



SINGLE-CELL COPY-NUMBER ANALYSIS

## **VIZBI 2015 WORKSHOP**

Robert Aboukhalil  
raboukha@cshl.edu

# What is Ginkgo?



# Outline

Experimental  
approaches

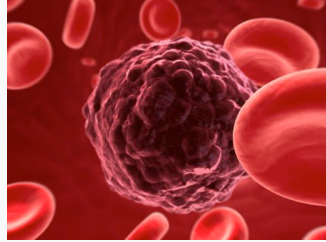
Computational  
approaches for  
CNV analysis

Demo

Hands-on



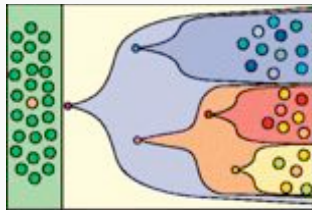
# Single-cell sequencing



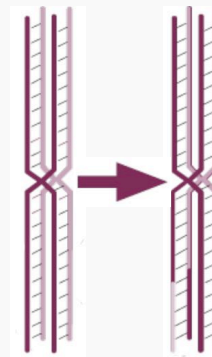
Circulating tumor cells



Neuronal mosaicism



Clonal evolution in tumors



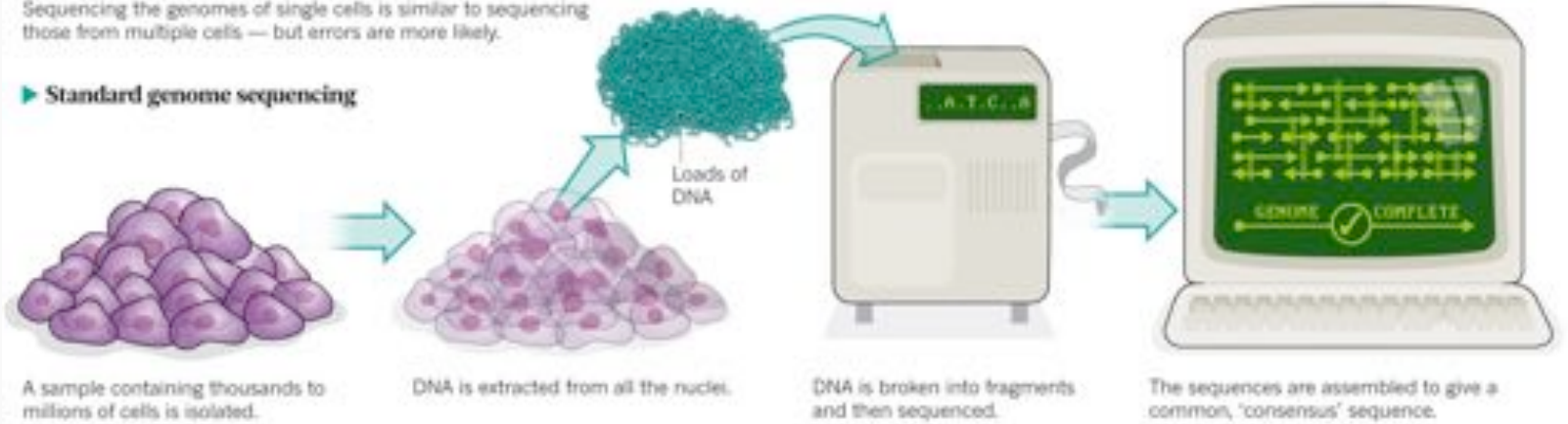
Recombination/  
crossover in germ cells

# Single-cell vs. bulk sequencing

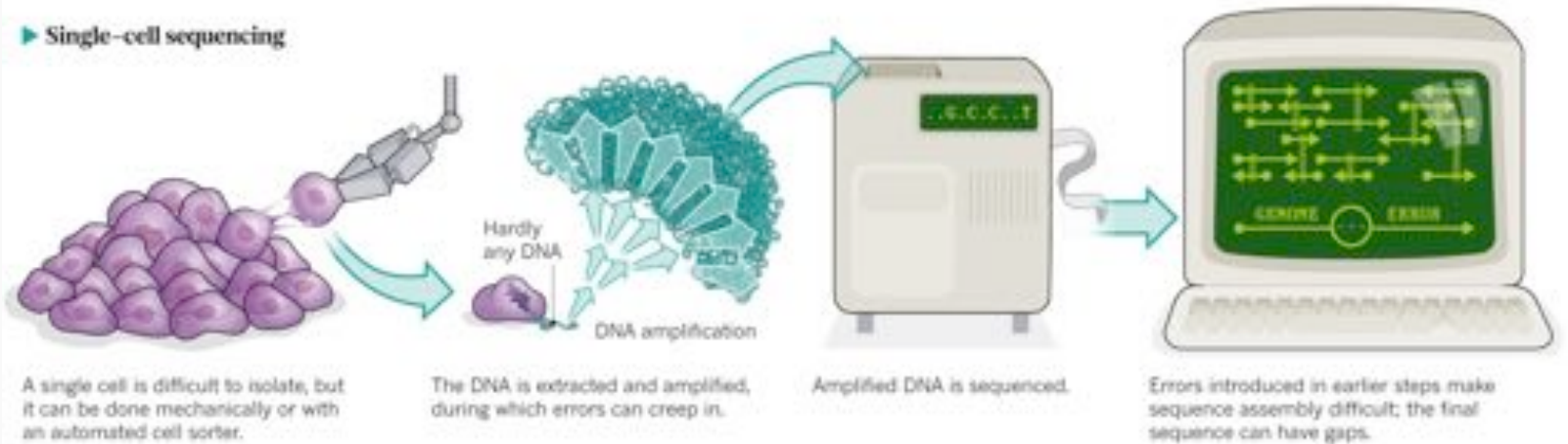
## ONE GENOME FROM MANY

Sequencing the genomes of single cells is similar to sequencing those from multiple cells — but errors are more likely.

### ► Standard genome sequencing



### ► Single-cell sequencing



# Single-cell vs. bulk sequencing

## ONE GENOME FROM MANY

Sequencing the genomes of single cells is similar to sequencing those from multiple cells — but errors are more likely.

### ► Standard genome sequencing



A sample containing thousands to millions of cells is isolated.



DNA is extracted from all the nuclei.



DNA is broken into fragments and then sequenced.



The sequences are assembled to give a common, 'consensus' sequence.

### ► Single-cell sequencing



A single cell is difficult to isolate, but it can be done mechanically or with an automated cell sorter.



The DNA is extracted and amplified, during which errors can creep in.



Amplified DNA is sequenced.

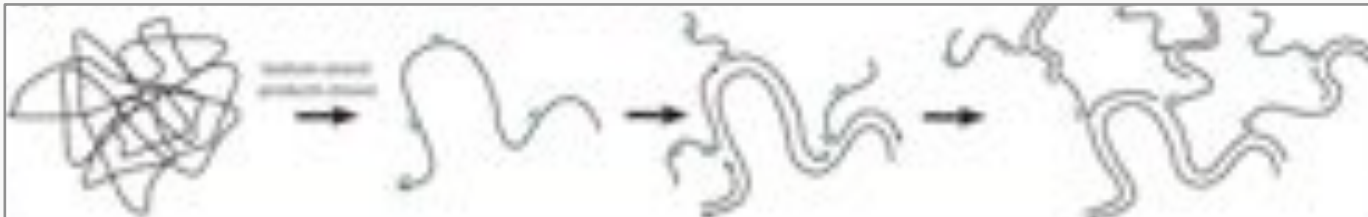


Errors introduced in earlier steps make sequence assembly difficult; the final sequence can have gaps.

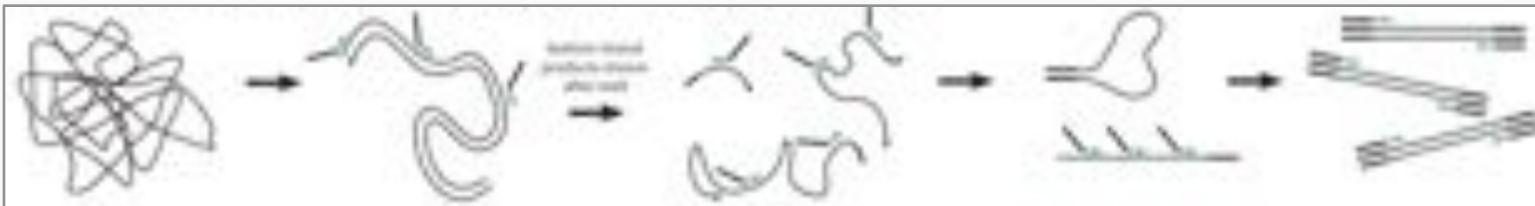
# Whole Genome Amplification (WGA) methods



DOP-PCR (Degenerate Oligonucleotide Primed PCR)



MDA (Multiple Displacement Amplification)



MALBAC (Multiple Annealing and Looping Based Amplification Cycles)



# Whole Genome Amplification (WGA) methods

The image displays three Whole Genome Amplification (WGA) methods, each represented by a schematic diagram and a label:

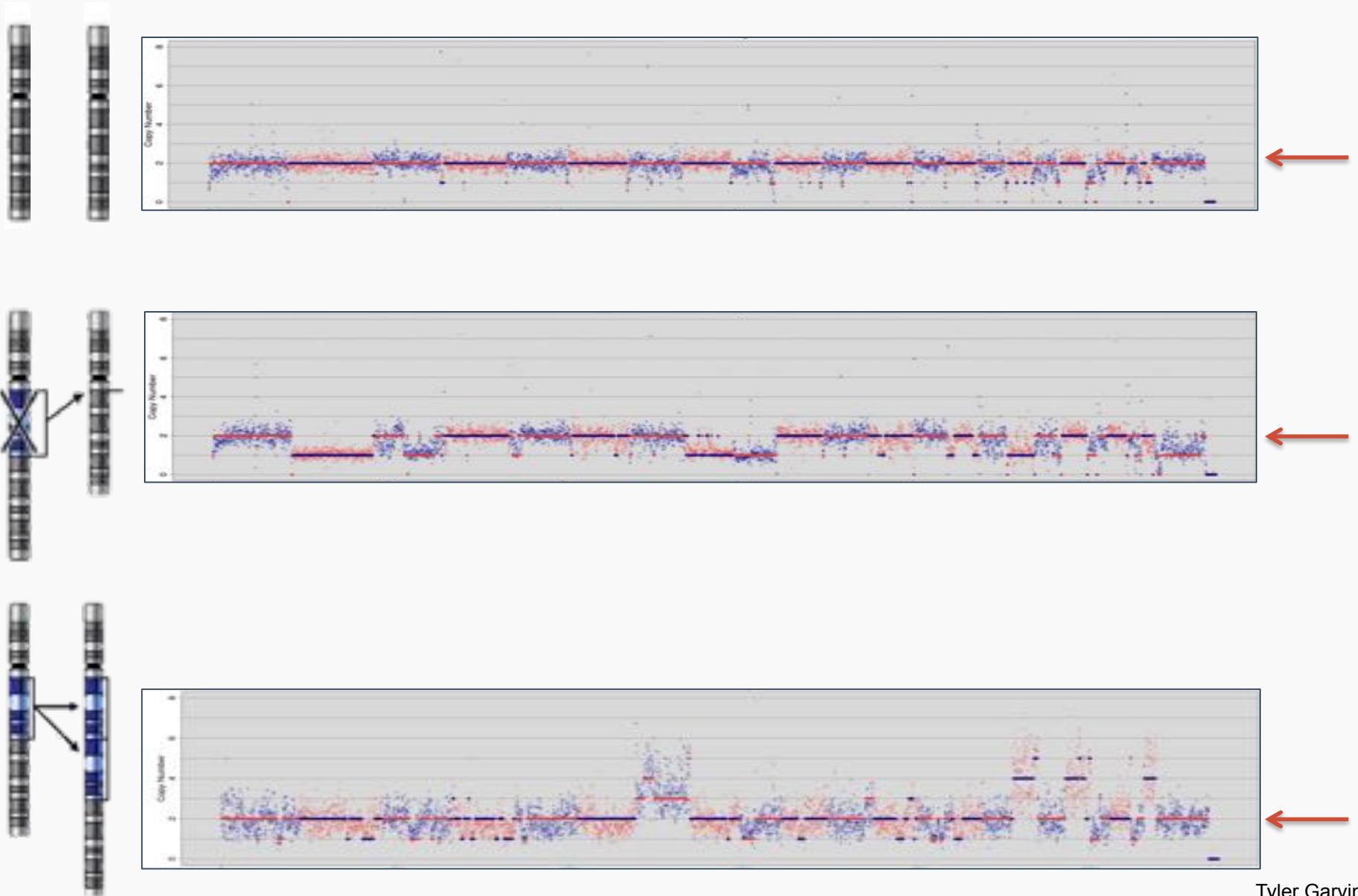
- DOP-PCR (Degenerate Oligonucleotide-primed PCR)**: Shown as a tangled network of lines.
- MDA (Multiple Displacement Amplification)**: Shown as a tangled network of lines.
- MALBAC (Multiple Annealing and Looping Based Amplification Cycles)**: Shown as a sequence of steps: a tangled network of lines, followed by a single strand, then a looped strand, and finally a double-stranded product.

The central bioRxiv preprint card contains the following information:

- Logos**: CSH Cold Spring Harbor Laboratory and bioRxiv beta.
- Text**: THE PREPRINT SERVER FOR BIOLOGY
- Link**: [bit.ly/1EkCCOO](https://bit.ly/1EkCCOO)
- Title**: **Interactive analysis and quality assessment of single-cell copy-number variations**
- Authors**: Tyler Garvin, Robert Aboukhalil, Jude Kendall, Timour Baslan, Gurinder S Atwal, James Hicks, Michael Wigler, Michael Schatz
- DOI**: <http://dx.doi.org/10.1101/011346>



# Copy-number variant analysis



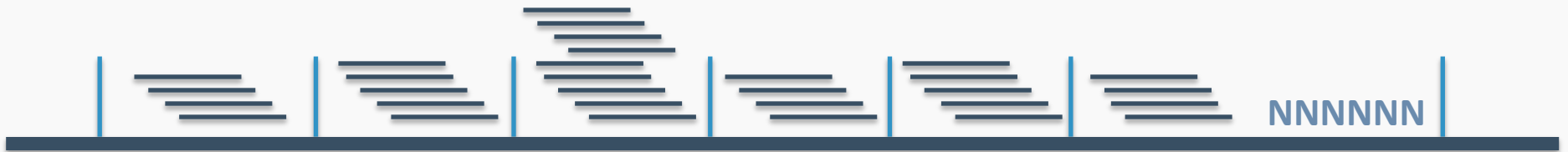
# Copy-number variant analysis



---

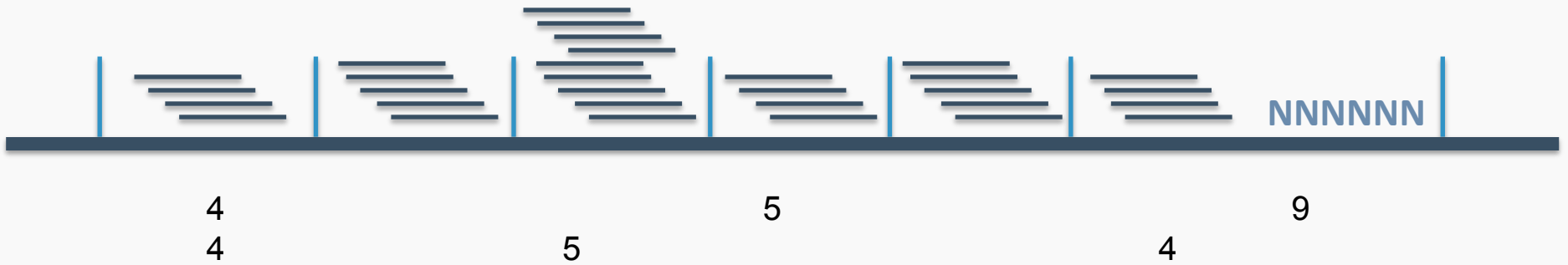
Low coverage allows us to study copy-number variants

# Copy-number variant analysis

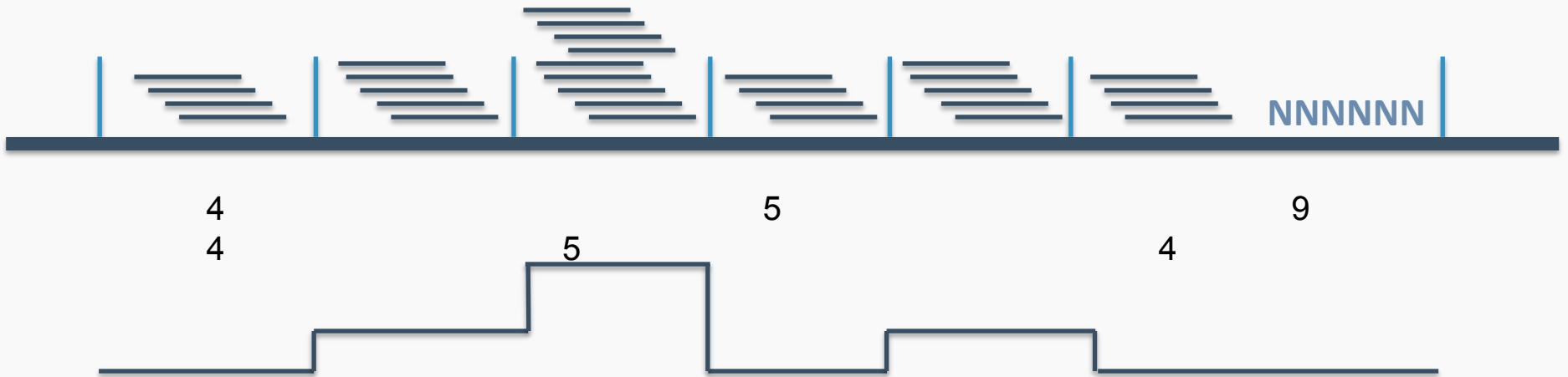


Divide genome into “bins” with ~50 – 100 reads / bin

# Copy-number variant analysis



# Copy-number variant analysis



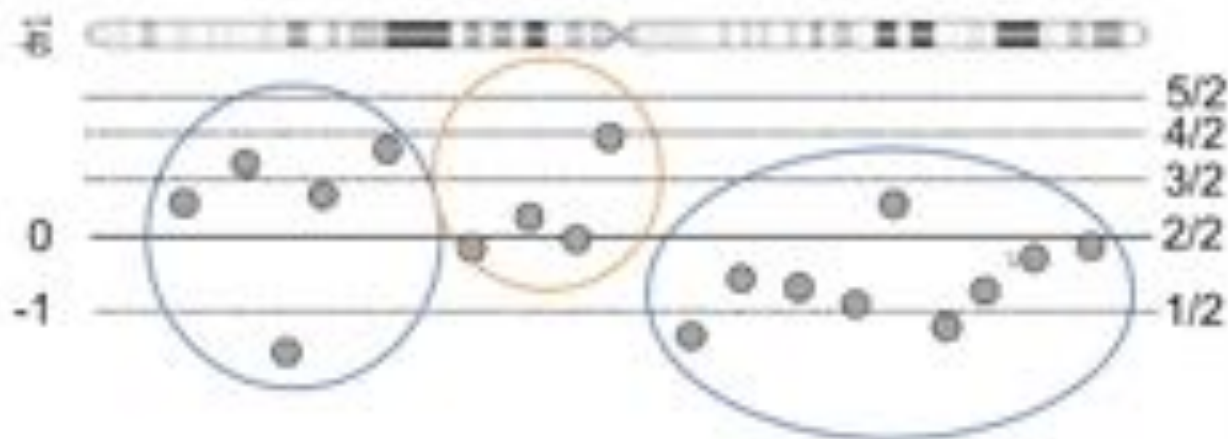
# Copy-number variant analysis



**Circular Binary Segmentation (CBS)**  
to reduce noise in data

## Circular Binary Segmentation

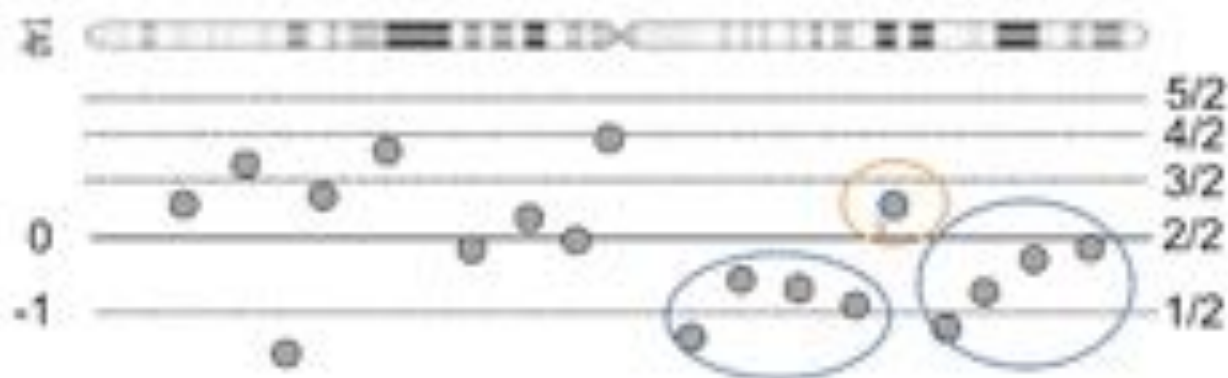
- Recursively divide up the genome until identify segments that have probe distribution different than neighbors





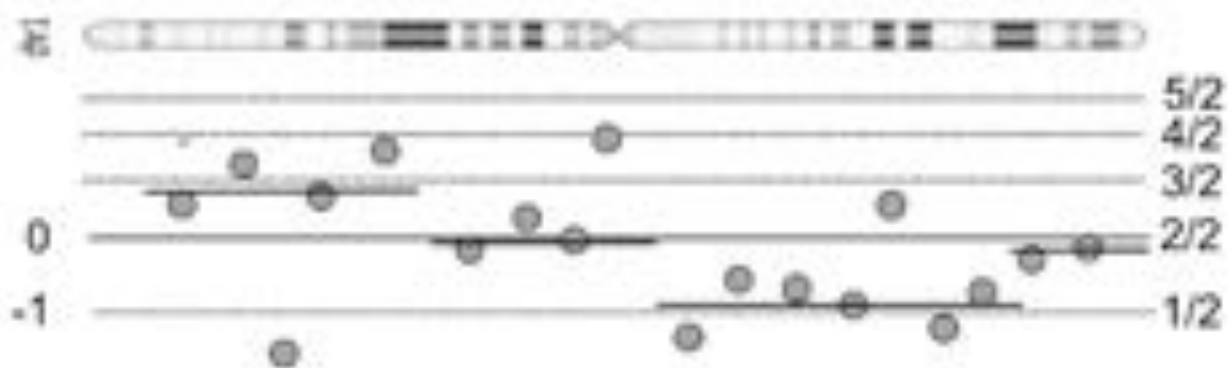
## Circular Binary Segmentation

- Recursively divide up the genome until identify segments that have probe distribution different than neighbors

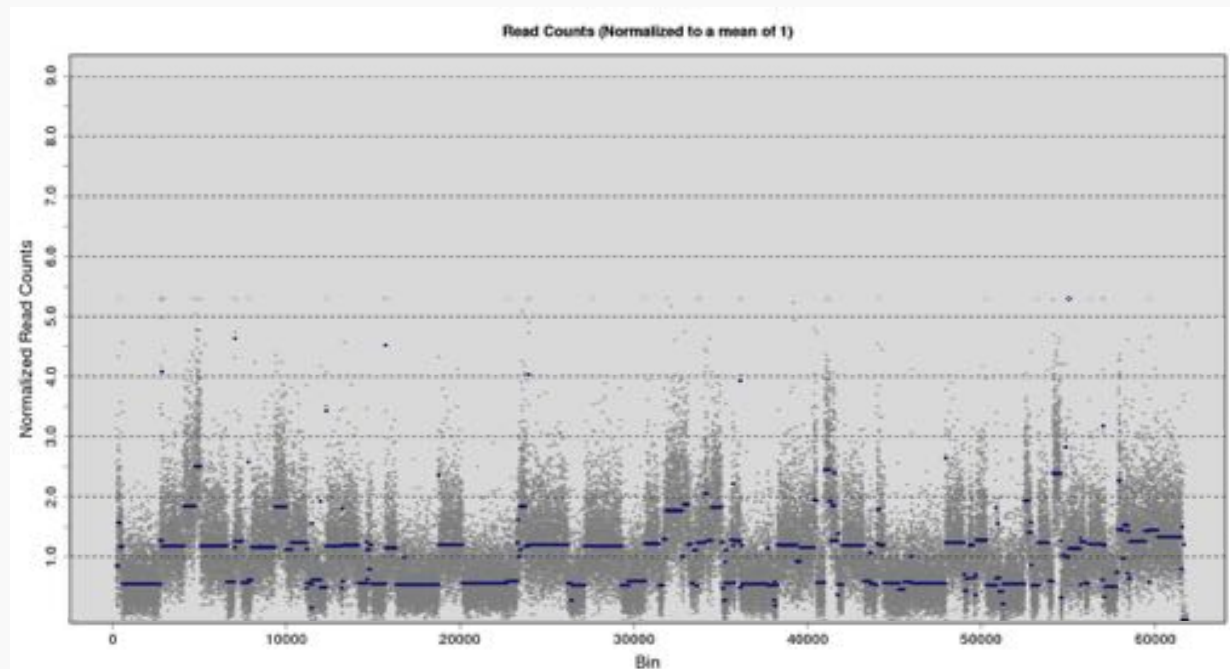


## Circular Binary Segmentation

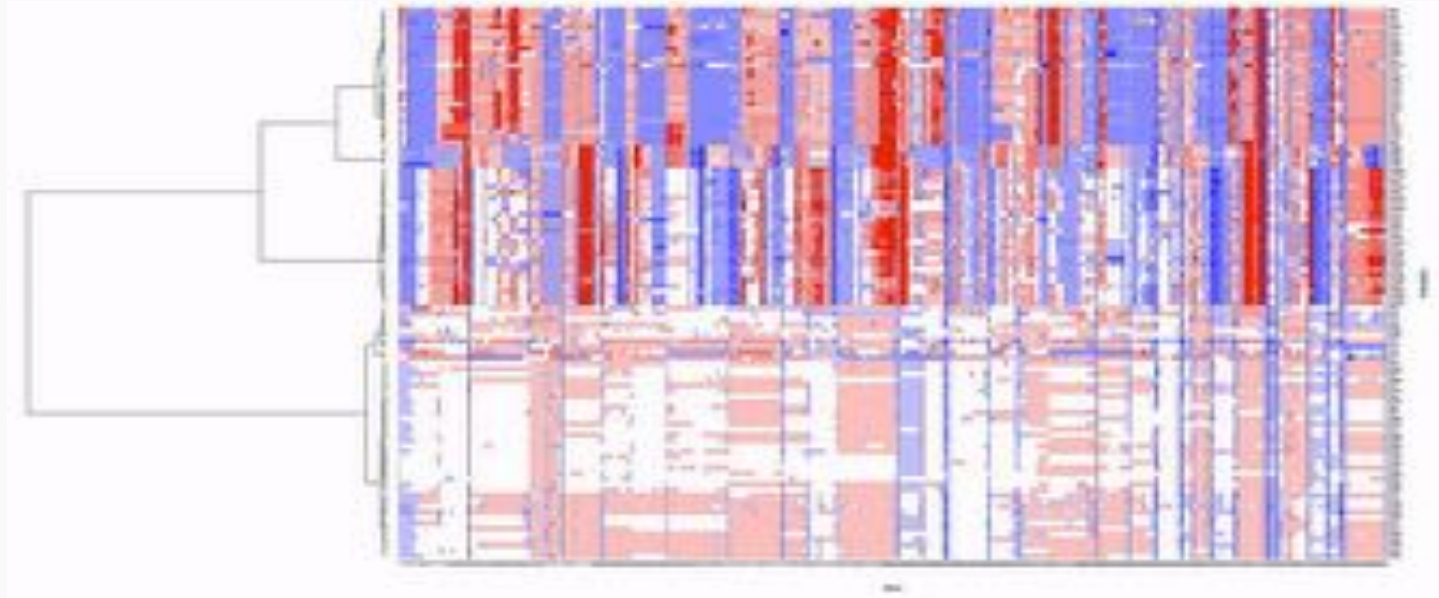
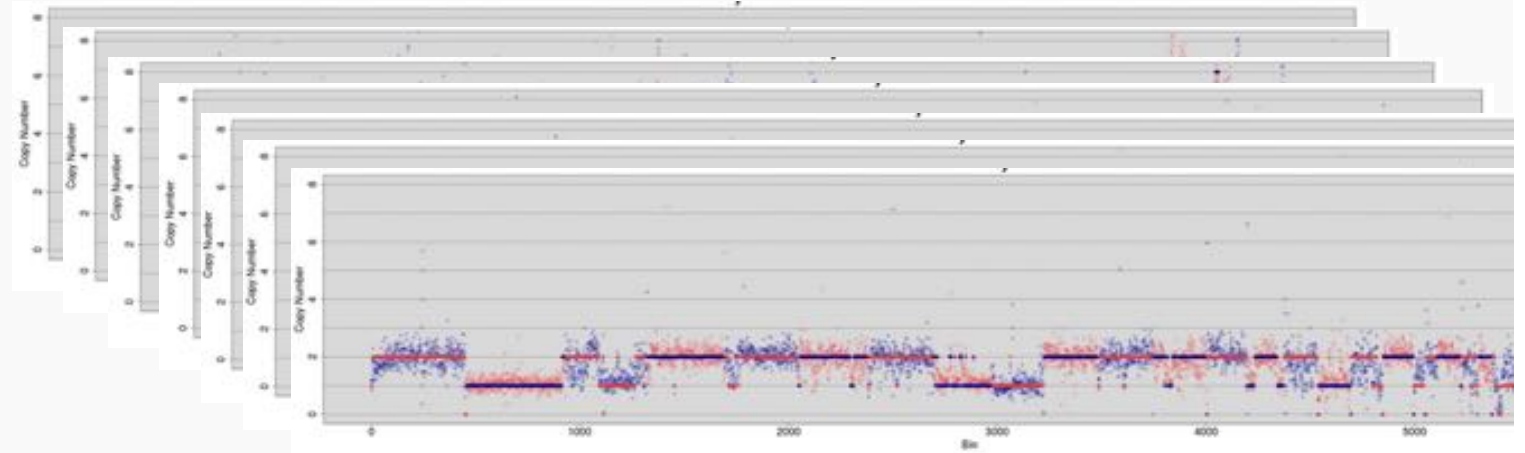
- Recursively divide up the genome until identify segments that have probe distribution different than neighbors



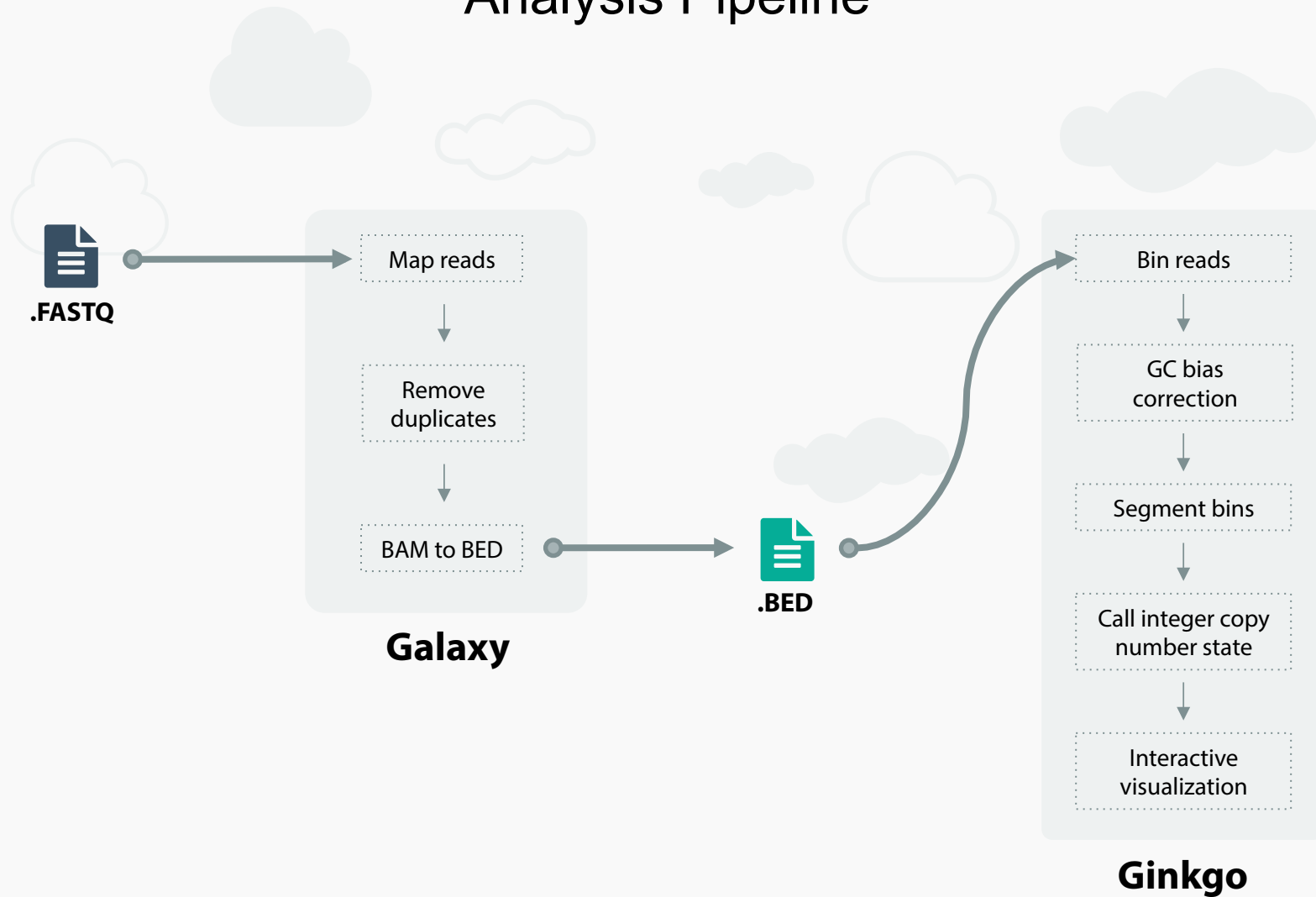
# Copy-number variant analysis



# Copy-number variant analysis



# Analysis Pipeline



# Galaxy intro



# Demo – Galaxy Pipeline



Galaxy

https://usegalaxy.org

Galaxy

Analyze Data | Workflow | Shared Data | Visualization | Cloud | Help | Home

Using 0 tools

Tools

search tools

- Get Data
- Lib-Over
- Text Manipulation
- Convert Formats
- Edit and Sort
- Join, Subtract and Cross
- NCA, QC and manipulation
- NCA Missing
- NCA, BAM Tools
- NCA Picard
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Growth, Disease Data
- Theory and Association
- sniffIT
- BETools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Motif analysis
- Motif Tools
- NLS TOOLBOX BETA
- NCA, SAM Tools

Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start [here](#) or consult our [help resources](#).

**Galaxy 101**  
Start small  
The very first tutorial you need

Tweets

PyPI Recent Updates [@py\\_updates](#) in [planemo 0.6.0: Command-line utilities to assist in building tools for the Galaxy project](#) [Pflp-/galaxytr](#) ... [8tLj7Y00CkYt](#)  
13 Retweeted by Galaxy Project  
Retweet

Galaxy Project [@galaxyproject](#) in [#usegalaxy Histories now support advanced searching and offer multiple other improvements.](#) [wiki.galaxyproject.org/histories](#) [go.twitter.com/1JRN7y5G](#)  
@ Show Photo

Yannick Worm [@yannick...](#) 12 Mar  
Fun genome Bioinformatics #job in  
Tweet to @galaxyproject

History

search histories

Unnamed history

0 tools

This history is empty. You can load your own data or get data from an external source.

PENNSYLVANIA STATE UNIVERSITY

JOHNS HOPKINS UNIVERSITY

TAGC

iPlant Collaborative

The Galaxy Team is a joint effort of the Center for Computational Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.

Galaxy

https://usegalaxy.org

Galaxy

Analyze Data | Workflow | Shared Data | Visualization | Cloud | Help | User

Tools

search tools

- Get Data
- Lib-Over
- Text Manipulation
- Convert Formats
- Edit and Sort
- Join, Subtract and Cross
- NCL QC and manipulation
- NCL Missing
- NCL BAM Tools
- NCL Picard
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Growth/Dispersal Data
- Theory/Association
- sniffIT
- BIGTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Motif/sequence analysis
- Motif Tools
- NGL TOOLBOX BETA
- NGL SAM Tools

Galaxy is an open source, web-based platform for data intensive bioinformatics research. If you are new to Galaxy start [here](#) or consult our [help resources](#).

Log in  
Register

History

search histories

Unnamed history

0 items

This history is empty. You can load your own data or get data from an external source.

Tweets

PyPI Recent Updates [@pyupio](#) in [@planetbeing](#)  
planetbeing 0.6.0: Command-line utilities to assist in building tools for the Galaxy project [https://galaxytoolkit.org/](#) [https://github.com/planetbeing/planetbeing](#)  
13 Retweeted by Galaxy Project  
Retweet

Galaxy Project [@galaxyproject](#)  
#usegalaxy Histories now support advanced searching and offer multiple other improvements. [wiki.galaxyproject.org/histories](#) [gal.twitter.com/1P8N7d5G](#)  
@ Show Photo

Yannick Worm [@yannick\\_worm](#) 12 Mar  
Fun genome Bioinformatics Job in  
Tweet to @galaxyproject

GXYcast 1  
2015

PENNSYLVANIA STATE UNIVERSITY

JOHNS HOPKINS UNIVERSITY

TAGC

iPlant Collaborative

The Galaxy Team is a joint effort of the Center for Computational Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.

https://usegalaxy.org/learn/create

Galaxy [Analyze Data](#) [Workflow](#) [Share Data](#) [Provisioning](#) [Cloud](#) [Help](#) [User](#) Using 0 bytes

Tools

- Get Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Edit and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Picard
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Growth/Disorder Data
- Theory/Association
- snRNAs
- BSITools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Metagenomic analyses
- Motif Tools
- NGS TOOLBOX BETA
- NGS SAM Tools

### Create account

**Email address**

**Password**

**Confirm password**

**Public name**

Your public name is an identifier that will be used to generate addresses for information you share publicly. Public names must be at least four characters in length and contain only lower-case letters, numbers, and the '-' character.

**Subscribe to mailing list**  
 [See all Galaxy project mailing lists.](#)

Please register only one account. The [usegalaxy.org](#) service is provided free of charge and has limited computational and data storage resources. Registration and usage of multiple accounts is tracked and such accounts are subject to termination and data deletion.

[Terms and Conditions for use of this service](#)

History

Unnamed history  
0 bytes

**i** This history is empty. You can load your own data or get data from an external source.

<https://usegalaxy.org>

Galaxy <https://usegalaxy.org> Using 0 tools

Analyze Data [Workflow](#) [Share Data](#) [Visualization](#) [Cloud](#) [Help](#) [Home](#)

Tools search tools

- Get Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Edit and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Picard
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Growth/Disorder Data
- Theory/Association
- sniffIT
- BSITools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Misassembly analyses
- Worf Tools
- NGS TOOLBOX BETA
- NGS SAM Tools

**Login**

Email address:

Password:

[Forgot password? Reset here](#)

**OpenID Login**

OpenID URL:

Or, authenticate with your [Google](#) account.

[Terms and Conditions for use of this service](#)

**History** search history

Unnamed history

0 tools

**i** This history is empty. You can load your own data or get data from an external source.



Galaxy

https://usegalaxy.org

Galaxy

Analyze Data | Workflow | Share Data | Visualization | Cloud | Help | Users


Tools

Search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Edit and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Visual
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phylogenetic Association
- snEff
- BEITools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Metatools
- NGS TOOLBOX BETA

Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start [here](#) or consult our [help resources](#).

Want help?  
Get answers.

 **BioStars**  
GALAXY EXPLAINED

Tweets

PyPI Recent Updates [@py\\_updates](#) in [planemo 0.6.0: Command-line utilities to assist in building tools for the Galaxy project](#) [PflFu/galaxydr](#) · [bit.ly/100CkxY](#)  
13 Retweeted by Galaxy Project  
Retweet

Galaxy Project [@galaxyproject](#) in [#usegalaxy Histories now support advanced searching and offer multiple other improvements. \[wiki.galaxyproject.org/histories\]\(#\) \[go.twitter.com/1JRN7y5G\]\(#\)](#)  
@ Show Photo

Yannick Worm [@yannick\\_worm](#) 12 Mar  
Fun genome Bioinformatics Job in  
Tweet to @galaxyproject

History

Search histories

Unnamed history

0 bytes

Annotations

Click here to edit annotations

This history is empty. You can [load your own data](#) or [get data from an external source](#).

PENNSYLVANIA STATE UNIVERSITY

JOHNS HOPKINS UNIVERSITY

TAGC

iPlant Collaborative

The Galaxy Team is a part of the Center for Computational Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.



Galaxy

https://usegalaxy.org

Galaxy

Analyze Data | Workflow | Share Data | Publications | Cloud | Help | Users

Tools

Search

Get Data

BLAST DNA STR

NGS Mapping

LRGE Map short reads against reference sequence

Workflows

• All workflows

Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start [here](#) or consult our [help resources](#).

**GXYcast 1**  
2015

Tweets

PyPi Recent Updates [@py\\_updates](#) in [planemo 0.6.0: Command-line utilities to assist in building tools for the Galaxy project](#) [@Pfls/galaxydr](#) [@SLYrY00CkYr](#) 11 Retweeted by Galaxy Project

Galaxy Project [@galaxyproject](#) in [usegalaxy Histories now support advanced searching and offer multiple other improvements.](#) [wiki.galaxyproject.org/histories](#) [go.twitter.com/1JRN7y5G](#) 11 Show Photo

Yannick Worm [@yannick...](#) 12 Mar [Fun genome Bioinformatics Job in](#)

Tweet to @galaxyproject

History

Search Histories

Unnamed history

0 jobs

Annotations

Click here to edit annotations

This history is empty. You can [load your own data](#) or [get data from an external source](#)

PENNSYLVANIA STATE UNIVERSITY

JOHNS HOPKINS UNIVERSITY

TAGC

iPlant Collaborative

The Galaxy Team is a part of the Center for Computational Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.



Galaxy <https://usegalaxy.org>

Galaxy [Analyze Data](#) [Workflow](#) [Share Data](#) [Visualization](#) [Cloud](#) [Help](#) [Login](#)

Tools [EMBL-EBI](#) [Services](#) [Research](#) [Training](#) [About us](#) [History](#) [Logout](#)

ENA  
European Nucleotide Archive

Home **Search & Browse** [Submit & Update](#) [About ENA](#) [Support](#)

Search:  [Search](#) [Advanced](#) [Sequence](#)

Database (440000) entries

Please subscribe to our announce mailing list from [ena@ebi.ac.uk](mailto:ena@ebi.ac.uk) or [ena@ebi.ac.uk](mailto:ena@ebi.ac.uk), to receive alerts about ENA services.

This history is empty. You can load your own data or get data from an external source.

EMBL-EBI [Home](#) [About us](#) [Contact us](#) [Internet](#)

Services	Research	Training	Industry	About us
By topic By name (A-Z) Help & Support	Overview Publications Research groups Postdocs & PhDs	Overview Train at EBI Train outside EBI Train online Contact organisers	Overview Members Area Workshops SME Forum Contact Industry programme	Overview Leadership Funding Background Collaboration Jobs People & groups News Events Visit us Contact us

EMBL-EBI, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SB, UK +44 (0)1223 49 44 44  
 Copyright © EMBL-EBI 2013 | EBI is an outstation of the European Molecular Biology Laboratory | [Privacy](#) | [Cookies](#) | [Terms of use](#)

[https://usegalaxy.org/feed\\_rss.html](https://usegalaxy.org/feed_rss.html) [source](https://usegalaxy.org/feed_rss.html) [redirect](https://usegalaxy.org/feed_rss.html) [feed](https://usegalaxy.org/feed_rss.html) [rss](https://usegalaxy.org/feed_rss.html)

SRR054622

The image shows a screenshot of a web browser displaying the Galaxy interface. The browser's address bar shows the URL <https://usegalaxy.org>. The main content area features the ENA (European Nucleotide Archive) logo and a search bar containing the text "SRR054622". Below the search bar, there are navigation links: "Home", "Search & Browse", "Submit & Update", "About ENA", and "Support". A blue banner below the navigation links contains a message: "Please subscribe to our announce mailing list from [enahelp@ebi.ac.uk](mailto:enahelp@ebi.ac.uk) to receive alerts about ENA services." The footer of the page includes the EMBL-EBI logo, a navigation menu with categories like "Services", "Research", "Training", "Industry", and "About us", and contact information: "EMBL-EBI, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK +44 (0)1223 49 44 44". Copyright information for 2015 is also present.

Galaxy <https://usegalaxy.org> Plug-in blocked Logout

Galaxy Analysis Data Workflow Share Data Visualization Cloud Help Log In

EMBL-EBI Services Research Training About us

Database (24,000,000) entries Advanced Search

Home **Search & Browse** [Submit & Update](#) [About ENA](#) [Support](#)

Please subscribe to our announce mailing list [ena-announce@ebi.ac.uk](mailto:ena-announce@ebi.ac.uk) to receive alerts about ENA services.

### Read: SRR054622

ILLUMINA  
European Nucleotide Archive

ILLUMINA  
European Nucleotide Archive

SCS Mapping  
LINK Map short reads against reference sequence

Workflows  
• All workflows

History  
Search history  
Unnamed history  
0 bytes  
Annotations  
Click here to edit annotations

**This history is empty. You can load your own data or get data from an external source.**

Send feedback to [Download: XML](#)

Submitting Centre	Run Date	Platform	Model	Read Count	Base Count
Cold Spring Harbor Laboratory		ILLUMINA	Illumina Genome Analyzer	2,799,858	131,913,184

Library Layout	Library Strategy	Library Source	Library Selection	Library Name
SCBCL	WGS	GENOMIC	RANDOM	SC.WGS.LIB

Broker Name  
NCBI

Navigation **Read Files**

This table contains the files for run SRR054622  
[Download files](#)

View: **TEXT** Download: **TEXT**

Galaxy <http://usegalaxy.org>

Galaxy Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users+ Using 0%

Tools

SMA

Get Data  
BLAST DNA SMA

NGS Mapping  
LRR2 Map short reads against reference sequence

Workflows  
• All workflows

### Read: SRR054622

Human Genome Analyzer sequencing; Single Cells from Breast Tumor T10

View: XML Send Feedback Download: XML

Submitting Centre	Run Date	Platform	Model	Read Count	Base Count
Cold Spring Harbor Laboratory		ILLUMINA	Human Genome Analyzer	2,789,858	133,913,184

Library Layout	Library Strategy	Library Source	Library Selection	Library Name
SINGLE	WGS	GENOMIC	RANDOM	SC-WGA-LIB

Broker Name  
NCBI

Navigation

Read Files

This table contains the files for run SRR054622

Download files View: TEXT Download: TEXT

Select columns

Showing results 1 - 1 of 1 results

Study accession	Secondary study accession	Sample accession	Secondary sample accession	Experiment accession	Run accession	Tax ID	Scientific name	Instrument model	Library layout	Fastq files (R1)	Fastq files (R2)
SRR002535	SRR002535	SAMN0004736	SRR074087	SRR021400	SRR054622	9606	Homo sapiens	Human Genome Analyzer	SINGLE	File 1	File 1

History

Search datasets

Unnamed history

0 bytes

Annotations  
[Click here to edit annotations](#)

This history is empty. You can load your own data or get data from an external source



Galaxy

https://usegalaxy.org

Galaxy

Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users

Tools

Search: SRA

Get Data

- 1. SRA DNA SRA

NCBI Mapping

1. SRA Map short reads against reference sequence

Workflows

- All workflows

History

Search history

Unnamed history

1 shown

0 bytes

Annotation

Click here to edit annotation

1. 1. SRA, SRR054622

2. File: ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR054/SRR054622/SRR054622.fastq.gz

A job has been successfully added to the queue - resulting in the following dataset:

- 1. 1. SRA, SRR054622 File: ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR054/SRR054622/SRR054622.fastq.gz

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

You are now being redirected back to Galaxy.

Galaxy

https://usegalaxy.org

Galaxy

Analyze Data | Workflow | Share Data | Visualization | Cloud | Help | Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Edit and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Filter
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phylogenetic Association
- snEff
- BEITools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Metatools
- NGS TOOLBOX BETA

Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start [here](#) or consult our [help resources](#).

Try Galaxy on the Cloud

Now you can host a personal Galaxy within the cloud

PyPI Recent Updates [@py\\_updates](#) in [planemo 0.6.0: Command-line utilities to assist in building tools for the Galaxy project](#) #PyPI/galaxytr... [bit.ly/100CkYI](#)  
13 Retweeted by Galaxy Project  
Retweet

Galaxy Project [@galaxyproject](#) in [#usegalaxy Histories now support advanced searching and offer multiple other improvements. \[wiki.galaxyproject.org/histories\]\(#\) \[go.twitter.com/1JRN7d5Q\]\(#\)](#)  
@ Show Photo

Yannick Worm [@yannick...](#) 12 Mar  
Fun genome Bioinformatics #job in  
Tweet to @galaxyproject

History

search histories

Unnamed history  
1 shown

0 items

Annotation  
Click here to edit annotation

1. ERI.MA.SRR05162  
2 File: ftp://ftp.ensembl.org/pub/FASTA/chr1/SRR051622/SRR051622.fasta.gz

PENNSYLVANIA STATE UNIVERSITY

JOHNS HOPKINS UNIVERSITY

TAGC

iPlant Collaborative

The Galaxy Team is a joint effort of the Center for Computational Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.



Galaxy

https://usegalaxy.org

Galaxy

Analyze Data | Workflow | Share Data | Visualization | Cloud | Help | Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Edit and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Visual
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phylogenetic Association
- snEff
- BEETools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Metagenomic analysis
- Metatools
- NGS TOOLBOX BETA

Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start [here](#) or consult our [help resources](#).

**Galaxy 101**  
Start small  
The very first tutorial you need

Tweets

PyPI Recent Updates @py\_updates in  
planemo 0.6.0: Command-line utilities to assist in building tools for the Galaxy project #fpga/galaxydr ... bit.ly/100CkYI  
13 Retweeted by Galaxy Project  
Retweet

Galaxy Project @galaxyproject in  
#usegalaxy Histories now support advanced searching and offer multiple other improvements. [wiki.galaxyproject.org/Histories](http://wiki.galaxyproject.org/Histories) [go.twitter.com/1JRN7NSQ](https://twitter.com/1JRN7NSQ)  
@ Show Photo

Yannick Worms @yannick... 12 Mar  
Fun genome Bioinformatics #job in  
Tweet to @galaxyproject

History

search histories

Unnamed history  
1 shown

0 items

Annotation  
Click here to edit annotation

1. EBLMA, SRR05162  
2. File: file:///data/eb/eb/eb/EBLMA/SRR05162/SRR05162.fasta.gz

PENNSYLVANIA STATE UNIVERSITY

JOHNS HOPKINS UNIVERSITY

TAGC

iPlant Collaborative

The Galaxy Team is a joint effort of the Center for Computational Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.

Galaxy

https://usegalaxy.org

Galaxy

Analyze Data | Workflow | Share Data | Visualization | Cloud | Help | Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Edit and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Pileup
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phylogenetic Association
- sniff
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Metatools
- NGS TOOLBOX BETA

Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start [here](#) or consult our [help resources](#).



**Tweets**

**PyPI Recent Updates** @py\_updates in planemo 0.6.0: Command-line utilities to assist in building tools for the Galaxy project #PyPI/galaxydr - 84.3k/100kxRT 13 Retweeted by Galaxy Project

**Galaxy Project** @galaxyproject in #usegalaxy Histories now support advanced searching and offer multiple other improvements. [wiki.galaxyproject.org/Histories](http://wiki.galaxyproject.org/Histories) [pdx.twitter.com/UPRN765Q](https://twitter.com/UPRN765Q) 82 Show Photo

**Yannick Wurm** @yannick\_wurm 12 Mar Fun genome Bioinformatics #job in [Tweet to @galaxyproject](#)

**History**

search histories

Unnamed history  
1 shown  
194.8 MB  
Annotation  
Click here to edit annotation

**L. BR. SA. SR054622.1** [+](#) [-](#) [x](#)  
IN: PHL: PHL334454646.uk  
TWO: Tasta.SR054622.1M  
R054622.fastq.gz

**PENNSYLVANIA STATE UNIVERSITY**

**JOHNS HOPKINS UNIVERSITY**

**TAGC**

**iPlant Collaborative**

The Galaxy Team is a part of the Center for Computational Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.



Galaxy

https://usegalaxy.org

Galaxy

Analyze Data | Workflow | Share Data | Visualization | Cloud | Help | Users | Logging Out

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Edit and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Visual
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phylogeny/Association
- snEff
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Metagenomic analysis
- Metatools
- NGS TOOLBOX BETA

Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start [here](#) or consult our [help resources](#).

**Running Your Own**  
**Understanding how Galaxy works**  
 An in-depth tutorial

PyPI Recent Updates [@py\\_updates](#) in [planemo 0.6.0: Command-line utilities to assist in building tools for the Galaxy project](#) #PyPI #Galaxy #Bioinformatics  
 11 Retweeted by Galaxy Project  
 Retweet

Galaxy Project [@galaxyproject](#) in [#usegalaxy Histories now support advanced searching and offer multiple other improvements. \[wiki.galaxyproject.org/Histories\]\(http://wiki.galaxyproject.org/Histories\) \[p.s. twitter.com/UPRNYSQ\]\(https://twitter.com/UPRNYSQ\)](#)  
 0 Retweets

Yannick Worm [@yannick\\_worm](#) 12 Mar  
 Fun genome Bioinformatics #ob in [@galaxyproject](#)  
 Tweet to @galaxyproject

History

search histories

Unnamed history  
 1 shown  
 194.8 MB  
 Annotation  
 Click here to edit annotation

L. BR. SA. MR014622.1  
 IN. PA. / TA. ST. A. BL. AC. U. H.  
 Y. W. L. / X. AA. S. AM. OS. S. S.  
 R. OS. 4622. FASTA.GZ

PENNSYLVANIA STATE UNIVERSITY

JOHNS HOPKINS UNIVERSITY

TAGC

iPlant Collaborative

The Galaxy Team is a part of the Center for Computational Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.

https://usegalaxy.org/instancs/inst0001/0000017046/UPRNYSQ/0001/0001

Galaxy <https://usegalaxy.org>

Galaxy Analyze Data Workflow Shared Data+ Visualization+ Cloud+ Help+ Users+ Logout

Tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Picard
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Stats
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BEITools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Metagenomic analysis
- Motif Tools
- NGS TOOLBOX BETA

Attributes Convert Format Database Permissions

### Edit Attributes

Name: EB: SRA: SRX014622 File: ftp://ftp.sra.ftp.ncbi.nlm.nih.gov/trace/sra/SRX014622/

Info:

Annotation / Notes:

Database / Build:

Additional Species Are Below

Save

Auto-detect

This will inspect the dataset and attempt to correct the above column values if they are not accurate.

History

Unnamed history  
1 shown  
194.8 MB  
Annotation  
Click here to edit annotation

- 1 EB: SRA: SRX014622  
File: ftp://ftp.sra.ftp.ncbi.nlm.nih.gov/trace/sra/SRX014622/

The screenshot displays the Galaxy web application interface. At the top, a browser window shows the URL <https://usegalaxy.org>. The main navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Inspections', 'Cloud', 'Help', and 'Users'. Below this is a 'Tools' sidebar on the left with a search bar and a list of categories including 'Get Data', 'Text Manipulation', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'NGS, QC and manipulation', 'Statistics', 'Genome Diversity', and 'NCBI TOOLS BETA'. The central workspace is currently showing the 'Change data type' tool, which has tabs for 'Attributes', 'Convert Format', 'Generate', and 'Distributions'. The 'Change data type' tool panel includes a 'New Type' dropdown menu set to 'String' and a 'Save' button. A warning message states: '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.' The right sidebar contains a 'History' panel with a search bar, showing one item named 'Unnamed history' with a size of '394.8 MB'. Below the history list, a dataset is visible with the following details: '1 EB 356,388024622', 'File Path: /Pa...46146', and 'URL: http://hg19.soe.ucsc.edu/cgi-bin/hgUiTable?table=chr11.geneTrack'. The bottom of the page shows the URL <https://usegalaxy.org/datasets/143480b1-3443-433c-8006-1704b448f28682/view/143480b1-3443-433c-8006-1704b448f28682>.

The screenshot shows the Galaxy web interface at <https://usegalaxy.org>. The main content area displays the 'Change data type' tool configuration. The 'New Type' dropdown menu is open, showing a list of data types with 'Samvanger' selected. A warning message states: 'Dataset but not modify its contents. Use this if Galaxy has incorrectly guessed the type of your dataset.' The left sidebar contains a 'Tools' menu with categories like 'Get Data', 'Text Manipulation', and 'Genomics'. The right sidebar shows the 'History' panel with a search bar and a list of datasets, including one with a green highlight.

Galaxy

Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users

Tools

search tools

Get Data  
Send Data  
Lift-Over  
Text Manipulation  
Convert Formats  
Filter and Sort  
Join, Subtract and Group  
NGS QC and manipulation  
NGS Mapping  
NGS BAM Tools  
NGS Read  
Extract Features  
Fetch Sequences  
Fetch Alignments  
Get Genomic Scores  
Operate on Genomic Intervals  
Statistics  
Graph/Display Data  
Phenotype Association  
sniff  
BEDTools  
Genome Diversity  
EMBOSS  
Regional Variation  
FASTA manipulation  
Evolution  
Multiple Alignments  
Microarray analysis  
Motif Tools  
NGS TOOLBOX BETA

Attributes Convert Format Database Reanalysis

Change data type

New Type

Text

Samvanger

Samvolumes

Samtools

Samtools

Text

Dataset but not modify its contents. Use this if Galaxy has incorrectly guessed the type of your dataset.

History

search datasets

Unnamed history  
1 shown

394.8 MB

Annotation  
Click here to full annotation

1. EN 356, 3800-622, File: /p1/7013140140, uk:vol1:fasta:380054,3800-622, 3800-622.fasta

Galaxy

https://usegalaxy.org

Galaxy

Analyze Data | Workflow | Share Data | Visualization | Cloud | Help | Users | Logging Out

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Edit and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Read
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Stats
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BSITools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Metagenomic analysis
- Motif Tools
- NGS TOOLBOX BETA

Attributes | Convert Format | **Dataset** | Permissions

Change data type

New Type:

Namesanger

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

Save

History

search datasets

Unnamed history

1 shown

194.8 MB

Annotation

Click here to edit annotation

1. EN 356, 588054622, File: /gal:/Toolshed/UCSC/ncbi/vstl/ncbi/588054,588054622/588054622.fa.gz

Galaxy

https://usegalaxy.org

Galaxy

Analyze Data | Workflow | Share Data | Visualization | Cloud | Help | Users | Logging Out

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- NCBI QC and manipulation
- NCBI Mapping
- NCBI BLAST Tools
- NCBI BLAST
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Stats
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BLAST Tools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Motif Tools
- NCBI TOOLBOX BETA

Changed the type of dataset EB\_SRA\_SRR054622 File: ftp://ftp.sra. nih.gov/pub/trace/SRR054622/SRR054622.fastq.gz to fastqsanger

Attributes | Convert Format | Database | Refinements

Edit Attributes

Name:  
EB\_SRA\_SRR054622 File: ftp://ftp.sra.

Info:

Annotation / Notes:

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

Database / Build:  
----- Additional Species Are Below -----

Save

Auto-annotate

This will inspect the dataset and attempt to correct the above column values, if they are not accurate.

History

search histories

Unnamed history  
1 shown

194.8 MB

Annotation:  
Click here to edit annotation

EB\_SRA\_SRR054622 File: ftp://ftp.sra.nih.gov/pub/trace/SRR054622/SRR054622.fastq.gz

Galaxy

https://usegalaxy.org

Galaxy

Analyze Data | Workflow | Share Data | Visualization | Cloud | Help | Users | Logging Out

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Edit and Sort
- Join, Subtract and Group
- NCBI QC and manipulation
- NCBI Mapping
- NCBI BLAST Tools
- NCBI BLAST
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Met. Tools
- NCBI TOOLBOX BETA

Changed the type of dataset EN SRA, SRX014622 File: ftp://ftp.sra. nih.gov/pub/ftp.sra/orig1/hsa/srpx/SRX014622/SRX014622.fastq.gz to fastq-compr

Attributes | Convert Format | Delete | Refresh

### Edit Attributes

**Name:**  
EN SRA, SRX014622 File: ftp://ftp.sra.

**Info:**

**Annotation / Notes:**

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

**Database / Build:**  
----- Additional Species Are Below -----

Save

Auto-convert

This will inspect the dataset and attempt to correct the above column values, if they are not accurate.

History

search histories

Unnamed history  
1 shown

194.8 MB

Annotation:  
Click here to edit annotation

1: EN SRA, SRX014622 (1) + - x  
In: ftp://ftp.sra.nih.gov/pub/ftp.sra/orig1/hsa/SRX014622/SRX014622.fastq.gz



Galaxy

https://usegalaxy.org

Galaxy

Analyze Data | Workflow | Shared Data | Visualizations | Cloud | Help | Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Edit and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Filter
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Tracks
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BEETools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Metagenomic analysis
- Met Tools
- NGS TOOLBOX BETA

Galaxy is an open source, web-research. If you are new to Galaxy, please visit our [data intensive biomedical](#) and [help resources](#).

**kmer-SVM**  
Galaxy server for identifying predictive regulatory sequence features in NGS data

**Galaxy / Beer Lab**

Yuan Le Bras (@yuanlebr) 29m  
With @cmorjeaud we love the new @galaxyproject release! Using Kmer-SVM from @khorling on @biogerovest Galaxy Instance [pic.twitter.com/NqkWhuqR](https://pic.twitter.com/NqkWhuqR)  
RT Retweeted by Galaxy Project  
Show Photo

Nate Conroy (@nconroy) 1h  
It's official, #usegalaxy finally has a new version! This is 15.03.1: RT @galaxyproject The March 2015 Release is here! [bit.ly/galaxy201503](https://bit.ly/galaxy201503)  
RT Retweeted by Galaxy Project  
Expand

Tweet to @galaxyproject

History

search history

Unnamed history  
2 shown, 1 hidden

721.5 MB

3. rmdup on data.3.16.8 EDI

1. RM\_SRA\_SRR054642.F1

10. 10.1.10.10.10.10.10.10

10.1.10.10.10.10.10.10

10.1.10.10.10.10.10.10

PENNSYLVANIA STATE UNIVERSITY

JOHNS HOPKINS UNIVERSITY

TACC

iPlant Collaborative

The Galaxy Team is a part of the Center for Collaborative Genomics and Bioinformatics at Penn State, and the Department of Biology and at Johns Hopkins University.

This instance of Galaxy is utilizing infrastructure generously provided by the iPlant Collaborative at the Texas Advanced Computing Center, with support from the National Science Foundation.

https://usegalaxy.org/workflows/tool\_published



Galaxy | Published Workflows

https://usegalaxy.org/workflow/list\_published

Galaxy [Analysis Tools](#) [Workflows](#) [Shared Data](#) [Visualization](#) [Cloud](#) [Help](#) [Users](#) Logging Out

### Published Workflows

Search name, annotation, owner, and tag

Advanced Search

Name	Annotation	Owner	Community Rating	Community Tags	Last Updated
<a href="#">Origo Preprocessing</a>		robertobouhail	☆☆☆☆☆		-3 minutes ago
<a href="#">Galaxy Variant 303</a>		galaxyproject	☆☆☆☆☆		Mar 06, 2015
<a href="#">Trim Reads Based on Quality (imported from uploaded file)</a>		jhm	☆☆☆☆☆		Feb 28, 2015
<a href="#">RNAseq/TRAFLINE</a>	RNA sequencing data analysis in a Transparent Reproducible and Automated Pipeline - TRAFLINE	maeffler	☆☆☆☆☆	<a href="#">RNA</a> <a href="#">FASTQ</a> <a href="#">FASTA</a> <a href="#">bioRxiv</a> <a href="#">summerburst</a>	Feb 25, 2015
<a href="#">TF_aylwin_302504940</a>		aylwin	☆☆☆☆☆		Feb 19, 2015
<a href="#">BioRxiv121_304</a>		ing-ef	☆☆☆☆☆	<a href="#">RNAseq</a>	Feb 09, 2015
<a href="#">Lab_3_Bioinformatics</a>		moreno	☆☆☆☆☆		Feb 07, 2015
<a href="#">Picard Quality stat on BAM</a>		umar	☆☆☆☆☆		Jan 26, 2015
<a href="#">Quality stat on BAM</a>		umar	☆☆☆☆☆		Jan 26, 2015
<a href="#">Quality Statistics work flow</a>		umar	☆☆☆☆☆		Jan 26, 2015
<a href="#">CATK 337 unified genoscope build</a>		umar	☆☆☆☆☆		Jan 26, 2015
<a href="#">Workflow: hg19 MQ 40 RQ 30</a>		umar	☆☆☆☆☆		Jan 26, 2015
<a href="#">SAM filter Picard duplicate removal CATK recalibration workflow hg19 successful run</a>		umar	☆☆☆☆☆		Jan 26, 2015
<a href="#">Workflow: BED test</a>		umar	☆☆☆☆☆		Jan 26, 2015

https://usegalaxy.org/libraries/index



Galaxy | Published Workflows

https://usegalaxy.org/u/nbertaboukhall/w/ginkgo-grep

Galaxy

Analysis Tools | Workflows | Shared Data | Visualization | Cloud | Help | Users

Published Workflows | nbertaboukhall | Ginkgo Preprocessing

Input workflow

### Galaxy Workflow 'Ginkgo Preprocessing'

Annotation

Step	Annotation
<b>Step 1: Input dataset</b>	
Input Dataset	
select at runtime	
<b>Step 2: Bowtie2</b>	
Is this library mate-paired?	
Single-end	
FASTQ file	
Output dataset 'output' from step 1	
Write unaligned reads to separate files?	
False	
Will you select a reference genome from your history or use a built-in index?	
Use a built-in index	
Select a reference genome	
hg19	
Specify the read group for this file?	
No	
Parameter Settings	
Use defaults	
<b>Step 3: Filter SAM or BAM, output SAM or BAM</b>	
SAM or BAM file to filter	
Output dataset 'output' from step 2	
Reader in output	
Include reader	

https://usegalaxy.org/workflows/imp/af/1/1024134463241344

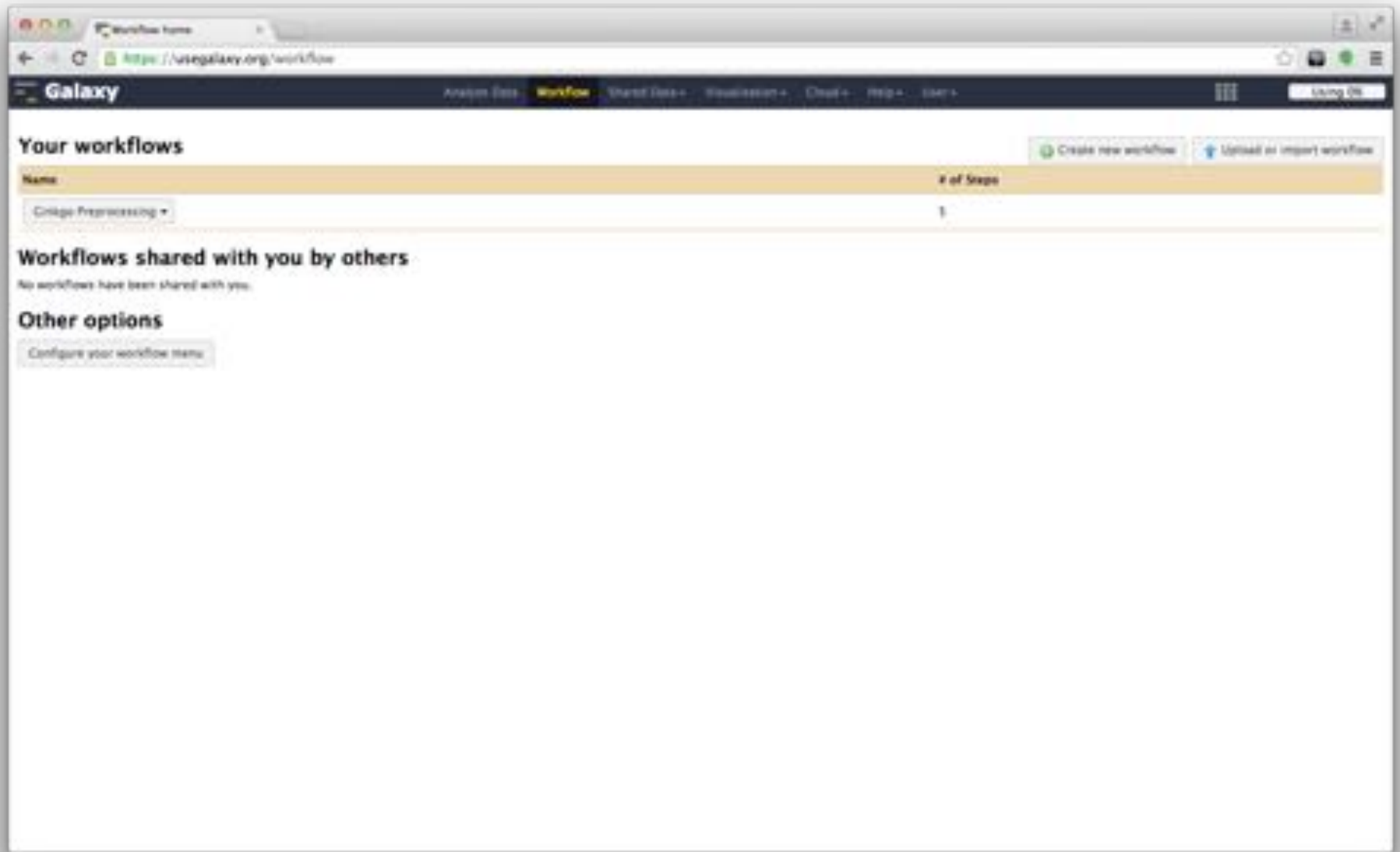
About this Workflow

Author: nbertaboukhall

Related Workflows: All published workflows, Published workflows by nbertaboukhall

Rating: Community (0 ratings, 00 empty), Years (0 ratings)

Tags: Community: none, Years



## Your workflows

[Create new workflow](#) [Upload as import workflow](#)

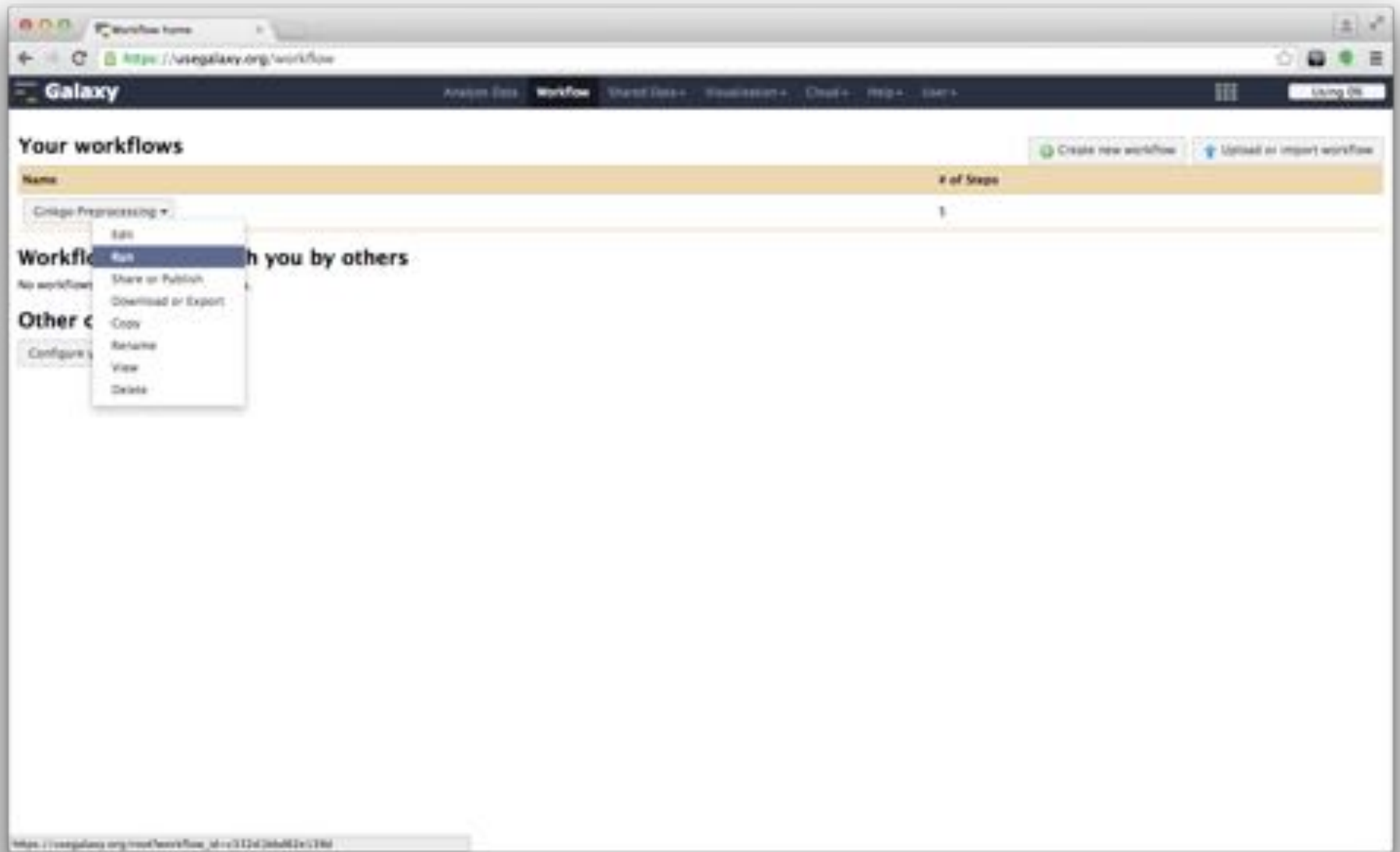
Name	# of Steps
Collage Preprocessing	1

## Workflows shared with you by others

No workflows have been shared with you.

## Other options

[Configure your workflow menu](#)



Galaxy

https://usegalaxy.org/workflow\_id=c332d1bbd62e139d

Galaxy Analyze Data Workflow Share Data+ Visualizations+ Cloud+ Help+ Users+ Logging Out

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Edit and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Picard
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Stats
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Motif Tools
- NGS TOOLBOX BETA

### Running workflow "Ginkgo Preprocessing"

Expand All Collapse

Step 1: Input dataset

Input Dataset

1: SRX154\_SRR054622.Fastq.gz File: <ftp://ftp.era4b.ac.uk/erf1/Tools/SRR054/SRR054622/SRR054622.Fastq.gz>

Fastq to Filter

Step 2: Bwa2 (version 0.2)

Step 3: Filter SAM or BAM, output SAM or BAM (version 1.2.0)

Step 4: rmdup (version 1.2.0)

Step 5: Convert from BAM to BED (version 0.1.0)

Send results to a new history

Run workflow

History

search histories

Unnamed history

1 shown

394.8 MB

Annotation

Click here to edit annotation

1: SRX154\_SRR054622.F1  
File: <ftp://ftp.era4b.ac.uk/erf1/Tools/SRR054/SRR054622/SRR054622.F1.fastq.gz>

Galaxy

https://usegalaxy.org/workflow\_id=c332d1bbd62e139d

Analyze Data Workflow Share Data+ Visualizations+ Cloud+ Help+ Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Edit and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Picard
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Stats
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Motif Tools
- NGS TOOLBOX BETA

### Running workflow "Ginkgo Preprocessing"

Expand All Collapse

Step 1: Input dataset

Input Dataset

1: SRX\_SRA\_SRR054622.Fix.Pa.1.Pa.oxa.afn.ac.uk.sm11.fastq.SRR054622.SRR054622.fastq.gz

View as File

Step 2: Bwa2 (version 0.7)

Step 3: Filter SAM or BAM, output SAM or BAM (version 1.2.0)

Step 4: rmdup (version 1.2.0)

Step 5: Convert from BAM to BED (version 0.1.0)

Send results to a new history

Run workflow

History

search histories

Unnamed history

1 shown

394.8 MB

Annotation

Click here to full annotation

1: SRX\_SRA\_SRR054622.Fix.Pa.1.Pa.oxa.afn.ac.uk.sm11.fastq.SRR054622.SRR054622.fastq.gz

Galaxy

https://usegalaxy.org/root/workflow\_id=c332d1bbd62e139d

Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Filter
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Metil Tools
- NGS TOOLBOX BETA

Successfully ran workflow "Omics Preprocessing". The following datasets have been added to the queue:

1. EB\_SRA\_SRR054622.fastq file: ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR054/SRR054622/SRR054622.fastq.gz
2. Bowtie2 on data 1: aligned reads
3. Filter SAM or BAM, output SAM or BAM on data 2: bam
4. rmdup on data 3
5. rmdup on data 3 (as BED)

History

search datasets

Unnamed history  
3 shown

394.8 MB

Annotation  
Click here to edit annotation

- 1. rmdup on data 3 (as BED)
- 5. rmdup on data 3
- 3. Filter SAM or BAM, output SAM or BAM on data 2: bam
- 2. Bowtie2 on data 1: aligned reads
- 1. EB\_SRA\_SRR054622 (1)  
file: ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR054/SRR054622/SRR054622.fastq.gz

Galaxy

https://usegalaxy.org/root/workflow\_id=c332d1bbd62e139d

Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Missing
- NGS BAM Tools
- NGS Read
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Metil Tools
- NGS TOOLBOX BETA

Successfully ran workflow "Omics Preprocessing". The following datasets have been added to the queue:

1. EB\_SRA\_SRR054622.fastq file: ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR054/SRR054622/SRR054622.fastq.gz
2. Bowtie2 on data 1: aligned reads
3. Filter SAM or BAM, output SAM or BAM on data 2: bam
4. rmdup on data 3
5. rmdup on data 3 (as BED)

History

search datasets

Unnamed history  
3 shown

394.8 MB

Annotation  
Click here to edit annotation

- 1. rmdup on data 3 (as BED)
- 4. rmdup on data 3
- 3. Filter SAM or BAM, output SAM or BAM on data 2: bam
- 2. Bowtie2 on data 1: aligned reads
- 1. EB\_SRA\_SRR054622 (1) file: ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR054/SRR054622/SRR054622.fastq.gz



Galaxy

https://usegalaxy.org/root/workflow\_id=c332d1bbd62e139d

Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Missing
- NGS BAM Tools
- NGS Picard
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Metil Tools
- NGS TOOLBOX BETA

Successfully ran workflow "Omics Preprocessing". The following datasets have been added to the queue:

- 1: SRX SRX SRX014622 File: ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR014/SRR014622/SRR014622.fastq.gz
- 2: Sorted on data 1: aligned reads
- 3: Filter SAM or BAM, output SAM or BAM on data 2: bam
- 4: rmdup on data 3
- 5: rmdup on data 3 (as BED)

History

search datasets

Unnamed history  
4 shown, 1 collapsed  
394.8 MB

Annotations  
Click here to edit annotations

- 1: rmdup on data 3 (as BED)
- 4: rmdup on data 3
- 3: Filter SAM or BAM, output SAM or BAM on data 2: bam
- 1: SRX SRX SRX014622 File: ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR014/SRR014622/SRR014622.fastq.gz

Galaxy

https://usegalaxy.org/root/workflow\_id=c332d1bbd62e139d

Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Filter
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Metil Tools
- NGS TOOLBOX BETA

Successfully ran workflow "Omics Preprocessing". The following datasets have been added to the queue:

- 1: SRX SRX SRX014622 File: ftp://ftp.sra.ftp.nih.gov/1/fastq/SRR014/SRR014622/SRR014622.fastq.gz
- 2: Sorted on data 1: aligned reads
- 3: Filter SAM or BAM, output SAM or BAM on data 2: bam
- 4: rmdup on data 3
- 5: rmdup on data 3 (a BGI)

History

search datasets

Unnamed history  
4 shown, 1 collapsed  
394.8 MB

Annotations  
Click here to edit annotations

- 1: rmdup on data 3 (a BGI)
- 1: Filter SAM or BAM, output SAM or BAM on data 2: bam
- 1: SRX SRX SRX014622 File: ftp://ftp.sra.ftp.nih.gov/1/fastq/SRR014/SRR014622/SRR014622.fastq.gz

Galaxy

https://usegalaxy.org/workflow\_id=c332d1bbd62e139d

Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Filter
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Stats
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Metil Tools
- NGS TOOLBOX BETA

Successfully ran workflow "Omics Preprocessing". The following datasets have been added to the queue:

- 1: SR: SRR14622.fastq.gz [v1.1/ftp.ensembl.org/pub/ftp/fastq/SRR14622/SRR14622.fastq.gz]
- 2: Sorted on data 1: aligned reads
- 3: Filter SAM or BAM, output SAM or BAM on data 2: bam
- 4: rmdup on data 3
- 5: rmdup on data 3 (in BDT)

History

search datasets

Unnamed history  
3 shown, 2 hidden

394.8 MB

Annotation  
Click here to edit annotation

- 1: rmdup on data 3 (in BDT)
- 2: rmdup on data 3
- 3: SR: SRR14622.fastq.gz [v1.1/ftp.ensembl.org/pub/ftp/fastq/SRR14622/SRR14622.fastq.gz]

Galaxy

https://usegalaxy.org/root/workflow\_id=c332d1bbd62e139d

Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Filter
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Stats
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Motif Tools
- NGS TOOLBOX BETA

Successfully ran workflow "Omics Preprocessing". The following datasets have been added to the queue:

- 1: EB\_SRA\_SRR054622.fastq.gz [v1.1/ftp.ncbi.nlm.nih.gov/trace/trace/SRA/SRR054622/SRR054622.fastq.gz]
- 2: Sorted on data 1: aligned reads
- 3: Filter SAM or BAM, output SAM or BAM on data 2: bam
- 4: rmdup on data 3
- 5: rmdup on data 3 (as BED)

History

search datasets

Unnamed history  
3 shown, 2 hidden

394.8 MB

Annotation  
Click here to edit annotation

- 1: rmdup on data 3 (as BED)
- 4: rmdup on data 3
- 1: EB\_SRA\_SRR054622.fastq.gz [v1.1/ftp.ncbi.nlm.nih.gov/trace/trace/SRA/SRR054622/SRR054622.fastq.gz]

Galaxy

https://usegalaxy.org/root/workflow\_id=c332d1bbd62e139d

Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Filter
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Stats
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Metil Tools
- NGS TOOLBOX BETA

Successfully ran workflow "Omics Preprocessing". The following datasets have been added to the queue:

- 1: SRX SRX SRX054622 file: ftp://ftp.ars.ubc.ca/ucsc/ver1/flags/SRR054622/SRR054622.fastq.gz
- 2: Sorted on data 1: aligned reads
- 3: Filter SAM or BAM, output SAM or BAM on data 2: bam
- 4: rmdup on data 3
- 5: rmdup on data 3 (in BQ)

History

search datasets

Unnamed history  
2 shown, 1 hidden  
394.8 MB

Annotation  
Click here to edit annotation

- 1: rmdup on data 3 (in BQ) x
- 1: SRX SRX SRX054622 (1) x  
file: ftp://ftp.ars.ubc.ca/ucsc/ver1/flags/SRR054622/SRR054622.fastq.gz

Galaxy

https://usegalaxy.org/root/workflow\_id=c332d1bbd62e139d

Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Filter
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Stats
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Metil Tools
- NGS TOOLBOX BETA

Successfully ran workflow "Omics Preprocessing". The following datasets have been added to the queue:

- 1: SRP SRX SRX054622 file: ftp://ftp.ncbi.nlm.nih.gov/trace/trace/SRR054622/SRR054622.fastq.gz
- 2: Sorted on data 1: aligned reads
- 3: Filter SAM or BAM, output SAM or BAM on data 2: bam
- 4: rmdup on data 3
- 5: rmdup on data 3 (as BED)

History

search datasets

Unnamed history  
2 shown, 1 hidden  
394.8 MB

Annotation  
Click here to add annotation

- 1: rmdup on data 3 (as BED)
- 1: SRP SRX SRX054622 (1)  
file: ftp://ftp.ncbi.nlm.nih.gov/trace/trace/SRR054622/SRR054622.fastq.gz

Galaxy

https://usegalaxy.org/root/workflow\_id=c332d1bbd62e139d

Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Missing
- NGS BAM Tools
- NGS Picard
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Stats
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray analysis
- Metil Tools
- NGS TOOLBOX BETA

Successfully ran workflow "Omics Preprocessing". The following datasets have been added to the queue:

- 1: SRX SRX:SRX014622 File: ftp://ftp.ensembl.org/pub/ftp/release/54.6/FASTA/SRR014622.SRR014622.fastq.gz
- 2: Sorted on data 1: aligned reads
- 3: Filter SAM or BAM, output SAM or BAM on data 2: bam
- 4: rmdup on data 3
- 5: rmdup on data 3 (as BED)

History

search datasets

Unnamed history  
2 shown, 1 hidden

721.5 MB

Annotation  
Click here to edit annotation

3: rmdup on data 3 (as B  
BED

1: SRX SRX:SRX014622.FI  
ftp://ftp.ensembl.org/pub/ftp/release/54.6/FASTA/SRR014622.SRR014622.fastq.gz



Galaxy [https://usegalaxy.org/root/workflow\\_id=332d1bbd62e139d](https://usegalaxy.org/root/workflow_id=332d1bbd62e139d) Logout

Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users

Tools 

search tools 

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- NGS QC and manipulation
- NGS Mapping
- NGS BAM Tools
- NGS Picard
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snPEP
- BBTools
- Genome Diversity
- EMBOSS
- Regional Variation
- FASTA manipulation
- Evolution
- Multiple Alignments
- Microarray Analysis
- Metatools
- NGS TOOLBOX BETA

 Successfully ran workflow "Omics Preprocessing". The following datasets have been added to the queue:

- 1: SRX SRX\_SRR254622.fastq file: <https://ftp.ncbi.nlm.nih.gov/seq/SRR/SRR254622/SRR254622.fastq.gz>
- 2: Sorted on data 1: aligned reads
- 3: Filter SAM or BAM, output SAM or BAM on data 2: bam
- 4: rmdup on data 3
- 5: rmdup on data 3 (as BED)

History   

search datasets 

Unnamed history   
2 shown, 1 hidden

721.5 MB   

Annotation:  
[Click here to edit annotation](#)

3: rmdup on data 3 (as BED)     
BED 

1: SRX SRX\_SRR254622.1     
file: [https://ftp.ncbi.nlm.nih.gov/seq/SRR/SRR254622.1.fastq.gz](https://ftp.ncbi.nlm.nih.gov/seq/SRR/SRR254622/SRR254622.1.fastq.gz)

<https://usegalaxy.org/helpcenter/faq/what-is-a-toolbox-beta/>



Galaxy [https://usegalaxy.org/workflow\\_id=332d1bbd62e139d](https://usegalaxy.org/workflow_id=332d1bbd62e139d)

Galaxy Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- NCL QC and manipulation
- NCL Missing
- NCL BAM Tools
- NCL Picard
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BEDTools
- Genome Diversity
- EMBOSS
- Regional Variation
- EASTA manipulation
- Evolution
- Multicoll. Alignments
- Metagenomic analysis
- Metil Tools
- NCL TOOLBOX BETA

	1	2	3	4	5	6
chr10	61988	62016	SRX054622.203867	30	+	
chr10	69090	69141	SRX054622.2024140	30	+	
chr10	69098	69146	SRX054622.481587	30	+	
chr10	69099	69147	SRX054622.739359	30	+	
chr10	72809	72907	SRX054622.1184885	30	-	
chr10	82075	82083	SRX054622.2384537	30	-	
chr10	101602	101610	SRX054622.2476079	30	-	
chr10	108416	108464	SRX054622.301574	35	-	
chr10	108416	108464	SRX054622.2321097	25	+	
chr10	110108	110116	SRX054622.1050882	34	-	
chr10	112529	112567	SRX054622.1459050	38	+	
chr10	112525	112573	SRX054622.651721	28	+	
chr10	113729	113771	SRX054622.2235401	30	-	
chr10	114074	114122	SRX054622.2190838	30	+	
chr10	114078	114126	SRX054622.1006253	30	+	
chr10	124019	124067	SRX054622.2083880	42	+	
chr10	130939	131007	SRX054622.213507	40	+	
chr10	132896	132944	SRX054622.1964746	34	+	
chr10	132905	132953	SRX054622.181640	35	+	
chr10	132910	132958	SRX054622.1188143	34	-	
chr10	135289	135337	SRX054622.1757379	42	-	
chr10	137966	138014	SRX054622.1142440	42	+	
chr10	141733	141781	SRX054622.1035785	42	+	
chr10	143964	144012	SRX054622.1722443	42	-	
chr10	148479	148527	SRX054622.879161	42	+	
chr10	148669	148717	SRX054622.1863231	26	+	
chr10	168517	168565	SRX054622.1815104	42	-	
chr10	168526	168574	SRX054622.1863154	42	-	
chr10	168611	168659	SRX054622.2720413	30	+	
chr10	173788	173836	SRX054622.2499973	42	-	
chr10	176020	176068	SRX054622.1812131	42	+	
chr10	179265	179313	SRX054622.1989607	40	-	
chr10	182939	182987	SRX054622.1731700	42	-	
chr10	182940	182988	SRX054622.1479478	42	-	
chr10	183960	184008	SRX054622.531468	48	+	

History

search history

Unnamed history  
2 items, 7 entries

721.5 MB

Annotations  
Click here to edit annotations

1. rmike on data 3 Jan. 8:10

1. SRX SRX054622.F1

In: [ftp://ftp.uswest1.illumina.com](http://ftp.uswest1.illumina.com)  
 /vsl1/seqs/SRX054622/SR  
 X054622.fastq.gz

[https://usegalaxy.org/workflow\\_id=332d1bbd62e139d](https://usegalaxy.org/workflow_id=332d1bbd62e139d)

Galaxy

https://usegalaxy.org/workflow\_id=332d1bbd62e139d

Galaxy Analyze Data Workflow Share Data+ Visualization+ Cloud+ Help+ Users

Tools

search tools

- Get Data
- Send Data
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- NCL QC and manipulation
- NCL Missing
- NCL BAM Tools
- NCL Filter
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- snEff
- BEDTools
- Genome Diversity
- EMBOSS
- Regional Variation
- EASTA manipulation
- Evolution
- Multicall Alignments
- Metagenomic analysis
- MetF Tools
- NCL TOOLBOX BETA

	1	2	3	4	5	6
chr10	61988	62016	SRX054622.2038647	30	+	
chr10	69090	69141	SRX054622.2024140	30	+	
chr10	69098	69146	SRX054622.481587	30	+	
chr10	69099	69147	SRX054622.739359	30	+	
chr10	72859	72907	SRX054622.1184885	30	-	
chr10	82075	82083	SRX054622.2384537	30	-	
chr10	101602	101610	SRX054622.2476079	30	-	
chr10	108416	108464	SRX054622.301574	25	-	
chr10	108416	108464	SRX054622.2321097	25	+	
chr10	110108	110116	SRX054622.1050882	34	-	
chr10	112529	112567	SRX054622.1459050	38	+	
chr10	112525	112573	SRX054622.601721	28	+	
chr10	113729	113771	SRX054622.2235401	30	-	
chr10	114074	114122	SRX054622.2190838	30	+	
chr10	114078	114126	SRX054622.1006253	30	+	
chr10	124019	124067	SRX054622.2083880	42	+	
chr10	130939	131007	SRX054622.213507	40	+	
chr10	132896	132944	SRX054622.1964746	34	+	
chr10	132905	132953	SRX054622.181640	35	+	
chr10	132910	132958	SRX054622.1188143	34	-	
chr10	135289	135337	SRX054622.1757379	42	-	
chr10	137966	138014	SRX054622.1142440	42	+	
chr10	141733	141781	SRX054622.1035785	42	+	
chr10	143964	144012	SRX054622.1722440	42	-	
chr10	148479	148527	SRX054622.879161	42	+	
chr10	148669	148717	SRX054622.1863231	26	+	
chr10	168517	168565	SRX054622.1815104	42	-	
chr10	168526	168574	SRX054622.1863154	42	-	
chr10	168611	168659	SRX054622.2720453	30	+	
chr10	173788	173836	SRX054622.2499973	42	-	
chr10	176020	176068	SRX054622.1812131	42	+	
chr10	179265	179313	SRX054622.1989607	40	-	
chr10	182939	182987	SRX054622.1731700	42	-	
chr10	182940	182988	SRX054622.1479478	42	-	
chr10	182960	183008	SRX054622.531464	40	+	
chr10	183111	183159	SRX054622.1114455	42	-	

History

search history

Unnamed history  
2 items, 7 objects

721.5 MB

Annotation:  
Click here to edit annotation

1. rmDup on data 3 (cat...  
BED  
+1,000,000 regions  
format: bed, chromosome: hg19  
[i] [o] [c] [w] [x] [y] [z] [a] [b] [c] [d] [e] [f] [g] [h] [i] [j] [k] [l] [m] [n] [o] [p] [q] [r] [s] [t] [u] [v] [w] [x] [y] [z] [aa] [ab] [ac] [ad] [ae] [af] [ag] [ah] [ai] [aj] [ak] [al] [am] [an] [ao] [ap] [aq] [ar] [as] [at] [au] [av] [aw] [ax] [ay] [az] [ba] [bb] [bc] [bd] [be] [bf] [bg] [bh] [bi] [bj] [bk] [bl] [bm] [bn] [bo] [bp] [bq] [br] [bs] [bt] [bu] [bv] [bw] [bx] [by] [bz] [ca] [cb] [cc] [cd] [ce] [cf] [cg] [ch] [ci] [cj] [ck] [cl] [cm] [cn] [co] [cp] [cq] [cr] [cs] [ct] [cu] [cv] [cw] [cx] [cy] [cz] [da] [db] [dc] [dd] [de] [df] [dg] [dh] [di] [dj] [dk] [dl] [dm] [dn] [do] [dp] [dq] [dr] [ds] [dt] [du] [dv] [dw] [dx] [dy] [dz] [ea] [eb] [ec] [ed] [ee] [ef] [eg] [eh] [ei] [ej] [ek] [el] [em] [en] [eo] [ep] [eq] [er] [es] [et] [eu] [ev] [ew] [ex] [ey] [ez] [fa] [fb] [fc] [fd] [fe] [ff] [fg] [fh] [fi] [fj] [fk] [fl] [fm] [fn] [fo] [fp] [fq] [fr] [fs] [ft] [fu] [fv] [fw] [fx] [fy] [fz] [ga] [gb] [gc] [gd] [ge] [gf] [gg] [gh] [gi] [gj] [gk] [gl] [gm] [gn] [go] [gp] [gq] [gr] [gs] [gt] [gu] [gv] [gw] [gx] [gy] [gz] [ha] [hb] [hc] [hd] [he] [hf] [hg] [hh] [hi] [hj] [hk] [hl] [hm] [hn] [ho] [hp] [hq] [hr] [hs] [ht] [hu] [hv] [hw] [hx] [hy] [hz] [ia] [ib] [ic] [id] [ie] [if] [ig] [ih] [ii] [ij] [ik] [il] [im] [in] [io] [ip] [iq] [ir] [is] [it] [iu] [iv] [iw] [ix] [iy] [iz] [ja] [jb] [jc] [jd] [je] [jf] [jg] [jh] [ji] [jj] [jk] [jl] [jm] [jn] [jo] [jp] [jq] [jr] [js] [jt] [ju] [jv] [jw] [jx] [jy] [jz] [ka] [kb] [kc] [kd] [ke] [kf] [kg] [kh] [ki] [kj] [kk] [kl] [km] [kn] [ko] [kp] [kq] [kr] [ks] [kt] [ku] [kv] [kw] [kx] [ky] [kz] [la] [lb] [lc] [ld] [le] [lf] [lg] [lh] [li] [lj] [lk] [ll] [lm] [ln] [lo] [lp] [lq] [lr] [ls] [lt] [lu] [lv] [lw] [lx] [ly] [lz] [ma] [mb] [mc] [md] [me] [mf] [mg] [mh] [mi] [mj] [mk] [ml] [mm] [mn] [mo] [mp] [mq] [mr] [ms] [mt] [mu] [mv] [mw] [mx] [my] [mz] [na] [nb] [nc] [nd] [ne] [nf] [ng] [nh] [ni] [nj] [nk] [nl] [nm] [nn] [no] [np] [nq] [nr] [ns] [nt] [nu] [nv] [nw] [nx] [ny] [nz] [oa] [ob] [oc] [od] [oe] [of] [og] [oh] [oi] [oj] [ok] [ol] [om] [on] [oo] [op] [oq] [or] [os] [ot] [ou] [ov] [ow] [ox] [oy] [oz] [pa] [pb] [pc] [pd] [pe] [pf] [pg] [ph] [pi] [pj] [pk] [pl] [pm] [pn] [po] [pp] [pq] [pr] [ps] [pt] [pu] [pv] [pw] [px] [py] [pz] [qa] [qb] [qc] [qd] [qe] [qf] [qg] [qh] [qi] [qj] [qk] [ql] [qm] [qn] [qo] [qp] [qq] [qr] [qs] [qt] [qu] [qv] [qw] [qx] [qy] [qz] [ra] [rb] [rc] [rd] [re] [rf] [rg] [rh] [ri] [rj] [rk] [rl] [rm] [rn] [ro] [rp] [rq] [rr] [rs] [rt] [ru] [rv] [rw] [rx] [ry] [rz] [sa] [sb] [sc] [sd] [se] [sf] [sg] [sh] [si] [sj] [sk] [sl] [sm] [sn] [so] [sp] [sq] [sr] [ss] [st] [su] [sv] [sw] [sx] [sy] [sz] [ta] [tb] [tc] [td] [te] [tf] [tg] [th] [ti] [tj] [tk] [tl] [tm] [tn] [to] [tp] [tq] [tr] [ts] [tt] [tu] [tv] [tw] [tx] [ty] [tz] [ua] [ub] [uc] [ud] [ue] [uf] [ug] [uh] [ui] [uj] [uk] [ul] [um] [un] [uo] [up] [uq] [ur] [us] [ut] [uu] [uv] [uw] [ux] [uy] [uz] [va] [vb] [vc] [vd] [ve] [vf] [vg] [vh] [vi] [vj] [vk] [vl] [vm] [vn] [vo] [vp] [vq] [vr] [vs] [vt] [vu] [vv] [vw] [vx] [vy] [vz] [wa] [wb] [wc] [wd] [we] [wf] [wg] [wh] [wi] [wj] [wk] [wl] [wm] [wn] [wo] [wp] [wq] [wr] [ws] [wt] [wu] [wv] [ww] [wx] [wy] [wz] [xa] [xb] [xc] [xd] [xe] [xf] [xg] [xh] [xi] [xj] [xk] [xl] [xm] [xn] [xo] [xp] [xq] [xr] [xs] [xt] [xu] [xv] [xw] [xx] [xy] [xz] [ya] [yb] [yc] [yd] [ye] [yf] [yg] [yh] [yi] [yj] [yk] [yl] [ym] [yn] [yo] [yp] [yq] [yr] [ys] [yt] [yu] [yv] [yw] [yx] [yz] [za] [zb] [zc] [zd] [ze] [zf] [zg] [zh] [zi] [zj] [zk] [zl] [zm] [zn] [zo] [zp] [zq] [zr] [zs] [zt] [zu] [zv] [zw] [zx] [zy] [zz]

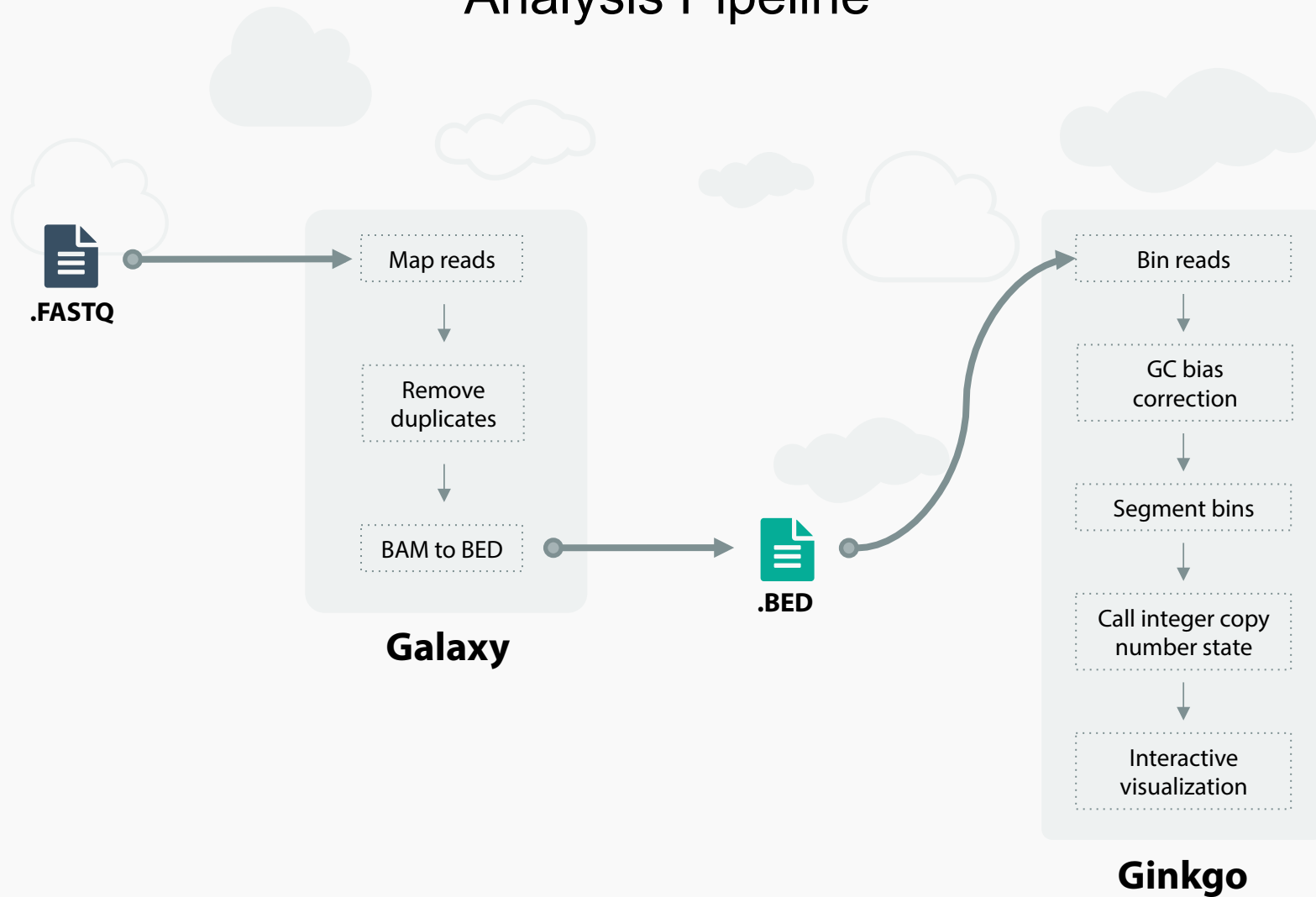
1. RM\_SNA\_SRX054622\_1  
W: 1001/1001001001001001  
/vol1/Tools/SRX054622/SRX054622.fasta.gz

The screenshot shows the Galaxy web interface at [https://usegalaxy.org/workflow\\_id=33261b6d62e139d](https://usegalaxy.org/workflow_id=33261b6d62e139d). The main content area displays a table of genomic data with columns for chromosome, start, end, ID, and other metrics. A sidebar on the left lists various tools, and a history panel on the right shows recent operations, including a BED file manipulation step.

chr10	start	end	ID	score	action
chr10	61388	62016	SRR054622.2139947	30	+
chr10	69090	69141	SRR054622.2024140	30	+
chr10	69098	69146	SRR054622.481587	30	+
chr10	69099	69147	SRR054622.739359	30	+
chr10	72809	72907	SRR054622.1184885	30	-
chr10	82075	82083	SRR054622.2384537	30	-
chr10	101842	101810	SRR054622.2476079	30	-
chr10	108416	108464	SRR054622.301574	15	-
chr10	108416	108464	SRR054622.2321097	25	+
chr10	110108	110116	SRR054622.1050882	34	-
chr10	112529	112567	SRR054622.1459050	38	+
chr10	112525	112573	SRR054622.651721	28	+
chr10	113729	113771	SRR054622.2235401	30	-
chr10	114074	114122	SRR054622.2190838	30	+
chr10	114078	114126	SRR054622.1006253	30	+
chr10	124019	124067	SRR054622.2083880	42	+
chr10	110919	111007	SRR054622.213507	40	+
chr10	112896	112944	SRR054622.1964746	34	+
chr10	112905	112953	SRR054622.181840	35	+
chr10	112910	112958	SRR054622.1188143	34	-
chr10	115289	115337	SRR054622.1757379	42	-
chr10	117966	118014	SRR054622.1142940	42	+
chr10	141733	141781	SRR054622.1035785	42	+
chr10	143964	144012	SRR054622.1722649	42	-
chr10	148479	148527	SRR054622.879161	42	+
chr10	148669	148717	SRR054622.1863231	28	+
chr10	168517	168565	SRR054622.1815104	42	-
chr10	168526	168574	SRR054622.1863258	42	-
chr10	168611	168659	SRR054622.2720413	30	+
chr10	171788	171836	SRR054622.2499973	42	-
chr10	176020	176068	SRR054622.1812131	42	+
chr10	179265	179313	SRR054622.1989607	40	-
chr10	182939	182987	SRR054622.1731700	42	-
chr10	182940	182988	SRR054622.1479478	42	-
chr10	183960	184008	SRR054622.531468	40	+

**History Panel:** Unnamed history (2 items, 7 objects). 721.5 MB. 1. rmbed on data 3.txt BED. -1,000,000 regions. Format: bed, database: hg19. A 'Download' button is visible.

# Analysis Pipeline



## Hands-on with Ginkgo



**Ginkgo:** [qb.cshl.edu/ginkgo](http://qb.cshl.edu/ginkgo)

**Sample Data:** [qb.cshl.edu/ginkgo/vizbi2015.tar](http://qb.cshl.edu/ginkgo/vizbi2015.tar)