

- Useful links/tutorials: [Information on mapping](#) [Fastq manipulation & SAM/BAM](#)
- 1) You are provided with 2 datasets (one contains forward reads, one contains reverse reads). Create a new history, unzip and upload the data to your history and load them into FastQC and MultiQC. Comment on the results.
 - a) How similar are the results for the two datasets?
 - b) What kind of problems there are in the datasets (share screenshots of the issues you find)?
 - c) Share a link to your history.

2) Create a new history. Repeat the steps (FastQC and MultiQC) in part 1 for the following datasets (see **Assignment3_part2.pdf** to help you with finding the data and the BWA-MEM step):

- ZR751 paired-end RNA-seq subsampled (end 1)
- ZR751 paired-end RNA-seq subsampled (end 2)
 - a) How similar are the results for the two datasets?
 - b) What kind of problems there are in the datasets (share screenshots of the issues you find)?
 - c) Use 'Map with BWA-MEM' tool to map the data on the human genome.
 - d) **Share a link to your history.**

3) Create a new history. Choose 6 Illumina iDEA datasets (like how you did in part 2). This time, use fastp instead of FastQC to quality-check your data. Share screenshots

of the graphs you get and explain what you see (You don't have to share a screenshot for 6 different datasets if they are all similar. Focus on the problems and differences you observe.). **Share a link to your history.**

- 4) Create a new history. Complete [this tutorial](#). Add screenshots to the steps you take. Shortly explain what you have done. **Share a link to your history.**

RPKM (Reads Per Kilobase Million) or FPKM (Fragments Per Kilobase Million) and TPM (Transcripts Per Kilobase Million): they are the gene expression units. the difference between them is the oreder of the operations. RPKM (Reads Per Kilobase Million) is used when the sequencing depends on the lenght, therefore used in the single end RNA seq. experiments, FPKM for paired end rna-seq experiments, it can two read as single process. in TPM , a normalization process are made on the data also. Therefore, the comparing the gene to the proportions in the gene is quite easy in this. In others, comparing samples is more harder because there is o normalization of the data.



Download from web or upload from disk

Regular

Composite

Collection

Rule-based

Name

Size

Type

Genome

Settings

Status



New File

255 b

Auto-det...



unspecified (?)



100%



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

https://zenodo.org/record/4541751/files/GSM461177_1.fastqsangerhttps://zenodo.org/record/4541751/files/GSM461177_2.fastqsangerhttps://zenodo.org/record/4541751/files/GSM461180_1.fastqsangerhttps://zenodo.org/record/4541751/files/GSM461180_2.fastqsanger

Type (set all):

Auto-detect



Genome (set all):

unspecified (?)

Choose local files

Choose remote files

Paste/Fetch data

Start

Pause

Reset

Close



Executed **FastQC** and successfully added undefined jobs to the queue.

The tool uses 4 inputs:

- 1: **GSM461177_1.fastqsanger**
- 2: **GSM461177_2.fastqsanger**
- 3: **GSM461180_1.fastqsanger**
- 4: **GSM461180_2.fastqsanger**

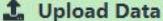
It produces 8 outputs:

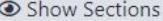
- 5: **FastQC on data 1: Webpage**
- 6: **FastQC on data 1: RawData**
- 7: **FastQC on data 2: Webpage**
- 8: **FastQC on data 2: RawData**
- 9: **FastQC on data 3: Webpage**
- 10: **FastQC on data 3: RawData**
- 11: **FastQC on data 4: Webpage**
- 12: **FastQC on data 4: RawData**

You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Multiqc

 Upload Data

 Show Sections

stat provides simple stats on BAM

QC aggregate results from Bioinformatics analyses into a single report

- fast all-in-one preprocessing for FASTQ files

Tools markup marks duplicate elements

S2 callpeak Call peaks from alignment results

oop - post-processing and QC of bunk analyses

 Executed **MultiQC** and successfully added 1 job to the queue.

The tool uses 4 inputs:

- 6: FastQC on data 1: RawData
- 8: FastQC on data 2: RawData
- 10: FastQC on data 3: RawData
- 12: FastQC on data 4: RawData

It produces this output:

- 14: MultiQC on data 12, data 10, and others: Webpage

You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

PHD Comics

 Random

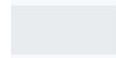


try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>



ces

ad



Executed **Cutadapt** and successfully added 1 job to the queue.

The tool uses 2 inputs:

- 1: **GSM461177_1.fastqsanger**
- 2: **GSM461177_2.fastqsanger**

It produces 3 outputs:

- 18: **Cutadapt on data 2 and data 1: Read 1 Output**
- 19: **Cutadapt on data 2 and data 1: Read 2 Output**
- 20: **Cutadapt on data 2 and data 1: Report**

You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

PHD Comics

Random



History

search

assign

17 shown

5.77 GB

20: C

a 2 and

19: C

a 2 and

utput

18: C

a 2 and

utput

14: M

data 10,

bpage

13: M

0, and o

BIN 508 Assignment #5_Leman | Reference-based RNA-Seq data | Galaxy

usegalaxy.eu

MPN navigator yön... Makale draft leman... GeoGebra - 100 mil... Middle East Technic... Probabilistic Graphi... Nizamettin Aydin Diğer yer işaretleri Okuma listesi

Galaxy Europe Workflow Visualize Shared Data Help User User Galaxy Using 2%

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools

cutada X

Upload Data

Show Sections

Cutadapt Remove adapter sequences from FASTQ/Fasta

MAGeCK count - collect sgRNA read counts from read mapping files

WORKFLOWS

All workflows

This is cutadapt 3.5 with Python 3.9.7
Command line parameters: -j=1 --output=out1.fq --paired-output=out2.fq --error-rate=0.1 --times=1 --overlap=3 --action=trim --minimum-length=20 --pair-filter=any --quality-cutoff=20 GSM461177_1_fastqsanger.fq
GSM461177_2_fastqsanger.fq
Processing reads on 1 core in paired-end mode ...
Finished in 187.29 s (18 µs/read; 3.39 M reads/minute).

==== Summary ====
Total read pairs processed: 10,575,821

== Read fate breakdown ==
Pairs that were too short: 147,810 (1.4%)
Pairs written (passing filters): 10,428,011 (98.6%)

Total basepairs processed: 782,610,754 bp
Read 1: 391,305,377 bp
Read 2: 391,305,377 bp

Quality-trimmed:
Read 1: 5,072,810 bp
Read 2: 8,648,619 bp

Total written (filtered): 763,282,345 bp (97.5%)
Read 1: 382,062,453 bp
Read 2: 381,219,892 bp

History

search datasets

assignment_5_508

17 shown
5.77 GB

20: Cutadapt on data 2 and data 1: Report

19: Cutadapt on data 2 and data 1: Read 2 Output

18: Cutadapt on data 2 and data 1: Read 1 Output

14: MultiQC on data 12, data 10, and others: Webpage

13: MultiQC on data 12, data 10, and others: Stats

15:27

https://zenodo.org/record/4541751/files/GSM461177_2.fastqsanger
https://zenodo.org/record/4541751/files/GSM461180_1.fastqsanger
https://zenodo.org/record/4541751/files/GSM461180_2.fastqsanger

New File

76 b

Auto-det...



unspecified (?)



100%



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

https://zenodo.org/record/4541751/files/Drosophila_melanogaster.BDGP6.87.gtf

Type (set all):

Auto-detect



Genome (set all):

unspecified (?)

Choose local files

Choose remote files

Paste/Fetch data

Start

Pause

Reset

Close

collect sgRNA read
mapping files

Annotation

Add an annotation or notes to a dataset; annotations are available when a history is viewed.

Database/Build

D. melanogaster Aug. 2014 (BDGP Release 6 + ISO1 MT/dm6) (dm6)

 Save

 Auto-detect

aster.BDGP6.87.

20: Cutadapt on
and data 1: Rep

19: Cutadapt on
and data 1: Read
ut

18: Cutadapt on
and data 1: Read
ut

14: MultiQC on
data 10, and oth
bpage

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools

cutada X

Upload Data

Show Sections

Cutadapt Remove adapter sequences from FASTQ/FASTA

MAGECK count - collect sgRNA read counts from read mapping files

WORKFLOWS

All workflows

Edit Dataset Attributes

Attributes updated.

Attributes Datatypes Permissions

New Type

gtf

This will change the datatype of the existing dataset but not modify its contents. Use this if Galaxy has incorrectly guessed the type of your dataset.

Save Auto-detect

History

search datasets ? X

assignment_5_508

18 shown, 3 hidden
5.91 GB

21: Drosophila_melanogaster.BDGP6.87.gtf ? P X

20: Cutadapt on data 2 and data 1: Report ? P X

19: Cutadapt on data 2 and data 1: Read 2 Output ? P X

large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

ad Data

Sections

ead mapper for

RNA-Seq FASTQ/FASTA file, forward reads

RNA-Seq FASTQ/FASTA file, reverse reads

Custom or built-in reference genome

Reference genome with or without an annotation

Select reference genome

History

search datasets

assignment_5_508

18 shown, 3 hidden

5.91 GB

21: Drosophila_melanogaster.BDGP6.87.gtf

20: Cutadapt on data 2 and data 1: Report

19: Cutadapt on data 2 and data 1: Read 2 Output

18: Cutadapt on data 2 and data 1: Read 1 Output

14: MultiQC on data 12, data 10, and others: Websc

RNA-Seq FASTQ/FASTA file, forward reads

RNA-Seq FASTQ/FASTA file, reverse reads

Custom or built-in reference genome

Reference genome with or without an annotation

Select reference genome

History

search datasets

assignment_5_508

18 shown, 3 hidden

5.91 GB

21: Drosophila_melanogaster.BDGP6.87.gtf

20: Cutadapt on data 2 and data 1: Report

19: Cutadapt on data 2 and data 1: Read 2 Output

18: Cutadapt on data 2 and data 1: Read 1 Output

14: MultiQC on data 12, data 10, and others: Websc



[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



rnastar

Upload Data

Show Sections

RNA STAR Gapped-read mapper for

NA-seq data

ORKFLOWS

l workflows



Executed **RNA STAR** and successfully added 1 job to the queue.

The tool uses 3 inputs:

- 18: Cutadapt on data 2 and data 1: Read 1 Output
- 19: Cutadapt on data 2 and data 1: Read 2 Output
- 21: Drosophila_melanogaster.BDGP6.87.gtf

It produces 3 outputs:

- 22: RNA STAR on data 21, data 19, and data 18: log
- 23: RNA STAR on data 21, data 19, and data 18: splice junctions.bed
- 24: RNA STAR on data 21, data 19, and data 18: mapped.bam

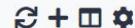
You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

We need your support ...

If Galaxy helped with the analysis of your data, please do not forget to **cite**:

Afsoon E et al. 2016 The Galaxy platform for accessible, reproducible and collaborative biomedical analyses:

History



search datasets

assignment_5_508

21 shown, 3 hidden

5.91 GB



24: RNA STAR on data 21, data 19, and data 18: mapped.bam



23: RNA STAR on data 21, data 19, and data 18: splice junctions.bed



22: RNA STAR on data 21, data 19, and data 18: log



21: Drosophila_melanogaster.BDGP6.87.gtf



20: Cutadapt on data 2 and data 1: Report



RNA star is an aligner to not an extension of DNA-mapper, aligns reads with the Maximal Mappable Prefix hits. it has high speed with lower error rates

BIN 508 Ass X | Reference-E X | Galaxy X | Gelen Kutu X | Grades: Vie X | Google Çev X | RNA star - G X | STAR X | + | - | X

usegalaxy.eu/workflows/list

MPN navigator yön... Makale draft leman... GeoGebra - 100 mil... Middle East Technic... Probabilistic Graphi... Nizamettin Aydin » Diğer yer işaretleri Okuma listesi

Galaxy Europe Workflow Visualize Shared Data Help User Using 2%

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools search tools Upload Data

Get Data

Send Data

Collection Operations

GENERAL TEXT TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

GENOMIC FILE MANIPULATION

Convert Formats

FASTA/FASTQ

Quality Control

SAM/BAM

	Started job on	Dec 29 13:31:55
Mapping speed, Million of reads per hour	Started mapping on	Dec 29 13:33:28
	Finished on	Dec 29 13:37:23
	Number of input reads	10428011
	Average input read length	73
	UNIQUE READS:	
	Uniquely mapped reads number	8666702
	Uniquely mapped reads %	83.11%
	Average mapped length	72.87
	Number of splices: Total	952323
	Number of splices: Annotated (sjdb)	943181
	Number of splices: GT/AG	942946
	Number of splices: GC/AG	7149
	Number of splices: AT/AC	288
	Number of splices: Non-canonical	1940
	Mismatch rate per base, %	0.77%
	Deletion rate per base	0.00%
	Deletion average length	1.48
	Insertion rate per base	0.00%
	Insertion average length	1.39
	MULTI-MAPPING READS:	
	Number of reads mapped to multiple loci	571324
	% of reads mapped to multiple loci	5.48%
	Number of reads mapped to too many loci	574223
	% of reads mapped to too many loci	5.51%
	UNMAPPED READS:	
	Number of reads unmapped: too many mismatches	0
	% of reads unmapped: too many mismatches	0.00%
	Number of reads unmapped: too short	599120
	% of reads unmapped: too short	5.75%
	Number of reads unmapped: other	16642
	% of reads unmapped: other	0.16%
	CUTADAPT READS:	

History search datasets

assignment_5_508

29 shown, 10 hidden

5.91 GB

1, data 19, and data 18: mapped.bam

23: RNA STAR on data 2
1, data 19, and data 18: splice junctions.bed

22: RNA STAR on data 21, data 19, and data 1 8: log

21: Drosophila_melanogaster.BDGP6.87.gtf

20: Cutadapt on data 2 and data 1: Report

results of rna star

Galaxy Europe



Workflow Visualize Shared Data Help User



Using 2%

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



search tools



Upload Data

Get Data

Send Data

Collection Operations

GENERAL TEXT TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

GENOMIC FILE MANIPULATION

Convert Formats

FASTA/FASTQ

Quality Control

SAM/BAM

chr2L	11345	11409	2	2	1	11	0	14
chr2L	11519	11778	2	2	1	21	0	17
chr2L	12222	12285	2	2	1	22	0	18
chr2L	12929	13519	2	2	1	38	0	17
chr2L	13626	13682	2	2	1	47	0	18
chr2L	14875	14932	2	2	1	49	0	15
chr2L	15712	17052	2	2	1	21	2	18
chr2L	15712	19879	2	2	1	0	1	5
chr2L	15712	19884	2	2	1	1	0	3
chr2L	17208	19879	2	2	0	1	0	14
chr2L	17213	18025	2	2	1	2	3	17
chr2L	17213	18260	2	2	1	4	6	17
chr2L	17213	18330	2	2	0	0	2	17
chr2L	17213	19879	2	2	1	11	0	17
chr2L	17213	21135	2	2	1	0	2	17
chr2L	20021	20830	2	2	0	0	2	17
chr2L	20021	21065	2	2	1	6	3	18
chr2L	20021	21135	2	2	1	3	0	17
chr2L	20974	21065	2	2	1	1	0	12
chr2L	26689	26765	2	2	1	1	0	8
chr2L	27491	28014	2	2	1	1	0	4
chr2L	38299	38534	2	2	1	1	0	8

History

search datasets



assignment_5_508

29 shown, 10 hidden

5.91 GB



Webpage

25: MultiQC on data 22: Stats

a list with 3 items

24: RNA STAR on data 2

1, data 19, and data 18:
mapped.bam

23: RNA STAR on data

21, data 19, and data 1
8: splice junctions.bed

22: RNA STAR on data 2

1, data 19, and data 18:
log

21: Drosophila melanog



parameters used for quality checking

usegalaxy.eu/workflows/list

MPN navigator yön... Makale draft leman... GeoGebra - 100 mil... Middle East Technic... Probabilistic Graphi... Nizamettin Aydin » Diğer yer işaretleri Okuma listesi

Galaxy Europe

Workflow Visualize Shared Data Help User Nizamettin Aydin Using 2%

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools	QNAME	FLAG
search tools	@HD VN:1.4 SO:coordinate	
	@SQ SN:chr2L LN:23513712	
	@SQ SN:chr2R LN:25286936	
	@SQ SN:chr3L LN:28110227	
	@SQ SN:chr3R LN:32079331	
	@SQ SN:chr4 LN:1348131	
	@SQ SN:chrM LN:19524	
Get Data	@SQ SN:chrUn_DS485919v1 LN:1021	
Send Data	@SQ SN:chrUn_DS483755v1 LN:6936	
Collection Operations	@SQ SN:chrUn_DS485425v1 LN:1143	
GENERAL TEXT TOOLS	@SQ SN:chrUn_DS484861v1 LN:1395	
Text Manipulation	@SQ SN:chrUn_DS484484v1 LN:2020	
Filter and Sort	@SQ SN:chrUn_DS483705v1 LN:27456	
Join, Subtract and Group	@SQ SN:chrUn_DS485490v1 LN:1127	
GENOMIC FILE MANIPULATION	@SQ SN:chrUn_DS485998v1 LN:1003	
Convert Formats	@SQ SN:chrUn_DS483873v1 LN:4222	
FASTA/FASTQ	@SQ SN:chrUn_DS485608v1 LN:1097	
Quality Control	@SQ SN:chrUn_DS485270v1 LN:1185	
SAM/BAM	@SQ SN:chrUn_DS485979v1 LN:1008	
	@SQ SN:chrUn_DS485398v1 LN:1148	
	@SQ SN:chrUn_DS484413v1 LN:3020	

History

search datasets

assignment_5_508
29 shown, 10 hidden
5.91 GB

Webpage

25: MultiQC on data 22: Stats
a list with 3 items

24: RNA STAR on data 21, data 19, and data 1
8: mapped.bam

23: RNA STAR on data 2 1, data 19, and data 18:
splice junctions.bed

22: RNA STAR on data 2 1, data 19, and data 18:
log

21: Drosophila melanog...



Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

iqc

Upload Data

Show Sections

t provides simple stats on BAM

C aggregate results from
rmatics analyses into a single

fast all-in-one preprocessing for
files

ols markup marks duplicate
ents

2 callpeak Call peaks from
ent results

op - post-processing and QC of
nk analyses

ols idxstats reports stats of the
dex file

MultiQC aggregate results from bioinformatics analyses into a single report (Galaxy Version 1.11+galaxy0)

Results

1: Results

Which tool was used generate logs?

STAR

Software name

STAR output

1: STAR output

Type of STAR output?

Log

STAR log output

23: RNA STAR on data 21, data 19, and data 18: splice junctions.bed
22: RNA STAR on data 21, data 19, and data 18: log
21: Drosophila_melanogaster.BDGP6.87.gtf
20: Cutadapt on data 2 and data 1: Report
19: Cutadapt on data 2 and data 1: Read 2 Output
18: Cutadapt on data 2 and data 1: Read 1 Output

+ Insert STAR output

History

search datasets

assignment_5_508

21 shown, 3 hidden

5.91 GB

24: RNA STAR on da
ta 21, data 19, and data
18: mapped.bam

23: RNA STAR on da
ta 21, data 19, and data
18: splice junctions.bed

22: RNA STAR on da
ta 21, data 19, and data
18: log

21: Drosophila_melanog
aster.BDGP6.87.gtf

20: Cutadapt on data 2
and data 1: Report

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



multiqc



Upload Data

Show Sections

flagstat provides simple stats on BAM files

MultiQC aggregate results from bioinformatics analyses into a single report

fastp - fast all-in-one preprocessing for FASTQ files

Samtools markup marks duplicate alignments

MACS2 callpeak Call peaks from alignment results

Allelopoop - post-processing and QC of Slamdunk analyses

Samtools idxstats reports stats of the BAM index file

Which tool was used generate logs?

STAR

Software name

STAR output

1: STAR output

Type of STAR output?

Log

STAR log output



23: RNA STAR on data 21, data 19, and data 18: splice junctions.bed

22: RNA STAR on data 21, data 19, and data 18: log

21: Drosophila_melanogaster.BDGP6.87.gtf

20: Cutadapt on data 2 and data 1: Report

19: Cutadapt on data 2 and data 1: Read 2 Output

18: Cutadapt on data 2 and data 1: Read 1 Output

Insert STAR output

Insert Results

Report title

History



search datasets



assignment_5_508

21 shown, 3 hidden

5.91 GB



24: RNA STAR on data 2

1, data 19, and data 18:
mapped.bam

23: RNA STAR on data 2

1, data 19, and data 18:
splice junctions.bed

22: RNA STAR on data 2

1, data 19, and data 18:
log

21: Drosophila_melanog

aster.BDGP6.87.gtf

20: Cutadapt on data 2

and data 1: Report

you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

QC

 Upload Data

Show Sections

provides simple stats on BAM

aggregate results from
natics analyses into a single

ast all-in-one preprocessing for
es

s markup marks duplicate
nts

callpeak Call peaks from
nt results

o - post-processing and QC of
k analyses

llectate converts state-of-the

 Executed **MultiQC** and successfully added 1 job to the queue.

The tool uses this input:

- 22: RNA STAR on data 21, data 19, and data 18: log

It produces this output:

- 26: MultiQC on data 22: Webpage

You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

Galactic News

News

History

search datasets

assignment_5_508

23 shown, 3 hidden

5.91 GB

  26: MultiQC on data 22: Webpage

  25: MultiQC on data 22: Sta ts

a list

  24: RNA STAR on data 21, data 19, and data 18: mapped.bam

  23: RNA STAR on data 21, data 19, and data 18: splice junctions.bed

  22: RNA STAR on data 21, data 19, and data 18: log

MPN navigator yön... Makale draft leman... GeoGebra - 100 mil... Middle East Technic... Probabilistic Graphi... Nizamettin Aydin » Diğer yer işaretleri Okuma listesi

Galaxy Europe

TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools

multiqc

Upload Data

Show Sections

agstat provides simple stats on BAM files

MultiQC aggregate results from bioinformatics analyses into a single report

stp - fast all-in-one preprocessing for FASTQ files

amtools markup marks duplicate alignments

ACS2 callpeak Call peaks from alignment results

leoop - post-processing and QC of Bamduck analyses

amtools idxstats reports stats of the BAM index file

Workflow Visualize Shared Data Help User

Using 2%

A modular tool to aggregate results from bioinformatics analyses across many samples into a single report.

Report generated on 2021-12-29, 13:45 based on data in: /data/jwd/main/038/056/38056694/working/multiqc_WDir

Welcome! Not sure where to start? Watch a tutorial video (6:06) don't show again

General Statistics

Copy table Configure Columns Plot Showing 1/1 rows and 2/2 columns.

Sample Name	% Aligned	M Aligned
RNA STAR on data 21_data 19_and data 18_log	83.1%	8.7

STAR

STAR is an ultrafast universal RNA-seq aligner.

Alignment Scores

Number of Reads Percentages

History

search datasets

assignment_5_508

23 shown, 3 hidden

5.91 GB

26: MultiQC on data 2 2: Webpage

25: MultiQC on data 22: Stats a list with 3 items

24: RNA STAR on data 2 1, data 19, and data 18: mapped.bam

23: RNA STAR on data 2 1, data 19, and data 18: splice junctions.bed

22: RNA STAR on data 2 1, data 19, and data 18: log

6°C Çok bulutlu 15:45 29.12.2021

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



multiqc



Upload Data

Show Sections

flagstat provides simple stats on BAM files

MultiQC aggregate results from bioinformatics analyses into a single report

fastp - fast all-in-one preprocessing for FASTQ files

Samtools markup marks duplicate alignments

MACS2 callpeak Call peaks from alignment results

Alleoop - post-processing and QC of Slamdunk analyses

Samtools idxstats reports stats of the BAM index file

STAR is an ultrafast universal RNA-seq aligner.

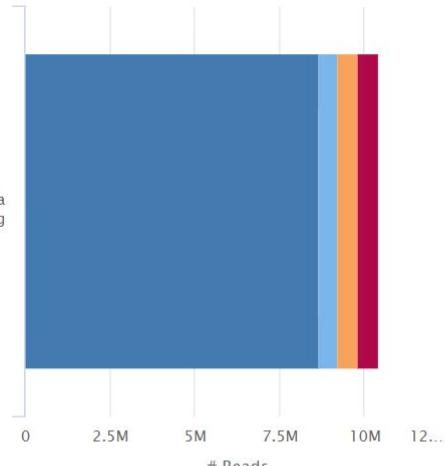
Alignment Scores

Number of Reads Percentages

STAR: Alignment Scores

Export Plot

RNA STAR on data 21_data 19_and data 18_log



Uniquely mapped Mapped to multiple loci

History

search datasets

assignment_5_508

23 shown, 3 hidden

5.91 GB

26: MultiQC on data 2

2: Webpage

25: MultiQC on data 22: Stats

a list with 3 items

24: RNA STAR on data 2

1, data 19, and data 18: mapped.bam

23: RNA STAR on data 2

1, data 19, and data 18: splice junctions.bed

22: RNA STAR on data 2

1, data 19, and data 18: log

Toolbox



A



 Galaxy Europe

 Workflow Visualize Shared Data ▾ Help ▾ User ▾

Using 2%

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools

 Upload Data

 Show Sections

flagstat provides simple stats on BAM files

MultiQC aggregate results from bioinformatics analyses into a single report

fastp - fast all-in-one preprocessing for FASTQ files

Samtools markup marks duplicate alignments

MACS2 callpeak Call peaks from alignment results

Alleloop - post-processing and QC of Slamdunk analyses

Samtools idxstats reports stats of the BAM index file

Sample

STAR_mqc-generalstats-star-uniquely_mapped_percent STAR_mqc-generalstats-star
83.11

History

[History](#)

MultiQC on data

a list with 3 items

general stats

sources

star

usegalaxy.eu

MPN navigator yön... Makale draft leman... GeoGebra - 100 mil... Middle East Technic... Probabilistic Graphi... Nizamettin Aydin

» Diğer yer işaretleri Okuma listesi

Galaxy Europe

[Workflow](#) [Visualize](#) [Shared Data](#) [Help](#) [User](#)

Using 2%

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>**Tools**

Module	Section	Sample Name	Source
STAR	SummaryLog	RNA STAR on data 21_data 19_and data 18_log	/data/jwd/main/038/056/38056694/working/multiqc_WDir

multiqc

[Upload Data](#)[Show Sections](#)**flagstat** provides simple stats on BAM files**MultiQC** aggregate results from bioinformatics analyses into a single report**fastp** - fast all-in-one preprocessing for FASTQ files**Samtools markup** marks duplicate alignments**MACS2 callpeak** Call peaks from alignment results**Alleyoop** - post-processing and QC of Slamdunk analyses**Samtools idxstats** reports stats of the BAM index file**History**[◀ Back to assignment_5_508](#)**MultiQC on data 22: Stats**

a list with 3 items

[general_stats](#)[sources](#)[star](#)

usegalaxy.eu

MPN navigator yön... Makale draft leman... GeoGebra - 100 mil... Middle East Technic... Probabilistic Graphi... Nizamettin Aydin » Diğer yer işaretleri Okuma listesi

Galaxy Europe

Workflow Visualize Shared Data Help User

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Sample	total_reads	avg_input_read_length	uniquely_mapped	uniquely_mapped_percent
RNA STAR on data 21_data 19_and data 18_log	10428011.0	73.0	8666702.0	83.1%

Tools

- multiqc

Upload Data

Show Sections

flagstat provides simple stats on BAM files

MultiQC aggregate results from bioinformatics analyses into a single report

fastp - fast all-in-one preprocessing for FASTQ files

Samtools markup marks duplicate alignments

MACS2 callpeak Call peaks from alignment results

Alleoop - post-processing and QC of Slamdunk analyses

Samtools idxstats reports stats of the BAM index file

History

Back to assignment_5_508

MultiQC on data 22: Stats
a list with 3 items

general_stats
sources
star

6°C Çok bulutlu 15:46 29.12.2021

Galaxy Europe



Workflow

Visualize

Shared Data ▾

Help ▾

User ▾



Using 2%

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



Infer Experiment

Upload Data**Show Sections****VCFprimers:** Extract flanking sequences for each VCF record**Map with BWA-MEM** - map medium and long reads (> 100 bp) against reference genome**StringTie merge** transcripts**DEXSeq-Count** Prepare and count exon abundances from RNA-seq data**plotDEXSeq** Visualization of the per gene DEXSeq results**MaxQuant (using mqpar.xml)****Kronik** processes Hardkor features to find peptides by chromatographic profiling

⚡ **Infer Experiment** speculates how RNA-seq were configured (Galaxy Version 2.6.4.1)

Input .bam file



24: RNA STAR on data 21, data 19, and data 18: mapped.bam



(--input-file)

Reference gene model



23: RNA STAR on data 21, data 19, and data 18: splice junctions.bed (as bed12)



(--refgene)

Number of reads sampled from SAM/BAM file (default = 200000)

200000

(--sample-size)

Minimum mapping quality

30

Minimum mapping quality for an alignment to be considered as "uniquely mapped" (--mapq)

Email notification



Send an email notification when the job completes.

History

search datasets



assignment_5_508

23 shown, 3 hidden

5.91 GB



26: MultiQC on data 22: [Webpage](#)

25: MultiQC on data 22: Stats

a list with 3 items



24: RNA STAR on data 2 1, data 19, and data 18: mapped.bam



23: RNA STAR on data 2 1, data 19, and data 18: splice junctions.bed



22: RNA STAR on data 2 1, data 19, and data 18: log

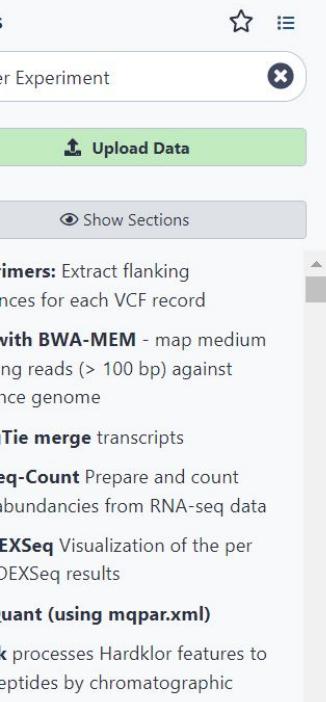


Galaxy Europe

 Workflow Visualize Shared Data ▾ Help ▾ User ▾

Using 2%

Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>



 Executed **Infer Experiment** and successfully added 1 job to the queue.

The tool uses 2 inputs:

- 24: RNA STAR on data 21, data 19, and data 18: **mapped.bam**
- 23: RNA STAR on data 21, data 19, and data 18: **splice junctions.bed** (as bed12)

It produces this output:

- 30: **Infer Experiment on data 23 and data 24**

You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

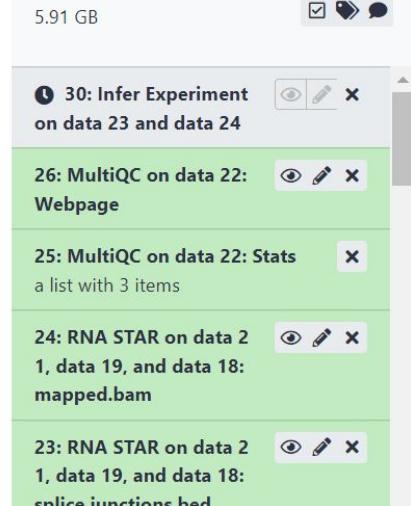


History

search datasets

Assignment_5_508

24 shown, 3 hidden



! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



featureCounts



Upload Data

Show Sections

goseq tests for overrepresented gene categories**featureCounts** Measure gene expression in RNA-Seq experiments from SAM or BAM files**DESeq2** Determines differentially expressed features from count tables**limma** Perform differential expression with limma-voom or limma-trend**goseq** tests for overrepresented gene categories**UMI-tools count** performs quantification of UMIs from BAM files**Remove Unwanted Variation from**Executed **featureCounts** and successfully added 1 job to the queue.

The tool uses 2 inputs:

- 24: RNA STAR on data 21, data 19, and data 18: mapped.bam
- 21: Drosophila_melanogaster.BDGP6.87.gtf

It produces 3 outputs:

- 31: featureCounts on data 21 and data 24: Counts
- 32: featureCounts on data 21 and data 24: Summary
- 33: featureCounts on data 21 and data 24: Feature lengths

You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

Upcoming Events

European Galaxy Events

History

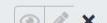


search datasets

**assignment_5_508**

27 shown, 3 hidden

5.91 GB

**33: featureCounts o n data 21 and data 24: F eature lengths****32: featureCounts o n data 21 and data 24: S ummary****31: featureCounts o n data 21 and data 24: C ounts****30: Infer Experiment on data 23 and data 24****26: MultiQC on data 22:**

feature count checks strand-specific read counting, to verify the alignment is made on genome not on the transcriptome.

Galaxy Europe

Workflow Visualize Shared Data Help User

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools

multicq

Upload Data

Show Sections

flagstat provides simple stats on BAM files

MultiQC aggregate results from bioinformatics analyses into a single report

fastp - fast all-in-one preprocessing for FASTQ files

amtools markup marks duplicate alignments

IACS2 callpeak Call peaks from alignment results

Alleoop - post-processing and QC of lamdunk analyses

amtools idxstats reports stats of the AM index file

MultiQC aggregate results from bioinformatics analyses into a single report (Galaxy Version 1.11+galaxy0)

Results

1: Results

Which tool was used generate logs?

featureCounts

Software name

Output of FeatureCounts

33: featureCounts on data 21 and data 24: Feature lengths
32: featureCounts on data 21 and data 24: Summary
31: featureCounts on data 21 and data 24: Counts
30: Infer Experiment on data 23 and data 24
26: MultiQC on data 22: Webpage
23: RNA STAR on data 21, data 19, and data 18: splice junctions.bed

+ Insert Results

Report title

It is printed as page header

Custom comment

History

search datasets

assignment_5_508

27 shown, 3 hidden

5.91 GB

33: featureCounts on data 21 and data 24: Feature lengths

32: featureCounts on data 21 and data 24: Summary

31: featureCounts on data 21 and data 24: Counts

30: Infer Experiment on data 23 and data 24

26: MultiQC on data 22: Webpage

15:55

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools	Status	RNA STAR on data 21	data 19	and data 18: mapped.bam	History
search tools	Status	RNA STAR on data 21, data 19, and data 18: mapped.bam			
	Assigned		6310		
	Unassigned_Unmapped		1189985		
	Unassigned_Read_Type		0		
	Unassigned_Singleton		0		
	Unassigned_MappingQuality		0		
	Unassigned_Chimera		0		
	Unassigned_FragmentLength		0		
	Unassigned_Duplicate		0		
	Unassigned_MultiMapping		3222785		
	Unassigned_Secondary		0		
	Unassigned_NonSplit		0		
	Unassigned_NoFeatures		8660432		
	Unassigned_Overlapping_Length		0		
	Unassigned_Ambiguity		4		

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools**Get Data****Send Data****Collection Operations****GENERAL TEXT TOOLS****Text Manipulation****Filter and Sort****Join, Subtract and Group****GENOMIC FILE MANIPULATION****Convert Formats****FASTA/FASTQ****Quality Control****SAM/BAM****Feature****Length**

Geneid	Length
FBgn0085804	252
FBgn0267431	6213
FBgn0039987	184
FBgn0058182	291
FBgn0267430	6420
FBgn0266747	497
FBgn0086917	1767
FBgn0010247	3112
FBgn0086378	1290
FBgn0263977	961
FBgn0069923	364
FBgn0039955	4390
FBgn0259821	5061
FBgn0027341	4155
FBgn0085812	614
FBgn0058198	1567
FBgn0037213	5461
FBgn0053294	2235
FBgn0000500	569
FBgn0037215	3247
FBgn0037217	1185



[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



search tools

Upload Data

Get Data

Send Data

Collection Operations

GENERAL TEXT TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

GENOMIC FILE MANIPULATION

Convert Formats

FASTA/FASTQ

Quality Control

SAM/BAM



Workflow

Visualize

Shared Data

Help

User



Geneid RNA STAR on data 21

Geneid RNA STAR on data 21, data 19, and data 18: mapped.bam

FBgn0085804	0
FBgn0267431	0
FBgn0039987	0
FBgn0058182	0
FBgn0267430	0
FBgn0266747	0
FBgn0086917	0
FBgn0010247	0
FBgn0086378	0
FBgn0263977	0
FBgn0069923	0
FBgn0039955	0
FBgn0259821	0
FBgn0027341	0
FBgn0085812	0
FBgn0058198	0
FBgn0037213	0
FBgn0053294	0
FBgn0000500	0
FBgn0037215	0
FBgn0037217	0

History

search datasets

assignment_5_508

29 shown, 10 hidden

5.91 GB

33: featureCounts on data 21 and data 24: Feature lengths

32: featureCounts on data 21 and data 24: Summary

31: featureCounts on data 21 and data 24: Counts

30: Infer Experiment on data 23 and data 24

26: MultiQC on data 22: Webpage

BIN 508 Assignment #5_Leman N X Reference-based RNA-Seq data X Galaxy X what is rna star - Google'da Ara X +

usegalaxy.eu MPN navigator yön... Makale draft leman... GeoGebra - 100 mil... Middle East Technic... Probabilistic Graphi... Nizamettin Aydin » Diğer yer işaretleri Okuma listesi

Galaxy Europe Workflow Visualize Shared Data Help User Using 2%

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools

- multiqc

Upload Data

Show Sections

flagstat provides simple stats on BAM files

MultiQC aggregate results from bioinformatics analyses into a single report

fastp - fast all-in-one preprocessing for FASTQ files

Samtools markup marks duplicate alignments

MACS2 callpeak Call peaks from alignment results

Alleyoop - post-processing and QC of Slamdunk analyses

Samtools idxstats reports stats of the BAM index file

Executed **MultiQC** and successfully added 1 job to the queue.

The tool uses this input:

- 32: featureCounts on data 21 and data 24: Summary

It produces this output:

- 35: MultiQC on data 32: Webpage

You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

We need your support ...

If Galaxy helped with the analysis of your data, please do not forget to **cite**:

Afgan E et al. 2016 The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res. 44, W3–W10 doi:10.1093/nar/gkw343

History search datasets assignment_5_508 29 shown, 3 hidden 5.91 GB

35: MultiQC on data 32: Webpage

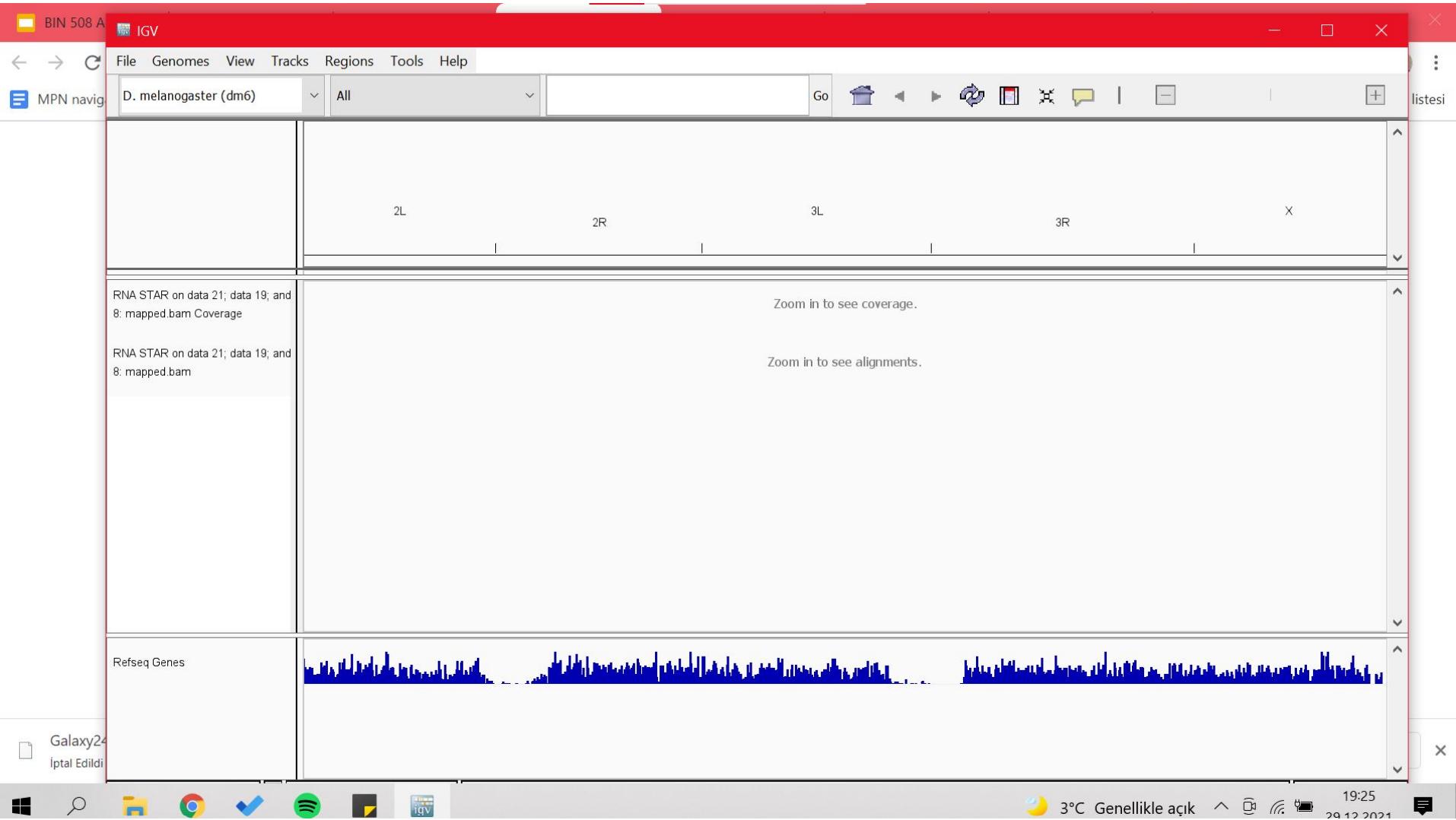
34: MultiQC on data 32: Stats a list

33: featureCounts on data 21 and data 24: Feature lengths

32: featureCounts on data 21 and data 24: Summary

31: featureCounts on data 21 and data 24: Count

6°C Çok bulutlu 15:56 29.12.2021



Intr

Dat

Qua

Ma

M

In

m

Cou

rea

gen

Ana

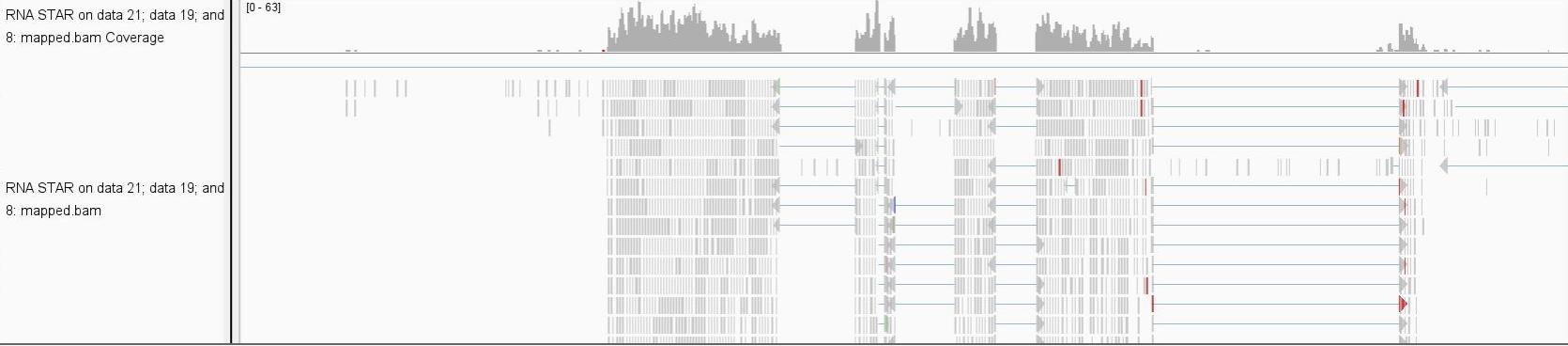
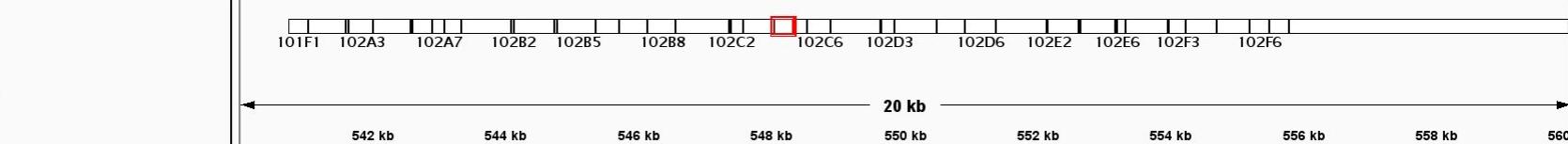
diff

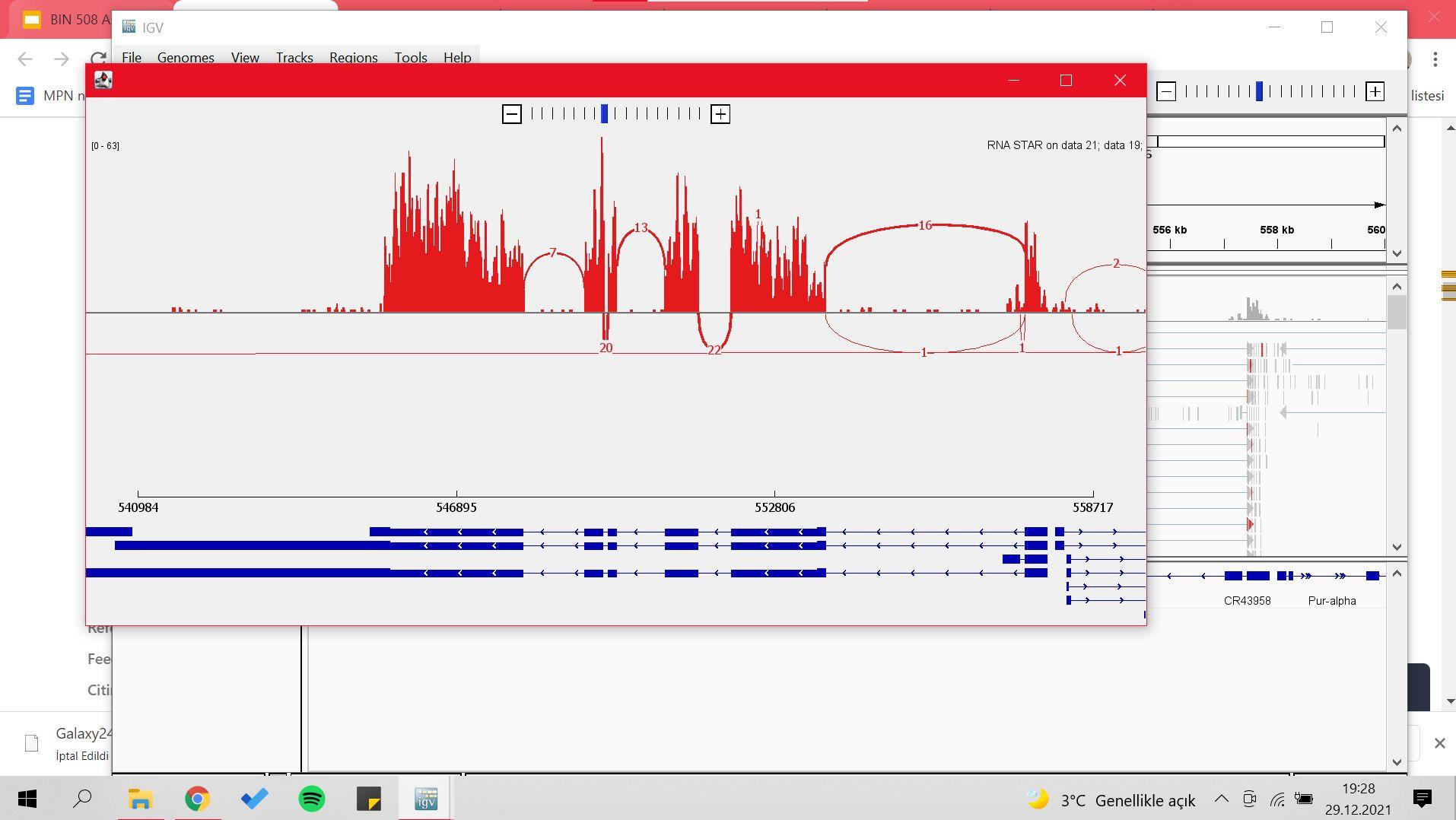
exp

Fun

ana

Cor





<https://usegalaxy.eu/u/lemannurnehri/h/assignmentpart2>

and faster. Instruction on <https://galaxyproject.eu/ftp>

• 32: featureCounts on data 21 and data 24: Summary

It produces this output:

• 35: MultiQC on data 32: Webpage

You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

We need your support ...

If Galaxy helped with the analysis of your data, please do not forget to **cite**:

Afgan E et al. 2016 The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update.

Nucleic Acids Res. 44, W3–W10
doi:10.1093/nar/gkw343

please **acknowledge** the Freiburg Galaxy server:

History			
assignment_part2			
search datasets			
7 shown			
(empty)			
7: GSM461182_untre at_single.counts			
6: GSM461181_treat _paired.counts			
5: GSM461180_treat _paired.counts			
4: GSM461179_treat _single.counts			
3: GSM461178_untre at_paired.counts			

usegalaxy.eu

MPN navigator yön... Makale draft leman... GeoGebra - 100 mil... Middle East Technic... Probabilistic Graphi... Nizamettin Aydin

Diger yer işaretleri Okuma listesi

Galaxy Europe Workflow Visualize Shared Data Help User Using 2%

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools

- deseq2

Upload Data

Show Sections

StringTie merge transcripts

DESeq2 Determines differentially expressed features from count tables

Annotate DESeq2/DEXSeq output tables Append annotation from GTF to differential expression tool outputs

StringTie transcript assembly and quantification

featureCounts Measure gene expression in RNA-Seq experiments from SAM or BAM files

TEntranscripts annotates reads to genes and transposable elements

DiffBind differential binding analysis of ChIP-Sea peak data

Factor

1: Factor

Specify a factor name, e.g. effects_drug_x or cancer_markers

Treatment

Only letters, numbers and underscores will be retained in this field

Factor level

1: Factor level

Specify a factor level, typical values could be 'tumor', 'normal', 'treated' or 'control'

treated

Only letters, numbers and underscores will be retained in this field

Counts file(s)

7: GSM461182_untreat_single.counts
6: GSM461181_treat_paired.counts
5: GSM461180_treat_paired.counts
4: GSM461179_treat_single.counts
3: GSM461178_untreat_paired.counts
2: GSM461177_untreat_paired.counts

History

search datasets

assignment_part2

7 shown

1.77 MB

7: GSM461182_untreat_single.counts
6: GSM461181_treat_paired.counts
5: GSM461180_treat_paired.counts
4: GSM461179_treat_single.counts
3: GSM461178_untreat_paired.counts
2: GSM461177_untreat_paired.counts

6°C Çok bulutlu 16:07 29.12.2021



! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools

deseq2

**Upload Data****Show Sections****StringTie merge transcripts****DESeq2** Determines differentially expressed features from count tables**Annotate DESeq2/DEXSeq output tables** Append annotation from GTF to differential expression tool outputs**StringTie** transcript assembly and quantification**featureCounts** Measure gene expression in RNA-Seq experiments from SAM or BAM files**TEtranscripts** annotates reads to genes and transposable elements

4: GSM461179_treat_single.counts

3: GSM461178_untreat_paired.counts

2: GSM461177_untreat_paired.counts

2: Factor level

Specify a factor level, typical values could be 'tumor', 'normal', 'treated' or 'control'

untreated

Only letters, numbers and underscores will be retained in this field

Counts file(s)



6: GSM461181_treat_paired.counts

5: GSM461180_treat_paired.counts

4: GSM461179_treat_single.counts

3: GSM461178_untreat_paired.counts

2: GSM461177_untreat_paired.counts

1: GSM461176_untreat_single.counts

+ Insert Factor level**+ Insert Factor****(Optional) provide a tabular file with additional batch factors to include in the model.**

Nothing selected

History

search datasets

**assignment_part2**

7 shown

1.77 MB

7: GSM461182_untreat_single.counts**6: GSM461181_treat_paired.counts****5: GSM461180_treat_paired.counts****4: GSM461179_treat_single.counts****3: GSM461178_untreat_paired.counts**

This image shows a screenshot of a bioinformatics software interface, likely RStudio or a similar environment, used for differential gene expression analysis.

Tools

- deseq2
- Upload Data**
- Show Sections

StringTie merge transcripts

DESeq2 Determines differentially expressed features from count tables

Annotate DESeq2/DEXSeq output tables Append annotation from GTF to differential expression tool outputs

StringTie transcript assembly and quantification

featureCounts Measure gene expression in RNA-Seq experiments from SAM or BAM files

TEntranscripts annotates reads to genes and transposable elements

DiffBind differential binding analysis of ChIP-Seq peak data

History

- 1: GSM4611/b_untreat_single.counts
- 2: Factor
- 3: Factor level
- 4: Factor level
- 5: Factor level
- 6: Factor level
- 7: Factor level

1: Factor level

+ Insert Factor level

Specify a factor name, e.g. effects_drug_x or cancer_markers

Sequencing

Only letters, numbers and underscores will be retained in this field

Factor level

1: Factor level

Specify a factor level, typical values could be 'tumor', 'normal', 'treated' or 'control'

PE

Only letters, numbers and underscores will be retained in this field

Counts file(s)

7: GSM461182_untreat_single.counts
6: GSM461181_treat_paired.counts
5: GSM461180_treat_paired.counts
4: GSM461179_treat_single.counts
3: GSM461178_untreat_paired.counts
2: GSM461177_untreat_paired.counts

2: Factor level

1: GSM4611/b_untreat_single.counts

2: Factor level

3: Factor level

4: Factor level

5: Factor level

6: Factor level

7: Factor level

1.77 MB

16:08 29.12.2021

 Upload Data

 Show Sections

StringTie merge transcripts

DESeq2 Determines differentially expressed features from count tables

Annotate DESeq2/DEXSeq output tables Append annotation from GTF to differential expression tool outputs

StringTie transcript assembly and quantification

featureCounts Measure gene expression in RNA-Seq experiments from SAM or BAM files

TEtranscripts annotates reads to genes and transposable elements

DiffBind differential binding analysis of ChIP-Seq peak data

6: GSM461181_treat_paired.counts
5: GSM461180_treat_paired.counts
4: GSM461179_treat_single.counts
3: GSM461178_untreat_paired.counts
2: GSM461177_untreat_paired.counts

2: Factor level

Specify a factor level, typical values could be 'tumor', 'normal', 'treated' or 'control'

SE

Only letters, numbers and underscores will be retained in this field

Counts file(s)

7: GSM461182_untreat_single.counts
6: GSM461181_treat_paired.counts
5: GSM461180_treat_paired.counts
4: GSM461179_treat_single.counts
3: GSM461178_untreat_paired.counts
2: GSM461177_untreat_paired.counts

+ Insert Factor level

Using 7 samples

7 shown

1.77 MB



7: GSM461182_untreat_single.counts   

6: GSM461181_treat_paired.counts   

5: GSM461180_treat_paired.counts   

4: GSM461179_treat_single.counts   

3: GSM461178_untreat_paired.counts   

2: GSM461177_untreat_paired.counts   

Galaxy Europe

Workflow Visualize Shared Data Help User

Using 2%

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



deseq2



Upload Data

Download from URL or upload files

Show Sections

StringTie merge transcripts

DESeq2 Determines differentially expressed features from count tables

Annotate DESeq2/DEXSeq output tables Append annotation from GTF to differential expression tool outputs

StringTie transcript assembly and quantification

featureCounts Measure gene expression in RNA-Seq experiments from SAM or BAM files

TEtranscripts annotates reads to genes and transposable elements

DiffBind differential binding analysis of ChIP-Seq peak data



Executed **DESeq2** and successfully added 1 job to the queue.

The tool uses 7 inputs:

- 4: GSM461179_treat_single.counts
- 5: GSM461180_treat_paired.counts
- 6: GSM461181_treat_paired.counts
- 1: GSM461176_untreat_single.counts
- 2: GSM461177_untreat_paired.counts
- 3: GSM461178_untreat_paired.counts
- 7: GSM461182_untreat_single.counts

It produces 3 outputs:

- 8: DESeq2 result file on data 7, data 4, and others
- 9: DESeq2 plots on data 7, data 4, and others
- 10: Normalized counts file on data 7, data 4, and others

You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

We need your support ...

History

search datasets

assignment_part2

10 shown

1.77 MB

10: Normalized counts file on data 7, data 4, and others

9: DESeq2 plots on data 7, data 4, and others

8: DESeq2 result file on data 7, data 4, and others

7: GSM461182_untreat_single.counts

6: GSM461181_treat_paired.counts

MA plot is an application for visualization of genomic data. However, the differences between measurements are visualized, in the form of M (log ratio) and A (mean average) scales

Galaxy Europe

Workflow Visualize Shared Data Help User

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools

pathview

Upload Data

Show Sections

Pathview for pathway based data integration and visualization

EGSEA easy and efficient ensemble gene set testing

WORKFLOWS

All workflows

R Graphics Output

1 / 5

65%

PC1 57% variance

PC2 27% variance

GSM461179_treat_single.counts

GSM461176_untreat_single.counts

GSM461182_untreat_single.counts

GSM461177_untreat_paired.counts

GSM461178_untreat_paired.counts

GSM461180_treat_paired.counts

GSM461181_treat_paired.counts

group

- PE:treated
- PE:untreated
- SE:treated
- SE:untreated

History

search datasets

assignment_part2

27 shown, 8 deleted

150.53 MB

11: Genes with significant adj p-value

10: Normalized counts file on data 7, data 4, and others

9: DESeq2 plots on data 7, data 4, and others

8: DESeq2 result file on data 7, data 4, and others

7: GSM461182_untreat_single.counts

6: GSM461181_treat_pai

7°C Çok bulutlu 17:22 29.12.2021

PCA plots are based on principle component analysis and represent and handle dimensionality in data.

usegalaxy.eu

MPN navigator yön... Makale draft leman... GeoGebra - 100 mil... Middle East Technic... Probabilistic Graphi... Nizamettin Aydin » Diğer yer işaretleri Okuma listesi

Galaxy Europe Workflow Visualize Shared Data Help User Using 2%

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools

pathview

Upload Data

Show Sections

Pathview for pathway based data integration and visualization

EGSEA easy and efficient ensemble gene set testing

WORKFLOWS

All workflows

R Graphics Output 3 / 5 65% Dispersion estimates

dispersion

mean of normalized counts

● gene-est
● fitted
● final

11: Genes with significa nt adj p-value

10: Normalized counts f ile on data 7, data 4, an d others

9: DESeq2 plots on dat a 7, data 4, and others

8: DESeq2 result file on data 7, data 4, and othe rs

7: GSM461182_untreat_ single.counts

6: GSM461181_treat_nai

https://usegalaxy.eu/tool_runner?tool_id=toolshed.g2.bx.psu.edu%2Frepos%

7°C Çok bulutlu 17:22 29.12.2021

Galaxy Europe

[Workflow](#) [Visualize](#) [Shared Data](#) [Help](#) [User](#) [Logout](#) [Grid](#)

Using 2%

[IP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

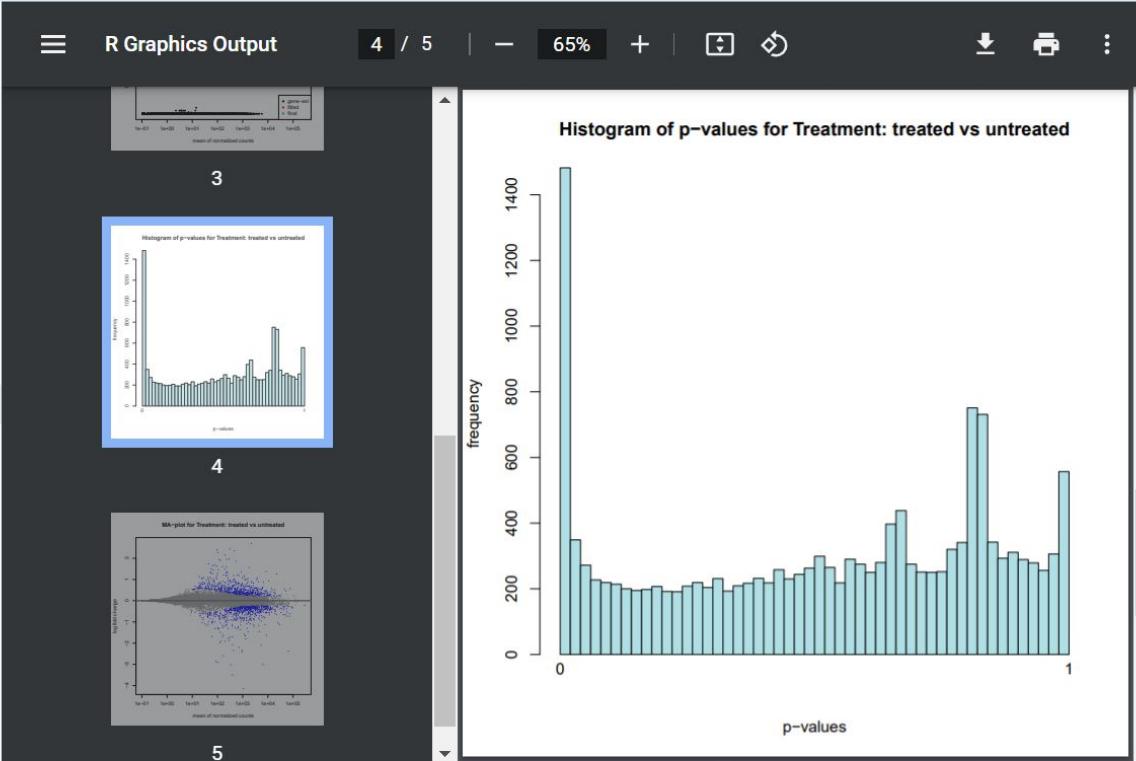
ols



pathview

[Upload Data](#)[Show Sections](#)**thview** for pathway based data integration and visualization**SEA** easy and efficient ensemble gene testing**ORKFLOWS**

workflows

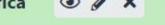
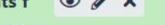
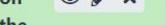
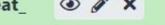
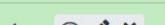
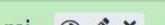
**History**

search datasets

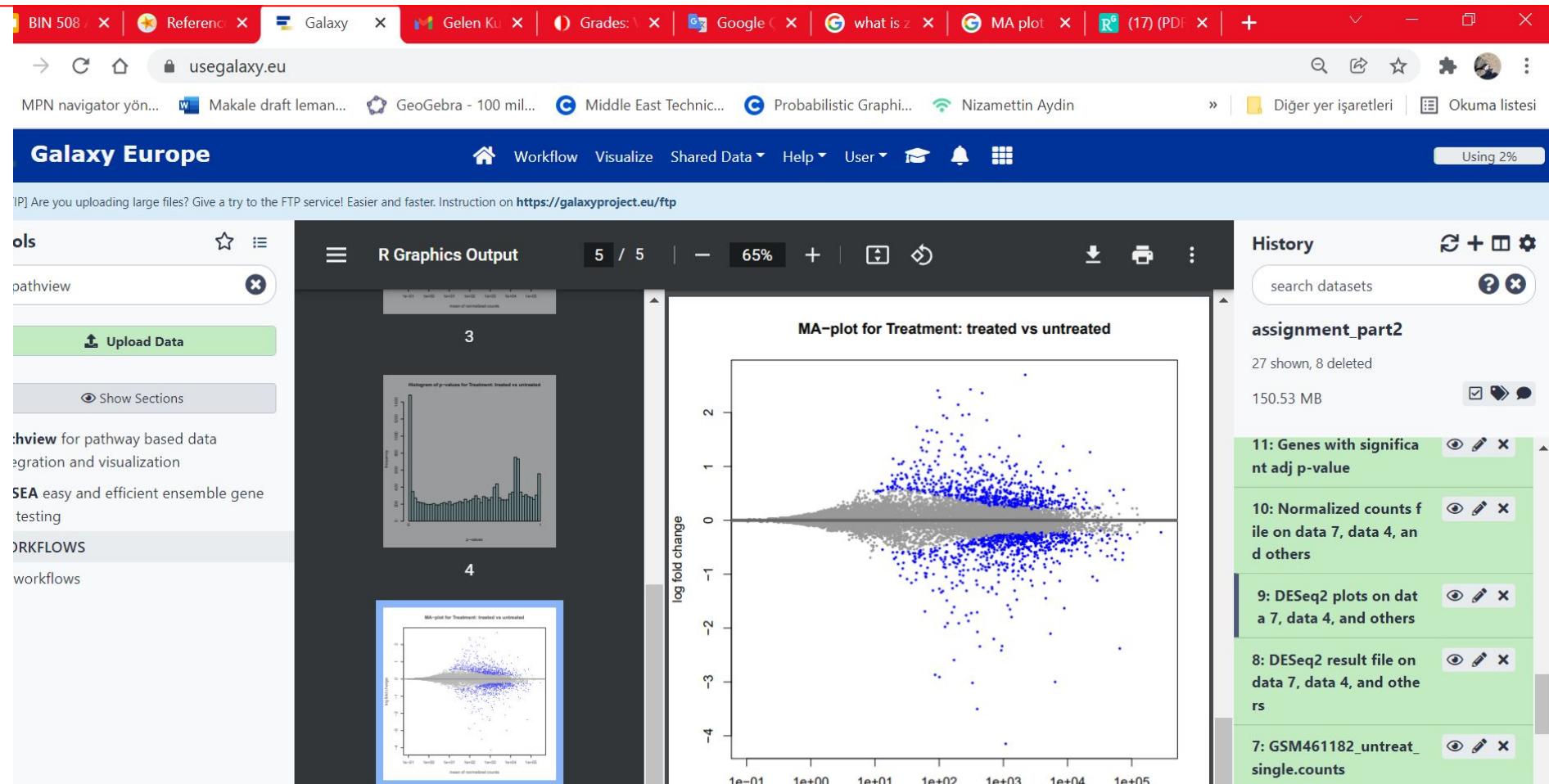
**assignment_part2**

27 shown, 8 deleted

150.53 MB

**11: Genes with significant adj p-value****10: Normalized counts file on data 7, data 4, and others****9: DESeq2 plots on data 7, data 4, and others****8: DESeq2 result file on data 7, data 4, and others****7: GSM461182_untreated_single.counts****6: GSM461181_treated_pai**

MA plot is an application for visualization of genomic data. However, the differences between measurements are visualized, in the form of M (log ratio) and A (mean average) scales



! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



Filter data on any column using simple expressions

Upload Data

Show Sections

tabular data

Create a model to recommend tools
using deep learning

Human Cell Atlas Matrix Downloader
retrieves expression matrices and
metadata from the Human Cell Atlas.

Filter data on any column using simple
expressions

Filter GFF data by attribute using
simple expressions

Filter GFF data by feature count
using simple expressions

SnpSift Filter Filter variants using
arbitrary expressions

W4m Data Subset Filter W4M data by

Filter data on any column using simple expressions (Galaxy Version 1.1.1)

Filter



8: DESeq2 result file on data 7, data 4, and others



Dataset missing? See TIP below.

With following condition

c7<0.05

Double equal signs, ==, must be used as shown above. To filter for an arbitrary string, use the Select tool.

Number of header lines to skip

0

Email notification



Send an email notification when the job completes.

Execute

Double equal signs, ==, must be used as "equal to" (e.g., c1 == 'chr22')

TIP: Attempting to apply a filtering condition may throw exceptions if the data type (e.g., string, integer) in every line of the

History



search datasets



assignment_part2

10 shown

1.77 MB



10: Normalized counts file on data 7, data 4, and others



9: DESeq2 plots on data 7, data 4, and others



8: DESeq2 result file on data 7, data 4, and others



7: GSM461182_untreat_single.counts



6: GSM461181_treat_pai_red.counts





! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



Filter data on any column using simple expressions

Upload Data

Show Sections

tabular data

Create a model to recommend tools
using deep learning

Human Cell Atlas Matrix Downloader
retrieves expression matrices and
metadata from the Human Cell Atlas.

Filter data on any column using simple
expressions

Filter GFF data by attribute using
simple expressions

Filter GFF data by feature count
using simple expressions

SnpSift Filter Filter variants using
arbitrary expressions

W4m Data Subset Filter W4M data by

Edit Dataset Attributes

Attributes

Convert

Datatypes

Permissions

Name

Genes with significant adj p-value

Info

Annotation

Add an annotation or notes to a dataset; annotations are available when a history is viewed.

Database/Build

unspecified (?)

Number of comment lines

History

search datasets

assignment_part2

11 shown

Rename history...

1.77 MB



11: Filter on data 8

10: Normalized counts file on data 7, data 4, and others

9: DESeq2 plots on data 7, data 4, and others

8: DESeq2 result file on data 7, data 4, and others

7: GSM461182_untreat_single.counts

6: GSM461181_treat_pai

usegalaxy.eu

MPN navigator yön... Makale draft leman... GeoGebra - 100 mil... Middle East Technic... Probabilistic Graphi... Nizamettin Aydin

Diger yer işaretleri Okuma listesi

Galaxy Europe

Workflow Visualize Shared Data Help User

Using 2%

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools

search tools

Upload Data

transcripts to a reference annotation

Filter FASTA on the headers and/or the sequences

Column arrange by header name

XPath compute xpath expressions on XML data

Sub-sample sequences files e.g. to reduce coverage

Filter sequences by ID from a tabular file

find in reference filter peptides that are present in proteins

Filter data on any column using simple expressions

Sort data in ascending or descending order

Filter data on any column using simple expressions (Galaxy Version 1.1.1)

11: Genes with significant adj p-value

Dataset missing? See TIP below.

With following condition

abs(c3)>1

Double equal signs, ==, must be used as shown above. To filter for an arbitrary string, use the Select tool.

Number of header lines to skip

0

Email notification

Send an email notification when the job completes.

Execute

Double equal signs, ==, must be used as "equal to" (e.g., c1 == 'chr22')

TIP: Attempting to apply a filtering condition may throw exceptions if the data type (e.g., string, integer) in every line of the columns being filtered is not appropriate for the condition (e.g., attempting certain numerical calculations on strings). If an exception is thrown when applying the condition to a line, that line is skipped as invalid for the filter condition. The number

History

search datasets

assignment_part2

11 shown

1.77 MB

11: Genes with significa nt adj p-value

10: Normalized counts f ile on data 7, data 4, an d others

9: DESeq2 plots on data 7, data 4, and others

8: DESeq2 result file on data 7, data 4, and othe rs

7: GSM461182_untreat_ single.counts

7°C Çok bulutlu 16:23 29.12.2021



! [TIP] Are you uploading large files? Give a try to

Tools

search tools

Upload Data

transcripts to a reference annotation

Filter FASTA on the headers and/or the sequences**Column arrange** by header name**XPath** compute xpath expressions on XML data**Sub-sample sequences files** e.g. to reduce coverage**Filter sequences by ID** from a tabular file**find in reference** filter peptides that are present in proteins**Filter** data on any column using simple expressions**Sort** data in ascending or descending order

Download from web or upload from disk

Regular

Composite

Collection

Rule-based

Name	Size	Type	Genome	Settings	Status
New File	76 b	Auto-det...	unspecified (?)		100%

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

```
https://zenodo.org/record/4541751/files/Drosophila_melanogaster.BDGP6.87.gtf
```

Type (set all):

Auto-detect



Genome (set all):

unspecified (?)

Choose local files

Choose remote files

Paste/Fetch data

Start

Pause

Reset

Close

history

Assignment_part2

3 shown

77 MB

13: Drosophila_melanogaster.BDGP6.87.gtf

12: Filter on data 11

1: Genes with significant adj p-value

0: Normalized counts for on data 7, data 4, and others

DESeq2 plots on data 4, and others

8: DESeq2 result file on

Using 2%

Galaxy Europe

Workflow Visualize Shared Data Help User

Using 2%

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



Annotate DESeq2/DEXSeq output

Upload Data

Show Sections

StringTie merge transcripts

Annotate DESeq2/DEXSeq output
tables Append annotation from GTF to differential expression tool outputs

DEXSeq Determines differential exon usage from count tables

DESeq2 Determines differentially expressed features from count tables

plotDEXSeq Visualization of the per gene DEXSeq results

Join HUMAnN2 generated tables

List spaln parameter tables Given a query species, list the spaln settings tables that exist, from closest related species to most different

>Annotate DESeq2/DEXSeq output tables Append annotation from GTF to differential expression tool outputs (Galaxy Version 1.1.0)

Tabular output of DESeq2/edgeR/limma/DEXSeq



12: Filter on data 11



(-in)

Input file type

DESeq2/edgeR/limma

(-m)

Reference annotation in GFF/GTF format



13: Drosophila_melanogaster.BDGP6.87.gtf



(-g)

Advanced options



Email notification



Send an email notification when the job completes.

Execute

History

search datasets

assignment_part2

13 shown

150.53 MB

13: Drosophila_melanogaster.BDGP6.87.gtf

12: Filter on data 11

11: Genes with significa
nt adj p-value

10: Normalized counts f
ile on data 7, data 4, an
d others

9: DESeq2 plots on data
7, data 4, and others

8: DESeq2 result file on
data 7, data 4, and othe

16:25

Upload Data

Show Sections

HiGlass an interactive Hi-C data visualizer.

MSI data exporter exports imzML and Analyze7.5 to tabular files

Faster Download and Extract Reads in FASTQ format from NCBI SRA

MSI plot spectra mass spectrometry imaging mass spectra plots

GROMACS energy minimization of the system prior to equilibration and production MD

MSI classification spatial classification of mass spectrometry imaging data

MSI segmentation mass spectrometry imaging spatial clustering

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

https://zenodo.org/record/4541751/files/Drosophila_melanogaster.BDGP6.87.gtf

New File 103 b tabular unspecified (?) 100%

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

ID	Base mean	log2(FC)	StdErr	Wald-Stats	P-value	P-adj	Chromosome	Start	End	Strand	Feature	Gene name

Type (set all): Genome (set all):

Paste/Fetch data Start Pause Reset Close

2010 update.
Nucleic Acids Res. 44, W3–W10
doi:10.1093/nar/gkw343

Assignment_part2

1 shown

50.53 MB

4: Annotate DESeq2/D
KSeq output tables on
data 13 and data 12

3: Drosophila_melanog
ster.BDGP6.87.gtf

2: Filter on data 11

1: Genes with significa
nt adj p-value

0: Normalized counts f
e on data 7, data 4, an
others

9: DESeq2 plots on data

16:26

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



Join two Datasets



Upload Data

Show Sections

DEV Datasets

VCFcombine: Combine multiple VCF datasets

Naive Variant Caller - tabulate variable sites from BAM datasets

Discriminant Analysis

Ensemble methods for classification and regression

Generate pileup from BAM dataset

Slice BAM by provided regions

Join two Datasets side by side on a specified field

Column join on multiple datasets

Column join on multiple datasets



Executed **Join two Datasets** and successfully added 1 job to the queue.

The tool uses 2 inputs:

- 10: Normalized counts file on data 7, data 4, and others
- 15: Genes with significant adj p-value & abs(FC) > 2

It produces this output:

- 16: Join two Datasets on data 15 and data 10

You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

PHD Comics

Random

History



search datasets



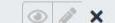
assignment_part2

16 shown

150.53 MB



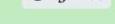
16: Join two Dataset s on data 15 and data 1 0



15: Genes with significa nt adj p-value & abs(FC) > 2



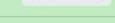
14: Annotate DESeq2/D EXSeq output tables on data 13 and data 12



13: Drosophila_melanog aster.BDGP6.87.gtf



12: Filter on data 11



7°C Çok bulutlu 16:33
29.12.2021

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



cut

Upload Data

Show Sections

enzymes

chicAggregateStatistic computes with a target file the to be tested regions for **chicDifferentialTest**

Conformer calculation for molecules (confab)

hicPCA compute the principal components for A / B compartment analysis

ExportToSpreadsheet export measurements into one or more files



Executed **Cut** and successfully added 1 job to the queue.

The tool uses this input:

- 16: Join two Datasets on data 15 and data 10

It produces this output:

- 17: Cut on data 16

You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

We need your support ...

If Galaxy helped with the analysis of your data, please do not forget to **cite**:

Afgan E et al. 2016 The Galaxy platform for accessible, reproducible and collaborative biomedical analyses:

History



search datasets



assignment_part2

17 shown

150.53 MB



17: Cut on data 16



16: Join two Datasets o n data 15 and data 10



15: Genes with significa nt adj p-value & abs(FC) > 2



14: Annotate DESeq2/D EXSeq output tables on data 13 and data 12



Galaxy Europe



Workflow Visualize Shared Data Help User



Using 2%

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



cut

**Upload Data****Show Sections**

enzymes

chicAggregateStatistic computes with a target file the to be tested regions for **chicDifferentialTest**

Conformer calculation for molecules (confab)

hicPCA compute the principal components for A / B compartment analysis

ExportToSpreadsheet export measurements into one or more files

Map with Bowtie for SOLiD

Kinwalker cotranscriptional folding of RNAs

Edit Dataset Attributes

Attributes updated.

Attributes

Convert

Datatypes

Permissions

Name

Normalized counts for the most differentially expressed genes

Info

Annotation

Add an annotation or notes to a dataset; annotations are available when a history is viewed.

Database/Build

History

search datasets



assignment_part2

17 shown

150.53 MB



17: Normalized counts f
or the most differenti
ally expressed genes

16: Join two Datasets o
n data 15 and data 10

15: Genes with significa
nt adj p-value & abs(FC)
> 2

14: Annotate DESeq2/D
EXSeq output tables on
data 13 and data 12

13: Drosophila_melanog
aster.BDGP6.87.gtf



! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



heatmap2



Upload Data

Show Sections

heatmap2

WORKFLOWS

All workflows

Cluster rows and columns

Distance method

Euclidean

The method used to compute the distance (dissimilarity) between both rows and columns. Default: Euclidean

Clustering method

Complete

The method used to compute the hierarchical clustering. Default: Complete

Labeling columns and rows

Label columns and not rows

Coloring groups

Blue to white to red

Data scaling

Do not scale my data

Email notification



Send an email notification when the job completes.

Execute

History



search datasets

assignment_part2

17 shown

150.53 MB



17: Normalized counts f
or the most differentia
lly expressed genes

16: Join two Datasets o
n data 15 and data 10

15: Genes with significa
nt adj p-value & abs(FC)
> 2

14: Annotate DESeq2/D
EXSeq output tables on
data 13 and data 12

13: Drosophila_melanog
aster.BDGP6.87.gtf



Tools



heatmap2



Upload Data

Show Sections

heatmap2

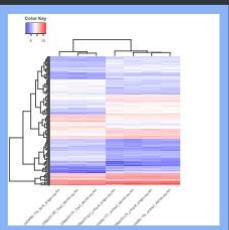
WORKFLOWS

All workflows

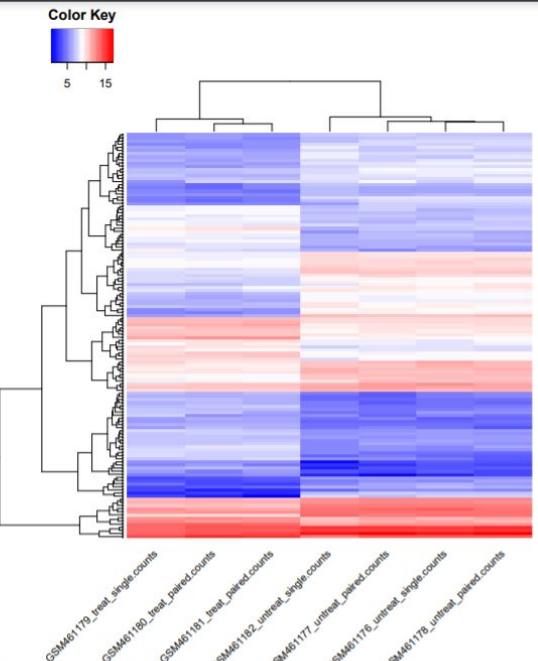
R Graphics Output

1 / 1

65%



1



History

search datasets



assignment_part2

18 shown

150.53 MB



18: heatmap2 on data

17

17: Normalized counts f
or the most differentia
lly expressed genes16: Join two Datasets o
n data 15 and data 1015: Genes with significa
nt adj p-value & abs(FC)
> 214: Annotate DESeq2/D
EXSeq output tables on
data 13 and data 12

Galaxy Europe



Workflow

Visualize

Shared Data

Help

User



Using 2%

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



table compute


[Upload Data](#)
[Show Sections](#)

goseq tests for overrepresented gene categories

Scanpy ComputeGraph to derive kNN graph

Convert SAM to interval

dada2: makeSequenceTable construct a sequence table (analogous to OTU table)

Table Compute computes operations on table data

Between-table Correlation

Correlation table between two tables and graphic representation

Calculate presence absence table



17: Normalized counts for the most differentially expressed genes



Input data has

 Select/Unselect all

- Column names on the first row
- Row names on the first column

[Advanced File Options](#)

Type of table operation

Perform a full table operation

Operation

Custom

See Examples 5, 7, and 8 for usage

Custom expression on 'table', along 'axis' (0 or 1)

```
table.sub(table.mean(1), 0)
```

The parameter name is `table` and `axis`, referring to the table being acted on and the column (0) or row (1) to perform the operation on. Numpy, math, Pandas DataFrame operators, and inline `if` `else` are supported (e.g. `np.log(table) - table.mean(0) / table.std(1)`). See Example #5 in the Help section.

[Output formatting options](#)

History

[search datasets](#)


assignment_part2

18 shown

150.53 MB



18: heatmap2 on data 1

7



17: Normalized counts f or the most differenti ally expressed genes



16: Join two Datasets o n data 15 and data 10



15: Genes with significa nt adj p-value & abs(FC) > 2



14: Annotate DESeq2/D EXSeq output tables on data 13 and data 12



Galaxy Europe



Workflow Visualize Shared Data Help User



Using 2%

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools

table compute

**Upload Data****Show Sections****goseq** tests for overrepresented gene categories**Scanpy ComputeGraph** to derive kNN graph**Convert SAM** to interval**dada2: makeSequenceTable** construct a sequence table (analogous to OTU table)**Table Compute** computes operations on table data

Between-table Correlation

Correlation table between two tables and graphic representation

Calculate presence absence table



Input data has

 Select/Unselect all Column names on the first row Row names on the first column

Advanced File Options

2: Tables



Table



19: Table Compute on data 17



Input data has

 Select/Unselect all Column names on the first row Row names on the first column

Advanced File Options

+ Insert Tables

Custom expression on 'tableN'

table2 div/table1 std(1) 0

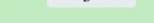
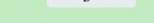
History

search datasets

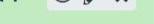
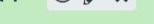
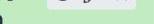
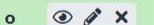
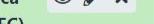
**assignment_part2**

19 shown

150.53 MB

**19: Table Compute on data 17****18: heatmap2 on data 1**

7

**17: Normalized counts for the most differentially expressed genes****16: Join two Datasets on data 15 and data 10****15: Genes with significant adj p-value & abs(FC) > 2**

7°C Çok bulutlu



16:40

29.12.2021

MPN navigator yön... Makale draft leman... GeoGebra - 100 mil... Middle East Technic... Probabilistic Graphi... Nizamettin Aydin » Diğer yer işaretleri Okuma listesi

Galaxy Europe

Workflow Visualize Shared Data Help User

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Edit Dataset Attributes

Attributes Convert Datatypes Permissions

Name: Z-scores for the most differentially expressed genes

Info:

Annotation:

Add an annotation or notes to a dataset; annotations are available when a history is viewed.

Database/Build: unspecified (?)

Number of comment lines:

History

search datasets

assignment_part2

20 shown
150.53 MB

20: Table Compute on data 19 and data 17

19: Table Compute on data 17

18: heatmap2 on data 17

17: Normalized counts for the most differentially expressed genes

16: Join two Datasets on data 15 and data 10

15: Genes with significa

Upload Data

Show Sections

goseq tests for overrepresented gene categories

Scanpy ComputeGraph to derive kNN graph

Convert SAM to interval

dada2: makeSequenceTable construct a sequence table (analogous to OTU table)

Table Compute computes operations on table data

Between-table Correlation

Correlation table between two tables and graphic representation

Calculate presence absence table

7°C Çok bulutlu 16:40 29.12.2021

Galaxy Europe

Workflow Visualize Shared Data Help User

Using 2%

TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools	GSM461176_untreat_single.counts	GSM461177_untreat_paired.counts	GSM461178_untreat_paired.counts	GSM
pathview	FBgn0037223	1.222027	0.674483	0.415875
	FBgn0037223	1.222027	0.674483	0.415875
	FBgn0037290	-0.822673	-0.815953	-0.869803
	FBgn0037290	-0.822673	-0.815953	-0.869803
	FBgn0000071	-0.801176	-0.748954	-0.699035
	FBgn0000071	-0.801176	-0.748954	-0.699035
	FBgn0037646	1.273448	0.333272	0.901618
	FBgn0037646	1.273448	0.333272	0.901618
	FBgn0037678	-0.707284	-0.387901	-1.017410
	FBgn0037678	-0.707284	-0.387901	-1.017410
	FBgn0261552	1.409067	0.855967	0.342847
	FBgn0261552	1.409067	0.855967	0.342847
	FBgn0002868	1.218196	-0.033963	0.225919
	FBgn0002868	1.218196	-0.033963	0.225919
	FBgn0037754	0.828921	0.438524	0.236678
	FBgn0037754	0.828921	0.438524	0.236678
	FBgn0265276	-0.608986	-0.706105	-0.900625
	FBgn0265276	-0.608986	-0.706105	-0.900625
	FBgn0038149	-0.843508	-0.792375	-0.672742
	FBgn0038149	-0.843508	-0.792375	-0.672742
	FRgn0038198	-0.795260	-0.634554	-1.011352

History
search datasets
assignment_part2
27 shown, 8 deleted
150.53 MB
ential expression
21: Compute on data 14
20: Z-scores for the most differentially expressed genes
19: Table Compute on data 17
18: heatmap2 on data 17
17: Normalized counts for the most differentially expressed genes

z-scores is the standard deviations from the mean, by using this, significant genes can be selected from the based on their deviations from the mean values

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

MPN navigator yon... Makale dritt Ieman... GeoGebra - 100 mil... Middle East Technic... Probabilistic Graphi... Nizamettin Aydin Diger yer işaretleri Okuma listesi

Galaxy Europe Workflow Visualize Shared Data Help User Using 2%

Tools

- table compute X
- Upload Data
- Show Sections

goseq tests for overrepresented gene categories

Scanpy ComputeGraph to derive kNN graph

Convert SAM to interval

dada2: makeSequenceTable construct a sequence table (analogous to OTU table)

Table Compute computes operations on table data

Between-table Correlation Correlation table between two tables and graphic representation

Calculate presence absence table

	GSM461176_untreat_single.counts	GSM461177_untreat_paired.counts	GSM461178_untreat_paired.counts	GSM
FBgn0037223	1.222027	0.674483	0.415875	
FBgn0037223	1.222027	0.674483	0.415875	
FBgn0037290	-0.822673	-0.815953	-0.869803	
FBgn0037290	-0.822673	-0.815953	-0.869803	
FBgn0000071	-0.801176	-0.748954	-0.699035	
FBgn0000071	-0.801176	-0.748954	-0.699035	
FBgn0037646	1.273448	0.333272	0.901618	
FBgn0037646	1.273448	0.333272	0.901618	
FBgn0037678	-0.707284	-0.387901	-1.017410	
FBgn0037678	-0.707284	-0.387901	-1.017410	
FBgn0261552	1.409067	0.855967	0.342847	
FBgn0261552	1.409067	0.855967	0.342847	
FBgn0002868	1.218196	-0.033963	0.225919	
FBgn0002868	1.218196	-0.033963	0.225919	
FBgn0037754	0.828921	0.438524	0.236678	
FBgn0037754	0.828921	0.438524	0.236678	
FBgn0265276	-0.608986	-0.706105	-0.900625	
FBgn0265276	-0.608986	-0.706105	-0.900625	
FBgn0038149	-0.843508	-0.792375	-0.672742	
FBgn0038149	-0.843508	-0.792375	-0.672742	
FBgn0038198	-0.795260	-0.634554	-1.011352	

History

- search datasets
- assignment_part2
 - 20 shown
 - 150.53 MB
- 20: Z-scores for the most differentially expressed genes
- 19: Table Compute on data 17
- 18: heatmap2 on data 17
- 17: Normalized counts for the most differentially expressed genes
- 16: Join two Datasets on data 15 and data 10

Galaxy Europe



Workflow

Visualize

Shared Data

Help

User



Using 2%

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



compute

Upload Data**Show Sections****hicPCA** compute the principal components for A / B compartment analysis**Flight curve** compute the regional expected pattern of abundance**SNP distance matrix** Compute distance in SNPs between all sequences in a FASTA file**bedtools OverlapBed** computes the amount of overlap from two intervals**Compute** an expression on every row**Samtools depth** compute the depth at each position or region**bedtools Compute both the depth and breadth of coverage** of features**Compute** an expression on every row (Galaxy Version 1.6)

Add expression

bool(c7<0.05)

as a new column to



14: Annotate DESeq2/DEXSeq output tables on data 13 and data 12



Dataset missing? See TIP below

Round result?



No

Avoid scientific notation



No

If yes, use fully expanded decimal representation when writing new columns (use only if expression produces decimal numbers).

Input has a header line with column names?

No

Select Yes to be able to specify a name for the new column and have it added to the header line. If you select No, the first line will be treated as a regular line: If it is empty or starts with a # character it will be skipped, otherwise the tool will attempt to compute the specified expression on it.

Email notification



History

search datasets



assignment_part2

20 shown

150.53 MB



15: Genes with significant adj p-value & abs(FC) > 2

14: Annotate DESeq2/DEXSeq output tables on data 13 and data 12

13: Drosophila_melanogaster.BDGP6.87.gtf

12: Filter on data 11

11: Genes with significant adj p-value

10: Normalized counts f



! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



cut

[Upload Data](#)[Show Sections](#)

~~cutseq~~ REMOVES A SPECIFIED SECTION from a sequence

seqtk cutN cut sequence at long N

Clearcut Generate a tree using relaxed neighbor joining

Cutadapt Remove adapter sequences from FASTQ/Fasta

Association of points in consecutive frames (slices) using the nearest neighbor algorithm

SpectraSTSearchAdapter Interface to the SEARCH Mode of the SpectraST executable

Cut columns from a table

Differential Cleavage : Select SNPs

Cut columns from a table (Galaxy Version 1.0.2)

Cut columns

c1,c8

Delimited by

Tab

From



21: Compute on data 14



Email notification



Send an email notification when the job completes.

[Execute](#)

WARNING: This tool breaks column assignments. To re-establish column assignments run the tools and click on the pencil icon in the latest history item.

i The output of this tool is always in tabular format (e.g., if your original delimiters are commas, they will be replaced with tabs). For example:

Cutting columns 1 and 3 from:

History

[search datasets](#)**assignment_part2**

21 shown

150.53 MB

21: Compute on data 14**20: Z-scores for the most differentially expressed genes****19: Table Compute on data 17****18: heatmap2 on data 17****17: Normalized counts for the most differentially expressed genes****16: Join two Datasets o**

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



change case



Upload Data

Show Sections

StringTie merge transcripts

goseq tests for overrepresented gene categories

Support vector machines (SVMs) for classification

Ensemble methods for classification and regression

Change Case of selected columns

SnpSift CaseControl Count samples are in 'case' and 'control' groups.

SnpSift CaseControl Count samples are in 'case' and 'control' groups.

Create InterMine Interchange Dataset

Intensity Check Statistical measures,



Change Case of selected columns (Galaxy Version 1.0.0)



From



22: Cut on data 21



Change case of columns

c1

Delimited by

Tab

To

Upper case

Email notification



Send an email notification when the job completes.

Execute

⚠ This tool breaks column assignments. To re-establish column assignments run the tool and click on the pencil icon in the resulting history item.

⚠ The format of the resulting dataset from this tool is always tabular.

History



search datasets



assignment_part2

22 shown

150.53 MB



22: Cut on data 21



21: Compute on data 14



20: Z-scores for the most differentially expressed genes



19: Table Compute on data 17



18: heatmap2 on data 7



17: Normalized counts for the most differentially expressed genes





[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



change case

Upload Data

Show Sections

StringTie merge transcripts**goseq** tests for overrepresented gene categories**Support vector machines (SVMs)** for classification**Ensemble methods** for classification and regression**Change Case** of selected columns**GnpSift CaseControl** Count samples are in 'case' and 'control' groups.**GnpSift CaseControl** Count samples are in 'case' and 'control' groups.**Create InterMine Interchange** Dataset**Intensity Check** Statistical measures,

Edit Dataset Attributes

Attributes

Convert

Datatypes

Permissions

Name

Gene IDs and differential expression

Info

Annotation

Add an annotation or notes to a dataset; annotations are available when a history is viewed.

Database/Build

unspecified (?)

Number of comment lines

History

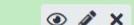
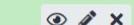


search datasets

**assignment_part2**

22 shown

150.53 MB

**22: Cut on data 21****21: Compute on data 14****20: Z-scores for the most differentially expressed genes****19: Table Compute on data 17****18: heatmap2 on data 17****17: Normalized counts for the most differentially expressed genes**

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

search histories (x)

search all datasets (?) (x) ...

Create new

Current History

assignment_part2

23 shown

150.53 MB

search datasets (?) (x)

23: featureCounts on data 21 and data 24: Feature lengths

22: Gene IDs and differential expression

21: Compute on data 14

20: Z-scores for the most differentially expressed genes

19: Table Compute on data 17

assignment_5_508

29 shown, 10 hidden

5.91 GB

search datasets (?) (x)

35: MultiQC on data 32: Webpage

34: MultiQC on data 32: Stats

a list with 3 items

33: featureCounts on data 21 and data 24: Feature lengths

32: featureCounts on data 21 and data 24: Summary

31: featureCounts on data 21 and data 24: Counts

30: Infer Experiment on data 23 and data 24

26: MultiQC on data 22: Webpage

25: MultiQC on data 22: Stats

Switch to

variant calling exercise

16 shown, 8 deleted

251.28 KB

search datasets (?) (x)

24: GEMINI query on data 15

23: GEMINI query on data 15

22: GEMINI query on data 15

21: GEMINI query on data 15

20: GEMINI query on data 15

19: GEMINI query on data 15

17: GEMINI database info on data 15

16: GEMINI query on data 15

Switch to



[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



change case



Upload Data

Show Sections

database.

Map with BWA-MEM - map medium and long reads (> 100 bp) against reference genome

StringTie merge transcripts**goseq** tests for overrepresented gene categories**Support vector machines (SVMs)** for classification**Ensemble methods** for classification and regression**Change Case** of selected columns**SnpSift CaseControl** Count samples are in 'case' and 'control' groups.**SnpSift CaseControl** Count samples

Change Case of selected columns (Galaxy Version 1.0.0)

From



23: featureCounts on data 21 and data 24: Feature lengths



Change case of columns

c1

Delimited by

Tab

To

Upper case

Email notification



Send an email notification when the job completes.

Execute

⚠ This tool breaks column assignments. To re-establish column assignments run the tool and click on the pencil icon in the resulting history item.

⚠ The format of the resulting dataset from this tool is always tabular.

History



search datasets



assignment_part2

23 shown

150.53 MB



23: featureCounts on data 21 and data 24: Feature lengths



22: Gene IDs and differential expression



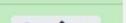
21: Compute on data 14



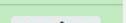
20: Z-scores for the most differentially expressed genes



19: Table Compute on data 17



18: heatmap2 on data 1



Galaxy Europe



Workflow

Visualize

Shared Data

Help

User



Using 2%

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools

change case

**Upload Data****Show Sections**

database.

Map with BWA-MEM - map medium and long reads (> 100 bp) against reference genome**StringTie merge** transcripts**goseq** tests for overrepresented gene categories**Support vector machines (SVMs)** for classification**Ensemble methods** for classification and regression**Change Case** of selected columns**SnpSift CaseControl** Count samples are in 'case' and 'control' groups.**SnpSift CaseControl** Count samples

Edit Dataset Attributes

Attributes**Convert****Datatypes****Permissions****Name**

Gene IDs and length

Info**Annotation**

Add an annotation or notes to a dataset; annotations are available when a history is viewed.

Database/Build

D. melanogaster Aug. 2014 (BDGP Release 6 + ISO1 MT/dm6) (dm6)

Number of comment lines**History**

search datasets

**assignment_part2**

24 shown

150.53 MB

24: Change Case on data 23**23: featureCounts on data 21 and data 24: Feature lengths****22: Gene IDs and differential expression****21: Compute on data 14****20: Z-scores for the most differentially expressed genes****19: Table Compute on d**

7°C Çok bulutlu



16:46

29.12.2021

Tools

goseq

**Upload Data****Show Sections****goseq** tests for overrepresented gene categories**goseq** tests for overrepresented gene categories**WORKFLOWS**

All workflows

You can obtain a mapping of genes to categories (for some genomes only) or you can provide your own category file.

Select a genome to use

Fruit fly (dm6)

Select Gene ID format

Ensembl Gene ID

Supported Gene IDs to automatically fetch categories should either be Entrez, Ensembl, or gene symbols.

Select one or more categories Select/Unselect all

- GO: Cellular Component
- GO: Biological Process
- GO: Molecular Function
- KEGG

By default, goseq tests all three major Gene Ontology branches; Cellular Component, Biological Process and Molecular Function. However, it is possible to limit testing to any combination and/or to also use KEGG pathways.

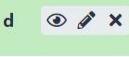
Method OptionsAdvanced OptionsOutput OptionsOutput Top GO terms plot?**History**

search datasets

**assignment_part2**

24 shown

150.53 MB

**24: Gene IDs and length****23: featureCounts on data 21 and data 24: Feature lengths****22: Gene IDs and differential expression****21: Compute on data 14****20: Z-scores for the most differentially expressed genes****19: Table Compute on data 17**

7°C Çok bulutlu 16:49
29.12.2021

error to create go and kegg pathways?

For this reason, I could not do the parts in the last part of the kegg pathway tutorial.

Tools

goseq X

Upload Data

Show Sections

goseq tests for overrepresented gene categories
goseq tests for overrepresented gene categories

WORKFLOWS

All workflows

24: Gene IDs and length

You can calculate the gene lengths using featureCounts or the Gene length and GC content tool.

Gene categories

Get categories

You can obtain a mapping of genes to categories (for some genomes only) or you can provide your own category file.

Select a genome to use

Fruit fly (dm6)

Select Gene ID format

Ensembl Gene ID

Supported Gene IDs to automatically fetch categories should either be Entrez, Ensembl, or gene symbols.

Select one or more categories

Select/Unselect all

GO: Cellular Component
 GO: Biological Process
 GO: Molecular Function
 KEGG

By default, goseq tests all three major Gene Ontology branches; Cellular Component, Biological Process and Molecular Function. However, it is possible to limit testing to any combination and/or to also use KEGG pathways.

Method Options

History

search datasets X

assignment_part2

27 shown
150.53 MB

27: goseq on data 24 and data 22: DE genes f or categories (GO/KEGG terms)

26: goseq on data 24 and data 22: Top over-r epresented GO terms plot

25: goseq on data 24 and data 22: Ranked cat egory list - Wallenius method

24: Gene IDs and length

23: featureCounts on da ta 21 and data 24: Featu re lengths

Galaxy Europe



Workflow

Visualize

Shared Data

Help

User



Using 2%

! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools

cut

Upload Data

Show Sections

SpectraSTSearchAdapter Interface to the SEARCH Mode of the SpectraST executable**Cut** columns from a table**Differential Cleavage** : Select SNPs differentially cut by specified restriction enzymes**chicAggregateStatistic** computes with a target file the to be tested regions for chicDifferentialTest**Conformer calculation** for molecules (confab)**hicPCA** compute the principal components for A / B compartment analysis

Cut columns from a table (Galaxy Version 1.0.2)

Cut columns

c1,c3

Delimited by

Tab

From



11: Genes with significant adj p-value



Email notification



Send an email notification when the job completes.

Execute

WARNING: This tool breaks column assignments. To re-establish column assignments run the tools and click on the pencil icon in the latest history item.

i The output of this tool is always in tabular format (e.g., if your original delimiters are commas, they will be replaced with tabs). For example:

Cutting columns 1 and 3 from:

History

search datasets

**assignment_part2**

25 shown, 8 deleted

150.53 MB



33: goseq on data 24
and data 22: Ranked category list - Wallenius method

24: Gene IDs and length



23: featureCounts on data 21 and data 24: Feature lengths



22: Gene IDs and differential expression



21: Compute on data 14



20: Z-scores for the most differentially expressed



7°C

Çok bulutlu



17:09



! [TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



cut

Upload Data**Show Sections**

SpectraSTSearchAdapter Interface to the SEARCH Mode of the SpectraST executable

Cut columns from a table

Differential Cleavage : Select SNPs differentially cut by specified restriction enzymes

chicAggregateStatistic computes with a target file the to be tested regions for chicDifferentialTest

Conformer calculation for molecules (confab)

hicPCA compute the principal components for A / B compartment

Edit Dataset Attributes

Attributes

Convert

Datatypes

Permissions

Name

Genes with significant adj p-value and their Log2 FC

Info

Annotation

Add an annotation or notes to a dataset; annotations are available when a history is viewed.

Database/Build

History

search datasets



assignment_part2

26 shown, 8 deleted

150.53 MB



34: Cut on data 11



33: goseq on data 24 and data 22: Ranked category list - Wallenius method

24: Gene IDs and length



23: featureCounts on data 21 and data 24: Feature lengths

22: Gene IDs and differential expression



Galaxy Europe

Workflow Visualize Shared Data Help User

Using 2%

[TIP] Are you uploading large files? Give a try to the FTP service! Easier and faster. Instruction on <https://galaxyproject.eu/ftp>

Tools



pathview

Upload Data**Show Sections****pathview** for pathway based data

integration and visualization

GSEA easy and efficient ensemble gene

set testing

WORKFLOWS

I workflows

FBgn0039155	-4.14844993707696
FBgn0003360	-2.99977727874955
FBgn0026562	-2.38016404990504
FBgn0025111	2.69993883051315
FBgn0029167	-2.10506155637674
FBgn0039827	-3.5030140832698
FBgn0035085	-2.41407407909701
FBgn0034736	-3.01817864352459
FBgn0264475	-2.33448625219322
FBgn0000071	2.36001723955655
FBgn0029896	-2.20240483862708
FBgn0038832	-2.34071110479949
FBgn0037754	-2.0719936164172
FBgn0264753	-2.59345911905122
FBgn0034897	-1.7951065076634
FBgn0035189	2.42860686966053
FBgn0011260	1.96746091024963
FBgn0027279	-1.16556778314061
FBgn0040099	-1.65424331975783
FBgn0024288	-2.86781820432587
FBgn0001226	1.56713548793081
FBgn0085359	-2.90913823328661
FBgn0039155	-4.14844993707696

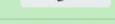
History

search datasets

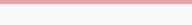
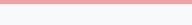
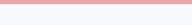
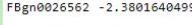
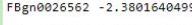
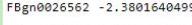
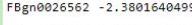
**assignment_part2**

27 shown, 8 deleted

150.53 MB

35: Pasted Entry**34: Genes with significant adj p-value and their Log2 FC**

1,091 lines

format: **tabular**, database: ?

usegalaxy.eu/datasets/edit

MPN navigator yön... Makale draft leman... GeoGebra - 100 mil... Middle East Technic... Probabilistic Graphi... Nizamettin Aydin Diğer yer işaretleri Okuma listesi

Galaxy Europe

Workflow Visualize Shared Data Help User

Using 2%

[TIP] Are you uploading large files? Give a try to

Tools

cut

Upload Data

Show Sections

SpectraSTSearchAdapter Interface to the SEARCH Mode of the SpectraST executable

Cut columns from a table

Differential Cleavage : Select SNPs differentially cut by specified restriction enzymes

chicAggregateStatistic computes with a target file the to be tested regions for chicDifferentialTest

Conformer calculation for molecules

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	11 b	tabular	unspecified (?)	0%	
Download data from the web by entering URLs (one per line) or directly paste content.					
00010 03040					

Type (set all): Auto-detect Q Genome (set all): unspecified (?)

history

search datasets

assignment_part2

5 shown, 8 deleted

50.53 MB

4: Genes with significa t adj p-value and their log2 FC

33: goseq on data 24 and data 22: Ranked category list - Wallenius method

4: Gene IDs and length

3: featureCounts on da 21 and data 24: Featu

references

<https://github.com/alexdobin/STAR>

https://hbctraining.github.io/Intro-to-rnaseq-hpc-O2/lessons/03_alignment.html

https://support.illumina.com/help/BS_App_RNASeq_Alignment OLH_1000000006112/Content/Source/Informatics/STAR_RNAseq.htm