

The National Center for Genome Analysis Support



INDIANA UNIVERSITY



Changing genomics analytical needs

- Next Gen sequencers are generating more data and getting cheaper
- Sequencing is:
 - Becoming commoditized at large centers and
 - Multiplying at individual labs
- Analytical capacity has not kept up
 - Bioinformatics support
 - Computational support (thousand points solution)
 - Storage support



NATIONAL CENTER FOR GENOME ANALYSIS SUPPORT

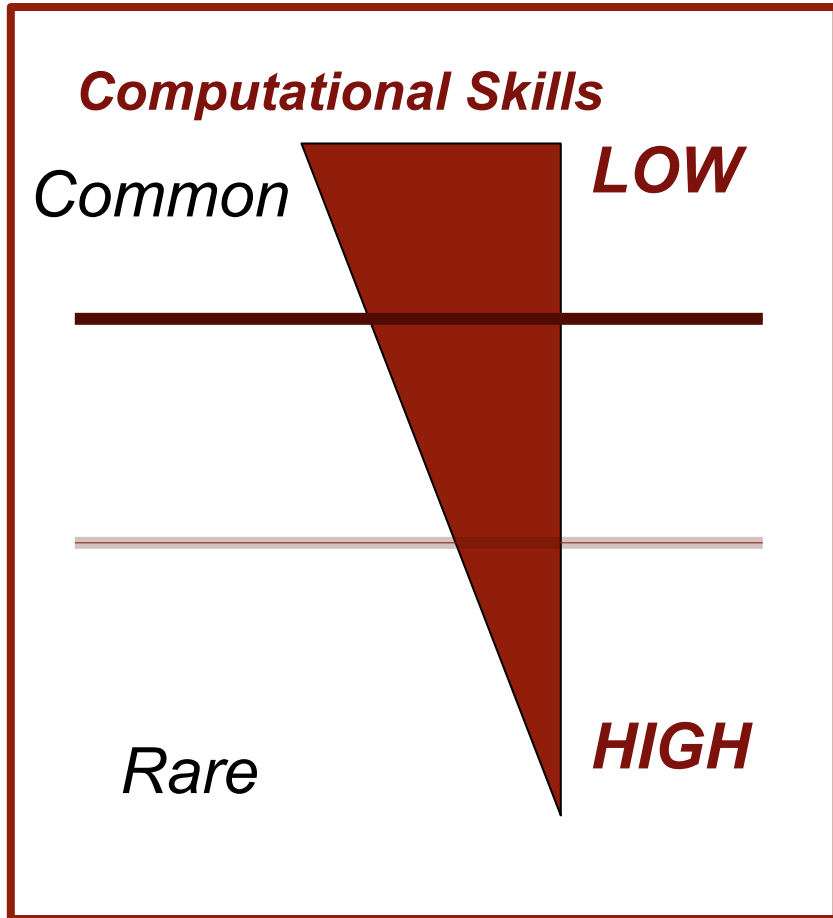
INDIANA UNIVERSITY

- Funded by National Science Foundation
 1. Large memory clusters for assembly
 2. Bioinformatics consulting for biologists
 3. Optimized software for better efficiency



- Collaboration across multiple institutions
- Open for business at: <http://ncgas.org>

Making it easier for Biologists



- Web interface to NCGAS resources
- Supports many bioinformatics tools
- Available for both research and instruction.



NCGAS Cyberinfrastructure at IU

- Mason large memory cluster (512 GB/node)
- Quarry cluster (16 GB/node)
- Data Capacitor (1 PB at 20 Gbps throughput)
- Research File System (RFS) for data storage
- Research Database Cluster for managing data sets.
- All interconnected with a high speed internal network (40 Gbps)

Tools

- [Import Data](#)
- [Sequence QC](#)
- [NGS: QC and manipulation](#)
- [NGS: Mapping](#)
- [NGS: RNA Analysis](#)
- [De novo Assembly](#)
- [Assembly QC](#)
- [NCBI Blast+](#)
- [Workflows](#)
- [Request Help](#)



NATIONAL CENTER FOR GENOME ANALYSIS SUPPORT

INDIANA UNIVERSITY



Welcome to the Galaxy Instance at Indiana University

Thank you for choosing Galaxy! NCGAS is committed to providing support for Indiana University research. Don't hesitate to contact help@ncgas.org if you find that you need a tool that is not supported by our current Galaxy or if you have questions or suggestions. When possible, we will have the requested tool up and running in two working days time, and failing that, will report the status of the request within that time.

This instance of the Galaxy is installed and maintained by National Center for Genome Analysis Support [NCGAS](#) (NSF Award #1062432)

The Computing power is provided by the Indiana University [Mason Compute Cluster](#)

The storage is provided by the Indiana University [Data Capacitor](#) (NSF Award # 0521433)

The web server is hosted on the Indiana University [Quarry Gateway Hosting](#)

The Galaxy project is supported in part by [NSF](#), [NHGRI](#), and the [Huck Institutes of the Life Sciences](#).

NCGAS is a center of the Pervasive Technology Institute at Indiana University

History

- 84: [Clustal run clustal log.txt](#)
- 83: [Clustal run output.clustal](#)
- 82: [clustalTest.txt](#)
- 81: [BAM-to-SAM on data](#)
80: [converted SAM](#)
- 80: [mapt.NA12156.altex.bam](#)
- 70: [D. grimshawi MRNA data](#)
- 69: [blastn on db](#)
- 68: [Drosophila grimshawi Gene data](#)
- 67: [FlyMine query](#)
- 66: [D. grimshawi Gene data](#)
2 lines
format: tabular, database: ?
Info: Gene data from <http://www.flymine.org/release-34.0/service/>.
Uploaded from <http://www.flymine.org/release-34.0/results.do>.
organisms: D. grimshawi



In Sum...

- NG Sequencing is creating a analytical problem that cannot be solved at sequencing centers
- NCGAS can provide a global scale infrastructure to better serve the needs of biologists who cannot become bioinformaticians to accomplish their research.
- NCGAS is improving existing commodity bioinformatic tools



INDIANA UNIVERSITY

Thank You

Questions?

Bill Barnett (barnettw@iu.edu)

Rich LeDuc (rleduc@iu.edu)

Le-Shin Wu (lew@iu.edu)

Carrie Ganote (cganote@iu.edu)



**NATIONAL CENTER FOR
GENOME ANALYSIS SUPPORT**

INDIANA UNIVERSITY

Acknowledgements & disclaimer

- This material is based upon work supported by the National Science Foundation under Grants No. ABI-1062432
- This work was supported in part by the Lilly Endowment, Inc. and the Indiana University Pervasive Technology Institute
- Any opinions presented here are those of the presenter(s) and do not necessarily represent the opinions of the National Science Foundation or any other funding agencies

License terms

- Please cite as: LeDuc, R.D., The National Center for Genome Analysis Support, presented at LOCATION OF TALK
- Items indicated with a © are under copyright and used here with permission. Such items may not be reused without permission from the holder of copyright except where license terms noted on a slide permit reuse.
- Except where otherwise noted, contents of this presentation are copyright 2011 by the Trustees of Indiana University.
- This document is released under the Creative Commons Attribution 3.0 Unported license (<http://creativecommons.org/licenses/by/3.0/>). This license includes the following terms: You are free to share – to copy, distribute and transmit the work and to remix – to adapt the work under the following conditions: attribution – you must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work). For any reuse or distribution, you must make clear to others the license terms of this work.