

Introduction to STRaM Analysis Workflow in Galaxy (STRaM Loci Ver. 18)

Li Binglin
2025.05

1. Apply for a web or lab Galaxy platform account

The screenshot shows the Galaxy web platform interface. At the top, there is a dark blue header with the Galaxy logo on the left and navigation links on the right: a grid icon, a graduation cap icon, a hexagon icon, a question mark icon, a user profile icon labeled "Using 0 b", and a "Login or Register" link.

The main content area is divided into two panels. The left panel, titled "Welcome to Galaxy, please log in", contains a login form with the following elements:

- A red box labeled "Login" with an arrow pointing to the "Public Name or Email Address" input field.
- A "Public Name or Email Address" input field.
- A "Password" input field.
- A small "or" text between the email and password fields.
- A link: "Forgot password? Click here to reset your password."
- A "Login" button.
- A link at the bottom: "Don't have an account? Register here."

The right panel, titled "Create a Galaxy account", contains a registration form with the following elements:

- A blue informational box at the top: "Please register only one account. The usegalaxy.org service is provided free of charge and has limited computational and data storage resources. **Registration and usage of multiple accounts is tracked and such accounts are subject to termination and data deletion.**"
- A "Create a Galaxy account" header.
- An "Email address" input field.
- A "Password" input field.
- A "Confirm password" input field.
- A "Public name" input field.
- A note: "Your public name is an identifier that will be used to generate addresses for information you share publicly. Public names must be at least three characters in length and contain only lower-case letters, numbers, dots, underscores, and dashes (., _ , -, ' ')."
- A checkbox: "Stay in the loop and join the galaxy-announce mailing list. This list is used for important Galaxy updates and newsletter access. We keep it streamlined, you should expect only 2-3 emails per month."
- A "Create" button.
- A link at the bottom: "Already have an account? Log in here."

[Galaxy](https://usegalaxy.org/) of Web Platform (<https://usegalaxy.org/>)

2. Import STRaM analysis workflow

Galaxy

Upload

Tools

Workflows

Workflow Invocations

Visualization

Histories

History Multiview

Datasets

Pages

Libraries

Workflows

My workflows

Workflows shared with me

Public workflows

Search my workflows by query or use the advanced filtering options

Sort by: Name Update time Filter: Show deleted Show bookmarked

Display: Grid List

No workflows found. You may create or import new workflows using the buttons above.

Create

Import

2.1 Click the "Workflow" button.

2.2 Click the "Import" button.

Import workflow

Archived file or url

GA4GH servers

TRS ID

Import from a Galaxy workflow export URL or a workflow file

Archived Workflow URL

If the workflow is accessible via a URL, enter the URL above and click Import.

Archived Workflow File

STRaM analysis workflow_v26.ga

Browse

If the workflow is in a file on your computer, choose it and then click Import.

Import workflow

2.3 Browse and select the latest STRaM workflow file and then click "Import workflow".

Workflows

My workflows

Workflows shared with me

Search my workflows by query or use the advanced filtering options

Sort by: Name Update time Filter: Show deleted Show bookmarked

edited in less than a minute

never run

STRaM analysis workflow v26 (imported from uploaded file)

STR analysis with sequence flank analysis

Add Tags

Share

Download

Edit

Run

2.4 Workflow file will be uploaded successfully and displayed.

3. Upload all datasets into a new history in Galaxy Server

Galaxy

Using 68% of 250.0 GB

libinglin

Upload

Tools

Workflows

Workflow Invocations

Visualization

Histories

History Multiview

Datasets

Pages

Libraries

Galaxy is an open source, web-based platform for data intensive biomedical research. If you are ne

You can install your own Galaxy by following the tutorial and choose from thousands of tools from

3.1 Create a new analysis history and assign a name, such as "STRaM analysis"

3.2 Click "Upload" button in the left menu bar..

Upload from Disk or Web to STRaM analysis

Regular Composite Collection Rule-based

You added 6 file(s) to the queue. Add more files or click 'Start' to proceed.

| | | | | | | |
|-------------------|----------|-------------|---|-----------------|----|--|
| AT104-11-LHD22477 | 372.8 MB | Auto-detect | Q | unspecified (?) | 0% | |
| AT104-11-LHD22477 | 357.6 MB | Auto-detect | Q | unspecified (?) | 0% | |
| GCA_000001405.1 | 832.5 MB | Auto-detect | Q | unspecified (?) | 0% | |
| STRaM_v18XY_flank | 2.7 KB | Auto-detect | Q | unspecified (?) | 0% | |
| STRaM_v18XY_flank | 2.6 KB | Auto-detect | Q | unspecified (?) | 0% | |
| STRaM_markers_v18 | 876 b | Auto-detect | Q | unspecified (?) | 0% | |

Type (set all): Auto-detect Reference (set all): unspecified (?)

Choose local file

Choose remote files

Paste/Fetch data

Start

Pause

Reset

Close

3.4 Click "Start" to begin the upload.

3.5 Files will upload successfully and appear in your "STRaM analysis" history.

History

search datasets

Unnamed history

0 B

STRaM analysis

0 B

This history is empty.

STRaM analysis

1.64 GB

6: STRaM_markers_v18_list_v21.tabular

5: STRaM_v18XY_flank5_v1.fasta

4: STRaM_v18XY_flank3_v1.fasta

3: GCA_000001405.15_GRCh38_no_alt_analysis_set.fna.gz

2: AT104-11-LHD22477_L2_2 fq.gz

1: AT104-11-LHD22477_L2_1 fq.gz

Note: The Raw data file should be uploaded before the reference genome file.

4. Run STRaM analysis workflow

The image shows two screenshots of the Galaxy web interface. The top screenshot displays the 'Workflows' page with a red box highlighting the 'STRaM analysis workflow v26' entry. A red arrow points to the 'workflow runs: 2' button, and another red arrow points to the 'Run' button. The bottom screenshot shows the workflow execution page with six input fields, each with a blue arrow pointing to a description of the input data.

Galaxy Workflows

My workflows | Workflows shared with me | Public workflows

Search my workflows by query or use the advanced filtering options

Sort by: Name | Update time | Filter: Show deleted | Show bookmarked

Display: [icon] [icon]

edited about 3 hours ago

STRaM analysis workflow v26

STR analysis with sequence flank analysis

Add Tags

Copy | Download | Share | Edit | Run

workflow runs: 2

Galaxy

Workflow: STRaM analysis workflow v26 (Version: 17)

edited about 3 hours ago

STR analysis with sequence flank analysis

Run Workflow

workflow runs: 2

Input paired-end read 1

1: AT104-11-LHD22477_L2_1.fq.gz

Raw data: Paired-end read 1

Input paired-end read 2

2: AT104-11-LHD22477_L2_2.fq.gz

Raw data: Paired-end read 2

GCA_000001405.15_GRCh38_no_alt_analysis_set.fna.gz

3: GCA_000001405.15_GRCh38_no_alt

Reference human genome hg38 (GRCh38_no_alt_analysis_set)

STRaM_v10XY_flank3_v3.fasta

4: STRaM_v18XY_flank3_v1.fasta

3' reference flanking sequences of loci (STRaM_v18XY_flank3_v1.fasta)

STRaM_v10XY_flank5_v3.fasta

5: STRaM_v18XY_flank5_v1.fasta

5' reference flanking sequences of loci (STRaM_v18XY_flank5_v1.fasta)

STRaM_markers_v10_list_v16.tabular

6: STRaM_markers_v18_list_v21.tabular

Marker name and genomic locations (STRaM_markers_v18_list_v21.tabular)

5. Running

Galaxy

Upload

Tools

Workflows

Workflow Invocations

Visualization

Histories

History Multiview

5.1 Running

Invoked Workflow: STRaM analysis workflow v26 (Version: 17)

invoked about 3 hours ago

History: STRaM analysis

44 of 44 steps successfully scheduled.

38 of 38 jobs complete.

Overview

Steps

Inputs

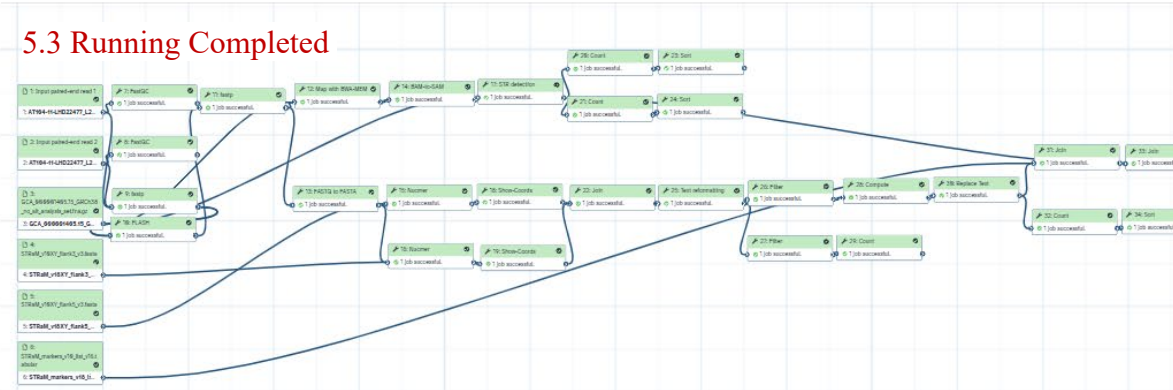
Outputs

Report

Export

Metrics

5.3 Running Completed



24: Nucmer on data 22 and data 5: alignment

23: BAM-to-SAM on data 2 1: converted SAM

22: FASTQ to FASTA on data 19

21: Map with BWA-MEM on data 19 and data 3 (mapped reads in BAM format)

20: fastp on data 14: HTML report

19: fastp on data 14: Read 1 output




29: Count on data 26

Add Tags

An error occurred with this dataset.

format **tabular**, database ?

Count of unique values in c1_cn_4_cn_5_cn_7_cn_10_cn_12 cut: invalid field value



Grey: Pending;
Orange: Running;
Green: Success;
Red: Error.

5.2 Troubleshooting

Note: Since the web platform is updated regularly, some tools may encounter errors. Troubleshooting is required for these tools according to error reports.

6. Result output

Galaxy

Upload

Tools

Workflows

Workflow Invocations

Visualization

| Column 1 | Column 2 | Column 3 | Column 4 | Column 5 | Column 6 | Column 7 | Column 8 | Column 9 | Column 10 | Column 11 | Column 12 | Column 13 |
|-----------|----------|----------|----------|----------|-----------|--------------|--------------|----------|-----------|-----------|-------------------------|-----------|
| 131468 | 44 | GGAA | 0 | chr1 | 77620496 | Chr1:D1S1611 | Chr1:D1S1611 | 44. | 77620484 | GGAA | 7.729917706007783e-08 | True |
| 24495 | 44 | GGAA | 0 | chr1 | 77620484 | Chr1:D1S1611 | Chr1:D1S1611 | 44. | 77620484 | GGAA | 0.0 | True |
| 5966 | 40 | GGAA | 0 | chr1 | 77620500 | Chr1:D1S1611 | Chr1:D1S1611 | 40. | 77620484 | GGAA | 1.0306556675780927e-07 | True |
| 169644 | 36 | TCTA | 0 | chr2 | 206150019 | Chr2:D2S1782 | Chr2:D2S1782 | 36. | 206150024 | TCTA | -1.2127090658586227e-08 | True |
| 15747 | 36 | TCTA | 0 | chr2 | 206150023 | Chr2:D2S1782 | Chr2:D2S1782 | 36. | 206150024 | TCTA | -2.4254181081866333e-09 | True |
| 144141 | 44 | TGAA | 0 | chr2 | 1489642 | Chr2:TPOX | Chr2:TPOX | 44. | 1489650 | TGAA | -2.6852017190661407e-06 | True |
| 24263 | 44 | TGAA | 0 | chr2 | 1489650 | Chr2:TPOX | Chr2:TPOX | 44. | 1489650 | TGAA | 0.0 | True |
| 6100 | 40 | TGAA | 0 | chr2 | 1489646 | Chr2:TPOX | Chr2:TPOX | 40. | 1489650 | TGAA | -1.3425990569584225e-06 | True |
| 5073 | 44 | TGAA | 0 | chr2 | 1489641 | Chr2:TPOX | Chr2:TPOX | 44. | 1489650 | TGAA | -3.0208529478993494e-06 | True |
| 17910908 | 40 | CTAT | 0 | chr2 | 17910900 | CTAT | CTAT | 40. | 17910908 | CTAT | 5.583186746765297e-07 | True |
| 17910908 | 56. | CTAT | 0 | chr2 | 17910908 | CTAT | CTAT | 56. | 17910908 | CTAT | 1.1166378481048702e-07 | True |
| 17910908 | 40. | CTAT | 0 | chr2 | 17910908 | CTAT | CTAT | 40. | 17910908 | CTAT | 0.0 | True |
| 31302797 | 36. | ATCT | 0 | chr2 | 31302797 | ATCT | ATCT | 36. | 31302797 | ATCT | 6.389205003680342e-08 | True |
| 31302797 | 36. | ATCT | 0 | chr2 | 31302797 | ATCT | ATCT | 36. | 31302797 | ATCT | 0.0 | True |
| 31302797 | 32. | ATCT | 0 | chr2 | 31302797 | ATCT | ATCT | 32. | 31302797 | ATCT | 1.2778409190921923e-07 | True |
| 150076321 | 44. | CTAT | 0 | chr2 | 150076321 | CTAT | CTAT | 44. | 150076321 | CTAT | 3.99796564850631e-08 | True |
| 150076321 | 52. | CTAT | 0 | chr2 | 150076321 | CTAT | CTAT | 52. | 150076321 | CTAT | 1.3326552516881693e-08 | True |
| 150076321 | 44. | CTAT | 0 | chr2 | 150076321 | CTAT | CTAT | 44. | 150076321 | CTAT | 0.0 | True |
| 123775551 | 52. | ATCT | 0 | chr2 | 123775551 | ATCT | ATCT | 52. | 123775551 | ATCT | 0.0 | False |
| 123775551 | 52. | ATCT | 0 | chr2 | 123775551 | ATCT | ATCT | 52. | 123775551 | ATCT | 4.0395699630552646e-09 | False |
| 123775551 | 48. | ATCT | 0 | chr2 | 123775551 | ATCT | ATCT | 48. | 123775551 | ATCT | 1.6158279656403555e-08 | False |
| 67841988 | 2 | TCTA | 0 | chr2 | 67841988 | TCTA | TCTA | 2 | 67841988 | TCTA | 5.8960532880579334e-08 | True |
| 67841988 | 2 | TCTA | 0 | chr2 | 67841988 | TCTA | TCTA | 2 | 67841988 | TCTA | 0.0 | True |

- Column 1 = reads count
- Column 2 = length of STR (bp)
- Column 3= repeat motif (bp)
- Column 4 = hamming distance
- Column 5 = chromosome
- Column 6 = left flanking region stop
- Column 7 = Marker name
- Column 8 = Marker name
- Column 9 = STR length of STR flank analysis
- Column 10 = the start coordinates of the STR sequence in the genome
- Column 11 = repeat motif
- Column 11 = The difference in the STR start positions
- Column 11 = STR length comparison of the STR length

6.1 Download the output tabular.

6.2 Perform subsequent analysis to distinguish between allele and stutter signals.

6.3 STRaM profiles collection and sample assessment