

# **Galactic Teaching: Genomics Virtual Lab Workshops**

Simon Gladman<sup>1,2</sup>, Clare Sloggett<sup>1</sup>, Mahtab Mirmomeni<sup>1</sup>, Andrew Lonie<sup>1</sup>

#### **Background**:

The Australian government has funded the development of a Genomics Virtual Laboratory (GVL): a set of analysis & visualisation platforms (currently Galaxy and UCSC genome browser) implemented on the Australian Research Cloud infrastructure (http://www.nectar.org.au/research-cloud), and community resources for best practice genomics, including protocols, workflows and tutorials for common genomics tasks. Resources for specific tasks (e.g. RNAseq-based differential gene expression) generally include: (i) a best-practice Galaxy workflow implementation; (ii) a document with background on the method; and most importantly (iii) a step-by-step tutorial that can be followed to learn how to use the workflow with real data, and may include tool suggestions and comparisons. These resources are an ideal platform for teaching real world bioinformatics analysis methods to users. An allocation exists for running workshops based on these resources on the Australian Research Cloud. See genome.edu.au for details.

GVL Launcher

+ → C 🏠 🗋 la

G\

Amaz

Cloud

Acces

Secret

Cluste

Instand

Show

Read the <u>backgrou</u>

## Infrastructure available for workshops/tutorials

- The NeCTAR Australian Research Cloud, which consists of configurable virtual computer resources available to the Australian Research Community.
- The Genomics Virtual Lab, which is an on-demand computer image with pre-installation of:
  - Bioinformatics software
  - Reference data sets and pre-computed indices
  - Web based interface using Galaxy
  - Command line available for power users
- Easy to start up and get going

×			drwxr-xr-x	5 root root 1 ubuntu ubuntu	4096 Jun 19 2013 6853 Mar 7 01:55
unch.genon _ Laur	ne.edu.au/launch ICher from BioCloudCentral.org	☆ 🔜 🚥 ≡	-rw-rr -rw-rr -rw-rr lrwxrwxrwx	1 ubuntu ubuntu 1 ubuntu ubuntu 1 ubuntu ubuntu 1 root root	220 Apr 3 2012 4549 Nov 14 20:27 3632 Jul 24 2013 21 Jun 19 2013
/ launch C	CloudBioLinux, CloudMan and Galaxy platforms on Cloud Computing resources (including		drwx lrwxrwxrwx drwx -rwxrwxrwx lrwxrwxrwx	2 ubuntu ubuntu 1 root root 3 ubuntu ubuntu 1 ubuntu ubuntu	4096 Jun 19 2013 20 Jun 19 2013 4096 Nov 29 03:13 992 Jun 19 2013 21 Jun 19 2013
on web:	NeCTAR (OpenStack)		lrwxrwxrwx lrwxrwxrwx drwxr-xr-x -rw-rw-r	2 u. 1 ubun 3 ubuntu	20 Jun 19 2013 20 Jun 19 2013 35 Jun 19 2013 Jun 19 2013 21 12:00
s key	Choose from the available clouds. The credentials you provide below must match (ie, exist on) the chosen cloud. Your Access Key ID. For the Amazon cloud, available from the security credentials page.		lrwxrwxrwx lrwxrwxrwx lrwxrwxrwx drwxr-xr-x	1 root roo 1 root root 1 root root 2 root root	409L
key	Your Secret Access Key. For the Amazon cloud, also available from the security credentials page.	≓ ni	× 🔁	ome.edu.au	
ord	Name of your cluster used for identification and relaunching. If creating a new cluster, type any name you like. Your choice of password, for the CloudMan web interface and accessing the instance via ssh or FreeNX.	y / G	VL-tut File type Encoding Total Seque	Conventional base Sanger / Illumina 1 2013533	Analyze Data Workflow ny a calls 1.9
e type	Extra-extra Large (16 vcpu / 64GB R/ Type (ie, virtual hardware configuration) of the instance to start.	ation ition t sd Gro	Filtered Sequence le %GC	uences 0 mgth 75 35	
w advanced <b>Irt an insta</b> website is a ig scalable i	an open service developed by the <b>GVL Launcher</b>	enton. ulation ( nanipula alysis alysis ils	tion 40 38 36 1	base sequence of	
n using the site can be dentials are id is not ava	Amazon cloud. The open source code is available on GitHub allowing you to also run this service locally. used for any of the available clouds. Note that you must have appropriate credentials for the chosen cloud. Provided not stored anywhere on this site and are used only to make the initial request to launch the desired instance. If a given ilable in the list and you would like to see it there, please contact us.	<u>Detectio</u> <u>lling</u>	n 32 111 30 28		
nching serve r sole respoi	ers on the Amazon cloud will incur <mark>usage fees</mark> from Amazon for their resources. By using this service you acknowledge nsibility for any costs accrued.	¥ K	24 22 20		
			18 16 14 12 10 8 6 4 2 0 1 2 3 4	56789 11 13 15	17 19 21 23 25 27 29



#### Using the GVL to teach Bioinformatics

- The steps involved in setting up and running a teaching session are:
  - Launch a GVL machine with Galaxy server
  - Distribute the learning materials and server address
- Tutorials are layered documents that contain step by step instructions, background on analysis methods and screenshots
- Sample datasets are included in each tutorial.
- There are tutorials and other learning materials available for the GVL on:
  - RNA-Seq and Digital Gene Expression analyses
  - Variant calling from NGS data
  - Genome assembly
  - And others

## **Availability**

• The NeCTAR Australian Research Cloud can be accessed via <u>www.nectar.org.au</u>

Figure 1: GVL launcher to a GVL customised Galaxy server in minutes



- The Genomics Virtual Lab launcher, tutorials and protocols are available at <u>www.genome.edu.au</u>
- Most of the tools for the pipelines are currently available in the Galaxy main toolshed and the rest will become available in the near future.

mmj	ch.Segantces ch.Algements	A which pabling	VisiverOptimizer Logithe
-	Oknomic Scores trate on Oknomic Intervals	wants ever the above range. Must be even	14: Yahen Optimizer visci on 🐵 🖉 20 Idata B. data 9. and data 7: Condig state
	Delica	April 1	13. Varbert Optimizer visus on
I to the workshop here	phyDisplay Date	Reput rest	The Transmustation on the has a distance of the
to the workshop here	utonal Variation	Stortweet +	1. And data 2: triemed reads
	Theoremis Analysis	Are the reads paired and to by	B. Trunctostudis: on data 3. data 40 // 32 1. and data 2. Div2.transmed parts
	Motif Tools	Tax the fact of the six a gained theory on the second se	7. Trimenostatic no data 3. data ob 2 30
	Multiple Alignments	Theremonate on each merid pars	1. and data 2: Doil trimmed para
	EASTA manipalation	Read disast for direction 2.	E mater ERED-HEDE Zhang Hered
	NOS: QC and manpulation NOS: Picanit (Insta)	Interviewed to do more part -	a
	NOS: Mapping	Element of transport of transport of the second sec	Partice Derivering Linear and Am
	NGS, RNA Anatysis	1 Remote AppLiesd Interest 1	ALECTIONARY AND OF A
	NOS: SAM Tools NOS: OATK Tools (beta)	Insurred literatus 2	A RECEIVED AND A WITH A RECEIVED AND A R
	NOS: Peak Calling	Fin type:	demonstration with a
	SNP/WGA: Data: Filtera		
	SNPWGA: QC; LD; Pots SNPWGA: Statistical Models		
	Phenotype Association	Red datase:	
	VCF.Toots test_tools	8. Tritmenonatic on dia.immedi readta	
	<u>Trensponate</u> Trim and otherwise     cleanup fasts files in a pair aware	Etracy is short mate paired:	
	matrice -	There is the advection of the advection	
	Automatically optimize a de-novo amagembic como Valvari		
	Easta Statistics Dountary summary	Add revel sport and therains.	
	BARNES for a family for.		
	MicrobialAssembly	Calmar werbon verver output in tog tre r:	
	• bitwastdawa	The advanced parameters?	
	< .		
	Figure 5: Scre	reenshot of assembled contigs (a), contig stats (b) and Velvet Optimiser	log (c)
	guite et ette		-9 (-)
	a		
	E Calany x		
	C 2 C 2 115.146.94.14		
	- Galaxy	Augustata Developer Venetation Anno Anno Anno Anno	Using 6.2 GB
	ENG-DRE.Toots	This chalavel is large and only the Brit mergethyle is shown tenser	
	Lift-Over Text Manipulation		Merning assembly sector? 15.08
	Eilter and Sort	ATTACK 1 UTTOT MAY DUE 314 LISTAGE (ATTACKTCCTTTMARTINTTALANGTAAAAAAACCCATATAACMITCCTTTACGETAC	17: Fasta Statutics on mata 13: ⊕ Ø 20
	Bole, Stobleart and Caram	THE ANTIFICTION ON ON ON ON ON ON ON OTHER ATTACTOR	The Party street has store with the party of



Figure 2: The different sections of a typical Genomics Virtual Lab tutorial document

www.genome.edu.au

Victorian **B**ioinformatics Consortium







The GVL acknowledges funding from the NeCTAR project http://www.nectar.org.au NeCTAR is an Australian Government project conducted as part of the Super Science initiative and financed by the Education Investment Fund.