



University of Iowa Custom Galaxy Deployment

Current Prototype Deployment and Next Steps

Ann Black
annblack@eng.uiowa.edu
Senior Application Developer

University of Iowa
Iowa Initiative in Human Genetics
Center for Bioinformatics and
Computational Biology

Bringing Galaxy to Iowa



■ Why host a local Galaxy deployment?

- No tight data quota caps
- Local data storage and transfer
- Customizable deployment, for example:
 - Control version of tools
 - Tune how tools are exposed
 - Expose additional custom tools and databases

■ Where will our local Galaxy reside?

- Directly with Helium, the University of Iowa's High Performance Computing Cluster
 - Over 3600 processing cores, 350 compute nodes
 - Connected to a high speed storage system (Lustre) over an Infiniband network capable of 40 Gb/sec connectivity
 - Highly available and fast.

Our local Galaxy: Where is it at now?

- **Current Status of our local Galaxy deployment:**
 - Alpha Deployment Phase, but It's Alive!!!!!!!
 - Hosted on a HPC test server and dispatching jobs to the compute cluster
 - Sub-set of tools exposed compared to the public Galaxy deployment
 - Capable of all public Galaxy server functionality
 - Happy to expose more tools/capabilities upon request
 - Human, hg19 reference genomes configured
 - Can view local Galaxy datasets
 - In IGV
 - ICTS local (and upto date) UCSC browser
 - Published Human exome analysis pipeline is available

- **Welcome more alpha users**
 - Please visit wiki for further details:
<http://wiki.uiowa.edu/display/galaxyusers>
 - Or send me an email: annblack@eng.uiowa.edu

Our local Galaxy: What's next?



■ Our future goals:

- Move off the test server and onto a galaxy dedicated HPC cluster submit server
 - The migration will not affect any current alpha use
- Polish current deployment to move into beta -> production
- Expose additional bioinformatics tools and Galaxy features based on community feedback
- Integrate with DNA Core sequencers to directly provide sequencer data into Galaxy for analysis
- Publish additional analysis pipelines

For More Information



■ For more information:

- Join our local galaxy-users listserv:
 - <https://list.uiowa.edu/scripts/wa.exe?SUBED1=galaxy-users&A=1>
- Visit our local galaxy deployment wiki page:
<http://wiki.uiowa.edu/display/galaxyusers>
- Explore the public Galaxy wiki and server:
 - Wiki: <http://wiki.g2.bx.psu.edu/>
 - Server: <http://use.galaxy.org>

■ Our local Galaxy deployment is brought to you by:

- [Iowa Initiative in Human Genetics](#)
- [Center for Bioinformatics and Computational Biology](#)
- [University of Iowa High Performance Compute Center](#)

Upcoming Events



- Bioinformatics Short Course August 1-3, 2012
 - **Mutation Detection Using Massively Parallel Sequencing: From Data Generation to Variant Annotation**

Massively parallel DNA sequencing technologies have ushered in the next wave of the genomics revolution. The clinical application of these technologies will make personalized genomic medicine a reality; in the research laboratory, these technologies are making innovative approaches to genome-wide investigations routine. In the summer of 2012, the Iowa Institute of Human Genetics will offer a Bioinformatics Short Course. The course will focus on most popular next-generation sequencing platforms - the Illumina HiSeq and MiSeq systems. Upon completion of this course, participants will understand the design of a next generation sequencing experiment and the workflow to achieve a particular result. A series of lectures to introduce basic concepts will be interwoven with practical examples of data generation and analysis. Hands-on sessions will enable participants to analyze data from its generation to interpretation. Participants will be required to bring their own laptops. Enrollment is limited.

- **Course will leverage Iowa's Galaxy Deployment**
- <http://www.medicine.uiowa.edu/humangenetics/bioinformaticscourse/>
- Brought to you by:
 - Iowa Initiative in Human Genetics