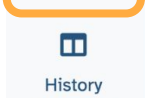
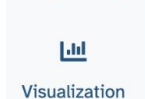
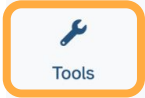
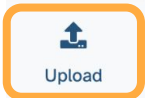


Summary - what we did

Galaxy tutorials

- Learned how to use Galaxy
- Used some common genomics tools for:
 - Data QC
 - Assembly
 - Assembly QC
 - Finding repeats
 - Annotation



Key Galaxy icons in the Activity bar

Activity bar


Home News People About Support Docs

Galaxy AUSTRALIA

Galaxy platform 2024 update published!

Read the latest developments supporting accessible, reproducible, and collaborative data analyses

With contributions from 130 authors representing 60 institutions
doi.org/10.1093/nar/gkae410



Galaxy Australia is an **open, web-based** platform for accessible, reproducible and transparent computational research. Galaxy supports thousands of

History

search datasets

Tutorial: Genome Annotation with Helixer

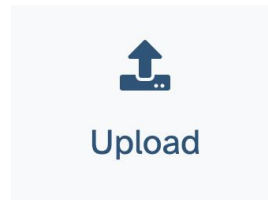
workshop-history

143 MB 26 1

- 27: Busco on data 6: GFF
- 26: Busco on data 6: summary image
- 25: Busco on data 6: missing buscos
- 24: Busco on data 6: full table

Getting data into Galaxy

- Upload data
 - e.g. Choose local file (from your computer)
 - e.g. Paste/Fetch data (e.g. a URL)
- For sequencing reads from NCBI Sequence Read Archive:
 - **Galaxy Tool:** `Faster Download and Extract Reads in FASTQ format from NCBI SRA`
 - You can find sequence reads you want on NCBI, copy the SRR number, and use it as input into the tool



Data QC

Extra links:

- For more about base qualities:
 - <https://training.galaxyproject.org/training-material/topics/introduction/tutorials/galaxy-intro-ngs-data-management/tutorial.html#what-are-base-qualities>
- For more about the plots produced by FastQC and what they mean:
 - <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

Assembly and annotation

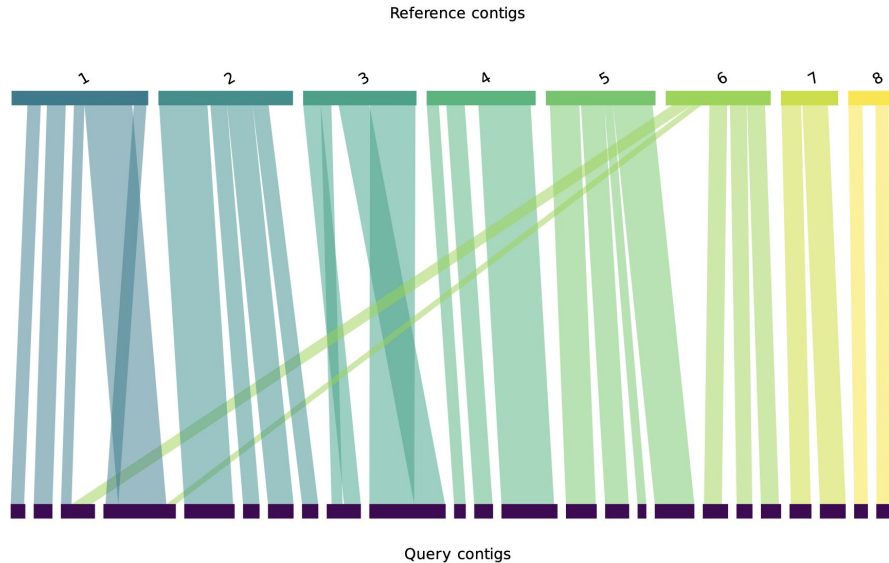
Extra links:

- More about assembling metagenomic data:
 - <https://training.galaxyproject.org/training-material/topics/microbiome/tutorials/metagenomics-assembly/tutorial.html>
- More about assembly decontamination:
 - <https://training.galaxyproject.org/training-material/topics/assembly/tutorials/assembly-decontamination/tutorial.html>

Synteny tools

- Align assembled contigs against a reference genome
 - Tool: RagTag, then plot with Plot RagTag output

Example



Metagenomics, proteomics, phylogenetics

Galaxy curated tutorial sets for:

- Metagenomics data processing and analysis for microbiome:
<https://training.galaxyproject.org/training-material/learning-pathways/metagenomics.html>
- Clinical metaproteomics workflows within Galaxy:
<https://training.galaxyproject.org/training-material/learning-pathways/clinical-metaproteomics.html>
- Proteogenomics:
<https://training.galaxyproject.org/training-material/learning-pathways/proteogenomics.html>

Biosynthetic gene clusters tutorial:

- https://training.galaxyproject.org/training-material/topics/ecology/tutorials/marine_omics_bgc/tutorial.html

Phylogenetics tutorial:

- https://training.galaxyproject.org/training-material/topics/evolution/tutorials/abc_intro_phylo/tutorial.html

What tools and workflows to use?

- Will depend on your data and research questions
- Taking some of the main tools we used or talked about this week:
 - We made some workflows to demonstrate how these tools could be used
 - Don't worry about the detail in these slides following, the idea is to show the range of tools available
 - You can import a workflow to your account, then zoom in, and change as needed (links provided)

Combined fungi assembly and annotation workflow



Assembly

1: Illumina R1 reads
output (input)

2: Illumina R2 reads
output (input)

6: Fungi: Illumina data QC and assembly
Illumina R1 reads
Illumina R2 reads
Assembly-report (html)
Assembled-contigs-from-
 Shovill (fasta)
 get-organelle-logfile (txt)
Fasta-statistics-on-
 assembly (tabular)
 Assembly-graph (jpg, png, svg)
 Mito-assembly-graph (jpg, png, svg)

subworkflow

Data QC and Assembly

Assembly QC

3: Reference_genome.fasta
output (input)

4: Proteins.fasta
output (input)

7: Fungi: Assembly QC, Blast, RagTag
assembly.fasta
Reference_genome.fasta
Ordered-contigs (fasta)
 blast-results (tabular, txt, html, blastxml)
 assembly-vs-ref-synteny (pdf)
 Chromeister-dotplot (png)

subworkflow

Assembly QC

Annotation

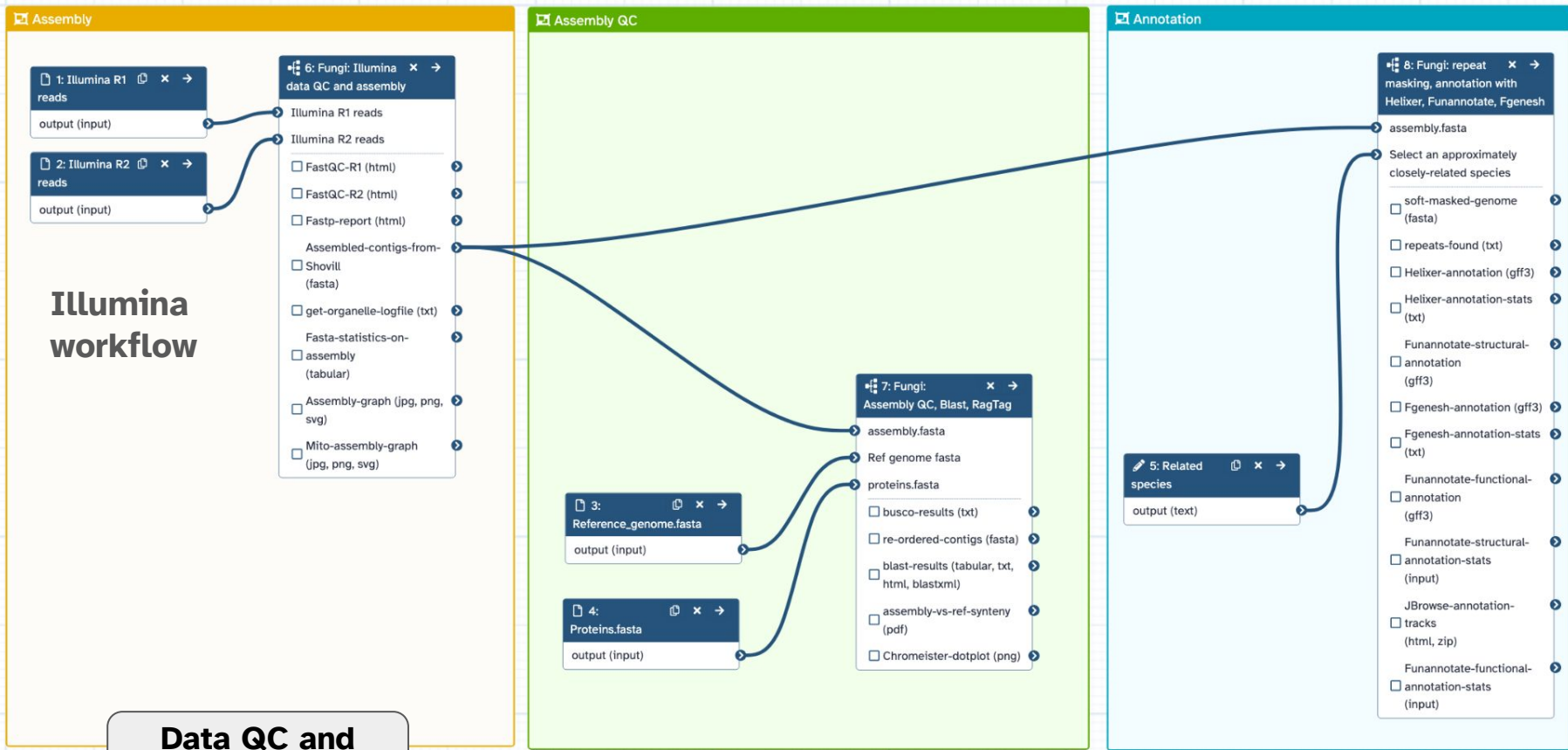
5: Related species
output (text)

8: Fungi: repeat masking, annotation with Helixer, Funannotate, Fgenesh
assembly.fasta
Select
repeats-found (txt)
 Helixer-annotation (gff3)
 Helixer-annotation-stats (txt)
 Funannotate-structural-annotation (gff3)
 Fgenesh-annotation (gff3)
 Fgenesh-annotation-stats (txt)
 Funannotate-functional-annotation (gff3)
 annotation-stats (input)
 JBrowse-annotation-tracks (html, zip)
 Funannotate-functional-annotation-stats (input)

subworkflow

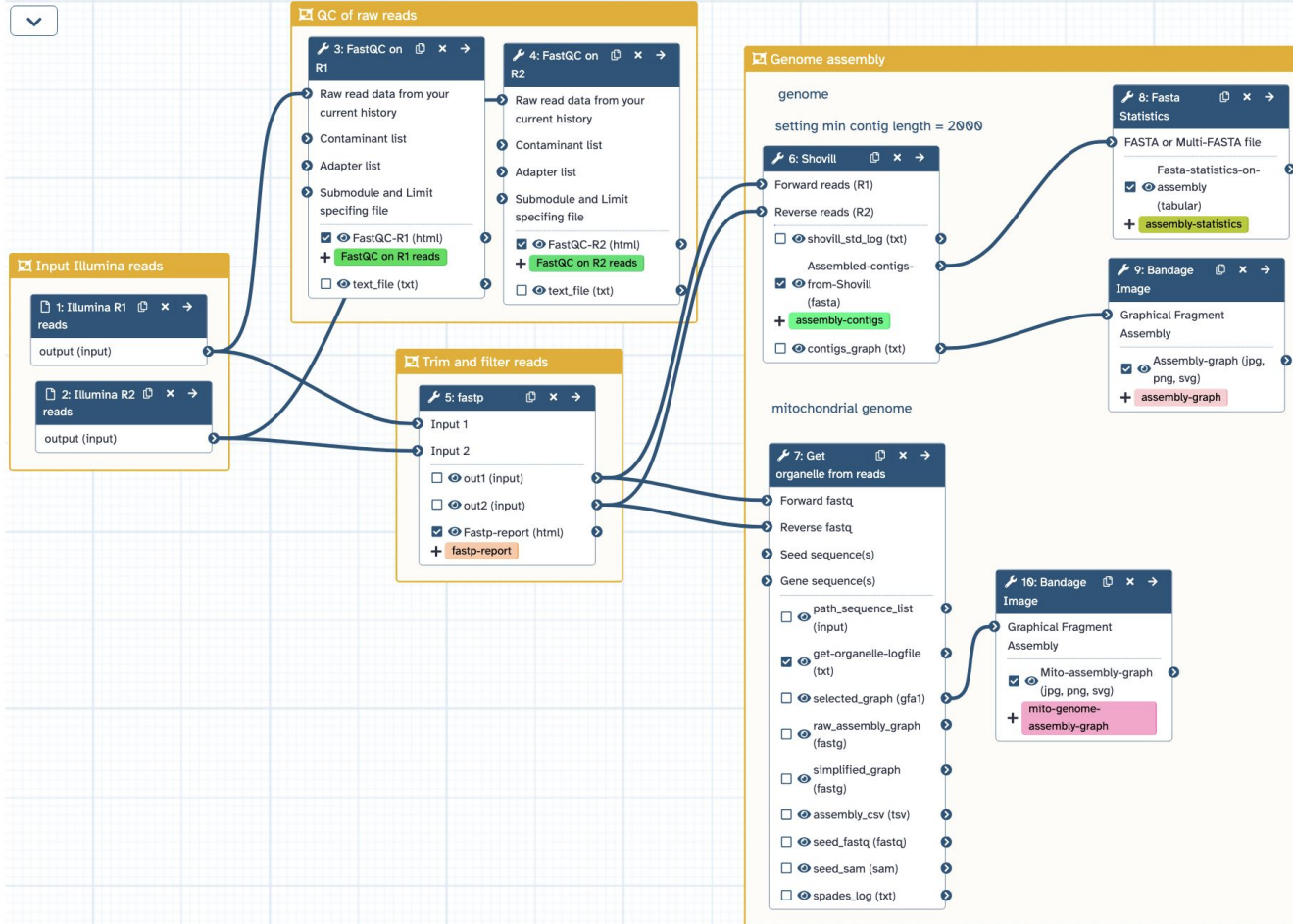
Annotation

Combined fungi assembly and annotation workflow



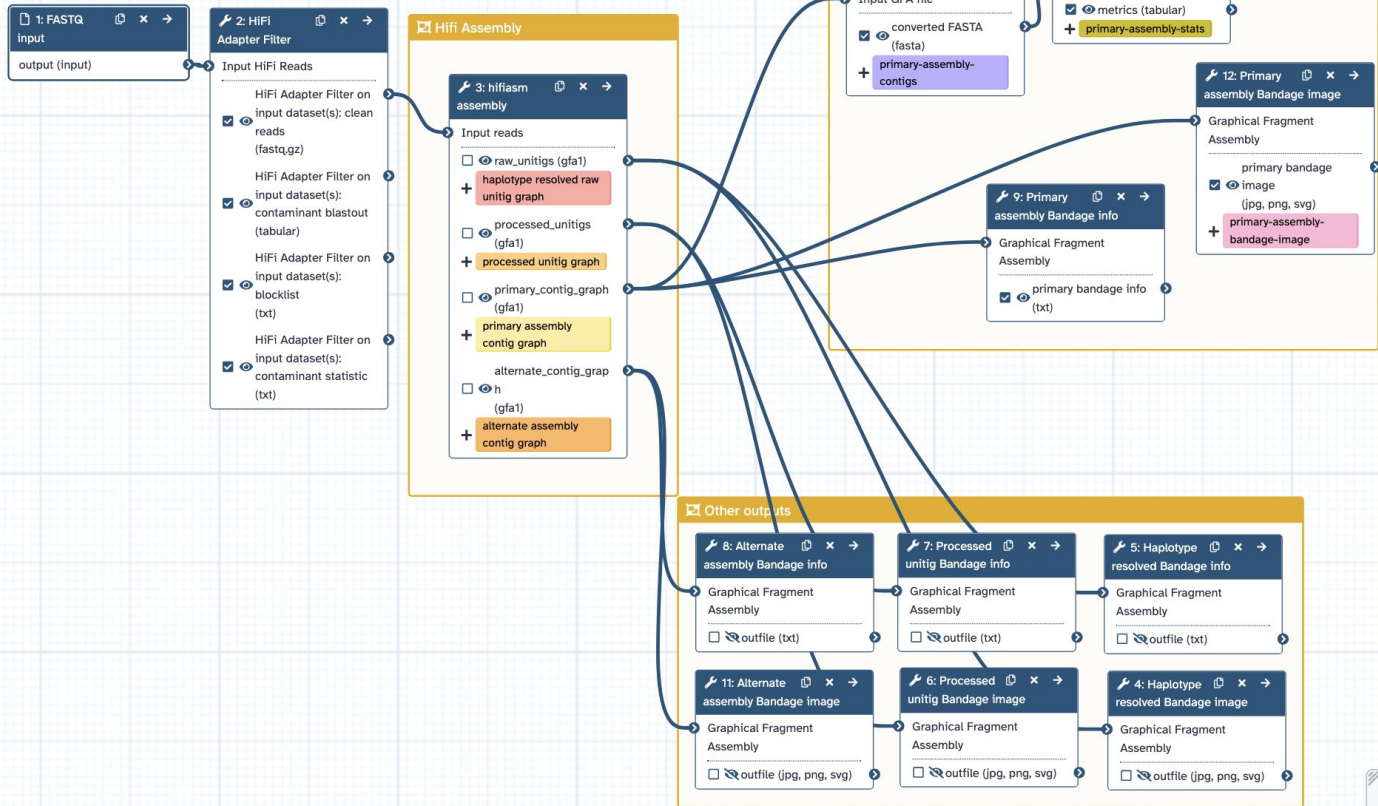
Data QC and Assembly

Fungi: Illumina data QC and assembly



Tools:
FastQC
Fastp
Shovill
get-organelle
Fasta Statistics
Bandage

Fungi: PacBio HiFi assembly (from TSI workflow)



Tools:
Hifi Adapter Filter
Hifiasm
Bandage

Fungi: Nanopore assembly (from TSI workflow)



1: Sequencing reads (in any of these formats: fastq, fastq.gz, fastqsanger, fastqsanger.gz)
output (input)

Trim and filter reads

Input file is split to make Porechop step faster

7: Split into separate files
FASTQ file to split
Number of new files
list_output_fastq (input)

2: How many new files to split into during read filtering stage?
output (integer)

3: Minimum average read quality score to filter on
output (integer)

4: Minimum read length to filter on
output (integer)

5: Trim this many nucleotides from start of read
output (integer)

Porechop adapter trimming
default settings; can be changed at runtime.
output is uncompressed, as required by nanofilt

8: Porechop on each file
Input FASTA/FASTQ
outfile (fasta, fastqsanger, fasta.gz, fastqsanger.gz)

Nanofilt default settings except for:
average read quality: 10
min read length: 200
trim from start of read: 50 bp
these are suggestions and can be changed at runtime, e.g., if reads are very high quality, you may wish to set a higher average read quality cutoff (or vice versa).

9: Nanofilt
uncompressed fastq file
Filter on a minimum average read quality score
Filter on a minimum read length
Trim n nucleotides from start of read
output (fastqsanger)
logfile (tabular)

Data QC

6: Raw reads FastQC
Raw read data from your current history
Contaminant list
Adapter list
Submodule and Limit specifying file
html_file (html)
fastqc-on-raw-reads
text_file (txt)

11: Trimmed, filtered reads FastQC
Raw read data from your current history
Contaminant list
Adapter list
Submodule and Limit specifying file
html_file (html)
fastqc-on-filtered-reads
text_file (txt)

Assembly

10: Collapse Collection
Collection of files to collapse into single dataset
output (input)
filtered-reads-fastqsanger

12: Flye
Input reads
consensus (fasta)
primary-assembly-contigs-fasta
assembly_graph (graph.dot)
assembly_gfa (txt)
primary-assembly-gfa
assembly_info (tabular)
flye_log (txt)

13: Primary assembly Fasta Statistics
fasta or multifasta file
metrics (tabular)
primary-assembly-fasta-stats

14: Primary assembly Bandage Info
Graphical Fragment Assembly
primary bandage info (txt)

15: Primary assembly Bandage Image
Graphical Fragment Assembly
primary bandage
image (jpg, png, svg)
primary-assembly-bandage-graph

Tools:
Porechop
Nanofilt
FastQC
Flye
Fasta Statistics
Bandage

Combined fungi assembly and annotation workflow



Assembly

1: Illumina R1 reads
output (input)

2: Illumina R2 reads
output (input)

6: Fungi: Illumina data QC and assembly
Illumina R1 reads
Illumina R2 reads
 FastQC-R1 (html)
 FastQC-R2 (html)
 Fastp-report (html)
Assembled-contigs-from-
 Shovill (fasta)
 get-organelle-logfile (txt)
Fasta-statistics-on-
 assembly (tabular)
 Assembly-graph (jpg, png, svg)
 Mito-assembly-graph (jpg, png, svg)

Illumina workflow

Assembly QC

3: Reference_genome.fasta
output (input)

4: Proteins.fasta
output (input)

7: Fungi: Assembly QC, Blast, RagTag
assembly.fasta
Ref genome fasta
proteins.fasta
 busco-results (txt)
 re-ordered-contigs (fasta)
 blast-results (tabular, txt, html, blastxml)
 assembly-vs-ref-syteny (pdf)
 Chromeister-dotplot (png)

Assembly QC

Annotation

5: Related species
output (text)

8: Fungi: repeat masking, annotation with Helixer, Funannotate, Fgenesh
assembly.fasta
Select an approximately closely-related species
 soft-masked-genome (fasta)
 repeats-found (txt)
 Helixer-annotation (gff3)
 Helixer-annotation-stats (txt)
 Funannotate-structural-annotation (gff3)
 Fgenesh-annotation (gff3)
 Fgenesh-annotation-stats (txt)
 Funannotate-functional-annotation (gff3)
 Funannotate-structural-annotation-stats (input)
 JBrowse-annotation-tracks (html, zip)
 Funannotate-functional-annotation-stats (input)

Fungi: Assembly QC, Blast, RagTag



Assess the assembly

4: Busco

Sequences to analyse

- busco-results (txt)
- + busco-output
- busco_table (tabular)
- busco_missing (tabular)
- summary_image (png)
- busco_gff (gff3)

Compare to a reference genome

re-order contigs by comparing to a reference

6: RagTag

Reference FASTA file

Query FASTA file

List of reference headers to ignore

List of query headers to leave uncorrected

- scaffold_paf (paf)
- scaffold_agp (agp)
- re-ordered-contigs (fasta)
- + re-ordered-contigs
- scaffold_log (txt)
- scaffold_stats (tabular)
- scaffold_confidence (tabular)

2: Ref genome fasta

output (input)

synteny plot

8: Plot RagTag output

PAF output from RagTag

AGP output from RagTag

assembly-vs-ref-

- synteny (pdf)
- + synteny-plot

dot plot of assembly vs ref

9: Chromeister

Query sequence

Reference sequence

- output (txt)
- Chromeister-dotplot (png)
- + dotplot
- output_csv (csv)
- output_events (txt)
- output_events_png (png)
- output_score (txt)

Blast the assembly

search genome for a protein/gene

3: proteins.fasta

output (input)

5: NCBI BLAST+ makeblastdb

Select input 1 > FASTA input

Optional ASN.1 file(s) containing masking data

- outfile (data, blastdbn, blastdbp)

7: NCBI BLAST+ tblastn

Protein query sequence(s)

Nucleotide BLAST database

- blast-results (tabular, txt, html, blastxml)
- + blast-results

Tools:
Busco
RagTag
Chromeister
Blast

Combined fungi assembly and annotation workflow

Assembly

1: Illumina R1 reads
output (input)

2: Illumina R2 reads
output (input)

6: Fungi: Illumina data QC and assembly

- Illumina R1 reads
- Illumina R2 reads
- FastQC-R1 (html)
- FastQC-R2 (html)
- Fastp-report (html)
- Assembled-contigs-from-
 - Shovill (fasta)
 - get-organelle-logfile (txt)
- Fasta-statistics-on-assembly (tabular)
- Assembly-graph (jpg, png, svg)
- Mito-assembly-graph (jpg, png, svg)

Illumina workflow

Assembly QC

3: Reference_genome.fasta
output (input)

4: Proteins.fasta
output (input)

7: Fungi: Assembly QC, Blast, RagTag

- assembly.fasta
- Ref genome fasta
- proteins.fasta
- busco-results (txt)
- re-ordered-contigs (fasta)
- blast-results (tabular, txt, html, blastxml)
- assembly-vs-ref-syteny (pdf)
- Chromeister-dotplot (png)

Annotation

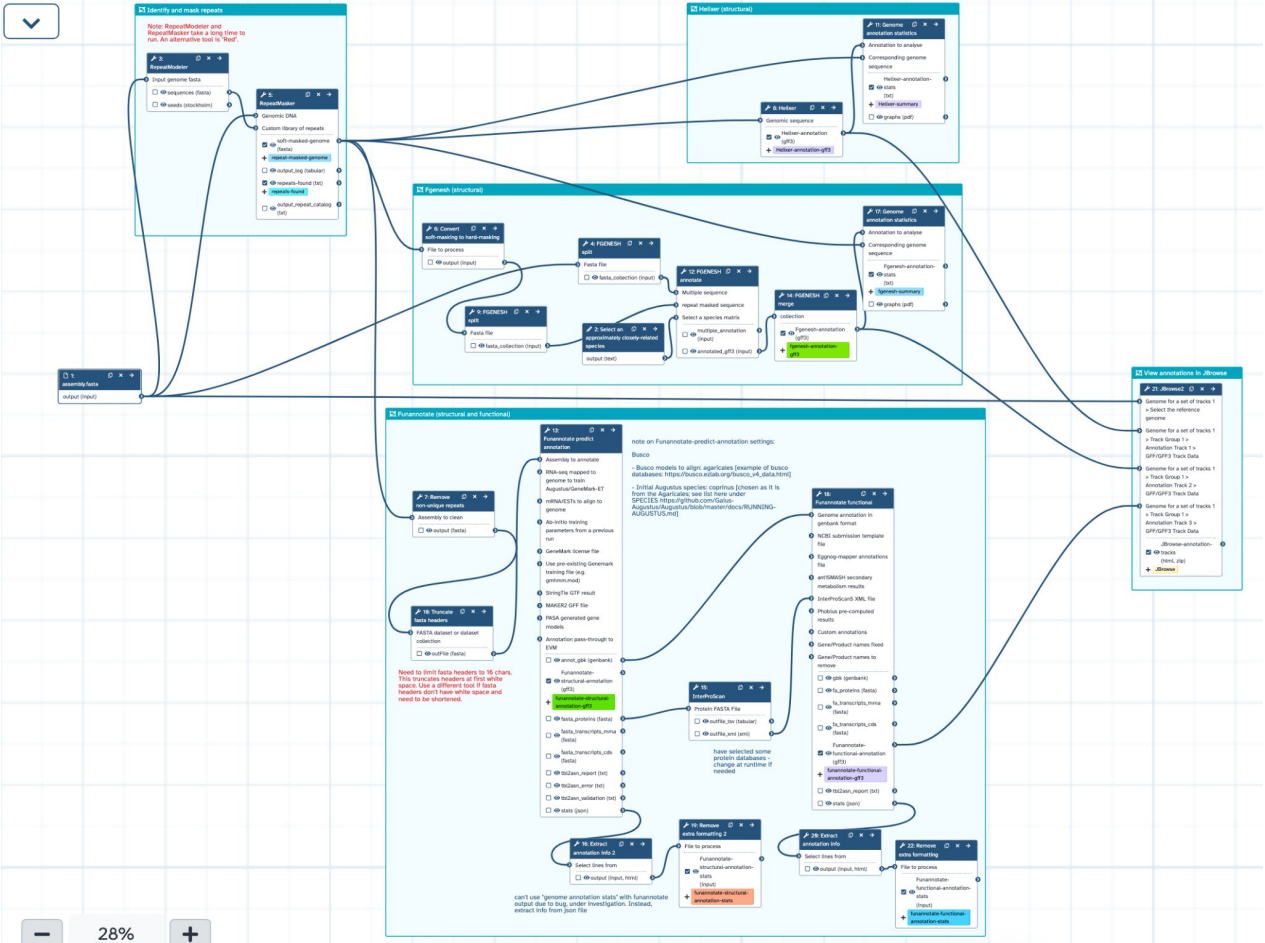
5: Related species
output (text)

8: Fungi: repeat masking, annotation with Helixer, Funannotate, Fgenesh

- assembly.fasta
- Select an approximately closely-related species
 - soft-masked-genome (fasta)
 - repeats-found (txt)
 - Helixer-annotation (gff3)
 - Helixer-annotation-stats (txt)
 - Funannotate-structural-annotation (gff3)
 - Fgenesh-annotation (gff3)
 - Fgenesh-annotation-stats (txt)
 - Funannotate-functional-annotation (gff3)
 - Funannotate-structural-annotation-stats (input)
 - JBrowse-annotation-tracks (html, zip)
 - Funannotate-functional-annotation-stats (input)

Annotation

Fungi: repeat masking, annotation with Helixer, Funannotate, Fgenesh



Tools:
RepeatModeler
RepeatMasker
Helixer
Fgenesh
Funannotate
InterProScan
JBrowse

Combined fungi assembly and annotation workflow



Assembly

1: Illumina R1 reads
output (input)

2: Illumina R2 reads
output (input)

6: Fungi: Illumina data QC and assembly
Illumina R1 reads
Illumina R2 reads
 FastQC-R1 (html)
 FastQC-R2 (html)
 Fastp-report (html)
Assembled-contigs-from-
 Shovill (fasta)
 get-organelle-logfile (txt)
Fasta-statistics-on-
 assembly (tabular)
 Assembly-graph (jpg, png, svg)
 Mito-assembly-graph (jpg, png, svg)

Data QC and Assembly

Assembly QC

3: Reference_genome.fasta
output (input)

4: Proteins.fasta
output (input)

7: Fungi: Assembly QC, Blast, RagTag
assembly.fasta
Ref genome fasta
proteins.fasta
 busco-results (txt)
 re-ordered-contigs (fasta)
 blast-results (tabular, txt, html, blastxml)
 assembly-vs-ref-syteny (pdf)
 Chromeister-dotplot (png)

Assembly QC

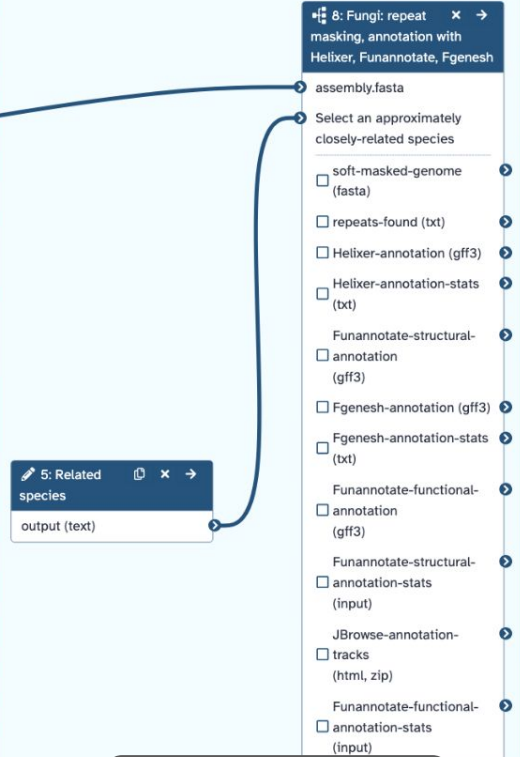
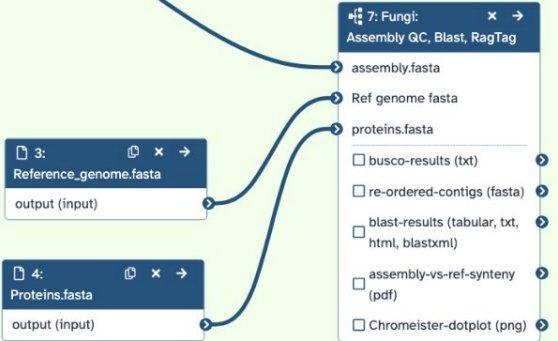
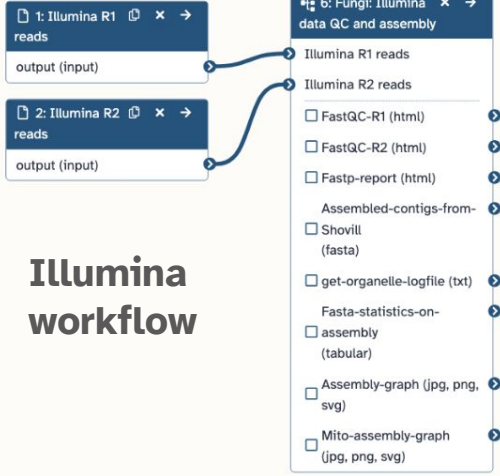
Annotation

5: Related species
output (text)

8: Fungi: repeat masking, annotation with Helixer, Funannotate, Fgenesh
assembly.fasta
Select an approximately closely-related species
 soft-masked-genome (fasta)
 repeats-found (txt)
 Helixer-annotation (gff3)
 Helixer-annotation-stats (txt)
 Funannotate-structural-annotation (gff3)
 Fgenesh-annotation (gff3)
 Fgenesh-annotation-stats (txt)
 Funannotate-functional-annotation (gff3)
 Funannotate-structural-annotation-stats (input)
 JBrowse-annotation-tracks (html, zip)
 Funannotate-functional-annotation-stats (input)

Annotation

Illumina workflow



Workflow tests

- We ran these workflows on some fungi data from BPA and NCBI (details next slide)
 - You can import/view the Galaxy histories to see how the tools worked and what the results look like
- (Note: these workflows and their results are just examples, so don't use the results as input data for real analyses)

Links to these workflows and histories

Workflow name	Workflow link	Data	Galaxy history
Combined workflow: Illumina assembly + QC + annotation	https://usegalaxy.org.au/u/anna/w/combined-fungi-assembly-and-annotation-workflow	<i>Psilocybe subaeruginosa</i> BPA sample 465877 Ref genome <i>P. cubensis</i> Protein: psilocybin synthase	https://usegalaxy.org.au/u/anna/h/combined-fungi-workflows-illumina-psilocybe
Fungi: Illumina data QC and assembly	https://usegalaxy.org.au/u/anna/w/fungi-illumina-assembly	<i>Psilocybe subaeruginosa</i> BPA sample 465877	https://usegalaxy.org.au/u/anna/h/psilocybe-illumina-assembly
Fungi: PacBio assembly	https://usegalaxy.org.au/u/anna/w/fungi-hifi-assembly	<i>Rhynchosporium commune</i> BPA sample 395386	https://usegalaxy.org.au/u/anna/h/rhynchosporium-pacbio-assembly
Fungi: Nanopore assembly	https://usegalaxy.org.au/u/anna/w/fungi-nanopore-assembly	<i>Aspergillus fumigatus</i> NCBI SRR23337894 SRR23337893 SRR23337895	https://usegalaxy.org.au/u/anna/h/aspergillus-nanopore-assembly
Fungi: Assembly QC, Blast, RagTag	https://usegalaxy.org.au/u/anna/w/fungi-assembly-qc	<i>Psilocybe subaeruginosa</i> assembly from Illumina workflow Ref genome <i>P. cubensis</i> (as in fewer contigs than ref genome for <i>P. subaeruginosa</i>). Protein: psilocybin synthase	https://usegalaxy.org.au/u/anna/h/psilocybe-assembly-qc
Fungi: repeat masking, annotation with Helixer, Funannotate, Fgenesh	https://usegalaxy.org.au/u/anna/w/fungi-annotation	<i>Psilocybe subaeruginosa</i> assembly from Illumina workflow	https://usegalaxy.org.au/u/anna/h/repeat-masking-and-annotation---psilocybe

Welcome to Galaxy Training!

Collection of tutorials developed and maintained by the worldwide Galaxy community

**Search for more
tutorials in the
Galaxy Training
Network**

Galaxy for Scientists

We have separated the tutorials into several categories based on field and technology. We are exploring other ways to organise the tutorials going forward!

Start Here

Topic	Tutorials
Introduction to Galaxy Analyses	13
Using Galaxy and Managing your Data	23

Not sure where to start?

Try the NGS Basic Learning Path!

Start Learning

Quickstart

Learning
Pathways



Galaxy for
SysAdmins





Galaxy for
Developers



Galaxy for
Teachers



Reminder on help:

- In Galaxy: top menu bar:  Help ▾
- Report bugs (bug icon under file) 
- Email: help@genome.edu.au
- Matrix: <https://gitter.im/galaxyproject/Lobby>
- Global help: <https://help.galaxyproject.org/>
- Galaxy Training Network: <https://training.galaxyproject.org/>
- Toolshed: <https://toolshed.g2.bx.psu.edu/>
- How-To-Guides: <https://australianbiocommons.github.io/how-to-hub/index>
- Australian BioCommons youtube channel:
<https://www.youtube.com/channel/UC5WIFNBSfmt3e8Js8o2fFqQ>

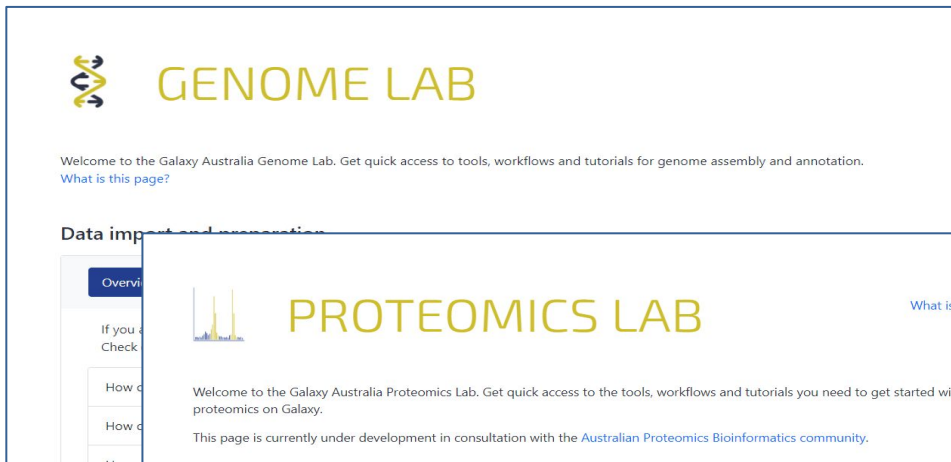
Galaxy or: How I Learned to Stop Worrying and Love the GUI

Galaxy Labs

User    

Jobs Admin

- Genome Lab
- Proteomics Lab
- Single Cell Lab



GENOME LAB

Welcome to the Galaxy Australia Genome Lab. Get quick access to tools, workflows and tutorials for genome assembly and annotation.
[What is this page?](#)

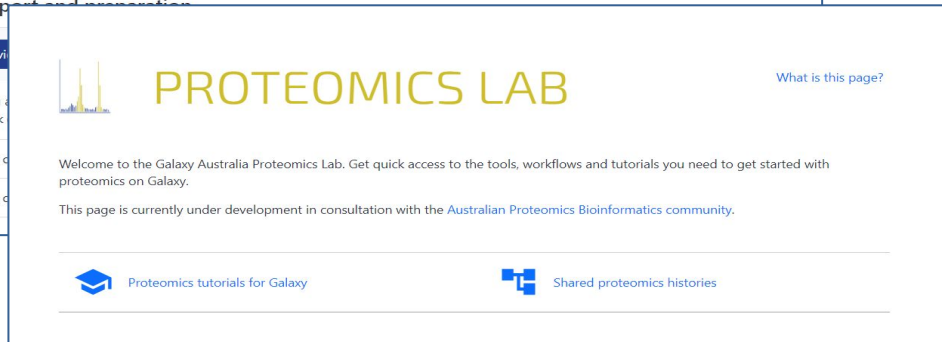
Data import and preparation

Overview

If you are a new user, please check the [Getting started](#) page.

How do I get started?

How do I use Galaxy?





PROTEOMICS LAB

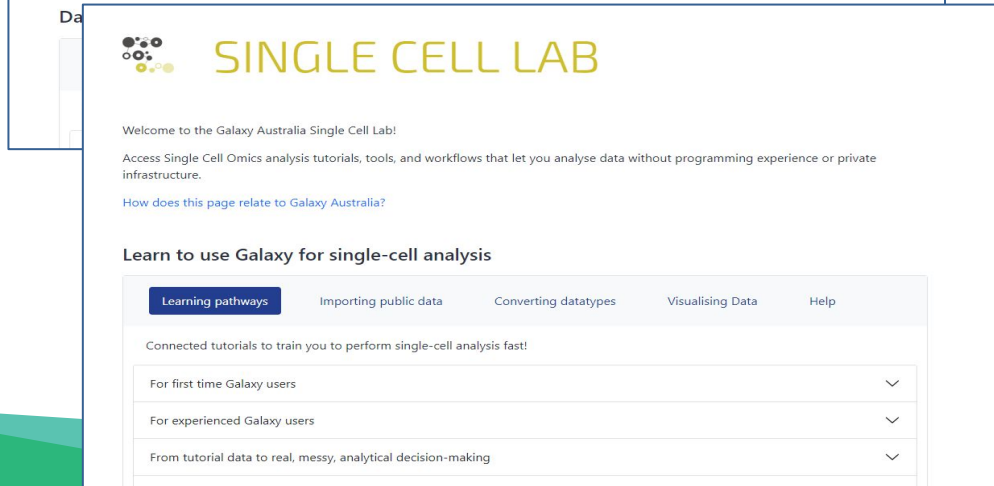
[What is this page?](#)

Welcome to the Galaxy Australia Proteomics Lab. Get quick access to the tools, workflows and tutorials you need to get started with proteomics on Galaxy.

This page is currently under development in consultation with the [Australian Proteomics Bioinformatics community](#).

 [Proteomics tutorials for Galaxy](#)

 [Shared proteomics histories](#)



SINGLE CELL LAB

Welcome to the Galaxy Australia Single Cell Lab!




Access Single Cell Omics analysis tutorials, tools, and workflows that let you analyse data without programming experience or private infrastructure.

[How does this page relate to Galaxy Australia?](#)

Learn to use Galaxy for single-cell analysis

[Learning pathways](#) [Importing public data](#) [Converting datatypes](#) [Visualising Data](#) [Help](#)

Connected tutorials to train you to perform single-cell analysis fast!

- For first time Galaxy users 
- For experienced Galaxy users 
- From tutorial data to real, messy, analytical decision-making 



High CPU

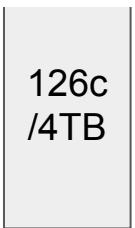
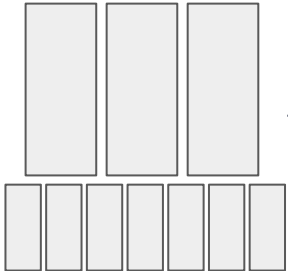
High memory

GPGPU

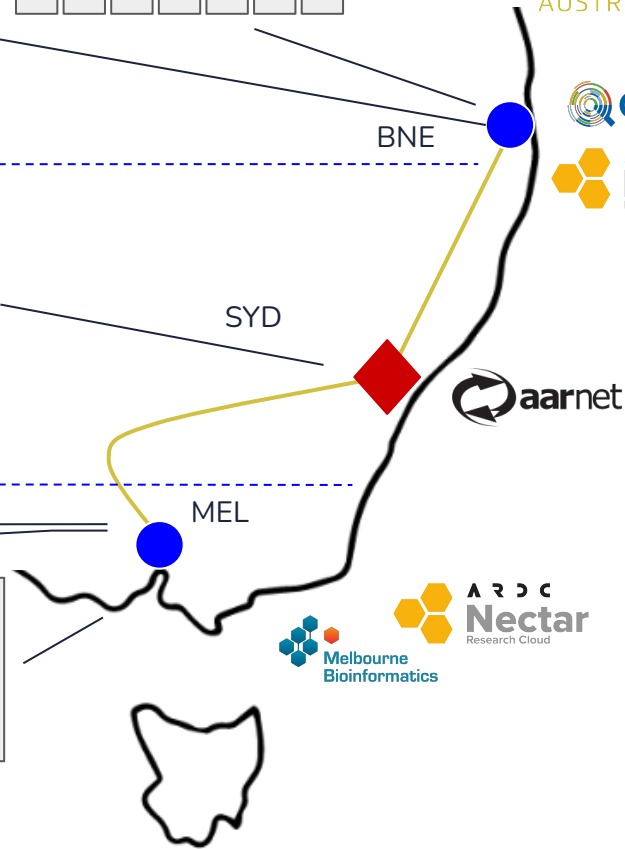
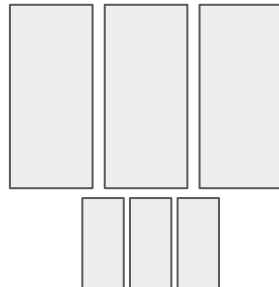
Rapid I/O

Commercial

Local slurm cluster:
7 x 16c/62GB +
3 x 32c/125GB



3 x 16c/62GB +
3 x 32c/125GB



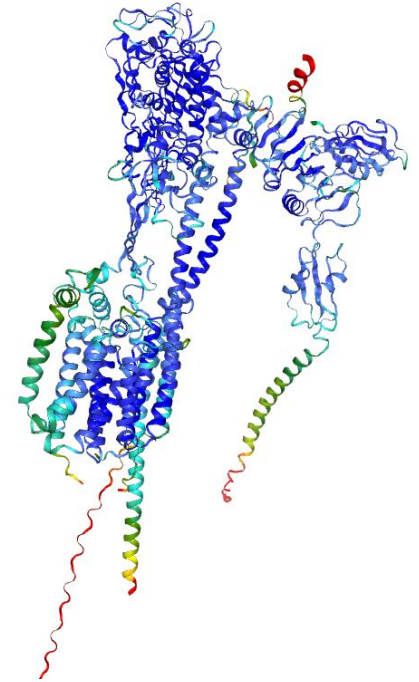
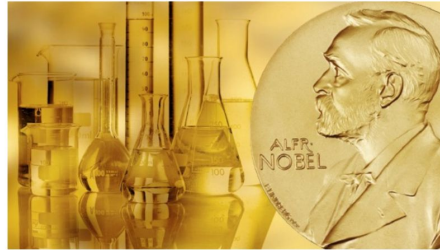
3D protein prediction in Galaxy Australia - Alphafold2

[nature](#) > collection

Collection 09 October 2024

Nobel Prize in Chemistry 2024

The 2024 Nobel Prize in Chemistry has been awarded to David Baker “for computational protein design” and to Demis Hassabis and John M. Jumper “for protein structure prediction”. Proteins are life’s essential building blocks, nature’s most ingenious molecular machines and the basis of all living organisms. Hassabis and Jumper have developed a series of artificial intelligence models to address the decades-long structural biology problem of how to predict the complex 3D structures of proteins solely from their linear amino acid sequences, while Baker has dedicated his scientific career to designing and constructing proteins that are not, and even can not, be found in nature. In recognition of this award, Nature Portfolio presents a collection of research, review and opinion articles that celebrates both contributions by the awardees and the advances they have inspired.

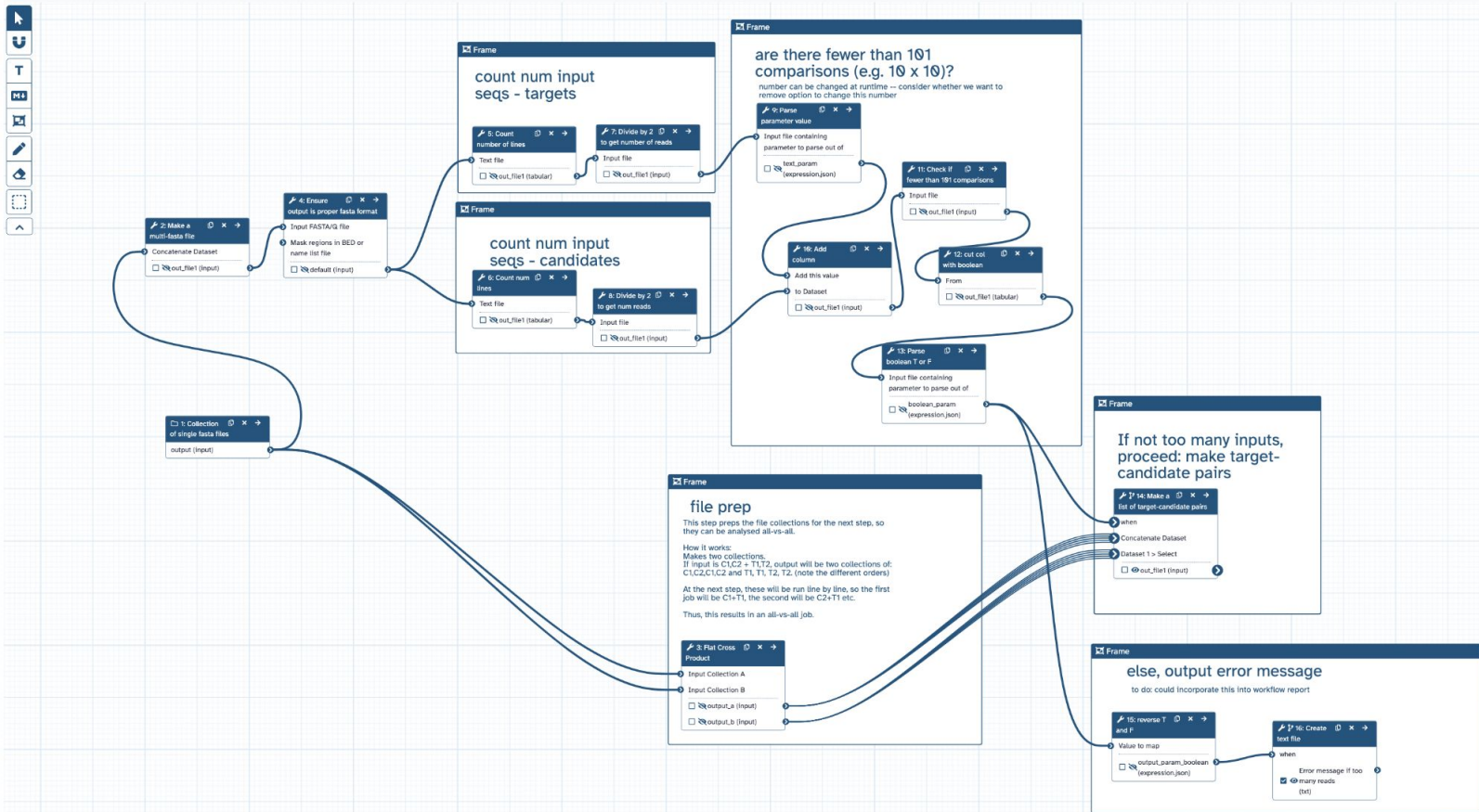


<https://www.nature.com/articles/d41586-024-03214-7>

B.subtilis_FtsL_DivIC_DivIB_PBP2B_SpoVE_complex
aa: 117, 125, 263, 713, 366

Alphafold as a screening tool

Pre-Alphafold workflow to create list of target-candidate pairs



Circos in Galaxy

Tools

circos

Show Sections

Circos visualizes data in a circular layout

Circos: Interval to Circos Text Labels reformats interval files to prepare for Circos text labels

Circos: Alignments to links reformats alignment files to prepare for Circos

Circos: Stack bigWigs as Histogram reformats for use in Circos stacked histogram plots

Circos: Table viewer easily creates circos plots from tabular data

Circos: Interval to Tiles reformats interval files to prepare for Circos tile plots

Circos: bigWig to Scatter reformats bigWig files to prepare for Circos 2d scatter/line/histogram plots

Circos: Resample 1/2D data reduce numbers of points in a dataset before plotting

Circos: Link Density Track reduce links to a density plot

Circos: Bundle Links reduce numbers of links in datasets before plotting

Circos Builder creates circos plots from standard bioinformatics datatypes.

Using Circos in Galaxy Australia

DESIGNING, MAKING AND CUSTOMIZING CIRCOS IMAGES WITH GALAXY

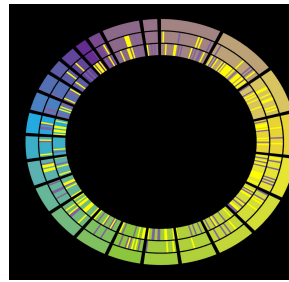
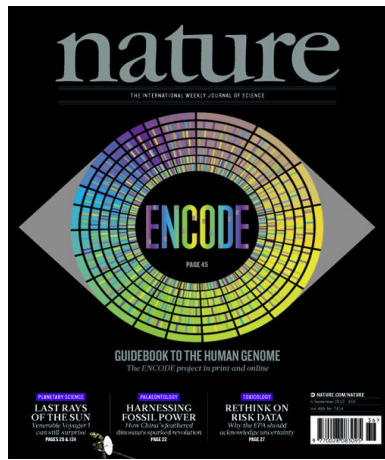
Martin Krzywinski
Canada's Michael Smith Genome Sciences Centre at BC Cancer
Vancouver, Canada

https://www.youtube.com/watch?v=j2R10doH5_4

Galaxy Training! Visualisation

Visualisation with Circos

<https://training.galaxyproject.org/training-material/topics/visualisation/tutorials/circos/tutorial.html>



Galaxy's response to that annoying virus: <https://galaxyproject.org/projects/covid19/>

PLOS PATHOGENS

OPINION

No more business as usual: Agile and effective responses to emerging pathogen threats require open data and open analytics

Abstract

The current state of much of the public-private sector is to hoard data and analysis, which is not only inefficient but also hinders the development of effective responses to emerging pathogen threats. We argue that open data and open analytics are essential for a more agile and effective response to emerging pathogen threats. We propose a set of principles for open data and open analytics, and discuss the challenges and opportunities associated with their implementation. We argue that open data and open analytics are essential for a more agile and effective response to emerging pathogen threats. We propose a set of principles for open data and open analytics, and discuss the challenges and opportunities associated with their implementation.

Timeline for 314,941 samples (Updated March 17, 2022)

The chart displays sample collection dates from January 2020 to December 2022. The y-axis lists five countries: EE (Estonia), GR (Greece), IE (Ireland), UK (United Kingdom), and ZA (South Africa). The x-axis shows months from Jan 2020 to Dec 2022. The top section, labeled 'Illumina', shows a high density of samples, particularly in the UK and ZA, with many samples collected in early 2020. The bottom section, labeled 'ONT' (Oxford Nanopore Technology), shows a lower density of samples, with a notable cluster in the UK and ZA in early 2021.

Cell

The emergence and ongoing convergent evolution of the SARS-CoV-2 N501Y lineages

Abstract

An analysis of phylogenetic and recombination events in SARS-CoV-2 sequences revealed the emergence of a major global clade in the SARS-CoV-2 lineage in late 2020. This clade is characterized by a major global shift in the SARS-CoV-2 lineage in late 2020. This clade is characterized by a major global shift in the SARS-CoV-2 lineage in late 2020. This clade is characterized by a major global shift in the SARS-CoV-2 lineage in late 2020.

Highlights

- Detected a major global shift in the SARS-CoV-2 lineage in late 2020
- Identified ongoing convergent evolution between the alpha, beta, and gamma lineages
- Defined the mutational multi-algorithm upon which these lineages are converging

Authors

Daniel P. Martin, Steven H. Whittam, Elizabeth Taylor, ... Daniel P. Martin, Steven H. Whittam, Elizabeth Taylor, ... Daniel P. Martin, Steven H. Whittam, Elizabeth Taylor, ...

Correspondence

dpmartin@cell.com

correspondence

Ready-to-use public infrastructure for global SARS-CoV-2 monitoring

Abstract

The COVID-19 pandemic has highlighted the need for a global infrastructure for monitoring SARS-CoV-2. We have developed a ready-to-use public infrastructure for global SARS-CoV-2 monitoring. This infrastructure includes a database of SARS-CoV-2 sequences, a pipeline for sequence analysis, and a web interface for data visualization. The infrastructure is designed to be scalable and flexible, allowing for the addition of new data and analysis tools as needed. The infrastructure is designed to be scalable and flexible, allowing for the addition of new data and analysis tools as needed.

The flowchart illustrates the infrastructure for global SARS-CoV-2 monitoring. It shows a central 'SARS-CoV-2 Sequences' database connected to 'Sequence Analysis' and 'Data Visualization' components. The 'Sequence Analysis' component includes 'Sequence Alignment', 'Variant Calling', and 'Phylogenetic Analysis'. The 'Data Visualization' component includes 'Sequence Viewer', 'Variant Viewer', and 'Phylogenetic Tree Viewer'. The infrastructure is designed to be scalable and flexible, allowing for the addition of new data and analysis tools as needed.

Evolution and Exceptional Conservation of ORF1a/b Overlap in Coronavirus

Abstract

The evolution and exceptional conservation of the ORF1a/b overlap in coronavirus genomes is discussed. The overlap is a region of the genome where the ORF1a and ORF1b genes overlap. This region is highly conserved across all coronavirus genomes, suggesting a functional constraint. The evolution of the overlap is discussed in the context of the overall evolution of coronavirus genomes. The overlap is a region of the genome where the ORF1a and ORF1b genes overlap. This region is highly conserved across all coronavirus genomes, suggesting a functional constraint.

Open Access

This article is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

Vertebrate Genome Project (VGP) on Galaxy

nature biotechnology

<https://doi.org/10.1038/s41587-023-02100-3>

<https://vgp.usegalaxy.org/>

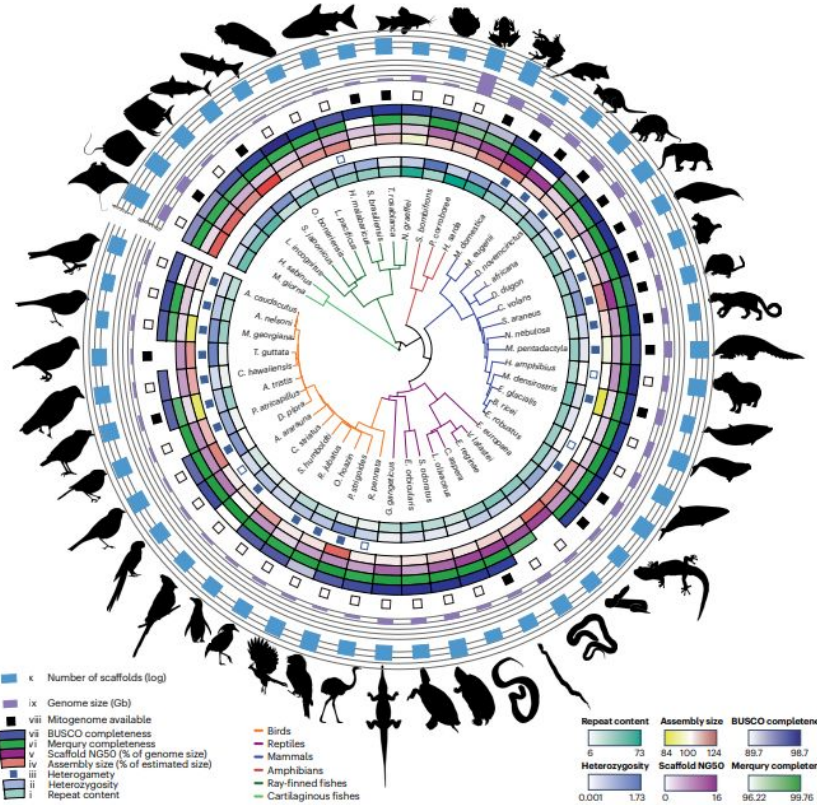
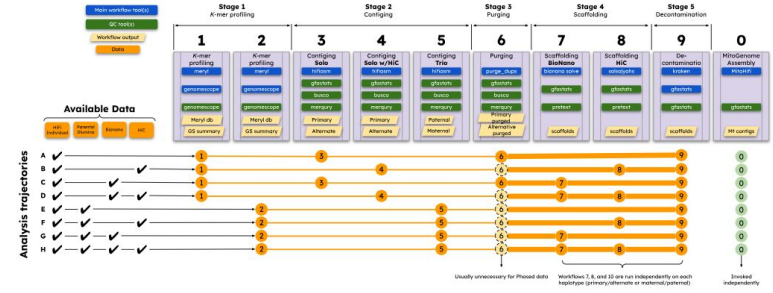
Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [nature biotechnology](#) > [correspondence](#) > [article](#)

Correspondence | Published: 26 January 2024

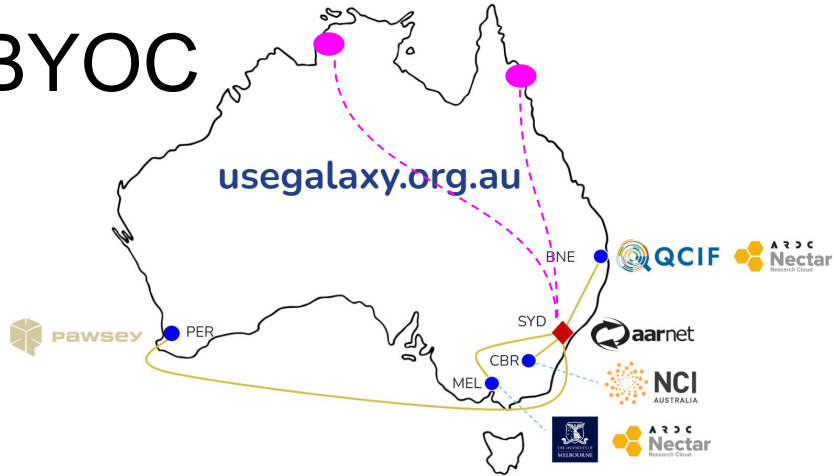
Scalable, accessible and reproducible reference genome assembly and evaluation in Galaxy

[Delphine Larivière](#), [Linelle Abueg](#), [Nadolina Brajuka](#), [Cristóbal Gallardo-Alba](#), [Bjorn Grüning](#), [Byung June Ko](#), [Alex Ostrovsky](#), [Marc Palmada-Flores](#), [Brandon D. Pickett](#), [Keon Rabbani](#), [Agostinho Antunes](#), [Jennifer R. Balacco](#), [Mark J. P. Chaisson](#), [Haoyu Cheng](#), [Joanna Collins](#), [Melanie Couture](#), [Alexandra Denisova](#), [Olivier Fedrigo](#), [Guido Roberto Gallo](#), [Alice Maria Gianì](#), [Grenville MacDonald Gooder](#), [Kathleen Horan](#), [Nivesh Jain](#), [Cassidy Johnson](#), ... [Giulio Formenti](#) ✉ [+ Show authors](#)



Bring Your Own Compute / Storage (BYOC / BYOS)

BYOC



BYOS

store_instances/create

Europe



Workflow

Visualize

Data ▾

Help ▾

User ▾



Azure Blob Storage

Amazon Web Services S3 Storage

Any S3 Compatible Storage

Google Cloud Storage

Select storage location template to create new storage location with. These templates are configured by your Galaxy administrator.

Galaxy Australia Team



Gareth Price



Catherine Bromhead



Justin Lee



Nuwan Goonasekera



Michael D'Silva



Igor Makunin



Michael Thang



Cameron Hyde



Tom Harrop



Anna Syme



A final word

Galaxy Australia Acknowledgment Statement

When you use Galaxy Australia to support your publication or project, please acknowledge its use with the following statement:

"This work is supported by Galaxy Australia, a service provided by Australian BioCommons and its partners. The service receives NCRIS funding through Bioplatforms Australia, as well as The University of Melbourne and Queensland Government RICF funding."

Please also cite the Galaxy publication:

The Galaxy Community, The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update, *Nucleic Acids Research*, Volume 52, Issue W1, 5 July 2024, Pages W83-W94,
<https://doi.org/10.1093/nar/gkae410/>

 **Galaxy** platform 2024 update published!

Read the latest developments supporting accessible, reproducible, and collaborative data analyses

With contributions from 130 authors representing 60 institutions
doi.org/10.1093/nar/gkae410



and Thanks for having us