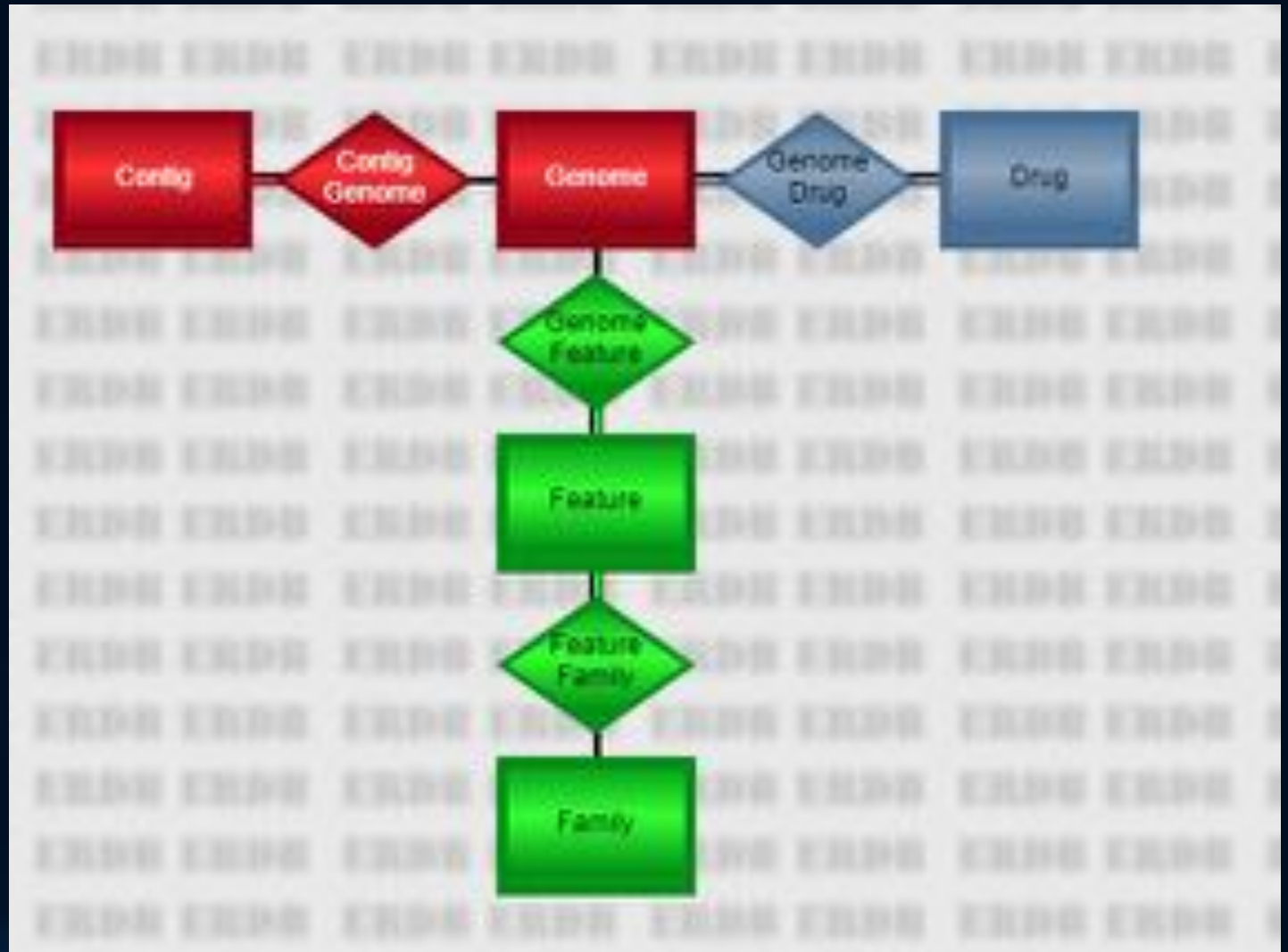
Use the --help option on the commands themselves.

```
C:\Users\Bruce\Documents\FIG\SEEDtk\Data> p3-match --help
p3-match.pl [-bchiv] [long options...] match-value
  -i STR --input STR      name of the input file (if not the
                           input)
  -c STR --col STR       column number (1-based) or name
  -b INT --batchSize INT input batch size
  --nohead              file has no headers
  -v --invert --reverse output non-matching records
  --discards STR       name of file to contain discarded
  -h --help             display usage information

C:\Users\Bruce\Documents\FIG\SEEDtk\Data>
```

The PATRIC Database

The command-line tools view the database as five heavily-indexed tables connected by relationships.



Genomes

- Our best estimate of the annotated DNA sequence for an organism.
- The gateway to all other tables.
 - Use **p3-all-genomes** to list all the genomes or a subset.
 - Use **p3-get-genome-features** to access the features of a genome.
 - Use **p3-get-genome-contigs** to access the genomes' sequences.
 - Use **p3-get-genome-drugs** to access drug resistance data about the genomes.

Drugs

- The secondary pathway into the tables.
- Represents an anti-microbial drug for which we have resistance data.
 - Use **p3-all-drugs** to get a list of drugs.
 - Use **p3-get-drug-genomes** to get the genomes known to be resistant or susceptible to a drug.

Common Notions

- All P₃ scripts operate on tab-delimited files with headers.
- Scripts are designed to be piped together.
- Common options include **--col** to specify the key column, **--nohead** if there are no headers, **--help** to get help.

Utility Scripts for Files with Headers

- `p3-extract` Select and re-order specific columns.
- `p3-sort` Sort by specified columns.
- `p3-match` Select records that possess (or do not possess) a particular value in a specified column.
- `p3-join` Horizontally join two files on a single key field.
- `p3-head` Display the first few lines of a file.
- `p3-echo` Create a small file.

p3-echo

- Helps to create small files without using a text editor or to start a pipe.
- Use the `-t` option to specify column headers. There will be one column per header.
- `p3-echo -t id 1313.7001 1313.7002`
produces one column with two records.
- `p3-echo -t id -t count 1313.7001 52`
produces two columns with one record.

A Sample Problem Solved by a Pipe

- Our query: How many genomes have a protein identical to “fig|210007.7.peg.1009”.
- This is Pantothenate kinase (coaA) in *Streptococcus mutans* UA159.
- How we solve this problem: every feature has an MD5 checksum for its protein. We find all proteins with the same checksum and count the genome IDs.

1: Get the MD5 for fig|210007.7.peg.1009

- `p3-echo "fig|210007.7.peg.1009" | p3-get-feature-data --attr aa_sequence_md5`
- The output

```
C:\Users\Bruce\Documents\FIG\SEEDtk\Data> p3-echo "fig|210007.7.peg.1009"  
id      feature.aa_sequence_md5  
fig|210007.7.peg.1009  6400069a6f7f32515c3a584ade0588d0
```

2: Count the Features with that MD5

- ```
p3-echo "fig|210007.7.peg.1009" |
p3-get-feature-data --attr aa_sequence_md5 |
p3-find-features aa_sequence_md5 --count
```

- The output

```
id feature.aa_sequence_md5 feature.count
fig|210007.7.peg.1009 6400069a6f7f32515c3a584ade0588d0 212
```

- This is not the right answer! We have the number of features with that protein, but not the number of genomes.

### 3: Count the Genomes with These Features

- ```
p3-echo "fig|210007.7.peg.1009" |  
p3-get-feature-data --attr aa_sequence_md5 |  
p3-find-features aa_sequence_md5  
                    --attr genome_id |  
p3-count genome_id
```

- The output

```
count  
153
```

- This is considerably less than 212.

Review of the Pipeline

- We used **p3-echo** to start with a feature ID.
- We used **p3-get-feature-data** to find the identifier for the feature's protein.
- We used **p3-find-features** to find the features with that identifier.
- We used **p3-count** to count the number of distinct genome IDs returned.

Accessing Your PATRIC Workspace

- Use **p3-login** to identify yourself.
- `p3-login rastuser25`
- `p3-login --status`
- `p3-login --logout`
- Logging in allows you to submit annotation jobs and access your genome and feature groups.

Creating a Genome Group

- ```
p3-echo penicillin |
p3-get-drug-genomes
 --eq "genome_name,Streptococcus pneumoniae"
 --resistant --attr genome_id,genome_name |
p3-put-genome-group --col genome_id
 strep.resist
```
- Puts all penicillin resistant Strep pneumoniae into the group strep.resist.
- ```
p3-get-genome-group strep.resist
```
- Lists all the genomes in the group.

Creating a Feature Group

- ```
p3-echo 171101.6 |
p3-get-genome-features
 --eq "product,hypothetical protein"
 --attr patric_id |
p3-put-feature-group sp.hypo
```
- Contains all the hypothetical proteins in Strep pneumoniae R6.
- ```
p3-get-feature-group sp.hypo
```
- Lists all the features in the group.

Annotating a Genome

- `p3-rast 100886`
 `"Catenibacterium from sample SRS576036"`
 `<sample.fna >test.gto 2>test.log`
- A new genome is created in your workspace.
- A copy, in JSON format, is put in **test.gto**.
- Any error and status messages appear in **test.log**.
- Note you have to give it a taxon ID and a name.

p3-rast Output

```
{
  "genetic_code" : "11",
  },
  "family_assignments" : [],
  "type" : "CDS",
  "id" : "fig|100886.26.peg.1540"
},
{
  "protein_translation" : "MLQIENASIAYGNDILFSGFNLQLERGEIASISGPGSGCGKSSLLNAILGFTPLKEGRIVLNGILLD",
  "aliases" : [],
  "location" : [
    [
      "100886.26.con.0010",
      "23684",
      "-",
      624
    ]
  ],
  "type" : "CDS"
},
"contigs" : [
  {
    "id" : "100886.26",
    "contigs" : [
      {
        "id" : "100886.26.con.0001",
```

The p3-rast Genomes

- Show up in **p3-all-genomes** output.
 - Use **--public** to avoid them.
 - Use **--private** to see only your private genomes.
- Can be viewed on the web site.