



Genifx GENOME INFORMATICS @ UAB

What we provide



Data storage. Free for "open data". Also a paid service for "closed" private data.



Software pipelines to analyze the data.



Shell access to run your own commands.



High-Performance Compute (HPC) cluster. >200 cores and counting.



Cloud-computing infrastructure.



Consultation on your NGS data analysis needs.



Reference datasets to analyze your data.

Cutting-edge technology Platform

- Transparent access to cloud platforms like Amazon AWS or DNANexus.
- NoSQL (MongoDB) based database store.
- Pegasus (<http://pegasus.isi.edu/>) pipeline software.
- Cutting-edge infrastructure to run Docker (<https://www.docker.com/>) containers.
- Distributed file system based on GlusterFS.
- Quad Infiniband network backbone
- High-memory servers for de novo genome assembly.

WHAT ?

We provide software and hardware infrastructure to store and analyze Next-Generation (NGS) sequence data.

WHY ?

Analysis of NGS data is resource intensive. Genifx fills a void at UAB in providing a dedicated facility for storage and analysis of NGS data. Genifx also promotes collaboration and data sharing among researchers at UAB.

HOW ?

The facility is free* for everyone willing to share their data with other researchers at UAB. You can of course pay to make your data private.

Genifx Pricing Model

Genifx operates in research collaboration model. The model is borrowed from the open-source software repository GitHub (<http://github.com>). As long as a researcher shares data with other researchers at UAB, the facility is free. This means that you control who can publish results derived from the analysis of your data and can see who downloaded it, but you cannot control the access.

We understand the need for keeping some data private. This can be done by paying for the service. The price is determined case by case basis and dictated by the market price of the hardware.

Available reference datasets

At Genifx we provide several databases and reference datasets to compare your data against. Some are commercial, some free. We preprocess most of these datasets for easy analysis. More databases can be added on request. Some of these datasets are as follows:

- The Cancer Genome Atlas (TCGA) - To compare your sequence data against thousands of cancer genomes.
- Kyoto Encyclopedia of Genes and Genomes (KEGG) - To analyze pathways.
- Pathway commons - A free alternative to KEGG
- Reactome and Irefindex - protein-protein interaction databases.
- Human Genome Mutation Database (HGMD) - A commercial database of human mutation.
- 1000 Genome project - The variations found in natural human populations.

Acknowledgements

Genifx has been created by grants from UAB Health Science Foundation (HSF) and Division of Informatics, Department of Pathology. Additional financial supports come from Center for Clinical and Translational Sciences (CCTS), Heflin Center for Genomic Sciences, and UAB Stem Cell Institute.

Genifx is also benefitted greatly from its integration with Choir Compute Cluster created by the Computational Molecular Biology Unit of the Division of Informatics, Department of Pathology.

Contact

<http://genifx.ifx.uab.edu/>

