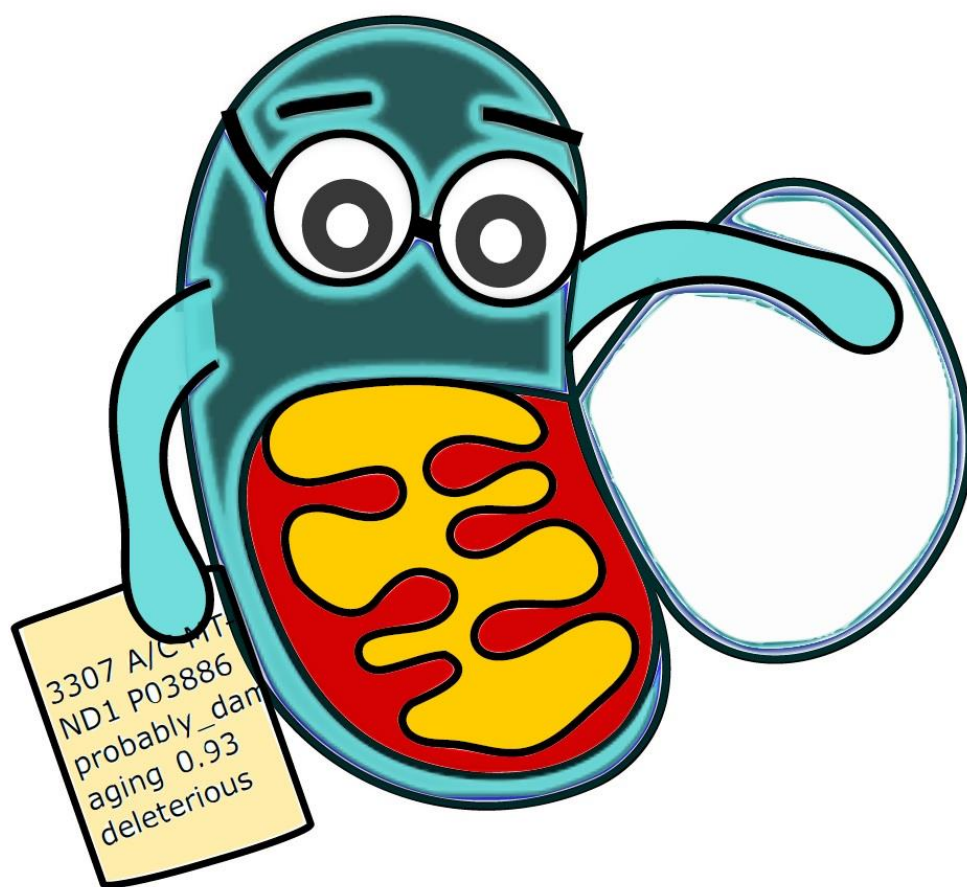


MITOLINK

User Manual version 1.0.2



Bioinformatics Centre, CSIR-Institute of Microbial Technology,
Chandigarh – 110036, India

Table of Contents

Introduction	3
I. Navigation Panel	5
1. Analyze Data	5
2. Workflow	5
3. Shared data library	7
4 Help	7
5. User	7
II. Tool Panel – MitoLink tools	8
1. Importing data to the MitoLink	8
2. Mito-LSDB Data –Query Builder	9
3. Variation analysis Tools	11
4. MtBrowse	15
5. Converters A. VCF like tabular to GFF.....	17
6. Mapping tools	20
7. Sequence Retrieval Tools	24
III. Detail Panel.....	27
IV. History Panel	28

Introduction

MitoLink is an integrated workflow system to facilitate understanding of genotype-phenotype correlations in cases of mitochondrial dysfunction. The workflow system is implemented using the open source workflow architecture, Galaxy. MitoLink is freely available and can be accessed at <http://ab-openlab.csir.res.in/mitolink>



Figure 1: MitoLink Overview page.

There are six modules in MitoLink, described in this manual. All the modules are accessible without registering to the system. However, for maintaining user-sessions it is recommended that anyone who is interested in creating data or task intensive workflow should register on https://ab-openlab.csir.res.in/mitolink_v1/login (Figure 2). The benefit of registration includes user sessions, saved histories, visualization, generation and execution of workflow and many others.

The screenshot shows the MitoLink Login Page. At the top, there is a navigation bar with links: Analyze Data, Workflow, Visualize, Shared Data, Help, and Login or Register. Below the navigation bar, there is a login form with the following elements:


- A welcome message: "Welcome to Galaxy, please log in".
- A text input field for "Public name or Email Address".
- A text input field for "Password".
- A link: "Forgot password? Click here to reset your password."
- A "Login" button.
- A link: "Don't have an account? Register here."

Figure 2: MitoLink Login Page

All the modules of MitoLink and some default modules by Galaxy are accessible through a web-based interface which has following components (Figure 3).

Navigation Panel: It provides the links to major components of the server like Tools Page (Analyze Data), Workflow System, Shared Libraries, Visualization, Help Section and User Login/Registration.

Tool Panel: This panel lists all the tools available in MitoLink along with default utilities in Galaxy.

Detail Panel (Canvas): This panel displays the interface of all the tools along with Input Parameters required to run a tool. It also provides help and examples to run a tool. This panel also displays the Output of a tool after its execution when user clicks on the eye  icon show in **History Panel**.

History Panel: This panel shows the information about the tools which are executed by a user. The information can include result after completion of a tool execution or error generated while running the tool. The workflow(s) are generated by extracting tasks from history panel.



Figure 3: MitoLink Galaxy Homepage

I. Navigation Panel

1. Analyze Data

The data analysis page is where everything happens. There, you can run any available tools on the data, run complete workflows, browse or download a result, and share files with other users. It is the default page when you open Galaxy in your browser, but you can also access it any time by clicking on "Analyze Data" in the Navigation Panel.



Figure 4: Analyze Data view and Tools

2. Workflow

Workflows are analyses that are intended to be executed (one or more times) with different user-provided input Datasets. Workflow can be reused over and over, not only reducing tedious work, but enhancing reproducibility by applying the same exact methods to all of your data. Workflow is nothing but creating pipeline, user can use it again and again or user can publish it.

Workflow can be created through navigation panel or from tool panel. In workflow section user can create workflow or can upload or import the workflow. The canvas is where inputs, tools, and noodles are added and connected as you build and modify your workflow by selecting 'Edit' opens the workflow editor view (Figure 5). The navigator provides a full view of your workflow in a condensed format (Figure 6) accessed by clicking on the gear icon on the right side of the center Workflow Canvas upper bar, the workflow editor menu (Figure 6) is for global editor actions. It consists of Save, Run, Edit Attributes, Auto re-layout, Close.

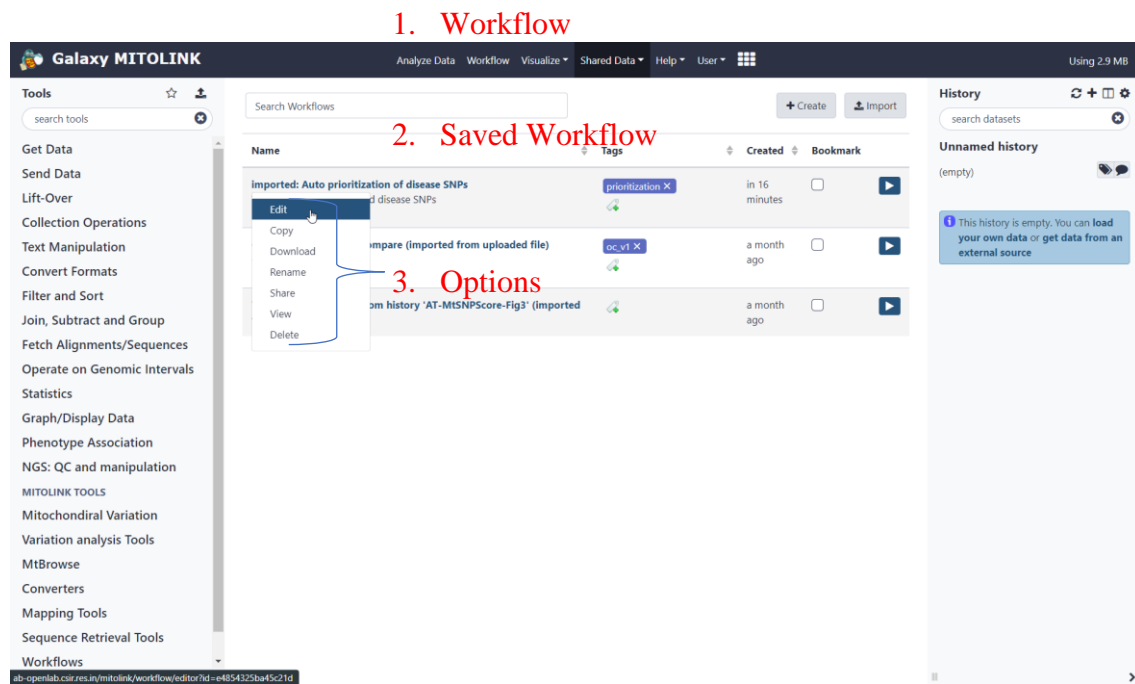


Figure 5: Options for workflow

The following example of workflow shows the “Auto prioritization of disease SNPs”. The Query Builder search result for Obese disease and Centenarian passed to the MtSNPscore. MtSNPscore generate result in five different form. Then we can do various analysis on generated result. Here we did filtration and comparison between Obese disease and Centenarian phenotype.

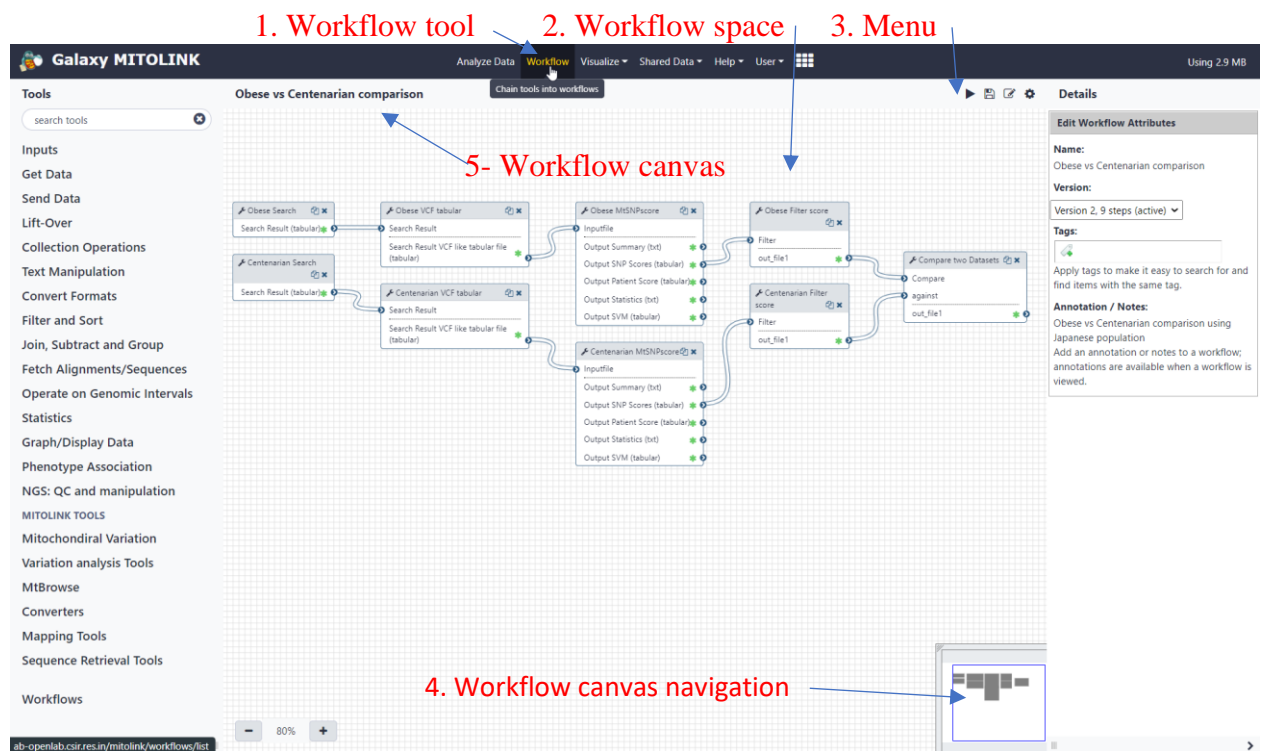


Figure 6: Workflow overview

3. Shared data library

Data libraries are collections of Datasets that are accessible from within a Galaxy instance. Libraries are designed for sharing datasets in between users or groups. The data library of

MitoLink consists of Mitochondrial protein list. Some of the actions that can be performed on data libraries are accessed by clicking the pop-up menu icon just right of the data library name.

- View Information –Shows the information about dataset.
- Import this dataset into your current history - this creates an item in your current history on which you can perform analysis. The item is a pointer to the library dataset disk file, so the file is not copied on disk.
- Download this dataset - this allows you to download a local copy of the dataset.

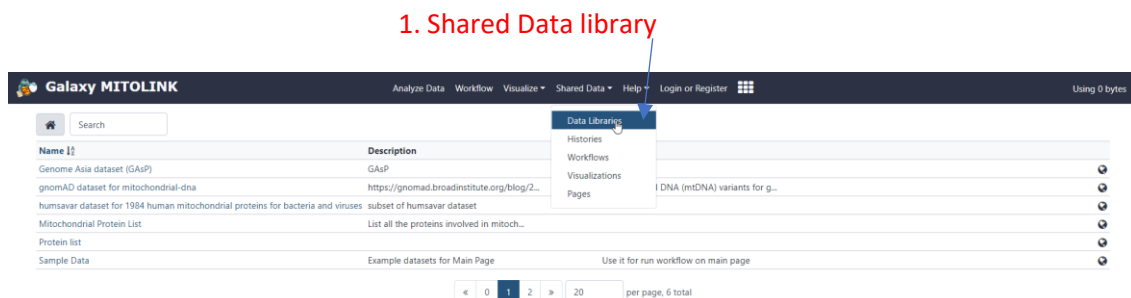


Figure 7: Data library view

4 Help

The help section of Galaxy consists of Support, Search, Mailing List, Videos, Wiki and How to cite Galaxy. User can find user manual in help section.

5. User

Login option and register option can get in user section. It is recommended that user register their account before using framework. Although unregistered users have access to tools available but their history is stored temporarily. On the other hand, registered users can save and retrieve their results in history panel later too.

II. Tool Panel – MitoLink tools

The user can get tools in tool panel. There are two categories of tools galaxy inbuilt tool and MitoLink tools. Galaxy tools consists of Data importing, Manipulation, Filtering, Sorting, Format conversion etc. MitoLink tool are specific for finding Genotype-Phenotype correlation.

1. Importing data to the MitoLink

A user can upload the data using the Galaxy tool **Get Data**. The uploaded data can be used for MtSNPscore analysis. There are many other services like UCSC browser, EBI ENA services, BioMart server, Flymine server etc. available in Get Data section which can be used for importing the data. If the data is not available in appropriate form, converter can be used. User can perform a query on the MitoLSDB data and pass the result to the MtSNPscore for further analysis. The following figure shows the file upload method.

1. Get data / Upload file

The screenshot displays the Galaxy MitoLink interface. On the left, a sidebar lists tool categories: Tools, Get Data, Send Data, Collection Operations, Lift-Over, Text Manipulation, Convert Formats, Filter and Sort, Join, Subtract and Group, Fetch Alignments/Sequences, Operate on Genomic Intervals, Statistics, Graph/Display Data, NGS: QC and manipulation, MITOLINK TOOLS, Mitochondrial Variation, Variation analysis Tools, Converters, Mapping Tools, Sequence Retrieval Tools, and DEPRECATED. The 'Get Data' section is active, showing 'Upload File from your computer' as the selected option. The central panel shows a table with one file: 'Sample_data_mtap6' (14.4 KB, Auto-detected, 100% status). Below the table are buttons for 'Choose local file', 'Paste/Fetch data', 'Pause', 'Reset', 'Start', and 'Close'. The right sidebar shows a 'History' panel with 'Unnamed history' and a list of datasets, including '1: Sample_data_mtap6_obese_step1.tsv'. A red text overlay 'Uploaded file' points to this dataset. The bottom of the screen contains a detailed description of the platform's purpose and a three-step pipeline for prioritizing SNPs.

Step 1 Input Data: The input data may either be searched using [Mito-LSDB query builder](#) tool or uploaded in VCF like format (please see file generated after Step2 for the accepted format). As a test case, search is performed using Mito-LSDB query builder to list variations in MTATP6 for obese phenotype: "Gene Like mtap6 AND Disease Equal to obese". The output of the search is: [Step 1 File](#)

Step 2 Format Conversion: In order to prioritize the variation data obtained in step one, we need to first convert the data to VCF like format (since MtSNPscore only takes VCF like format files). For conversion, [Search result to VCF like format](#) converter module is used. The following file contains the output after conversion to VCF like format. [Step 2 File](#)

Step 3 MtSNPscore Analysis: [MtSNPscore tool](#) (under Variation Analysis section), can be used to score SNPs using the file generated in previous step as input. This tool generates five output files. The file named "Output SNP scores" may be referred for accessing the scores for variations submitted to MtSNPscore. [Step 3 File](#)

Developed by: Rakesh Kumar, Neeraj Kumar Rajput, Bani Jolly, Amol Narwade, Anshu Bhardwaj, CSIR-Institute of Microbial Techno

Figure 8: Get Data screen

2. Mito-LSDB Data –Query Builder

MitoLSDB is an integrated platform to catalogue disease association studies on mtDNA. MitoLSDB work according to the Gene, variant DNA position, variant DNA wild-type, variant DNA muted, variant protein position, variant protein wild-type, variant protein muted, variant codon change wild-type, variant codon change mutated, variant frequency, disease and population. The generated result can be used as an input for MtSNPscore for further analysis.

1. Search

Galaxy MITOLINK

Analyze Data Workflow Visualize Shared Data Help User Using 246.7 KB

Tools

search tools

Convert Formats

Filter and Sort

Join, Subtract and Group

Fetch Alignments/Sequences

Operate on Genomic Intervals

Statistics

Graph/Display Data

NGS: QC and manipulation

MITOLINK TOOLS

Mitochondrial Variation

Query Builder perform simple or advance query on mitoLSDB data

Find SNP in Mito-LSDB searches the presence of given SNP in Mito-LSDB database

Variation analysis Tools

Converters

Mapping Tools

Sequence Retrieval Tools

DEPRECATED

Phenotype Association

Workflows

All workflows

Query Builder perform simple or advance query on mitoLSDB data (Galaxy Version 1.0.1)

Field: Gene

Field Operator: Equal to

Keyword: mitatp6

Add more Condition

1: Add more Condition

Connector Operator: AND

Field: Disease

Field Operator: Equal to

Keyword: Obese

Insert Add more Condition

Execute

TIP: Please use LIKE Field Operator when using Gene, Disease and Population in the Field value

Accepted Gene Names: MTND1, MTND2, MTND3, MTND4, MTND4L, MTND5, MTND6, MTCYB, MTCO1, MTCO2, MTCO3, MTATP6, MTATP8, MTND1, MTND2, MTND3, MTND4, MTND4L, MTND5, MTND6, MTCYB, MTCO1, MTCO2, MTCO3, MTATP6, MTATP8, MTND1, MTND2, MTND3, MTND4, MTND4L, MTND5, MTND6, MTCYB, MTCO1, MTCO2, MTCO3, MTATP6, MTATP8

History

search datasets

imported: Example Workflow for 'Obese_AZ_MTATP6_comparison'

17 shown

232.08 KB

8: Output Statistics

7: Output Patient Score

6: Output SNP Scores

5: Output Summary

4: Search Result

3: Search Result VCF like tabular file

2: Search Result

1: Search Result

213 lines

format: tabular, database: ?

1	2	3
Gene Name	Variant Pos	Position Variant Pos
MTATP6	8563	A
MTATP6	8563	A
MTATP6	8563	A
MTATP6	8563	A

Figure 9: Query Builder tool view

3. Variation analysis Tools

Variation analysis tools can be used for Nuclear Genome and Mitochondrial Genome Variation Analysis.

A. PolyPhen

PolyPhen is a nuclear genome variant analysis tool. PolyPhen is a tool which predict possible impact of amino acid substitution using structural and evolutionary consideration for all mitochondrial proteins. User can get list mitochondrial protein which are mitochondrial encoded and mitochondrial protein which are nuclear encoded in the shared data library.

1. Polyphen tool

The screenshot shows the Galaxy MITOLINK interface for the Polyphen Mitochondrial mutation tool (Galaxy Version 0.1.3). The interface is divided into three main sections:

- Left Sidebar (Tools):** Contains a search bar and a list of tools categorized under "MITOLINK TOOLS" and "NUCLEAR GENOME VARIATION ANALYSIS". The "Polyphen Mitochondrial mutation" tool is highlighted.
- Main Form Area:** Contains the following fields:
 - UniProt ID***: A text input field with the example "A6NC05". A red annotation "2. UniProt ID is mandatory" points to this field.
 - Substitution position**: A text input field with the example "92".
 - Wild-type amino acid**: A text input field with the example "L". A red annotation "3. Optional user input" points to this field.
 - Mutant amino acid**: A text input field with the example "M".
 - Execute Button**: A blue button with a checkmark icon and the text "Execute".
 - INFO**: A blue information icon followed by the text "This tool gives the polyphen result on all mitochondrial proteins present in shared library".
- Right Sidebar (History):** Contains a search bar and a section titled "Example1" with the text "(empty)". A red annotation "4. Output" points to this section. A blue information icon followed by the text "This history is empty. You can load your own data or get data from an external source" is also present.

Figure 11: PolyPhen tool view

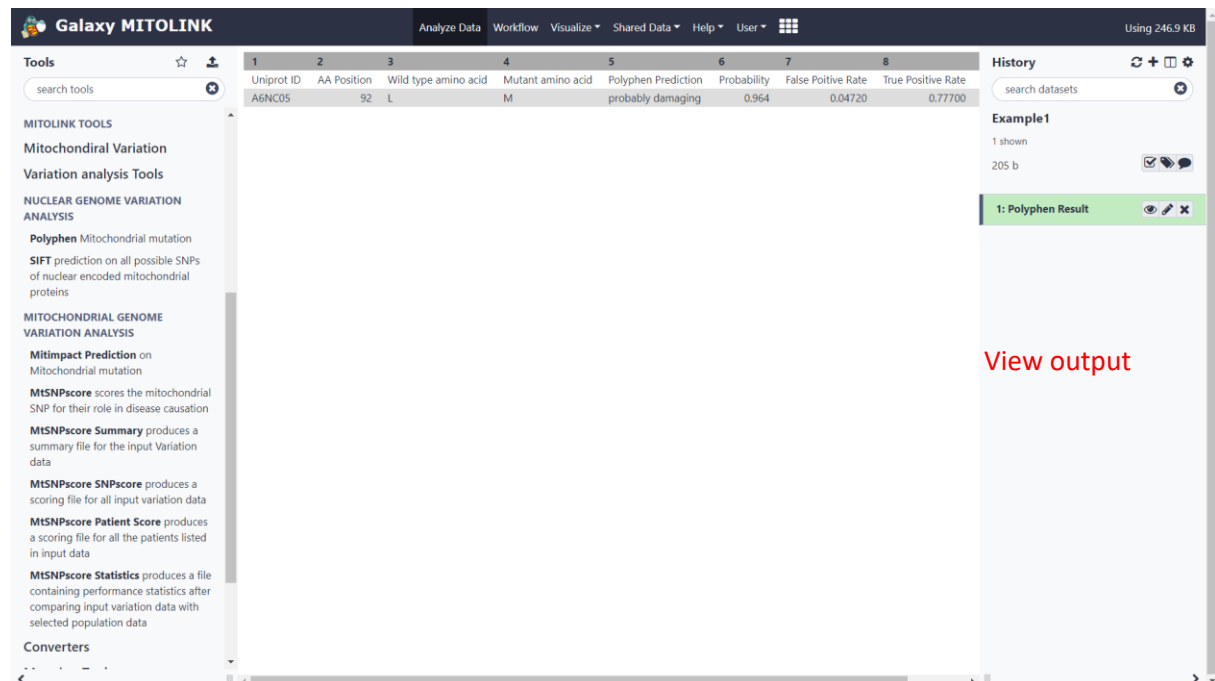


Figure 12: Polyphen output

B. MitImpact Prediction

MitImpact is a mitochondrial variant analysis tool. MitImpact is a collection of pre-computed pathogenicity predictions for all nucleotide changes that cause non-synonymous substitutions in human mitochondrial protein coding genes. It counts 24,115 amino acid variations in all the 13 coding genes of the Homo sapiens mitochondrion (cf. NCBI Reference Sequence: NC_012920.1). The input option for this tool is genomic position. According to the genomic position it show result for nucleotide substitution, and gives score like PolyPhen2, SIFT, FatHmm score, CADD score etc.

1. MitImpact prediction tool

The screenshot shows the Galaxy MITOLINK interface. The top navigation bar includes 'Analyze Data', 'Workflow', 'Visualize', 'Shared Data', 'Help', and 'User'. The left sidebar lists various tools under 'MITOLINK TOOLS', including 'Mitochondrial Variation', 'Variation analysis Tools', 'NUCLEAR GENOME VARIATION ANALYSIS', 'MITOCHONDRIAL GENOME VARIATION ANALYSIS', and 'Converters'. The main panel displays the 'MitImpact Prediction on Mitochondrial mutation (Galaxy Version 0.1.2)' tool. The 'Genomic Position' field is set to 9038. Below the field, it says 'Allowed single positions (Insert a number between 3307 and 15887)'. An 'Execute' button is present. An 'INFO' message states: 'This tool gives the MitImpact prediction on protein encoding genes of mitochondrial genome'. The right sidebar shows the 'History' panel with a search bar and a list of datasets. The 'Example1' dataset is selected, showing '2 shown' and '1.6 KB'. Below the history, there are two tabs: '2: MitImpact Result' and '1: Polyphen Result'.

Figure 13: MitImpact prediction tool view

The screenshot shows the Galaxy MITOLINK interface with the 'MitImpact Prediction on Mitochondrial mutation (Galaxy Version 0.1.2)' tool. The 'Genomic Position' field is set to 9038. The 'Execute' button is present. An 'INFO' message states: 'This tool gives the MitImpact prediction on protein encoding genes of mitochondrial genome'. The right sidebar shows the 'History' panel with a search bar and a list of datasets. The 'Example1' dataset is selected, showing '2 shown' and '1.6 KB'. Below the history, there are two tabs: '2: MitImpact Result' and '1: Polyphen Result'. The '2: MitImpact Result' tab is active, showing '3 lines, 2 comments' and 'format: tabular, database: ?'. The output is displayed in a table with columns: Genomic_position, NT_sub, Gene_symbol, OXPHOS_Complex, Ensembl_Gene_ID, Ensembl_Protein_ID, Uniprot_Name, Uniprot_ID, and NCBI. The output shows two rows of data for the position 9038.

1	2	3	4	5	6	7	8	9
Genomic_position	NT_sub	Gene_symbol	OXPHOS_Complex	Ensembl_Gene_ID	Ensembl_Protein_ID	Uniprot_Name	Uniprot_ID	NCBI
9038	T/C	MT-ATP6	V	ENSG00000198899	ENSP00000354632	ATP6_HUMAN	P00846	
9038	T/A	MT-ATP6	V	ENSG00000198899	ENSP00000354632	ATP6_HUMAN	P00846	

Figure 14: MitImpact prediction output

C. MtSNPscore tool

MtSNPscore is comprehensive weighted scoring system used for identification of mtDNA variations that can impact pathogenicity and would likely be associated with disease. The tool generates five outputs as follows: MtSNPscore summary which summarize the input data. MtSNPscore SNP scores provides detailed analysis result. It consists of Sample ID, SNPsites, Locus, local position, protein position, SIFT tolerated or not etc. MtSNPscore Patients score provide score for patient, MtSNPscore statistics provide general statistics of the output, statistics only reported when rCRS is not selected as Normal Population. MtSNPscore SVM file provide output in SVM readable format, SVM file also only generated when rCRS is not selected as Normal Population

1. MtSNPscore tool

The screenshot displays the Galaxy MITOLINK interface for the MtSNPscore tool. The main configuration area includes the following fields:

- Inputfile:** 3: Search Result VCF like tabular file (Annotated with red text: 2. Search result or uploaded VCF file)
- Select normal population from this list:** Cent (Annotated with red text: 3. Population)
- SIFT Intolerant:** 5
- PolyPhen Possibly damaging:** 3
- PolyPhen Probably damaging:** 5
- Invariant Peptides Present:** 6 (Annotated with red text: 4. Parameters)
- HRE Present:** 8
- HRE Absent:** 1
- Allele frequency/ Mutation D > N:** 1
- Allele frequency/ Mutation D < N:** 1
- H-strand Promoter Present:** (checkbox)

The right sidebar shows the **History** panel with a list of datasets. The top entry is 'Imported: Example Workflow for 'Obese_AZ_MTATP6_comparison'' with 17 shown datasets and a size of 232.08 KB. Below this, a list of outputs is shown, including '8: Output Statistics', '7: Output Patient Score', '6: Output SNP Scores', '5: Output Summary', '4: Search Result VCF like tabular file', '3: Search Result VCF like tabular file', '2: Search Result', and '1: Search Result'. The '5: Output Summary' is highlighted with a red box and annotated with red text: 5. Output.

Figure 15: MitSNPscore tool view

4. MtBrowse

MtBrowse is a combination of database and interactive web pages for displaying annotations on genomes. It is used to display detailed view of genome. The analysis result generated by MitoLSDB, variation analysis tools can be display in MtBrowse by importing it in GFF format.

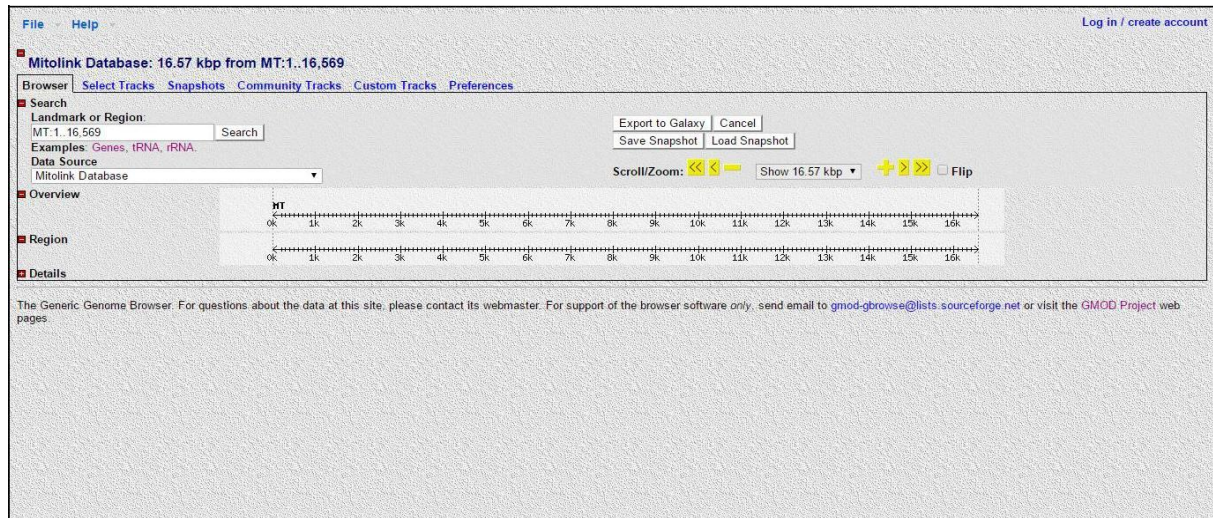


Figure 16: MtBrowse view

For Annotation user have to upload data from Custom track. Data can be upload by entering the data in text box or from URL of by direct file upload option. GBrowse accepts custom track data in a variety of formats including BED, GFF, and GFF3. For dense quantitative data use Wiggle (WIG) format.

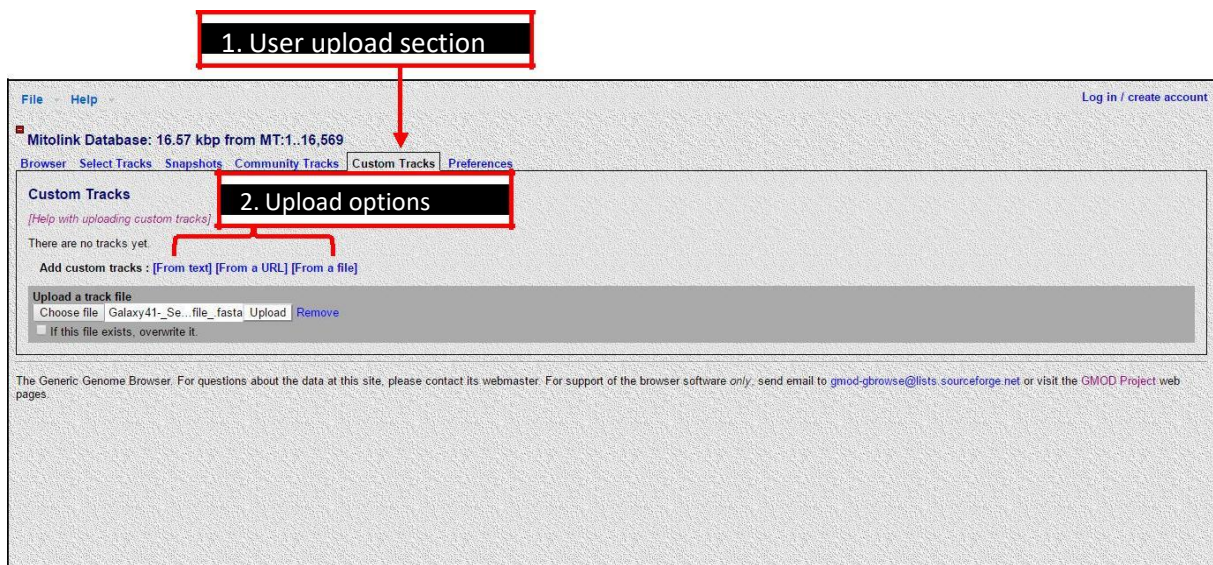


Figure 17: Custom track options

1. Browse the annotation on genome

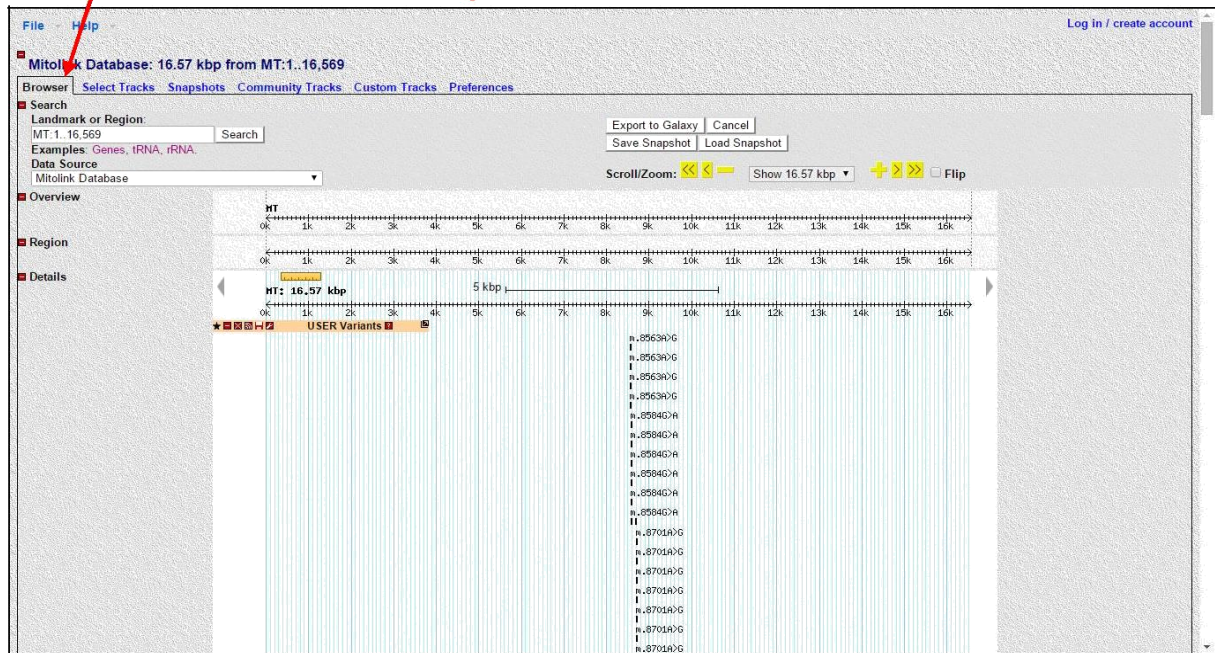


Figure 18: Browse panel view

To select the tracks you would like to have displayed in your MtBrowse genome viewer, go to the 'Select Tracks' tab at the top of the MtBrowse window and choose your tracks of interest. Then hit the "Back to Browser" links at the top and bottom of the page take you back to the MtBrowse genome viewer, which will now have all of the selected tracks displayed.

1. Customize track selection

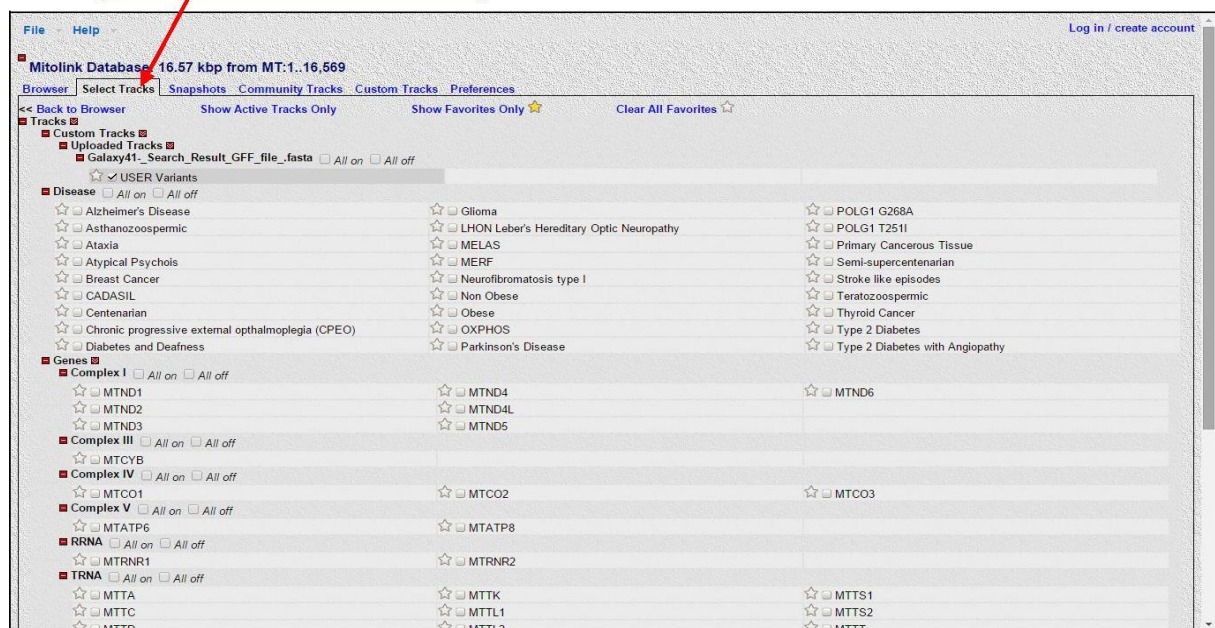


Figure 19: Select track view

5. Converters A. VCF like tabular to GFF

Converts a VCF file to GFF file format. Converted GFF file can be downloaded and used as a custom track on the MtBrowse for visualization.

1. VCF to GFF converter

The screenshot displays the Galaxy / MITOLINK interface for the VCF to GFF converter. The left sidebar lists various tools, with 'Converters' expanded to show 'VCF to GFF converts a vcf file to gff file format'. The central panel shows the tool's execution interface, including an 'Inputfile VCF file' dropdown menu with '1: test_1.vcf' selected. The right sidebar shows the 'History' panel with '2: Output GFF file' listed. Red annotations highlight these key elements: '1. VCF to GFF converter' points to the tool in the left panel; '2. User Input' points to the 'Inputfile VCF file' dropdown; '3. Output' points to the '2: Output GFF file' entry in the history panel.

Figure 20: VCF to GFF converter view

B. Search result to GFF

Converts the output of MitoLSDB search into GFF file format. It takes, output of a search on MitoLSDB data as the input. Converted GFF file can be downloaded and used as a custom track on the MtBrowse for visualization.

1. Search output to GFF converter

The screenshot displays the Galaxy / MITOLINK web interface. The left sidebar lists various tool categories: Get Data, Lift-Over, Text Manipulation, Filter and Sort, Join, Subtract and Group, Convert Formats, Extract Features, Fetch Sequences, Fetch Alignments, Statistics, Graph/Display Data, MITOLINK TOOLS, Mito-LSDB Data, Variation analysis Tools, MBrowse, Converters, Mapping Tools, Sequence Retrieval Tools, and Workflows. The 'Converters' section is expanded, showing 'VCF to GFF converts a vcf file to gff file format', 'Search result to GFF converts the output of Search into gff file format', and 'Search result to VCF converts the output of Search into VCF file format'. A red arrow points from the 'Search result to GFF' tool to the main panel. The main panel shows the tool's title, a description, and an input field labeled '1: Search Result' with an 'Execute' button. A red box labeled '2. User Input' points to this field. The right sidebar shows a history of runs, with '2: Search Result GFF file' and '1: Search Result' listed. A red box labeled '3. Output' points to the '1: Search Result' entry.

Figure 21: Search result to GFF converter view

C. Search result to VCF like tabular format

Converts the output of Search into VCF like tabular file format. It takes, output of a search on MitoLSDB data as the input. Converted VCF file can be used as input to the MtSNPscore tool for analysis.

1. Search output to VCF converter

The screenshot displays the Galaxy / MITOLINK web interface. On the left, a 'Tools' sidebar lists various categories including 'Get Data', 'Text Manipulation', 'Filter and Sort', 'Join, Subtract and Group', 'Convert Formats', 'Fetch Sequences', 'Fetch Alignments', 'Statistics', 'Graph/Display Data', 'MITOLINK TOOLS', 'Mito-LSDB Data', 'Variation analysis Tools', 'MBrowse', 'Converters', 'Mapping Tools', 'Sequence Retrieval Tools', and 'Workflows'. A red arrow points from the '1. Search output to VCF converter' label to the 'Search result to VCF' tool in the 'Converters' section.

The main panel shows the configuration for the 'Search result to VCF' tool (Galaxy Tool Version 0.1.0). The 'Search Result' input field is set to '3: Search Result VCF file', which is highlighted by a red box and labeled '2. User input'. Below this is an 'Execute' button and an information message: 'INFO: Please provide output of a search on MitoLSDB data as the input. Converted VCF file can be used as input to the MtSNPscore tool for analysis.'

On the right, the 'History' panel shows a list of datasets. The top entry is '3: Search Result VCF file', which is highlighted by a red box and labeled '3. Output'. Below it is '1: Search Result'. A red arrow points from the '3. Output' label to the '3: Search Result VCF file' entry.

Figure 22: Search result to VCF converter view

6. Mapping tools

Mapping tools is useful for gene search, genomic position from gene location, reference nucleotide search, and mitochondrial protein search.

A. Gene Mapping

Gene mapping tool searches a genomic position from a gene location for mitochondrial protein.

1. Gene mapping tool

The screenshot shows the Galaxy MITOLINK interface. The main workspace contains the following elements:

- Header:** Galaxy MITOLINK, Analyze Data, Workflow, Visualize, Shared Data, Help, User, and a grid icon. The user is using 248.3 KB.
- Left Sidebar:** Tools (search tools), Fetch Alignments/Sequences, Operate on Genomic Intervals, Statistics, Graph/Display Data, NGS: QC and manipulation, MITOLINK TOOLS, Mitochondrial Variation, Variation analysis Tools, Converters, Mapping Tools (Gene Mapping, Genomic to Gene Mapping, Find reference nucleotide, Find Mitochondrial Protein), Sequence Retrieval Tools, DEPRECATED, Phenotype Association, Workflows, All workflows.
- Main Workspace:**
 - Gene Mapping searches a genomic position from a gene location (Galaxy Version 1.0.1)**
 - Select a gene:** A dropdown menu with "MT-ATP6" selected. An annotation "2. Select gene from drop down menu" points to this field.
 - Gene Location:** An input field with "681" entered. An annotation "3. Location of gene" points to this field.
 - Execute:** A button to run the tool.
 - USAGE:** Select a gene and provide gene position for which you want the genomic position.
 - TIP:** Use the following table for gene sizes. Please enter Gene location <= Gene Size.
 - Gene Name Gene Size:** A list of mitochondrial genes and their sizes:
 - MT-RNR1 954
 - MT-RNR2 1559
 - MT-TF 71
 - MT-TV 69
 - MT-TL1 75
 - MT-TI 69
 - MT-TQ 72
 - MT-TM 68
 - MT-TW 68
 - MT-TA 69
 - MT-TN 73
 - MT-TC 66
 - MT-TV 66
 - MT-TS1 69
 - MT-TD 68
 - MT-TK 70
 - MT-TG 68
 - MT-TR 65
 - MT-TH 69

- Right Sidebar:**
- History:** search datasets, 2 shown, 1.6 KB.
- Example1:** 2: MitImpact Result, 1: Polyphen Result.

Figure 23: Gene mapping tool view

Galaxy MITOLINK

Analyze Data Workflow Visualize Shared Data Help User

Using 266.1 KB

Tools

search tools

Get Data

Send Data

Collection Operations

Lift-Over

Text Manipulation

Convert Formats

Filter and Sort

Join, Subtract and Group

Fetch Alignments/Sequences

Operate on Genomic Intervals

Statistics

Graph/Display Data

NGS: QC and manipulation

MITOLINK TOOLS

Mitochondrial Variation

Variation analysis Tools

Converters

Mapping Tools

Sequence Retrieval Tools

DEPRECATED

Phenotype Association

Workflows

Genomic Position: 9287

Reference Nucleotide on this position from rCRS sequence: A

History

search datasets

Example1

8 shown, 1 deleted

4.96 KB

9: Reference Sequence

8: Reference Gene Sequence

7: Mitochondrial Protein Search

6: Reference Nucleotide

5: Genomic Location

3: Genomic Location

83 bytes

format: fasta, database: ?

display with IGV local

Genomic Position: 9287

Reference Nucleotide on this position from rCRS sequence: A

2: Mitimpact Result

1: Polyphen Result

View output

Figure 24: Gene mapping tool output

B. Genomic to Gene Mapping

Genomic to Gene mapping tool uses nucleotide genomic position along with Wild type and Mutated nucleotide in HGVS format and tool will display its gene location, gene name and other information.

1. Genomic to gene mapping

Galaxy MITOLINK

Analyze Data Workflow Visualize Shared Data Help User

Using 266.1 KB

Tools

search tools

Get Data

Send Data

Collection Operations

Lift-Over

Text Manipulation

Convert Formats

Filter and Sort

Join, Subtract and Group

Fetch Alignments/Sequences

Operate on Genomic Intervals

Statistics

Graph/Display Data

NGS: QC and manipulation

MITOLINK TOOLS

Mitochondrial Variation

Variation analysis Tools

Converters

Mapping Tools

Sequence Retrieval Tools

DEPRECATED

Phenotype Association

Workflows

Genomic to Gene Mapping using the Genomic position (Galaxy Version 1.0.1)

Enter Genomic position of Mitochondrial DNA

m.A8527G

Use HGVS format. Ex. m.T14484C

Execute

USAGE: Please provide the genomic position of a nucleotide along with Wild type and Mutated nucleotide in HGVS format and tool will display its gene location, gene name and other information

History

search datasets

Example1

8 shown, 1 deleted

4.96 KB

9: Reference Sequence

8: Reference Gene Sequence

7: Mitochondrial Protein Search

6: Reference Nucleotide

5: Genomic Location

3: Genomic Location

2: Mitimpact Result

1: Polyphen Result

2. User input

3. Output

Figure 25: Genomic to gene mapping tool view

Galaxy MITOLINK

Tools

search tools

Get Data

Send Data

Collection Operations

Lift-Over

Text Manipulation

Convert Formats

Filter and Sort

Join, Subtract and Group

Fetch Alignments/Sequences

Operate on Genomic Intervals

Statistics

Graph/Display Data

NGS: QC and manipulation

MITOLINK TOOLS

Mitochondrial Variation

Variation analysis Tools

Converters

Mapping Tools

Sequence Retrieval Tools

DEPRECATED

Phenotype Association

Workflows

172.16.1.44:8080/datasets/7566c8b77c2a3e5b/display?preview=True

Entered Position: 8527
Entered Wild Type Nucleotide: A
Entered Mutated Nucleotide: G
['MATP6', 'MATP6'] [162, 1]

Gene Name: MATP6, MATP6
Position of Nucleotide on Gene: 162, 1
Position of Amino Acid: 55, 1
Nucleotide on Reference Sequence: A, A
Reported Mutated Nucleotide: G
Codon before Mutation: AAA, ATG
Codon after Mutation: AAG, GTG
Amino Acid before Mutation: K, M
Amino Acid after Mutation: K, V

History

search datasets

Example1

8 shown, 1 deleted

4.96 KB

9: Reference Sequence

8: Reference Gene Sequence

7: Mitochondrial Protein Search

6: Reference Nucleotide

405 bytes
format: fasta, database: ?
display with IGV local

Entered Position: 8527
Entered Wild Type Nucleotide: A
Entered Mutated Nucleotide: G
['MATP6', 'MATP6'] [162, 1]

3: Genomic Location

2: MitImpact Result

1: Polyphen Result

View output

Figure 26: Genomic to gene mapping tool output

C. Reference Nucleotide search

Find reference nucleotide tool searches the Nucleotide present in rCRS sequence. It takes input as genomic position and returns the Nucleotide present at that position in Reference (rCRS) genome

1. Find reference nucleotide tool

Galaxy MITOLINK

Find reference nucleotide searches the Nucleotide (Version 1.0.1)

Enter Genomic position of Mitochondrial DNA

8527

Ex. 1000. Please enter position between 1 to 16569

Execute

USAGE: Takes the genomic position and returns the Nucleotide present at that position in Reference (rCRS) genome

History

search datasets

Example1

5 shown, 1 deleted

2.52 KB

6: Reference Nucleotide

43 bytes
format: fasta, database: ?
display with IGV local

Nucleotide at 8527 in reference genome is A

5: Gene Position

3: Genomic Location

2: MitImpact Result

1: Polyphen Result

2. Genomic position

3. Output

Figure 27: Reference nucleotide searching tool

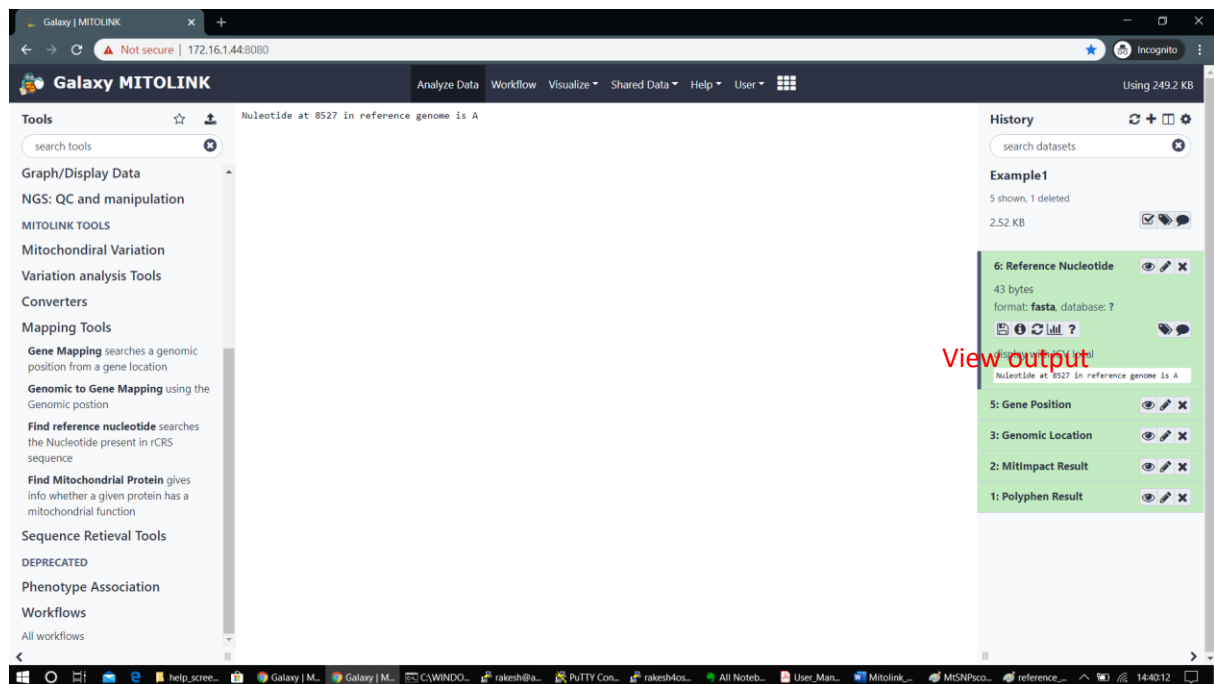


Figure 28: Reference nucleotide search output

D. Find Mitochondrial Protein

The tool searches the entered protein with in-house mitochondrial protein library and gives the various evidence for mitochondrial function along with various other information.

1. Find Mitochondrial protein

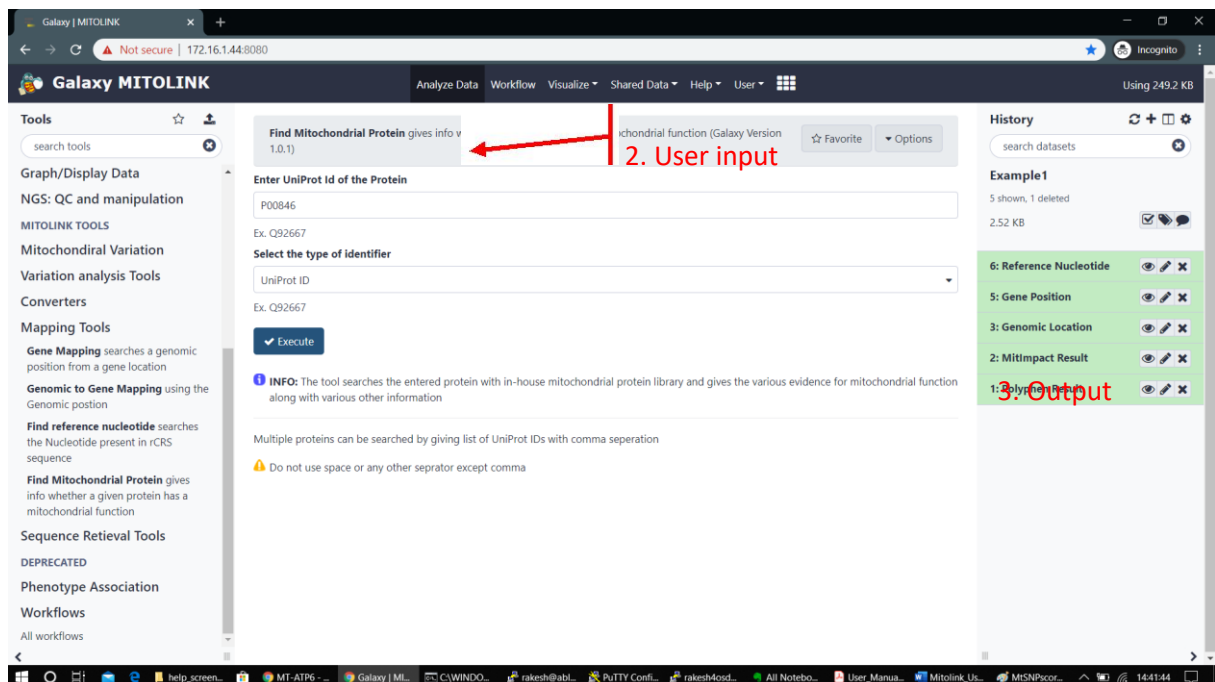


Figure 29: Mitochondrial protein search tool view

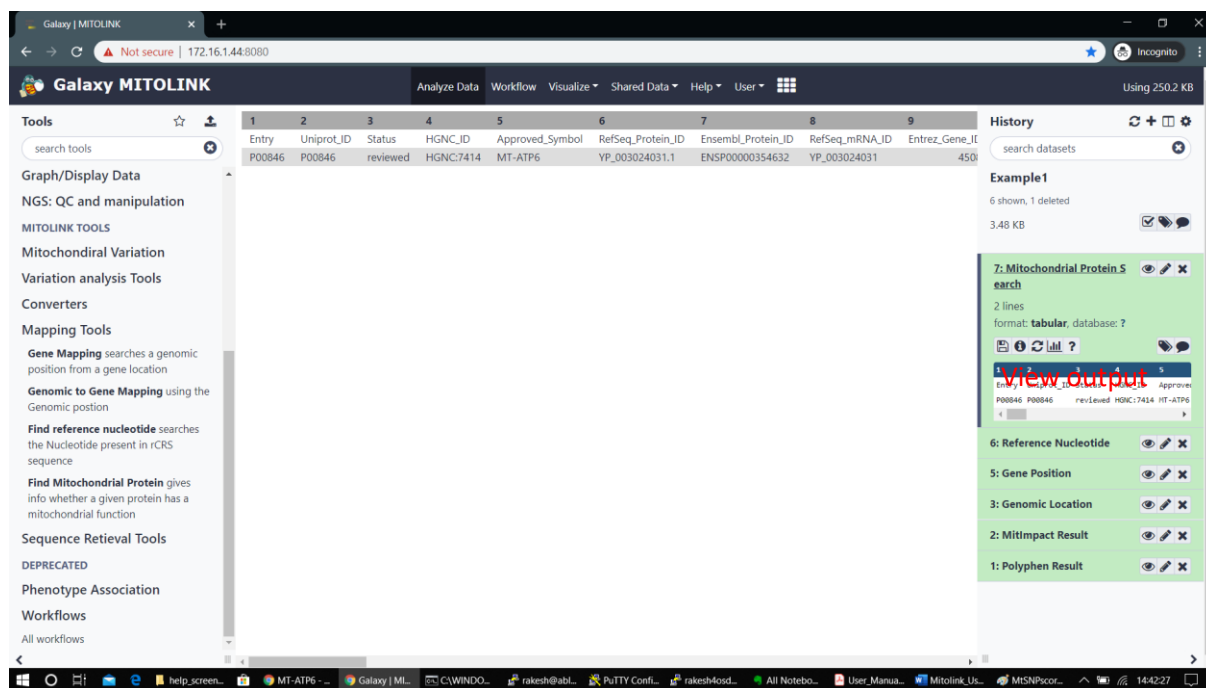


Figure 30: Mitochondrial protein search output

7. Sequence Retrieval Tools

A. Gene sequence retrieval

The tool will return the gene sequence from reference (rCRS) genome for selected gene.

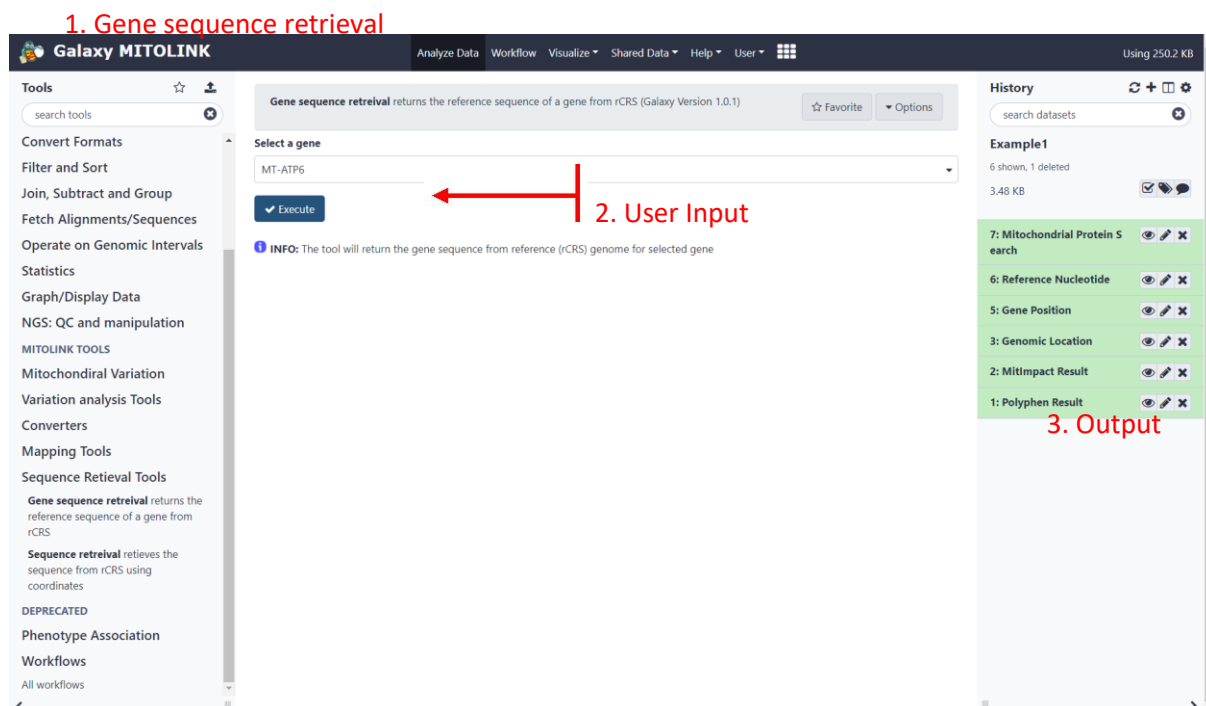


Figure 31: Gene sequence retrieval tool

Galaxy MITOLINK

Analyze Data Workflow Visualize Shared Data Help User

Using 250.9 KB

Tools

search tools

Convert Formats

Filter and Sort

Join, Subtract and Group

Fetch Alignments/Sequences

Operate on Genomic Intervals

Statistics

Graph/Display Data

NGS: QC and manipulation

MITOLINK TOOLS

Mitochondrial Variation

Variation analysis Tools

Converters

Mapping Tools

Sequence Retrieval Tools

Gene sequence retrieval returns the reference sequence of a gene from rCRS

Sequence retrieval retrieves the sequence from rCRS using coordinates

DEPRECATED

Phenotype Association

Workflows

All workflows

History

search datasets

Example 1

7 shown, 1 deleted

4.21 KB

8: Reference Gene Sequence

1 sequences

format: fasta, database: ?

display with IGV local

View output

Figure 32: Gene sequence retrieval output

B. Sequence retrieval

The tool will retrieve the sequence starting and ending with the given coordinates from reference (rCRS) genome

1. Sequence retrieval tool

Galaxy MITOLINK

Analyze Data Workflow Visualize Shared Data Help User

Using 250.9 KB

Tools

search tools

Convert Formats

Filter and Sort

Join, Subtract and Group

Fetch Alignments/Sequences

Operate on Genomic Intervals

Statistics

Graph/Display Data

NGS: QC and manipulation

MITOLINK TOOLS

Mitochondrial Variation

Variation analysis Tools

Converters

Mapping Tools

Sequence Retrieval Tools

Sequence retrieval retrieves the sequence from rCRS using coordinates (Galaxy Version 1.0.1)

Starting coordinates

8527

End coordinates

9207

Execute

INFO: The tool will retrieve the sequence starting and ending with the given coordinates from reference (rCRS) genome

History

search datasets

Example 1

7 shown, 1 deleted

4.21 KB

8: Reference Gene Sequence

7: Mitochondrial Protein Sequence

6: Reference Nucleotide

5: Gene Position

3: Genomic Location

2: MitImpact Result

1: Polyphen Result

Output

Figure 33: Sequence retrieval tool view

Galaxy MITOLINK

Analyze Data Workflow Visualize Shared Data Help User

Using 251.6 KB

Tools

search tools

Convert Formats

Filter and Sort

Join, Subtract and Group

Fetch Alignments/Sequences

Operate on Genomic Intervals

Statistics

Graph/Display Data

NGS: QC and manipulation

MITOLINK TOOLS

Mitochondrial Variation

Variation analysis Tools

Converters

Mapping Tools

Sequence Retrieval Tools

DEPRECATED

Phenotype Association

Workflows

All workflows

>NC_012920.1 NC_012920.1 Homo sapiens mitochondrion, complete genome

ATGAGCAAAATCTGTTGGCTTCATTGATGCCCCCAAACTCTAGGCTACCGCCGGA
GTACTGATCATTCTATTTCCCTCTATTGATCCCACTTCAAAATATCTCATCAACAG
CGACTAATCACCCCAACATGACTAATCAAACTAACCTCAAAACAAATGATAACCAT
CACAACTAAAGGACGAACCTGATCTCTTATAGTATGCTTAAATCATTTTATGEC
ACACATAACCTCTCGGACTCTGCTCTACTATTACACCAACCACTATCTATA
AACCTAGCCATGEGCATCCCTTATAGCGGGCACAGTGATTATAGGCTTTGGCTAAG
ATTAATAATGCCCTAGGCCACTTCTACCAAGGACACCTACACCTTATCCCAT
CTAGTTATATCGAAACCATCAGCTACTCTAACCAATAGCCCTGGCGGTACGCTA
ACCGTAAACATTACTGAGGACCTACTCATGCACTAATTGGAGGCGCACCTAGCA
ATATCAACCATTAACCTTCCCTCTACACTTATCATCTTCACAATCTAATCTACTGACT
ATCTAGAAATCGCTGCTGCTTAAATCAAGCTTACGTTTTACACTCTAGTAAGGCTC
TACCTGCACGACACACATAA

History

search datasets

Example 1

8 shown, 1 deleted

4.96 KB

9: Reference Sequence

1 sequences

format: fasta, database: ?

display with IGV local

NC_012920.1 NC_012920.1 Homo sapiens mitocho
ATGAGCAAAATCTGTTGGCTTCATTGATGCCCCCAAACTCTAGGCTACCGCCGGA
GTACTGATCATTCTATTTCCCTCTATTGATCCCACTTCAAAATATCTCATCAACAG
CGACTAATCACCCCAACATGACTAATCAAACTAACCTCAAAACAAATGATAACCAT
CACAACTAAAGGACGAACCTGATCTCTTATAGTATGCTTAAATCATTTTATGEC
ACACATAACCTCTCGGACTCTGCTCTACTATTACACCAACCACTATCTATA
AACCTAGCCATGEGCATCCCTTATAGCGGGCACAGTGATTATAGGCTTTGGCTAAG
ATTAATAATGCCCTAGGCCACTTCTACCAAGGACACCTACACCTTATCCCAT
CTAGTTATATCGAAACCATCAGCTACTCTAACCAATAGCCCTGGCGGTACGCTA
ACCGTAAACATTACTGAGGACCTACTCATGCACTAATTGGAGGCGCACCTAGCA
ATATCAACCATTAACCTTCCCTCTACACTTATCATCTTCACAATCTAATCTACTGACT
ATCTAGAAATCGCTGCTGCTTAAATCAAGCTTACGTTTTACACTCTAGTAAGGCTC
TACCTGCACGACACACATAA

8: Reference Gene Sequence

7: Mitochondrial Protein Search

6: Reference Nucleotide

5: Gene Position

3: Genomic Location


2: Mitimpact Result

1: Polyphen Result

View output

Figure 34: Sequence retrieval output

III. Detail Panel

This panel displays the interface of all the tools along with Input Parameters required to run a tool. It also provides help and examples to run a tool. This panel also displays the Output of a tool after its execution when user clicks on the eye  icon show in **History Panel**.

IV. History Panel

When data is uploaded from your computer or analysis is done on existing data using Galaxy, each output from those steps generates a dataset. These datasets (and the output datasets from later analysis on them) are stored by Galaxy in **Histories**.

Users that have registered an account and logged in can have many histories and the history panel allows switching between them and creating new ones.

The screenshot displays the Galaxy MITOLINK interface. The top navigation bar includes 'Analyze Data', 'Workflow', 'Visualize', 'Shared Data', 'Help', and 'User'. The left sidebar lists various tools and workflows. The main area is titled 'Saved Histories' and contains a table of history entries. A search bar is at the top of this section. The table has columns for 'Name', 'Items', 'Datasets', 'Tags', and 'Sharing'. The entries are:

Name	Items	Datasets	Tags	Sharing
Unnamed history	1	1		
Example1	9	8	1	
imported: Example Workflow for 'Obese_AZ_MTATP6_comparison'	17	17		history
Unnamed history	1	1		14.4 KB, 4 hours ago, 3 hours ago

Below the table are buttons for 'Rename', 'Delete', 'Delete Permanently', and 'Undelete'. A note states: 'Histories that have been deleted for more than a time period specified by the Galaxy administrator(s) may be permanently deleted.'

On the right, the 'History' panel shows a search bar and a list of datasets. The selected history is 'imported: Example Workflow for 'Obese_AZ_MTATP6_comparison'', which contains 17 datasets. The first dataset is '8: Output Statistics' (232.08 KB). Below the list, a table shows the first few lines of the selected dataset:

Gene Name	Variant	DBA	Position	Variant	DBA
MTATP6	8563		A		
MTATP6	8563		A		
MTATP6	8563		A		
MTATP6	8563		A		

Figure 35: History Panel view