

## **BIOPLATFORMS NATIONAL INITIATIVES**

# **Workshop: Fungal Genomics with Galaxy**

October 29<sup>th</sup>-31<sup>st</sup> 2024

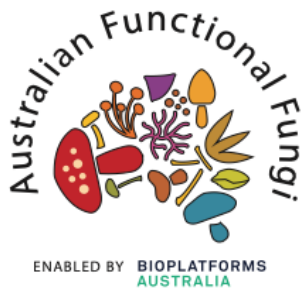
### **WORKSHOP LEADERS**

Prof. Benjamin Schwessinger (Australian National University)

Dr Alistair McTaggart (Psymbiotika Lab)

Dr Gareth Price / Dr Anna Syme (Galaxy Australia)

We would like to acknowledge the contribution of the Australian Functional Fungi Initiative and Plant Pathogen Omics Initiative in the generation of data and support in delivery of this bioinformatics training workshop. These National Initiatives are supported by Bioplatforms Australia and the Australian BioCommons, enabled by the Commonwealth's National Collaborative Research Infrastructure Strategy (NCRIS.)



### **Acknowledgement of Country**

We acknowledge the traditional custodians of Australia's vast and diverse terrestrial, marine and estuarine ecosystems. We respect their deep and continuing connection and knowledge of Australia's rich biodiversity. We pay our respects to Elders of the past and present.

<b>PROGRAM OVERVIEW</b>		
<b>DAY 1 - Tues 29<sup>th</sup> Oct</b>		<b>Facilitator</b>
9am	Welcome & Scene Setting	Benjamin and Alistair
10.30am	Break	
11am	Theory - Data input and quality control	Benjamin
12pm	How does Galaxy work?	Gareth and Anna
1pm	Lunch	
2pm	Galaxy - Data input and QC	Gareth
3pm	Theory - Genome assembly	Mareike
4pm	Galaxy - Genome Assembly	Anna
5pm	Close of Day 1	
6pm	Workshop dinner	Bioplatforms
<b>DAY 2 - Wed 30<sup>th</sup> Oct</b>		
9am	Recap Day 1 and revise genome assembly output	Anna
9.45am	Theory - QC of genome assembly	Rita
10.30am	Break	
11am	Galaxy - Genome Assembly QC	Gareth
12pm	Lunch	
1pm	Theory - TE annotation	Yan
1.45pm	Galaxy - Finding/masking repeats (incl. TEs)	Anna
2.30pm	Coffee	
3pm	Q&A session	Benjamin
3:30pm	Project talks and discussion (10 min talks, 5 mins questions)	Alistair (Chair)
5:30pm	Workshop BBQ and social	Chef, Benjamin and Alistair
<b>DAY 3 - Thurs 31<sup>st</sup> Oct</b>		
9am	Theory - Gene prediction and functional annotation	Rita and Yan
10.30am	Break	
11am	Galaxy - Gene prediction, functional annotation	Gareth and Anna
12.30pm	Lunch	
1:30pm	Q&A Session	Benjamin & Alistair
2:30pm	Where to from here, and how else can I use Galaxy?	Gareth and Anna
4pm	Workshop close	

<b>WORKSHOP DATA RESOURCES</b>	
<b>Data available in Bioplatforms Data Portal</b>	<p><a href="#">2024 Fungi Bioinformatics Workshop</a> - Please see above for the registration and login details.</p> <p><i>Acknowledgement: Sincere thanks to Alistair McTaggart, Tara Garrad and Kelly Hill for sharing their data for training purposes in this workshop.</i></p>
<b>Long read (PacBio HiFi)</b>	<p><i>Rhynchosporium commune</i></p> <p><a href="https://data.bioplatforms.com/organization/bpa-bioinformatics-workshop?tags=pacbio-hifi">https://data.bioplatforms.com/organization/bpa-bioinformatics-workshop?tags=pacbio-hifi</a></p>
<b>Short read (Illumina)</b>	<p><i>Psilocybe subaeruginosa</i></p> <p><a href="https://data.bioplatforms.com/organization/bpa-bioinformatics-workshop?tags=illumina-short-read">https://data.bioplatforms.com/organization/bpa-bioinformatics-workshop?tags=illumina-short-read</a></p>

<b>WORKSHOP GUIDE AND MATERIALS</b>	
DAY 1 - Tue 29 Oct	
9am	<p>Welcome (Alistair and Benjamin)</p> <p>Slides: BioinformaticsOpeningSession.pdf (see Zenodo folder of slides for PDF)</p>
10.30am	Break (All)
11am	<p>Theory Session - Data input and Quality Control (Benjamin)</p> <p>Slides: IntroAndDataQC.pdf (see Zenodo folder of slides for PDF)</p>
12pm	<p>How does Galaxy work (Gareth)</p> <ul style="list-style-type: none"> <li>• Slides: HowDoesGalaxyWork.pdf</li> <li>• (see Zenodo folder of slides for PDF)</li> </ul> <p>Workshop overview and Galaxy-intro tutorial (Anna)</p> <ul style="list-style-type: none"> <li>• Slides: GalaxyTutorialsOverview.pdf</li> <li>• Tutorial: <a href="#">A short introduction to Galaxy</a></li> <li>• Galaxy history of tutorial: <a href="https://usegalaxy.org.au/u/anna/h/tutorial-a-short-introduction-to-galaxy">https://usegalaxy.org.au/u/anna/h/tutorial-a-short-introduction-to-galaxy</a></li> </ul>

1pm	Lunch (All)
2pm	<p>Galaxy - Data QC (Gareth)</p> <p>Tutorial:</p> <ul style="list-style-type: none"> <li>• <a href="#">Sequence analysis / Quality Control / Hands-on</a> [skip FastQE]</li> <li>• Galaxy history of tutorial with tutorial data: <a href="https://usegalaxy.org.au/u/g.price123/h/bpa-fungi---quality-control">[https://usegalaxy.org.au/u/g.price123/h/bpa-fungi---quality-control]</a></li> </ul> <p>Other information:</p> <ul style="list-style-type: none"> <li>• Galaxy history of tutorial with fungi data: <a href="https://usegalaxy.org.au/u/g.price123/h/bpa-fungi-data---psilocybe-sample-465836">[https://usegalaxy.org.au/u/g.price123/h/bpa-fungi-data---psilocybe-sample-465836]</a> and <a href="https://usegalaxy.org.au/u/g.price123/h/bpa-fungi-data---rhynchosporium-sample-395386">[https://usegalaxy.org.au/u/g.price123/h/bpa-fungi-data---rhynchosporium-sample-395386]</a></li> <li>• Other tutorial: <a href="#">QC of Illumina data</a></li> </ul>
3pm	<p>Theory - Genome assembly (Mareike)</p> <p>Slides: GenomeAssembly.pdf (see Zenodo folder of slides for PDF)</p>
4pm	<p>Galaxy - Genome Assembly (Anna)</p> <p>Tutorial:</p> <ul style="list-style-type: none"> <li>• <a href="#">Genome assembly</a> [demonstrates assembly, with nanopore reads]</li> <li>• Galaxy history of tutorial with tutorial data: <a href="https://usegalaxy.org.au/u/anna/h/tutorial-chloroplast-genome-assembly">https://usegalaxy.org.au/u/anna/h/tutorial-chloroplast-genome-assembly</a></li> </ul> <p>Other information:</p> <ul style="list-style-type: none"> <li>• Galaxy history of tutorial with fungi data: <i>Aspergillus</i> from NCBI: <a href="https://usegalaxy.org.au/u/anna/h/test-chloro-genome-assembly-tutorial-on-aspergillus-data">https://usegalaxy.org.au/u/anna/h/test-chloro-genome-assembly-tutorial-on-aspergillus-data</a></li> <li>• <b>Illumina data:</b> Shovill: Example with <i>Psilocybe</i> data: <a href="https://usegalaxy.org.au/u/anna/h/assembly-of-bpa-fungi-data---psilocybe-with-shovill">https://usegalaxy.org.au/u/anna/h/assembly-of-bpa-fungi-data---psilocybe-with-shovill</a></li> <li>• Tutorial: Illumina data <a href="#">Genome Assembly of a bacterial genome (MRSA) sequenced using Illumina MiSeq Data</a></li> <li>• <b>PacBio Hifi data:</b> Hifiasm: Example with <i>Rhynchosporium</i> data: <a href="https://usegalaxy.org.au/u/anna/h/assembly-of-bpa-fungi-data---rhynchosporium-with-hifiasm">https://usegalaxy.org.au/u/anna/h/assembly-of-bpa-fungi-data---rhynchosporium-with-hifiasm</a></li> <li>• Tutorial: Hifi data <a href="#">Using the VGP workflows to assemble a vertebrate genome with HiFi and Hi-C data</a> [includes HiC]</li> <li>• <b>Nanopore data:</b> Flye: Tutorial: <a href="#">Large genome assembly and polishing</a></li> </ul>

5pm	Day 1 Closed
6pm	<b>Workshop dinner</b>
DAY 2 - Wed 30 Oct	
9am	<p><b>Recap</b> - Day 1. Work through genome assembly output and Q&amp;A (Anna)</p> <p><b>How to get data from BPA into Galaxy</b></p> <ul style="list-style-type: none"> <li>• Go to the BPA data portal: <a href="https://data.bioplatforms.com/">https://data.bioplatforms.com/</a></li> <li>• Log in</li> <li>• Bottom menu: Organisations</li> <li>• Click on "2024 Fungi Bioinformatics Workshop"</li> <li>• Find sample (everyone pick a random one - so not to overload)</li> <li>• <b>click on the actual file</b></li> <li>• Right click on "Access", then "copy download URL"</li> <li>• In Galaxy, upload data, paste/fetch, paste in this URL</li> </ul> <p><b>Assembly of Illumina data</b></p> <p>Ideas:</p> <ul style="list-style-type: none"> <li>• Get the data: two Illumina files (a pair)</li> <li>• QC: e.g. FastQC</li> <li>• Assembly: e.g. Shovill, using Spades</li> <li>• While waiting for assembly:</li> <li>• You could look at the shared history and the outputs:</li> <li>• <b>Illumina data:</b> Shovill: Example with <i>Psilocybe</i> data: <a href="https://usegalaxy.org.au/u/anna/h/assembly-of-bpa-fungi-data---psilocybe-with-shovill">https://usegalaxy.org.au/u/anna/h/assembly-of-bpa-fungi-data---psilocybe-with-shovill</a></li> <li>• Copy of Galaxy history:</li> <li>• What do the outputs mean?</li> <li>• Do they differ from your results?</li> </ul>
9.45am	<p>Theory - QC of genome assembly (Rita)</p> <p>Slides: GenomeAssemblyQC.pdf</p> <p>(see Zenodo folder of slides for PDF)</p>
10.30am	Break (All)

11am	<p>Galaxy - Genome Assembly QC (Gareth)</p> <p>Tutorials:</p> <ul style="list-style-type: none"> <li>• <a href="#">Genome Assembly Quality Control</a></li> <li>• Galaxy history of tutorial with tutorial data: <a href="https://usegalaxy.org.au/u/g.price123/h/bpa-fungi---genome-assembly-quality-control">https://usegalaxy.org.au/u/g.price123/h/bpa-fungi---genome-assembly-quality-control</a></li> </ul> <p>[have moved the Blast tutorial into the TE-tutorial session later in the day]</p> <p>Other information</p> <ul style="list-style-type: none"> <li>• Galaxy history of tutorial with fungi data: Example with <i>Psilocybe</i> data / <i>Rhynchospodium</i> data: Galaxy history <a href="https://usegalaxy.org.au/u/g.price123/h/bpa-fungi-data-for-gtn-gneome-assembly-qc">https://usegalaxy.org.au/u/g.price123/h/bpa-fungi-data-for-gtn-gneome-assembly-qc</a></li> <li>• Other useful tools: RagTag, Plot RagTag output, Circos</li> <li>• Other tutorial: <a href="#">ERGA post-assembly QC</a></li> </ul>
12pm	Lunch (All)
1pm	<p>Theory - TE annotation (Yan)</p> <p>Slides: TEannotation.pdf (see Zenodo folder of slides for PDF)</p>
1.45pm	<p>Galaxy - TE annotation and masking (Anna)</p> <p>Quick recap and blast</p> <ul style="list-style-type: none"> <li>• Document: BlastGenomeAssembly.pdf</li> <li>• (see Zenodo folder of slides for PDF)</li> <li>• Galaxy history for blast: <a href="https://usegalaxy.org.au/u/anna/h/blast-the-psilocybe-genome">https://usegalaxy.org.au/u/anna/h/blast-the-psilocybe-genome</a></li> </ul> <p>Tutorial:</p> <ul style="list-style-type: none"> <li>• <a href="#">Masking repeats with RepeatMasker</a> [skip section on masking with Red]</li> <li>• Galaxy history of tutorial with tutorial data: <a href="https://usegalaxy.org.au/u/anna/h/tutorial-repeat-masker">https://usegalaxy.org.au/u/anna/h/tutorial-repeat-masker</a></li> </ul> <p>Other information:</p> <ul style="list-style-type: none"> <li>• Galaxy history of tutorial with fungi data: <a href="https://usegalaxy.org.au/u/anna/h/repeat-masking-of-psilocybe-genome">https://usegalaxy.org.au/u/anna/h/repeat-masking-of-psilocybe-genome</a></li> <li>• Red repeat masking</li> </ul>
2.30pm	Coffee break (All)
3pm	Q&A Session (All)

3:30pm	<p>Initiative project presentations by participants</p> <p><b>Dr James Hane</b> – MycoProcessor: Cost-effective genomic tools for regional-scale diagnostics and surveillance of fungal pathogen effector and fungicide resistance profiles, offering new capabilities for recommendation of optimal variety and chemical applications. (Centre for Crop and Disease Management, Curtin University)</p> <p><b>Dr Camille Truong</b> – Comparative genomics in native and exotic <i>Amanita</i> species. (Royal Botanic Gardens, Melbourne, Victoria)</p> <p><b>Dr Edward Gilding</b> – Reducing the turnaround time to Panama Disease diagnosis. (Department of Agriculture and Fisheries (Queensland))</p> <p><b>Dr Matt Barrett</b> – Fungarium genomics: Gearing up for a ‘sequence everything’ model of fungarium curation. (Australian Tropical Herbarium, James Cook University)</p> <p><b>Emma Corro</b> – Wild Fungi. (MYCOmmunity Applied Mycology)</p> <p><b>Dr Kelly Hill</b> – <i>Rhynchosporium commune</i>. (South Australian Research and Development Institute)</p> <p><b>Dr Alistair Mctaggart</b> – Population genomics of magic mushrooms. (Psymbiotika Lab)</p>
5:30pm	<b>Workshop BBQ and Social (All)</b>
Day 3 - Thurs 31 Oct	
9am	<p>Theory - Gene prediction, functional annotation (Rita and Yan)</p> <p>Slides: Gene_annotation_theory_combined.pdf</p> <p>(see Zenodo folder of slides for PDF)</p>
10.30am	Break (All)

11am	<p>Galaxy - Gene prediction, functional annotation (Anna and Gareth)</p> <p>Tutorials:</p> <ul style="list-style-type: none"> <li>● Helixer: <a href="#">Genome annotation with Helixer</a> [skip section on Busco]</li> <li>● Galaxy history of tutorial with tutorial data (<i>Mucor</i>): <a href="https://usegalaxy.org.au/u/anna/h/tutorial-genome-annotation-with-helixer">https://usegalaxy.org.au/u/anna/h/tutorial-genome-annotation-with-helixer</a></li>   <li>● Document: FgenesH_tutorial.pdf</li> <li>● (see Zenodo folder of slides for PDF)</li> <li>● Galaxy history of tutorial with tutorial data (<i>Mucor</i>):</li> <li>● New copy of input data: import this one: <a href="https://usegalaxy.org.au/u/anna/h/input-data-for-fgenesh-tutorial-new-copy">https://usegalaxy.org.au/u/anna/h/input-data-for-fgenesh-tutorial-new-copy</a></li> <li>● History of the steps run with tutorial data: <a href="https://usegalaxy.org.au/u/anna/h/fgenesh-tutorial-on-mucor">https://usegalaxy.org.au/u/anna/h/fgenesh-tutorial-on-mucor</a></li> <li>● A note on bug reporting / troubleshooting in Galaxy</li> <li>● A note on grep</li> </ul> <p>Other information:</p> <ul style="list-style-type: none"> <li>● Galaxy history: Fgenesh with fungi data: <i>Psilocybe</i>: <a href="https://usegalaxy.org.au/u/anna/h/fgenesh-of-psilocybe">https://usegalaxy.org.au/u/anna/h/fgenesh-of-psilocybe</a></li> <li>● Other tools: Maker, Braker</li> <li>● Other tutorials <ul style="list-style-type: none"> <li>○ <a href="#">Genome annotation with Funannotate</a></li> <li>○ <a href="#">Genome annotation with Maker (short)</a></li> <li>○ <a href="#">Genomic Data Visualisation with JBrowse</a></li> </ul> </li> </ul>
12.30pm	Lunch (All)
1:30pm	Q&A Session (All)
Following	<p>Where from here? More resources on Galaxy and beyond. (Anna, Gareth, and others)</p> <p>Slides: Summary.pdf (see Zenodo folder of slides for PDF)</p>
4pm	Workshop Close (Goodbye and so long)