

Logowanie

Welcome to Galaxy, please log in

Public Name or Email Address

Password

[Forgot password? Click here to reset your password.](#)

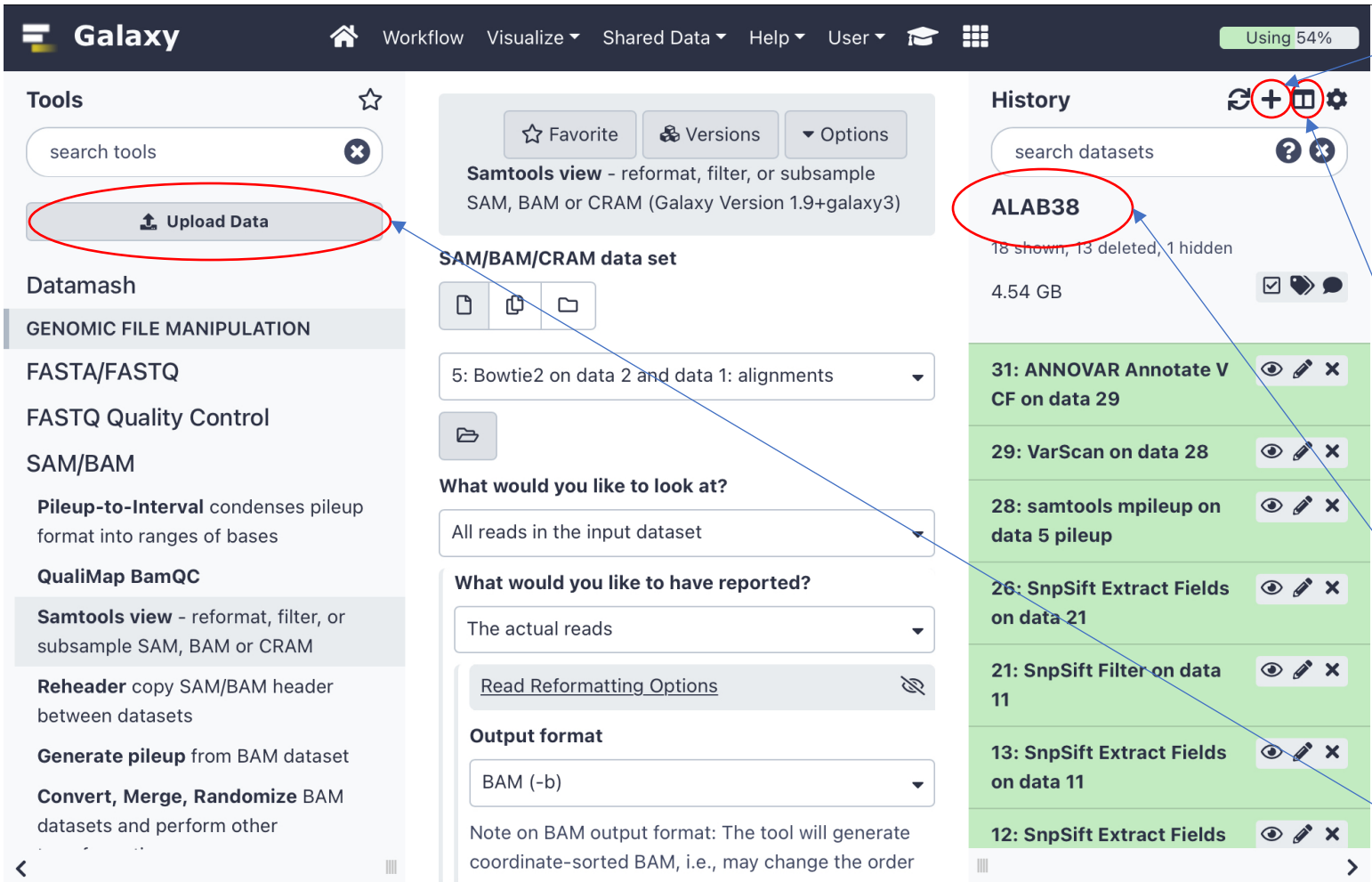
Don't have an account? [Register here.](#)

← Login

← Hasło

← Rejestracja nowego użytkownika

Obszary robocze



Nowa historia

Lista historii

Nazwa historii

Import nowych danych

Narzędzia

Obszar roboczy opcji

Historia

Import danych

Galaxy

Workflow Visualize Shared Data Help User

Using 54%

Download from web or upload from disk

Regular Composite Collection Rule-based

You added 1 file(s) to the queue. Add more files or click 'Start' to proceed.

Name	Size	Type	Genome	Settings	Status
New File	37 b	Auto-de...	----- Additional ...		0%

Download data from the web by entering URLs (one per line) or directly paste content.

Http://przyklad.pl/sekwencja.fastq.gz

Type (set all): Auto-detect Genome (set all): ----- Additional ...

Choose local files Choose remote files Paste/Fetch data Start Pause Reset Close

12: SnpSift Extract Fields

1. Naciśnij „Paste/Fetch data”.
2. Skopiuj link do danych.
3. Naciśnij „START”.
4. Zamknij okno „Close”

Analiza jakości (fastqc)

FastQC Read Quality reports (Galaxy Version 0.72+galaxy1)

Favorite

Versions

Options

Short read data from your current history

3: Bowtie2 on 95

Contaminant list

3: Bowtie2 on 95

95_R2.fastq.gz

tab delimited file with

1: 95_R1.fastq.gz

Adapter list

Nothing selected

list of adapters adapter sequences which will be explicitly searched against the library, tab delimited file with 2 columns: name and sequence. (--adapters)

Submodule and Limit specifying file

Nothing selected

a file that specifies which submodules are to be executed (default=all) and also specifies the thresholds for the each submodules warning parameter

Disable grouping of bases for reads >50bp

No

Using this option will cause fastqc to crash and burn if you use it on really long reads, and your plots may end up a ridiculous size. You have been warned! (--nogroup)

Lower limit on the length of the sequence to be shown in the report

As long as you set this to a value greater or equal to your longest read length then this will be the sequence length used to create your read groups. This can be useful for making directly comparable statistics from datasets with somewhat variable read lengths. (--min_length)

length of Kmer to look for

7

note: the Kmer test is disabled and needs to be enabled using a custom Submodule and limits file (--kmers)

Email notification

No

Send an email notification when the job completes.

Execute

FastQC

FastQC aims to provide a simple way to do some quality control checks on raw sequence data coming from high throughput sequencing pipelines. It provides a set of analyses which you can use to get a quick impression of whether your data has any problems of which you should be aware before doing any further analysis.

The main functions of FastQC are:

- Import of data from BAM, SAM or FastQ/FASTQ.gz files (any variant).
- Providing a quick overview to tell you in which areas there may be problems
- Summary graphs and tables to quickly assess your data
- Export of results to an HTML based permanent report
- Offline operation to allow automated generation of reports without running the interactive app

16: FastQC on data 1: Raw Data

Eye

Pencil

X

15: FastQC on data 1: Web page

Eye

Pencil

X

1. Wybierz z listy plik fastq.gz do oceny jakości.

2. Wciśnij "Execute".

3. Oglądaj wyniki analizy jakości po naciśnięciu podglądu pliku typu Webpage

Trimming (trimmomatic)

WorkflowVisualizeShared DataHelpUser

Trimmomatic flexible read trimming tool for Illumina NGS data (Galaxy Version 0.38.0)

FavoriteVersionsOptions

Single-end or paired-end reads?

Paired-end (two separate input files)

Input FASTQ file (R1/first of pair)

1: 95_R1.fastq.gz

Input FASTQ file (R2/second of pair)

2: 95_R2.fastq.gz

Perform initial ILLUMINACLIP step?

No

Cut adapter and other illumina-specific sequences from the read

Trimmomatic Operation

1: Trimmomatic Operation

Select Trimmomatic operation to perform

Sliding window trimming (SLIDINGWINDOW)

Number of bases to average across

4

Average quality required

20

+ Insert Trimmomatic Operation

Output trimlog file?

No

(-trimlog)

Output trimmomatic log messages?

No

these are the messages written to stderr (eg. for use in MultiQC)

Job Resource Parameters

Use default job resource parameters

Email notification

No

Send an email notification when the job completes.

Execute

Zaznaczyć: "Paired-end"

Wybrać plik forward (R1)

Wybrać plik reverse (R2)

Wybrać usuwanie adapterów Illuminy (opcjonalnie)

Wybrać metodę trimmingu (na początek np. SLIDINGWINDOW), opcje:
Number of bases: 4
Average quality: 30

Wybrać: "Execute"

Allignment – mapowanie odczytów do sekwencji referencyjnej (bowtie2)

WorkflowVisualizeShared DataHelpUser

Bowtie2 - map reads against reference genome (Galaxy Version 2.4.2-galaxy0)

FavoriteVersionsOptions

Is this single or paired library?

Paired-end

FASTAQ file #1

2_W6_R2_Seq.gz

Must be of datatype "fastqsanger" or "fastq"

FASTAQ file #2

2_W6_R2.fastq.gz

Must be of datatype "fastqsanger" or "fastq"

Write unaligned reads (in fastq format) to separate file(s)

No

-unl-un-conc (possibly with -gz or -bz2). This triggers -un parameter for single reads and -un-conc for paired reads

Write aligned reads (in fastq format) to separate file(s)

No

-all-al-conc (possibly with -gz or -bz2). This triggers -al parameter for single reads and -al-conc for paired reads

Do you want to set paired-end options?

No

See "Alignment Options" section of Help below for information

Will you select a reference genome from your history or use a built-in index?

Use a built-in genome index

Sequences were indexed using default options. See "Indexes" section of help below

Select reference genome

Baboon (Papio anubis) [pap-anub1]

If your genome of interest is not listed, contact the Galaxy team

Set read groups information?

Do not set

Specifying read group information can greatly simplify your downstream analyses by allowing combining multiple datasets.

Select analysis mode

1: Default setting only

Do you want to use presets?

☒ Yes, just use defaults

☐ Very fast end-to-end (-very-fast)

☐ Fast end-to-end (-fast)

☐ Sensitive end-to-end (-sensitive)

☐ Very sensitive end-to-end (-very-sensitive)

☐ Very fast local (-very-fast-local)

☐ Fast local (-fast-local)

☐ Sensitive local (-sensitive-local)

☐ Very sensitive local (-very-sensitive-local)

Allow selecting among several preset parameter settings. Choosing between these will result in dramatic changes in runtime. See help below to understand effects of these presets.

Do you want to tweak SAMBAM Options?

No

See "Output Options" section of Help below for information

Save the bowtie2 mapping statistics to the history

No

Job Resource Parameters

Use default job resource parameters

Email notification

No

Send an email notification when the job completes.

Execute

Zaznaczyć: "Paired-end"

- Wybrać plik forward (R1)

Wybrać plik reverse (R2)

Wybrać sekwencję referencyjną z lokalnej bazy danych (hg19 lub hg38)

Wybrać: "Execute"

Alignent – mapowanie odczytów do sekwencji referencyjnej (bowtie2)

Informacje o składaniu:

The screenshot displays the Bowtie2 web interface. At the top, a tab is labeled '3: Bowtie2 on 95'. Below this, the file size '47.8 MB' and format 'format: bam, database: hg38' are shown. A large section contains alignment statistics: '442984 reads; of these: 442984 (100.00%) were paired; of these: 13085 (2.95%) aligned concordantly 0 times, 429845 (97.03%) aligned concordantly exactly 1 time, and 54 (0.01%) aligned concordantly >1 times'. Below the statistics, it says '13085 pairs aligned concord'. At the bottom of this section, there is a toolbar with icons for download, share, info, refresh, and help. A 'Download' button is highlighted, and a dropdown menu is open showing options: 'load dataset', 'Download bam_index', and 'display at bam.iobio'. Below the menu, there is a text input field with 'Binary bam alignments file' and a button '2: 95_R2.fastq.gz'.

3: Bowtie2 on 95

47.8 MB

format: bam, database: hg38

442984 reads; of these:
442984 (100.00%) were paired; of these:
13085 (2.95%) aligned concordantly 0 times
429845 (97.03%) aligned concordantly exactly 1 time
54 (0.01%) aligned concordantly >1 times

13085 pairs aligned concord

Download

load dataset

Download bam_index

display at bam.iobio

Binary bam alignments file

2: 95_R2.fastq.gz

Informacja o formacie i rozmiarze pliku

Informacja o liczbie zmapowanych odczytów

Pobieranie plików bam oraz indeksów na lokalny komputer

Allignment – detekcja wariantów (FreeBayes)

usegalaxy.org

Manuals SnpSift Filter -... FTP Download BigBlueButton... COVID-19 gen... ngs-blog.pl Pisanie pierwsz...

Workflow Visualize Shared Data Help User

FreeBayes bayesian genetic variant detector (Galaxy Version 1.3.1)

Choose the source for the reference genome

Locally cached

Run in batch mode?

☒ Run individually
☐ Merge output VCFs

Selecting individual mode will generate one VCF dataset for each input BAM dataset. Selecting the merge option will produce one VCF dataset for all input BAM datasets

BAM dataset

3: Bowtie2 on 95

Using reference genome

Human (Homo sapiens): hg38

Limit variant calling to a set of regions?

Do not limit

Sets --targets or --region options

Read coverage

Use defaults

Sets --min-coverage, --limit-coverage, and --skip-coverage

Choose parameter selection level

1. Simple diploid calling

Select how much control over the freebayes run you need

Job Resource Parameters

Use default job resource parameters

Email notification

☐ No

Select an email notification when the job completes.

✓ Execute

Wybrać plik w formacie bam

Wybrać sekwencję referencyjną z lokalnej bazy danych (hg19 lub hg38)

Wybrać: "Execute"

Alignment – filtrowanie wariantów (SnpSift Filter)

SnpSift Filter Filter variants using arbitrary expressions (Galaxy Version 4.3+1.galaxy1)

Input variant list in VCF format

4: FreeBayes on 95

Type of filter expression

Simple expression

Filter criteria

[QUAL>=30]&(DP>=10)

Need help? See the tool help below for some examples.

Invert filter

No

Select variants that do not match the filter expression. (--inverse)

Filter mode

Retain selected variants, remove others

Email notification

No

Send an email notification when the job completes.

✓ Execute

Wybrać plik w formacie VCF, np. Po detekcji wariantów (FreeBayes)

Wpisać kryteria filtrowania, np. (QUAL>=30)&(DP>=10), co oznacza średnia jakość odczytów o PhredScore (QUAL) powyżej 30 oraz pokrycie (DP) powyżej 10.

Wybrać: "Execute"

Alligment – ściągnięcie bazy danych anotacji (SnpEff download)

SnpEff download: download a pre-built database (Galaxy Version 4.3+T.galaxy2)

Select the annotation database you want to download (e.g. GRCh38.86, mm10 etc.)

hg38

The list of available databases can be obtained with 'SnpEff databases' tool

Email notification

☐ No

Send an email notification when the job completes.

✓ Execute

Wpisać nazwę bazy danych anotacji, np. "hg38" lub "hg19"

Wybrać: "Execute"

Annotacje wariantów (Snpeff eff)

Workflow: Workflow • Shared Data • Help • User • Settings

Snpeff eff: annotate variants (Snpeff Version 4.3+7 (github))

Download Snpeff database from history

Input format: VCF

Output format: VCF (only if input is VCF)

Create CSV report, useful for downstream analysis (csvOut): No

Genome source: Downloaded snpeff database in your history

Download Snpeff database from history

Download Snpeff database from history

Upstream / Downstream length: 5000 bases

Set size for splice sites (donor and acceptor) in bases: 2 bases

SpliceRegion Settings: Use Defaults

Annotation options: Select/Unselect all

Use custom interval file for annotation: No (interval file not found)

Only use the transcripts in this file: Nothing selected

Filter output: Select/Unselect all

Filter out specific Effects: No

Chromosomal position: Use default (based on input type)

Sort by prepared to chromosome name

Produce Summary Stats: No

Suppress reporting usage statistics to server: No

Email notifications: No

Execute

Wybrać plik csv z przefiltrowanymi danymi

Wybrać opcję "Download Snpeff database from history" i poniżej wybrać plik ze ściągniętą bazą danych anotacji.

Wybrać: "Execute"

Annotacje wariantów (SnpEff eff) - wyniki

6 and data 4

8: SnpEff eff: on data 6 and data 5 - HTML stats

37.7 KB

format: **html**, database: **hg38**

Picked up _JAVA_OPTIONS: -Djava.io.tmpdir=/galaxy-repl/main/jobdir/036/815/36815885/_-Xmx7g -Xms256m

HTML file

7: SnpEff eff: on data 6 and data 5

46 lines, 523 comments

format: **vcf**, database: **hg38**

Picked up _JAVA_OPTIONS: -Djava.io.tmpdir=/galaxy-repl/main/jobdir/036/815/36815885/_-Xmx7g -Xms256m

display at UCSC main
display with IGV local
display at vcf.iobio vcf.iobio.io

1. Chrom

```
##fileformat=VCFv4.2
##fileDate=20210724
##source=freeBayes v1.3.1-dirty
##reference=/cvmfs/data.galaxyproject
##contig=<ID=chr1, length=248956422>
```

6: hg38_annot

5: SnpSift Filter on data 4

Plik z raportem anotacji w formacie html.

Plik z danymi w formacie vcf.

Annotacje wariantów - ściąganie pliku z referencją anotacji

Ze strony:

http://ftp.ensembl.org/pub/release-104/variation/vcf/homo_sapiens/

Należy ściągnąć plik z referencją anotacji dla odpowiedniego chromosomu (w formacie vcf.gz).

Następnie należy zrobić upload tego pliku na serwer Galaxy używając narzędzia Upload Data.

Annotacja wariantów (SnpSift Annotate)

SnpSift Annotate SNPs from dbSnp (Galaxy Version 4.3+t.galaxy1)

Favorite

Versions

Options

Variant input file in VCF format

12: SnpSift Annotate on data 11 and data 7

VCF File with ID field annotated (e.g. dbSNP.vcf)

12: SnpSift Annotate on data 11 and data 7

The ID field for a variant in input will be assigned from a matching variant in this file.

Fields to annotate

Only annotate ID field (no INFO fields will be added)

Email notification

No

Send an email notification when the job completes.

Execute

Wybrać plik w formacie VCF, po użyciu narzędzia SnpEff

Wybrać plik z referencją anotacji

Wybrać: "Execute"

Ekstrakcja pól (Snpsift Extract Field)

Workflow Visualize Shared Data Help User

SnpSift Extract Fields from a VCF file into a tabular file (Galaxy Version 4.3+t.galaxy0) Favorite Options

Variant input file in VCF format

12: SnpSift Annotate on data 11 and data 7

Fields to extract

CHROM POS ID REF ALT FILTER

Separated by spaces. See help below for an explanation

One effect per line

☒ Yes

When variants have more than one effect, lists one effect per line, while all other parameters in the line are repeated across multiple lines

multiple field separator

Separate multiple fields in one column with this character, e.g. a comma, rather than a column for each of the multiple values (-s)

empty field text

Represent empty fields with this value, rather than leaving them blank (-e)

Email notification

☐ No

Send an email notification when the job completes.

✓ Execute

Wybrać plik w formacie VCF, po użyciu narzędzia SnpSift Annotate

Wybrać pola do ekstrakcji, np. CHROM POS ID REF ALT FILTER ANN[*].IMPACT ANN[*].GENE ANN[*].GENEID GEN[0].GT[0]

Zaznaczyć YES

Wybrać: "Execute"

Przykładowy plik wynikowy

CHROM	POS	ID	REF	ALT	IMPACT	GENE	ALLELE
chr4	154567669	rs6056	C	T	LOW	FGB	0/1
chr4	154567669	rs6056	C	T	LOW	FGB	0/1
chr4	154567933	rs2227411	A	G	MODIFIER	FGB	0/1
chr4	154567933	rs2227411	A	G	MODIFIER	FGB	0/1
chr4	154569680	rs4681	C	T	MODERATE	FGB	0/1
chr4	154569680	rs4681	C	T	LOW	FGB	0/1
chr4	154570607	rs4220	G	A	LOW	FGB	0/1
chr4	154571269	rs2227439	T	C	MODIFIER	FGB	1/1
chr4	154571269	rs2227439	T	C	HIGH	FGB	1/1