

# Here's one we prepared earlier

(re)creating bioinformatics methods and  
workflows with Galaxy Australia

Gareth Price  
Johan Gustafsson

Galaxy Australia

Home News Events People About Support Docs

# Galaxy AUSTRALIA

★ **Alphafold 2.0**  
Galaxy Australia is now deploying Alphafold 2.0, an AI program for predicting protein structure. This tool is available to all Australian researchers from recognized institutes. [Apply now](#) to access this tool.

Galaxy Australia is an **open, web-based** platform for accessible, reproducible and transparent computational research. Galaxy supports thousands of documented and maintained tools that are free to use. We facilitate on-demand training capacities and provision **600GB** for Australian institutional (and 100GB for other) users.

History  
search datasets  
Unnamed history  
(empty)  
This history is empty. You can **load your own data** or **get data from an external source**

open, web-based platform for  
accessible, reproducible and transparent  
computational research

# A world of Galaxies

usegalaxy.\*



# Galaxy numbers

3

Main instances

160+

Total implementations

# Galaxy numbers

3

Main instances

160+

Total implementations

1,500+

Tools available on Galaxy  
Australia

8,000+

Tools available in total



# Galaxy numbers

3

Main instances

160+

Total implementations

1,500+

Tools available on Galaxy  
Australia

61,000+

Workflow invocations

4.7M+

Jobs runs

8,000+

Tools available in total

# Galaxy numbers

3

Main instances

160+

Total implementations

1,500+

Tools available on Galaxy  
Australia

61,000+

Workflow invocations

4.7M+

Jobs runs

8,000+

Tools available in total

**>600 GB**

Storage available to  
each Australian users

**11M+**

User datasets

# Galaxy numbers

3

Main instances

160+

Total implementations

1,500+

Tools available on Galaxy  
Australia

61,000+

Workflow invocations

4.7M+

Jobs runs

8,000+

Tools available in total

>600 GB

Storage available to  
each Australian users

20,000+

Users on Galaxy  
Australia

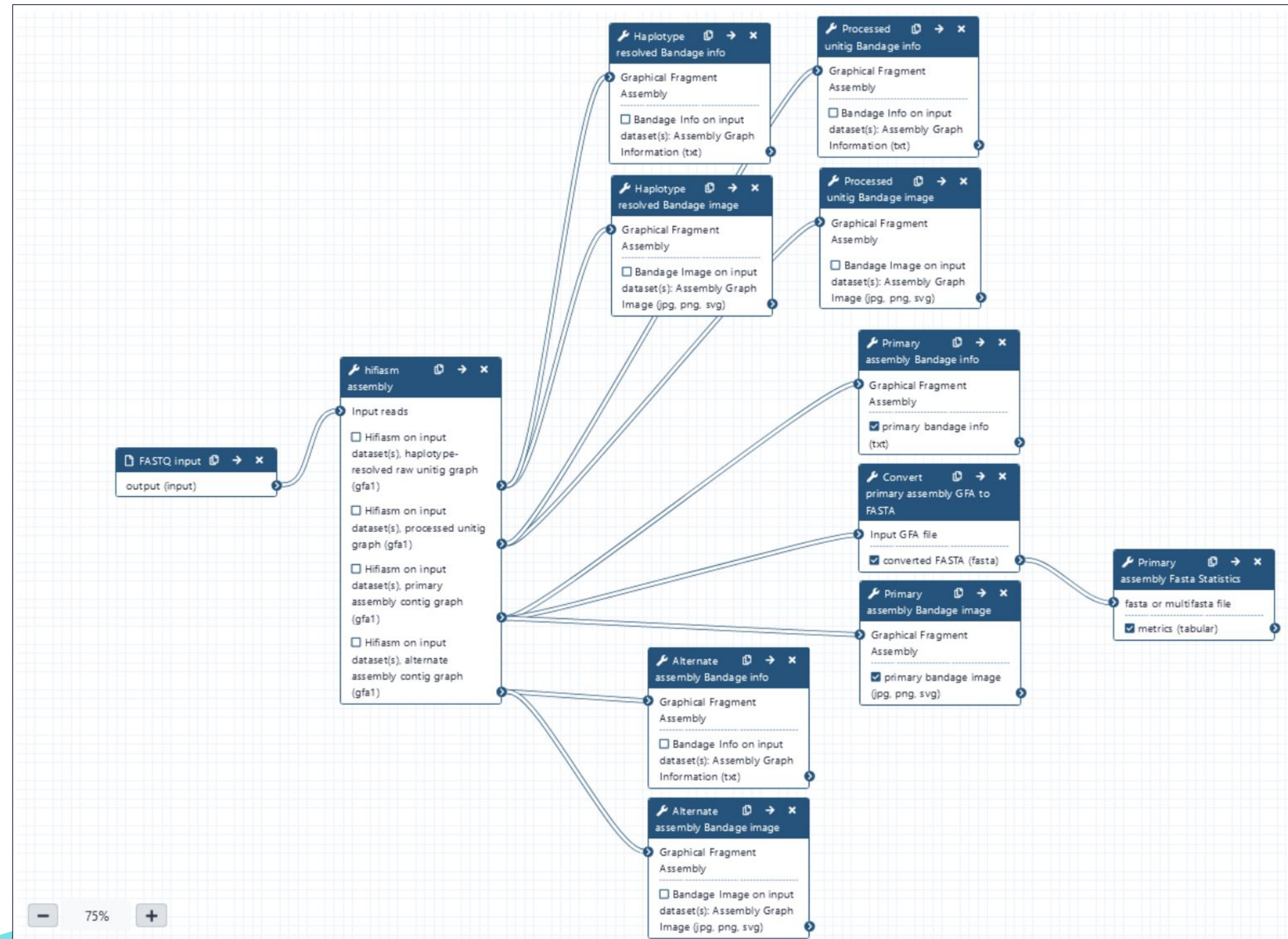
11M+

User datasets



# Workflows

# Workflows are important





# Workflows are important

321

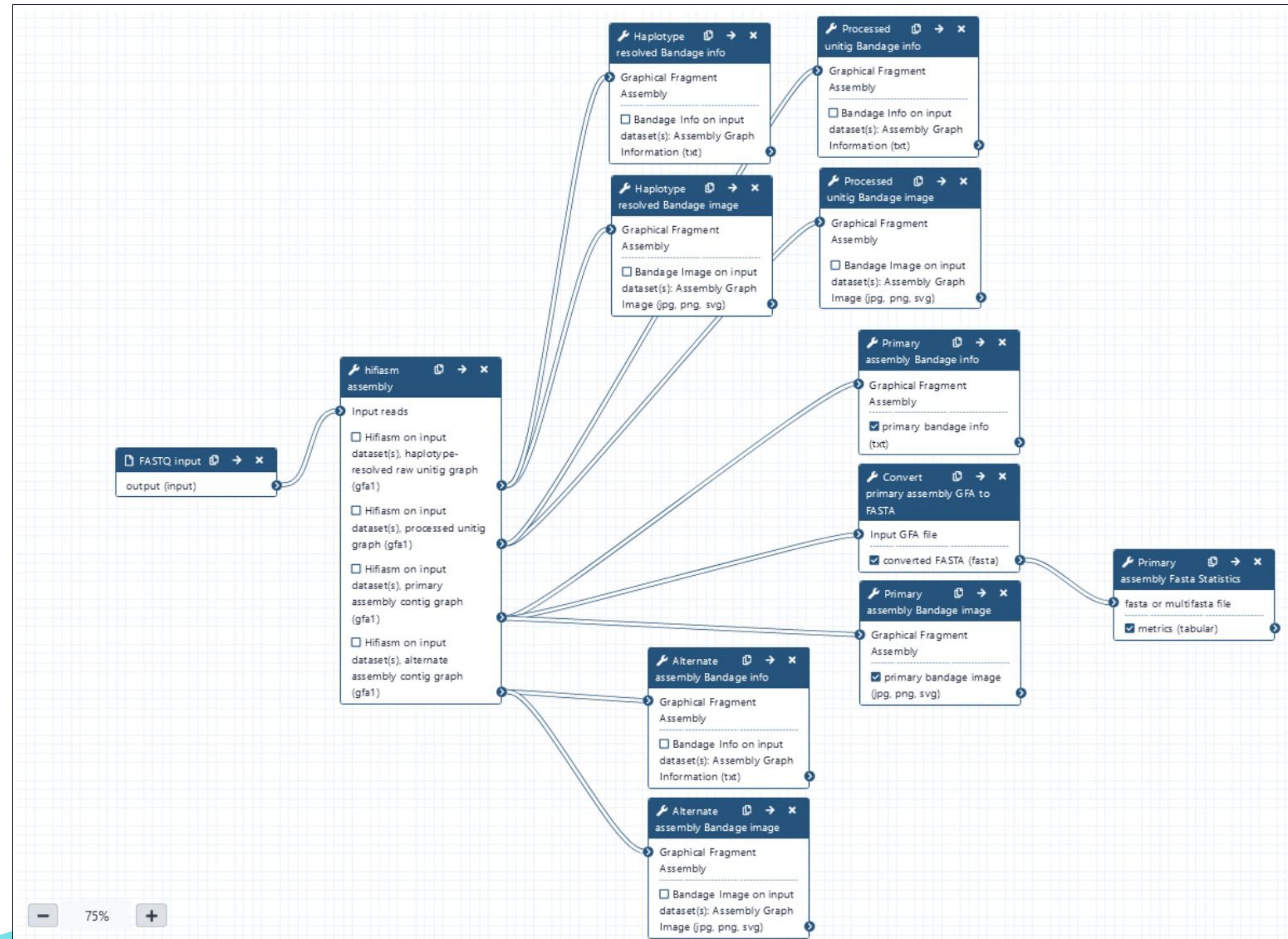
# workflow languages

<https://s.apache.org/existing-workflow-systems>

440,068

# clones for nf-core workflows

<https://nf-co.re/stats>



# Workflows exist in significant numbers

1,829

# workflows in Dockstore  
(WDL, CWL, NextFlow, Galaxy)

<https://dockstore.org/search?entryType=workflows&searchMode=files>

268

# workflows in WorkflowHub  
(all public)

<https://workflowhub.eu/workflows>

# Workflows exist in significant numbers

1,829

# workflows in Dockstore  
(WDL, CWL, NextFlow, Galaxy)

<https://dockstore.org/search?entryType=workflows&searchMode=files>

753

# of workflows on  
Galaxy Europe instance

<https://usegalaxy.eu/api/workflows>

268

# workflows in WorkflowHub  
(all public)

<https://workflowhub.eu/workflows>



# Including locally

1,829

# workflows in Dockstore  
(WDL, CWL, NextFlow, Galaxy)

<https://dockstore.org/search?entryType=workflows&searchMode=files>

753

# of workflows on  
Galaxy Europe instance

<https://usegalaxy.eu/api/workflows>

268

# workflows in WorkflowHub  
(all public)

<https://workflowhub.eu/workflows>

303

# of workflows on  
Galaxy Australia instance

<https://usegalaxy.org.au/api/workflows>



# Why should I create a workflow?

# Workflows can be ...

- ✓ Complex
- ✓ Time consuming
- ✓ Maintenance heavy
- ✓ Difficult to find
- ✓ Difficult to redeploy
- ✓ Duplicated by many

# Workflows can be ... should be ...

- ☑ Complex
- ☑ Time consuming
- ☑ Maintenance heavy
- ☑ Difficult to find
- ☑ Difficult to redeploy
- ☑ Duplicated by many

- ☑ Findable
- ☑ Reusable
- ☑ Citable

# Workflows can be ... should be ...

## ... and we would say ...

- ☑ Complex
- ☑ Time consuming
- ☑ Maintenance heavy
- ☑ Difficult to find
- ☑ Difficult to redeploy
- ☑ Duplicated by many

- ☑ Findable
- ☑ Reusable
- ☑ Citable
- ☑ Straightforward to construct
- ☑ Compute agnostic
- ☑ User friendly

# Why should I create a workflow?

Short term investment in intellect  
Long term gain in reproducibility  
Long term gain in visibility & citation



# Galaxy can be part of the solution to these challenges

The screenshot displays the Galaxy web interface for an "Unnamed workflow". On the left, a "Tools" sidebar is visible, containing a search bar and a list of tool categories and specific tools. The categories include "FILE AND META TOOLS", "GENERAL TEXT TOOLS", "GENOMIC FILE MANIPULATION", "COMMON GENOMICS TOOLS", and "GENOMICS ANALYSIS". The "GENOMICS ANALYSIS" category is currently selected, showing tools like Assembly, Annotation, Mapping, Variant Calling, ChIP-seq, RNA-seq, Multiple Alignments, Bacterial Typing, Phenotype Association, Phylogenetics, Epigenetics, and Expression Tools. The main workspace is a large grid. On the right, a metadata panel includes fields for "Name" (Unnamed workflow), "Version" (1: Oct 25th 2022, 0 steps), and "Annotation". Below these are sections for "License" (Specify a license for this workflow), "Creator" (Add a new creator - either a person or an organization), and "Tags" (Apply tags to make it easy to search for and find items with the same tag).



# Galaxy can be part of the solution to these challenges

Going from idea to successful workflow isn't necessarily intuitive

The screenshot displays the Galaxy web interface for creating a workflow. On the left, the 'Tools' panel is visible, featuring a search bar and a list of tool categories: 'FILE AND META TOOLS', 'GENERAL TEXT TOOLS', 'GENOMIC FILE MANIPULATION', and 'GENOMICS ANALYSIS'. The main workspace is a grid for building a workflow. On the right, a metadata panel shows fields for Name, Version, Annotation, License, Creator, and Tags.

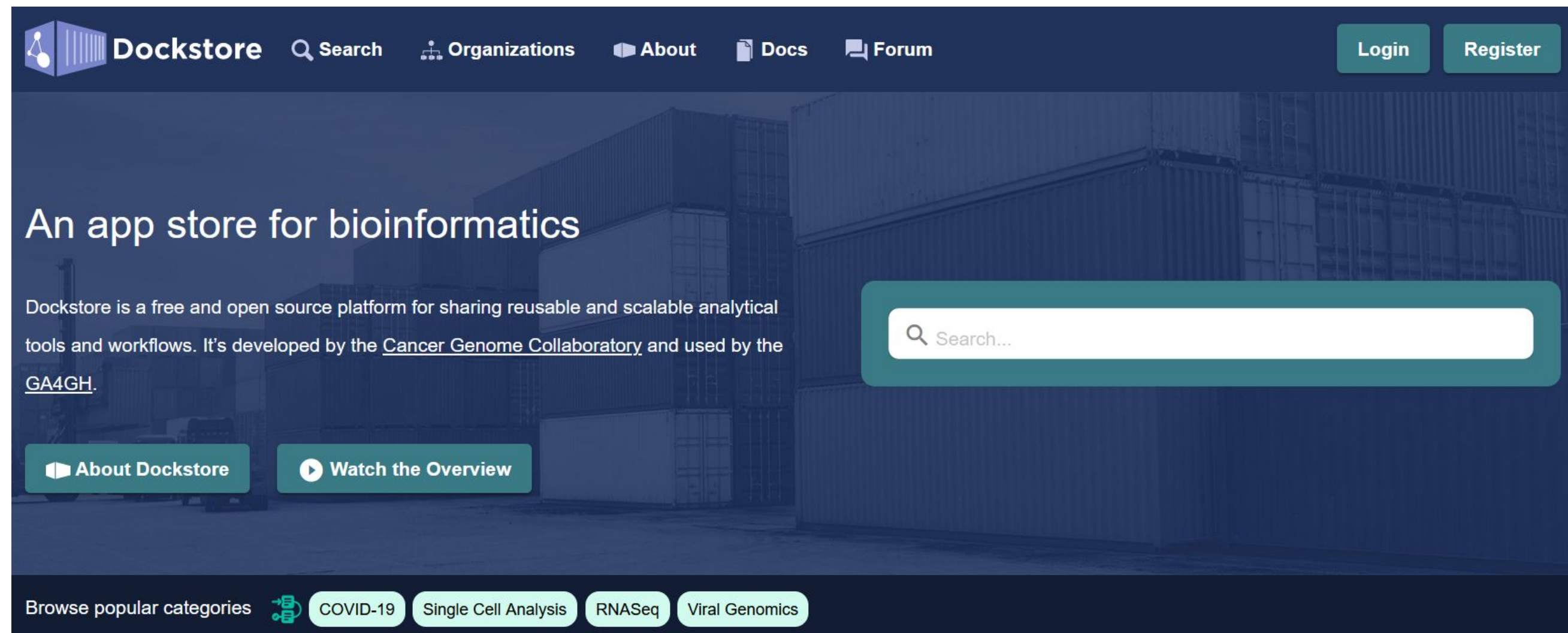
# How do I get started?

Option 1 might be  
trying to find a  
workflow to reuse



# Two key international options

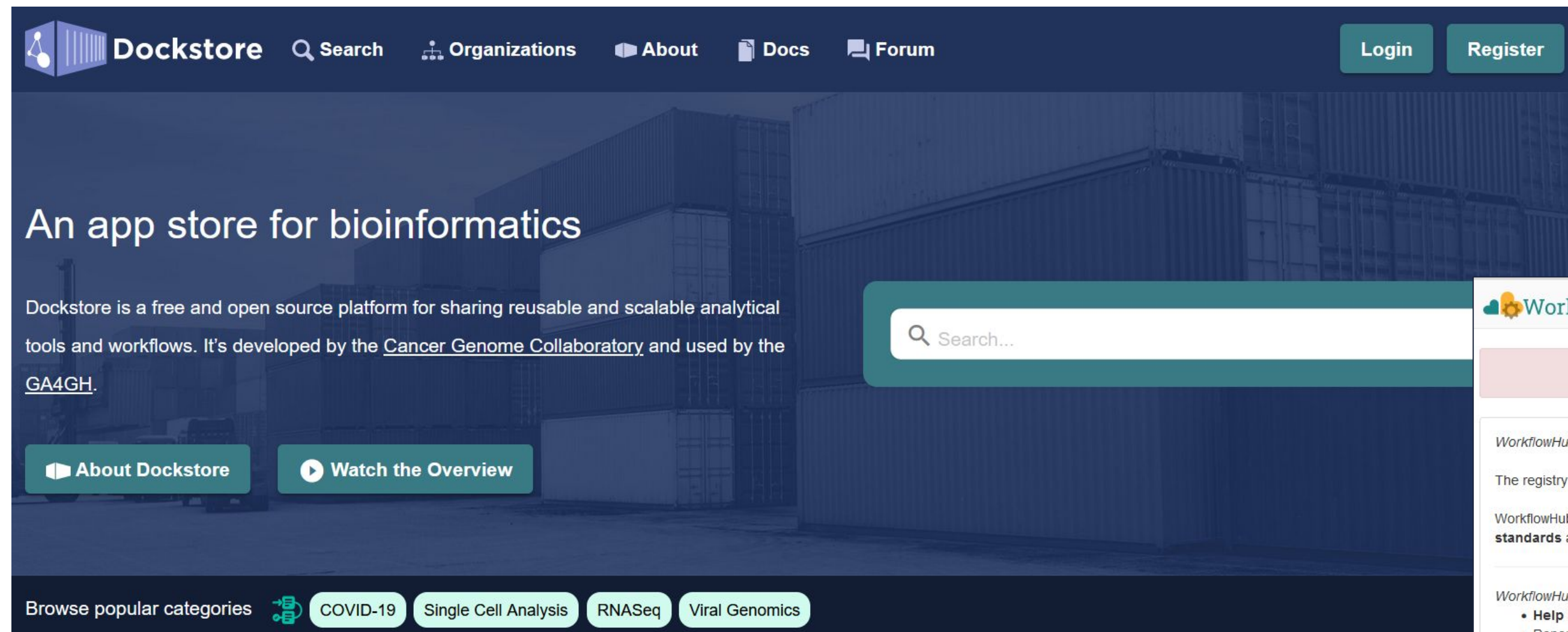
## Dockstore



<https://dockstore.org/>

# Two key international options

## Dockstore

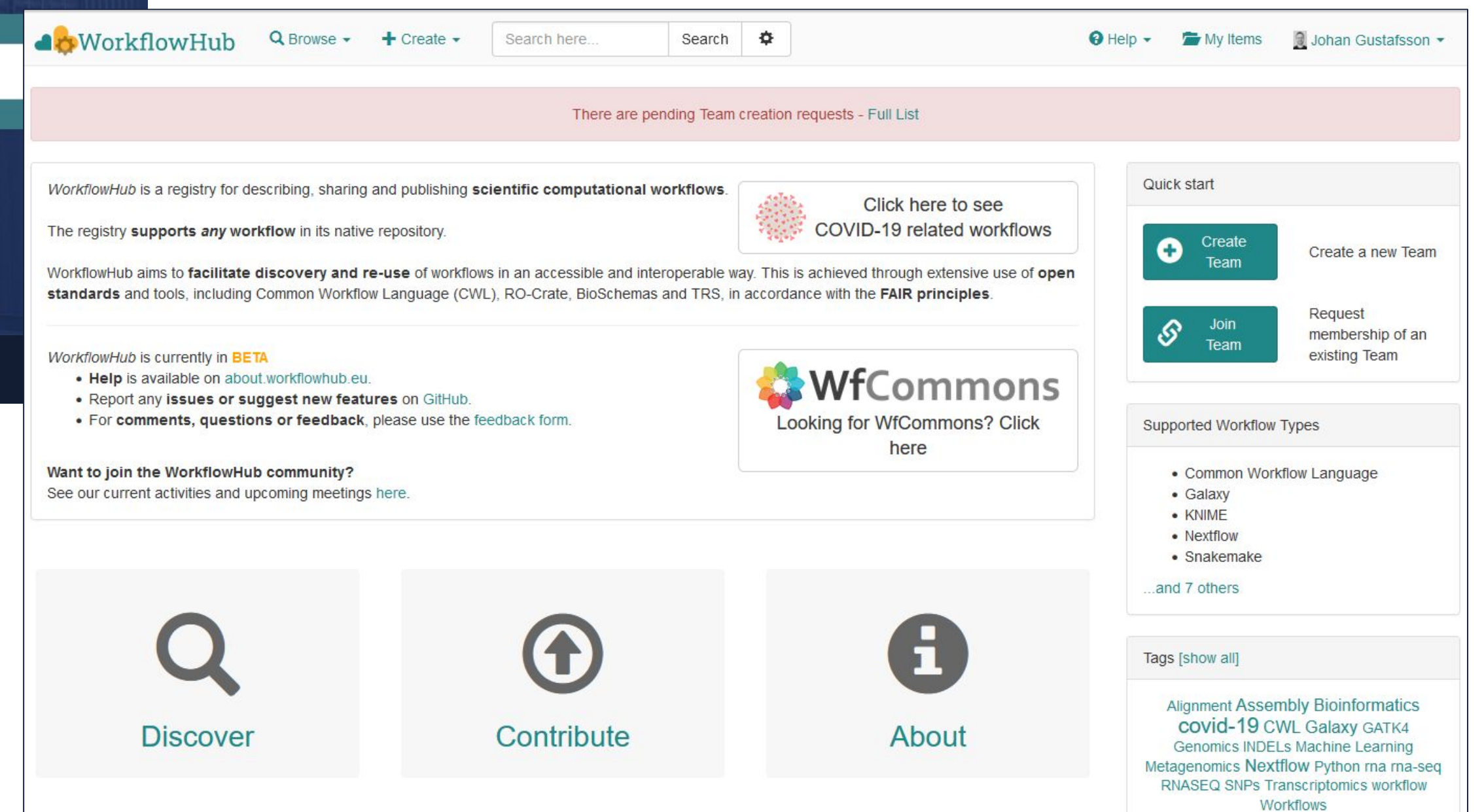


The screenshot shows the Dockstore website homepage. At the top, there is a navigation bar with the Dockstore logo, a search bar, and links for Organizations, About, Docs, and Forum. There are also Login and Register buttons. The main content area features the headline "An app store for bioinformatics" and a sub-headline: "Dockstore is a free and open source platform for sharing reusable and scalable analytical tools and workflows. It's developed by the Cancer Genome Collaboratory and used by the GA4GH." Below this, there are two buttons: "About Dockstore" and "Watch the Overview". At the bottom, there is a section for "Browse popular categories" with tags for COVID-19, Single Cell Analysis, RNASeq, and Viral Genomics.

<https://dockstore.org/>

<https://about.workflowhub.eu/docs/getting-started/>

## WorkflowHub









The screenshot shows the WorkflowHub website homepage. At the top, there is a navigation bar with the WorkflowHub logo, a search bar, and links for Browse, Create, and Help. There is also a user profile for Johan Gustafsson. A notification banner at the top right says "There are pending Team creation requests - Full List". The main content area features the headline "WorkflowHub is a registry for describing, sharing and publishing scientific computational workflows." and a sub-headline: "The registry supports any workflow in its native repository." Below this, there is a section for "WorkflowHub is currently in BETA" with a list of links: "Help is available on about.workflowhub.eu.", "Report any issues or suggest new features on GitHub.", and "For comments, questions or feedback, please use the feedback form." There is also a "WfCommons" logo and a link "Looking for WfCommons? Click here". At the bottom, there are three buttons: "Discover", "Contribute", and "About". On the right side, there is a "Quick start" section with "Create Team" and "Join Team" buttons, and a "Supported Workflow Types" section with a list of workflow types: Common Workflow Language, Galaxy, KNIME, Nextflow, and Snakemake, followed by "...and 7 others". There is also a "Tags [show all]" section with a list of tags: Alignment Assembly Bioinformatics, covid-19 CWL Galaxy GATK4, Genomics INDELS Machine Learning, Metagenomics Nextflow Python rna rna-seq, RNASEQ SNPs Transcriptomics workflow, and Workflows.

<https://workflowhub.eu/>



# WorkflowFinder

Copy CSV Search:

Workflow information					Citation & credits	Deployment		
Workflow name (+ links)	Topic(s)	Operation(s)	License	Updated	DOI	Contributing team(s)	User guide	Open workflow (if available)
<b>Assembly polishing</b>  WorkflowHub	Sequence assembly	Sequence assembly	Apache-2.0	2022-09-16	<a href="https://doi.org/10.48546/workflowhub.workflow.226.1">10.48546/workflowhub.workflow.226.1</a>	Galaxy Australia Australian BioCommons	<a href="#">See user guide</a>	<a href="#">Open workflow on Galaxy Australia</a>
<b>Assembly with Flye</b>  WorkflowHub	Sequence assembly	De-novo assembly	Apache-2.0	2022-09-16	<a href="https://doi.org/10.48546/workflowhub.workflow.225.1">10.48546/workflowhub.workflow.225.1</a>	Galaxy Australia Australian BioCommons	<a href="#">See user guide</a>	<a href="#">Open workflow on Galaxy Australia</a>
<b>Assess genome quality</b>  WorkflowHub	Sequence assembly		Apache-2.0	2022-09-16	<a href="https://doi.org/10.48546/workflowhub.workflow.229.1">10.48546/workflowhub.workflow.229.1</a>	Galaxy Australia Australian BioCommons	<a href="#">See user guide</a>	<a href="#">Open workflow on Galaxy Australia</a>
<b>BAM to FASTQ + QC v1.0</b>  WorkflowHub	Sequence assembly	Conversion Sequencing quality control	GPL-3.0	2022-10-17	<a href="https://doi.org/10.48546/workflowhub.workflow.220.2">10.48546/workflowhub.workflow.220.2</a>	Australian BioCommons Galaxy Australia	<a href="#">See user guide</a>	<a href="#">Open workflow on Galaxy Australia</a>
<b>Combined workflows for large genome assembly</b>  WorkflowHub			Apache-2.0	2022-09-16	<a href="https://doi.org/10.48546/workflowhub.workflow.230.1">10.48546/workflowhub.workflow.230.1</a>	Galaxy Australia Australian BioCommons	<a href="#">See user guide</a>	<a href="#">Open workflow on Galaxy Australia</a>
<b>Data QC</b>  WorkflowHub	Sequence assembly	Sequencing quality control	Apache-2.0	2022-09-16	<a href="https://doi.org/10.48546/workflowhub.workflow.222.1">10.48546/workflowhub.workflow.222.1</a>	Galaxy Australia Australian BioCommons	<a href="#">See user guide</a>	<a href="#">Open workflow on Galaxy Australia</a>

[https://australianbiocommons.github.io/2\\_1\\_workflows.html](https://australianbiocommons.github.io/2_1_workflows.html)



# From within Galaxy Australia Tool Registry Server

GA4GH Tool Registry Server (TRS) Workflow Search

TRS Server: [workflowhub.eu](#)

hifi genome assembly

? x

Name	Description	Organization
PacBio HiFi genome assembly using hifiasm v2.1	# PacBio HiFi genome assembly using hifiasm v2.1 ## General usage recommendations Please see the [Genome assembly with hifiasm on Galaxy Australia](https://australianbiocommons.github.io/how-to-guides/genome_assembly/hifi_assembly) guide. ## See [change log](./change_log.md) ## Acknowledgements The workflow & the [doc_guidelines template used](https://github.com/AustralianBioCommons/doc_guidelines) are supported by the Australian BioCommons via Bioplatforms Australia funding, the Australian Research Data Commons ( <a href="https://doi.org/10.47486/PL105">https://doi.org/10.47486/PL105</a> ) and the Queensland Government RICF programme. Bioplatforms Australia and the Australian Research Data Commons are enabled by the National Collaborative Research Infrastructure Strategy (NCRIS).	

**Name:** PacBio HiFi genome assembly using hifiasm v2.1

**Description:** # PacBio HiFi genome assembly using hifiasm v2.1 ## General usage recommendations Please see the [Genome assembly with hifiasm on Galaxy Australia](https://australianbiocommons.github.io/how-to-guides/genome\_assembly/hifi\_assembly) guide. ## See [change log](./change\_log.md) ## Acknowledgements The workflow & the [doc\_guidelines template used](https://github.com/AustralianBioCommons/doc\_guidelines) are supported by the Australian BioCommons via Bioplatforms Australia funding, the Australian Research Data Commons (<https://doi.org/10.47486/PL105>) and the Queensland Government RICF programme. Bioplatforms Australia and the Australian Research Data Commons are enabled by the National Collaborative Research Infrastructure Strategy (NCRIS).

**Organization:** Australian BioCommons, Galaxy Australia

**Versions**

- Version 1
- master @ c4797d8
- v2.1.0

# From within Galaxy Australia Tool Registry Server

GA4GH Tool Registry Server (TRS) Workflow Search

TRS Server: **workflowhub.eu**

hifi genome assembly

? x




Name	Description	Organization
PacBio HiFi genome assembly using hifiasm v2.1	# PacBio HiFi genome assembly using hifiasm v2.1 ## General usage recommendations Please see the [Genome assembly with hifiasm on Galaxy Australia](https://australianbiocommons.github.io/how-to-guides/genome_assembly/hifi_assembly) guide. ## See [change log](./change_log.md) ## Acknowledgements The workflow & the [doc_guidelines template used](https://github.com/AustralianBioCommons/doc_guidelines) are supported by the Australian BioCommons via Bioplatforms Australia funding, the Australian Research Data Commons (https://doi.org/10.47486/PL105) and the Queensland Government RICF programme. Bioplatforms Australia and the Australian Research Data Commons are enabled by the National Collaborative Research Infrastructure Strategy (NCRIS).	

**Name:** PacBio HiFi genome assembly using hifiasm v2.1

**Description:** # PacBio HiFi genome assembly using hifiasm v2.1 ## General usage recommendations Please see the [Genome assembly with hifiasm on Galaxy Australia](https://australianbiocommons.github.io/how-to-guides/genome\_assembly/hifi\_assembly) guide. ## See [change log](./change\_log.md) ## Acknowledgements The workflow & the [doc\_guidelines template used](https://github.com/AustralianBioCommons/doc\_guidelines) are supported by the Australian BioCommons via Bioplatforms Australia funding, the Australian Research Data Commons (https://doi.org/10.47486/PL105) and the Queensland Government RICF programme. Bioplatforms Australia and the Australian Research Data Commons are enabled by the National Collaborative Research Infrastructure Strategy (NCRIS).

**Organization:** Australian BioCommons, Galaxy Australia

**Versions**

- Version 1 
- master @ c4797d8 
- v2.1.0 



# From within Galaxy Australia Tool Registry Server

GA4GH Tool Registry Server (TRS) Workflow Search

TRS Server: **workflowhub.eu**

hifi genome assembly

? x

Name	Description	Organization
PacBio HiFi genome assembly using hifiasm v2.1	# PacBio HiFi genome assembly using hifiasm v2.1 ## General usage recommendations Please see the [Genome assembly with hifiasm on Galaxy Australia](https://australianbiocommons.github.io/how-to-guides/genome_assembly/hifi_assembly) guide. ## See [change log](./change_log.md) ## Acknowledgements The workflow & the [doc_guidelines template used](https://github.com/AustralianBioCommons/doc_guidelines) are supported by the Australian BioCommons via Bioplatforms Australia funding, the Australian Research Data Commons (https://doi.org/10.47486/PL105) and the Queensland Government RICF programme. Bioplatforms Australia and the Australian Research Data Commons are enabled by the National Collaborative Research Infrastructure Strategy (NCRIS).	

**Name:** PacBio HiFi genome assembly using hifiasm v2.1

**Description:** # PacBio HiFi genome assembly using hifiasm v2.1 ## General usage recommendations Please see the [Genome assembly with hifiasm on Galaxy Australia](https://australianbiocommons.github.io/how-to-guides/genome\_assembly/hifi\_assembly) guide. ## See [change log](./change\_log.md) ## Acknowledgements The workflow & the [doc\_guidelines template used](https://github.com/AustralianBioCommons/doc\_guidelines) are supported by the Australian BioCommons via Bioplatforms Australia funding, the Australian Research Data Commons (https://doi.org/10.47486/PL105) and the Queensland Government RICF programme. Bioplatforms Australia and the Australian Research Data Commons are enabled by the National Collaborative Research Infrastructure Strategy (NCRIS).

**Organization:** Australian BioCommons, Galaxy Australia

**Versions**

- Version 1
- master @ c4797d8
- v2.1.0

# Option 2

## What if you need to *create a workflow?*

# Using Galaxy to build workflows

1

*De novo*

Add tools sequentially to  
create a workflow

# Using Galaxy to build workflows

1

## *De novo*

Add tools sequentially to  
create a workflow

2

## From a history

Use a set of tools  
sequentially

Use history to create,  
curate and finalise a  
workflow file



# Using Galaxy to build workflows

1

## *De novo*

Add tools sequentially to create a workflow

2

## From a history

Use a set of tools sequentially

Use history to create, curate and finalise a workflow file

3

## From a publication

Read publication

Find / request tools for Galaxy

Create workflow using *de novo* build approach

# Using Galaxy to build workflows

1

## *De novo*

Add tools sequentially to create a workflow

Paint the sistine chapel from scratch

2

## From a history

Use a set of tools sequentially

Use history to create, curate and finalise a workflow file

“Re-paint” the sistine chapel from previous experience

3

## From a publication

Read publication

Find / request tools for Galaxy

Create workflow using de novo build approach

Paint the sistine chapel based on someone else painting an equivalent chapel

# Time for some demos!

Now for some final  
suggestions

# Follow workflow best practices

**Best Practices Review**

- ⚠️ This workflow is not annotated. Providing an annotation helps workflow executors understand the purpose and usage of the workflow.  
✎ Annotate your Workflow.
- ⚠️ This workflow does not specify creator(s). This is important metadata for workflows that will be published and/or shared to help workflow executors know how to cite the workflow authors.  
✎ Provide Creator Details.
- ⚠️ This workflow does not specify a license. This is important metadata for workflows that will be published and/or shared to help workflow executors understand how it may be used.  
✎ Specify a License.
- ✅ Workflow parameters are using formal input parameters.
- ✅ All non-optional inputs to workflow steps are connected to formal input parameters.
- ✅ All workflow inputs have labels and annotations.
- ✅ This workflow has outputs and they all have valid labels.
- ⚠️ This workflow has no labeled outputs, please select and label at least one output.

Annotations

Creators

License

Parameters

Output labels



# Follow workflow best practices

Publish your workflows

**Best Practices Review**

- ⚠️ This workflow is not annotated. Providing an annotation helps workflow executors understand the purpose and usage of the workflow.
  - Annotate your Workflow.
- ⚠️ This workflow does not specify creator(s). This is important metadata for workflows that will be published and/or shared to help workflow executors know how to cite the workflow authors.
  - Provide Creator Details.
- ⚠️ This workflow does not specify a license. This is important metadata for workflows that will be published and/or shared to help workflow executors understand how it may be used.
  - Specify a License.
- ✅ Workflow parameters are using formal input parameters.
- ✅ All non-optional inputs to workflow steps are connected to formal input parameters.
- ✅ All workflow inputs have labels and annotations.
- ✅ This workflow has outputs and they all have valid labels.
- ⚠️ This workflow has no labeled outputs, please select and label at least one output.

Annotations

Creators

License

Parameters

Output labels

Published Workflows

search name, annotation, owner, and

Advanced Search

Name	Annotation	Owner	Community Rating	Community Tags	Last Updated
exploring iris dataset with stats and plots		delaney_burnard	★★★★★		Aug 18, 2022
VGP assembly: training workflow (imported from GTN materials August 4 2022)	VGP assembly tutorial	anna	★★★★★	assembly	Aug 04, 2022
PacBio HiFi genome assembly using hifiasm v1.0	Assembly, visualisation and quality control workflow for high fidelity reads built from circular consensus sequence (PacBio HiFi) data.	johan	★★★★★	fastq genome_assembly hifi hifiasm	Jul 19, 2022
Generic Assembly	This workflow has been designed to take paired end reads from an Illumina platform, perform some QC and then assemble them into contigs. It is mainly for...	simongladman	★★★★★	assembly bacterial	Jul 18, 2022
GA-virReport-Stats (Public)	Downstream workflow for positive detections from GA-virReport to compute mapping statistics to a reference genome and to generate reference guided genome...	lelwala	★★★★★	ga-virreport-confidence	Jun 28, 2022
GA-VirReport-Single_input_files (Public)	Bioinformatics workflow for detecting viruses and viroids from host plant small RNA sequencing data	lelwala	★★★★★	ga-virreport	Jun 28, 2022
GA-VirReport-Multiple_input_files_(Public)	Bioinformatics workflow for detecting viruses and viroids from host plant small RNA sequencing data	lelwala	★★★★★	ga-virreport	Jun 28, 2022
Generic consensus construction from VCF calls	Build a consensus sequence from FILTER PASS variants with intrasample allele-frequency above a configurable consensus threshold. Hard-mask regions with...	galaxy-australia	★★★★★	generic mls	Jun 06, 2022
Genetic variation analysis reporting	This workflow takes a VCF dataset of variants produced by any of the variant calling workflows in...	galaxy-australia	★★★★★	generic mpvx	Jun 06, 2022
Generic variation analysis on WGS PE data	This workflows performs paired end read mapping with bwa-mem followed by sensitive variant calling across a wide range of AFs with lofreq	galaxy-australia	★★★★★	mpvx generic	Jun 06, 2022
Di Marzio Workflow		dimarziogabriel	★★★★★		Jun 04, 2022
Stacks RAD-seq de novo workflow	Version 1	anna	★★★★★		May 31, 2022
Stacks RAD-seq reference-guided workflow	Version 1	anna	★★★★★		May 31, 2022
Partial ref-guided workflow - bwa mem only	Version 1	anna	★★★★★		May 19, 2022
Partial ref-guided workflow - gstacks and pops	Version 1	anna	★★★★★		May 18, 2022
Partial de novo workflow: c-s-g-pops only	Version 1	anna	★★★★★		May 18, 2022
Partial de novo workflow: ustacks only	Version 1	anna	★★★★★		May 18, 2022
QC of RADseq reads	Version 1	anna	★★★★★		May 18, 2022



# Register your workflows

WorkflowHub

Browse Create Search here... Search

## New Workflow

Currently there are three ways to register a workflow:

- **Upload/Import Files** - upload a workflow file (or import from elsewhere on the web), along with an optional diagram and CWL description if you have one
- **Import Git Repository** - import a Git repository from GitHub, GitLab or any other publicly accessible git repository on the web.
- **Upload/Import Workflow RO-Crate** - upload or import a *Workflow RO-Crate* (Note: this format is under development, click [here](#) to find out more)

Currently supported workflow types are: Bpipe, Common Workflow Language, Galaxy, Janis, Jupyter, KNIME, Nextflow, PyCOMPSs, Python, Scipion, Shell Script, Snakemake.

**Upload/Import Files**

Import Git Repository

Upload/Import Workflow RO-Crate

**Workflow\***

The main executable workflow.

Local file Remote URL

Browse... No file selected.

**Workflow Type\*** New workflow type

Other

The type of the above workflow.

**Abstract CWL**

(Optional) The abstract CWL that describes the workflow above.

Local file Remote URL

Browse... No file selected.

**Diagram**

(Optional) A diagram that illustrates the main workflow.

Local file Remote URL

Browse... No file selected.

Register or Cancel

WorkflowHub

Home About Documentation GitHub Search WorkflowHub pr

## How to get started with WorkflowHub

- 1. Register yourself on WorkflowHub
- 2. Decide which space you would like to create a team in
- 3. Create a team or join an existing team
- 4. Register workflow(s)

WorkflowHub will help make your workflows findable and citable, and this guide will help you get started with that process.

**Tip:** if you want to test WorkflowHub and its features, there is a development instance available here.

### 1. Register yourself on WorkflowHub

Register yourself on the sign up page of WorkflowHub if you do not have an account. You can also log in using your LS Login or GitHub credentials.

### 2. Decide which space you would like to create a team in

Figure 1 below highlights the basic structure of WorkflowHub, and examples of the three types of spaces where you could create a team.

You can request to create a team in any or all of the following spaces:

- **Independent teams:** this is where you can create a team if you do not require a specific space
- **A pre-existing space,** for example:
  - Australian BioCommons
  - BioExcel
  - EOSC Life



# Reference your workflows

## Purge duplicates from hifiasm assembly v1.0 v1.0.0 (latest)

[View on GitHub](#)[Request Contact](#)[Unsubscribe](#)[Download RO Crate](#)[Run on usegalaxy.eu](#)[Add new](#)[Actions](#)[Overview](#)[Files](#)[Related items](#)

Workflow Type: Galaxy

Stable

### Purge-duplicates-from-hifiasm-assembly

General recommendations for using [Purge-duplicates-from-hifiasm-assembly](#)

Please see the [Genome assembly with hifiasm on Galaxy Australia](#) guide.

#### Acknowledgements

The workflow & the [doc\\_guidelines template used](#) are supported by the Australian BioCommons via Bioplatforms Australia funding, the Australian Research Data Commons (<https://doi.org/10.47486/PL105>) and the Queensland Government RICF programme. Bioplatforms Australia and the Australian Research Data Commons are enabled by the National Collaborative Research Infrastructure Strategy (NCRIS).

SEEK ID: <https://workflowhub.eu/workflows/237?version=2>

DOI: [10.48546/workflowhub.workflow.237.2](https://doi.org/10.48546/workflowhub.workflow.237.2)

## Inputs

ID	Name	Description	Type
HiFiASM 1o assembly	HiFiASM 1o assembly	n/a	n/a
HiFi reads as FASTQ	HiFi reads as FASTQ	n/a	n/a

#### Creators and Submitter

##### Creator

[Gareth Price](#)

##### Additional credit

Gareth Price

##### Submitter

[Johan Gustafsson](#)

#### Citation

[Copy](#)

Price, G. (2022). *Purge duplicates from hifiasm assembly v1.0*. WorkflowHub. <https://doi.org/10.48546> /WORKFLOWHUB.WORKFLOW.237.2

American Psychological Association 7th

#### License

GNU General Public License 3.0

#### Citation

[Copy](#)

Price, G. (2022). *Purge duplicates from hifiasm assembly v1.0*. WorkflowHub. <https://doi.org/10.48546> /WORKFLOWHUB.WORKFLOW.237.2



# Staying Connected through Community

## Welcome to the Galaxy Community Hub

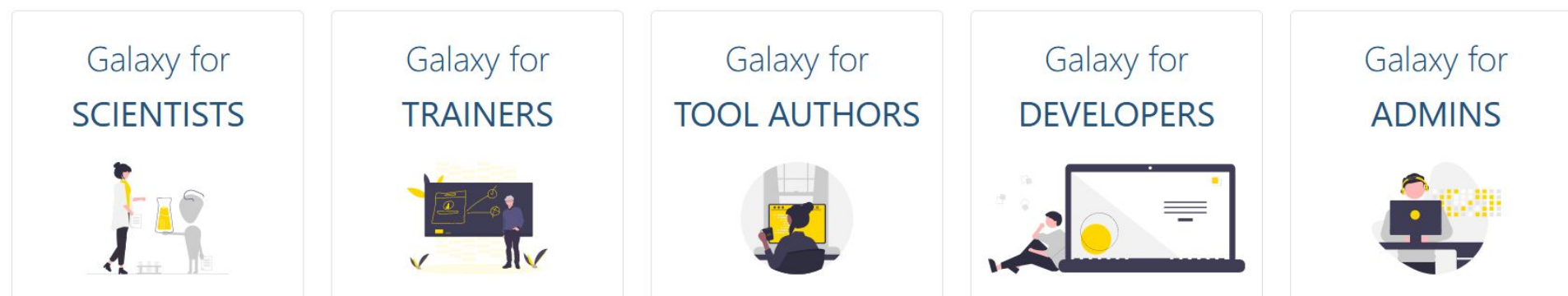
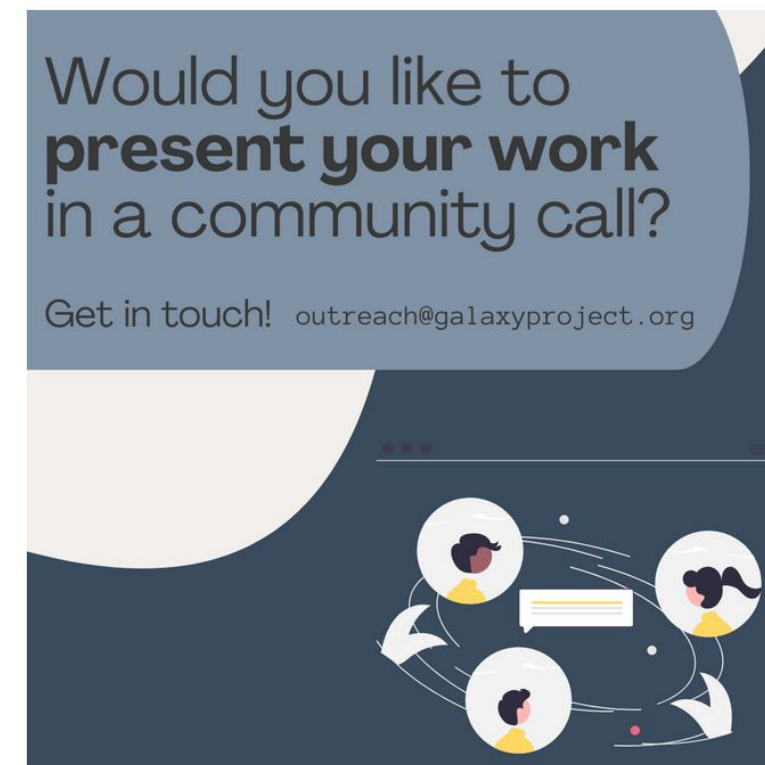
The meeting point where you can find curated documentation for all things Galaxy

Galaxy is an **open-source** platform for **FAIR data analysis** that enables users to:

- Use **tools** from various domains (that can be plugged into **workflows**) through its graphical web interface.
- Run code in **interactive environments** (RStudio, Jupyter...) along with other tools or workflows.
- **Manage data** by sharing and publishing results, workflows, and visualizations.
- **Ensure reproducibility** by capturing the necessary information to repeat and understand data analyses.

The **Galaxy Community** is actively involved in helping the ecosystem improve and sharing scientific discoveries.

Get Started: First Steps with Galaxy



<https://galaxyproject.org>

### Bioinformatics workflows

As the scale and complexity of biological data continues to grow, computational workflows can facilitate high throughput analysis by allowing researchers to automate various tedious and repetitive tasks such as managing data, running different tools, managing intermediate files, and dealing with computational job schedulers.

Various modern workflow management systems and languages (e.g. NextFlow, Snakemake, Galaxy, CWL, WDL etc) can help by offering a high level of reproducibility, portability and computing platform independence which enables researchers to focus more on developing new methods and the interpretation of the results.

<https://www.biocommons.org.au/workflows>



# Thanks!

Any questions?

You can email me at: [johan@biocommons.org.au](mailto:johan@biocommons.org.au)

You can contact Galaxy Australia at: [help@genome.edu.au](mailto:help@genome.edu.au)