

The Genomics Virtual Lab at work

Andrew Lonie², Ron Horst¹, Igor Makunin¹, Derek Benson¹, Clare Sloggett², Simon Gladman²,
Nuwan Goonasekera², Yousef Kowsar², Mark Crowe³, Enis Afgan⁴

¹University of Queensland, r.horst@uq.edu.au

²Victorian Life Sciences Computation Initiative, University of Melbourne, alonie@unimelb.edu.au

³QFAB Bioinformatics, Brisbane m.crowe@qfab.org

⁴ Center for Informatics and Computing at the Ruđer Bošković Institute, Zagreb afgane@gmail.com

SHORT ABSTRACT

The Genomics Virtual Lab (GVL) provides tools and resources for computationally intensive bioinformatics research, with emphasis on analysis of high-throughput sequencing data.

Galaxy [1] is an open, web-based platform for data intensive biomedical research; with a user-friendly web interface and no requirement for knowledge of bash scripting or programming for analysis of genome scale data. The GVL maintains a production scale Galaxy Server, a tutorial Galaxy Server and a full mirror of UCSC [2] Genome Browser; all implemented on the NeCTAR [3] Australian Research Cloud.. We offer easy to follow step-by-step tutorials for Galaxy, such as RNA-Seq (Differential gene expression using NGS) or Variant Detection. For advanced users we provide images for deployment of a scalable virtual computer cluster on the NeCTAR cloud, or a private customisable Galaxy server with extra features such as built-in R Studio [4] and iPython Notebook [5].

GVL resources can be used by novice and skilled Biologists, Bioinformaticians and Bioinformatics Tool Developers.

Resources are available in several platform versions: personal, aimed at individual researchers; server, aimed at small research groups; and managed, aimed at large teams or institutions.

GVL has been under development for around two years and several of the GVL resources now have a track record of real life use. This presentation will summarise the status of the GVL and present some case studies of the GVL in action.

GVL RESOURCES AND PLATFORM

GENOME.EDU.AU

GVL provides Researchers and Developers with a number of **resources**....

- **Galaxy**: perform reproducible analyses via a friendly user interface
- **Command Line Notebook**: perform reproducible analyses with full command line access and a windows interface
- **Tutorials**: learn analysis techniques
- **CloudMan** [6]: manage your cluster on the cloud, dynamically add and remove nodes
- **UCSC browser**: mirror of the UCSC Genome Browser

...on several **platforms** on the NeCTAR Australian Research Cloud

- **Personal**: Launch your own instance with your default NeCTAR allocation
- **Server**: Launch your own analysis cluster on your research NeCTAR allocation, tailored to your needs
- **Managed**: Use the GVL managed resources

GVL Resource Features

	Highlights	Platform	Biologist	Bioinformatician	Developer	Learner
Galaxy	Friendly interface, Reproducible analysis	GVL USE, Server, Personal	♥♥♥♥♥	♥♥		
Notebook	Reproducible command line analysis	Server, Personal		♥♥♥♥♥	♥♥♥♥♥	

Tutorials	Learn analysis techniques	GVL LEARN	♥			♥♥♥♥♥
CloudMan	Manage your platform	Server, Personal	♥	♥♥♥♥♥	♥♥♥♥♥	
UCSC	Australian Mirror	GVL BROWSE	♥♥♥♥♥	♥♥♥♥♥		♥♥♥

GVL Platform Features

	Easy to Set up	Command access	Line	Large Datasets	Install Tools	Galaxy	Cluster/Virtual Lab
Personal: GVL GET	♥♥♥♥	♥♥♥♥♥		♥♥	♥♥♥♥♥		
Server: GVL GET	♥♥	♥♥♥♥♥		♥♥♥	♥♥♥♥♥		♥♥♥♥♥ [†]
Managed: GVL USE	♥♥♥♥♥			♥♥♥♥♥			
Managed: GVL LEARN	♥♥♥♥♥			♥			

REFERENCES

- Giardine, B. et al. *Galaxy: A platform for interactive large-scale genome analysis*. Genome Res. 2005. 15: 1451-1455
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. *The human genome browser at UCSC*. Genome Res. 2002 Jun;12(6):996-1006.
- NeCTAR Virtual Laboratories <https://nectar.org.au/virtual-laboratories-1>, accessed 4 June 2014
- rStudio <http://www.rstudio.com/>, accessed 4 June 2014
- iPython Notebook <http://ipython.org/notebook.html>, accessed 4 June 2014
- CloudMan <http://wiki.galaxyproject.org/CloudMan>, accessed 4 June 2014

ABOUT THE AUTHOR(S)

A/Prof Andrew Lonie is a faculty member of the department of Computing and Information Systems at the University of Melbourne and foundation head of the Life Sciences Computation Centre, a cross-institutional centre of bioinformatics and computational biology expertise and infrastructure support within the Victorian Life Sciences Computation Initiative. Genomics, proteomics and biomedical image analysis experts within the LSCC collaborate with and support life sciences researchers in a variety of research projects across Victoria; the centre is also responsible for implementing and disseminating best practice methods and techniques, advising on experimental design and interpretation, and resourcing and maintaining informatics analysis platforms. A/Prof Lonie is also academic coordinator of the MSc (Bioinformatics) at the University of Melbourne.