Introduction to HPRC Galaxy
Your Login Password

- Both state of Texas law and TAMU regulations prohibit the sharing and/or illegal use of computer passwords and accounts
- Don’t write down passwords
- Don’t choose easy to guess/crack passwords
- Change passwords frequently
For More Help...

Website: hprc.tamu.edu
Email: help@hprc.tamu.edu
Telephone: (979) 845-0219
Visit us in person: Henderson Hall, Room 114A

Help us, help you -- we need more info
• Which Cluster
• UserID/NetID
• Job id(s) if any
• Location of your job file, input/output files
• Application used if any
• Module(s) loaded if any
• Error messages
• Steps you have taken, so we can reproduce the problem
What is the Galaxy Project?

HPRC Galaxy is not a public Galaxy server

- reproducible workflow
- shared data and workflows
- many popular bioinformatic tools are available
- no programming knowledge required

usegalaxy.org
Galaxy 101 - galaxyproject.org/tutorials

Table of contents

- **Galaxy 101** - explains basics of Galaxy use. This tutorial takes you through downloading human genome annotation from UCSC Table Browser and manipulation of these data to count the number of single nucleotide polymorphisms in exons of protein-coding genes. It ends with the creation of a workflow and its use for the analysis of new data. This tutorial consists of two parts:
  - Part 1
  - Part 2
- **Introduction to NGS technologies** - a quick overview of next-generation sequencing technologies currently present on the market.
- **Processing many samples at once** - this tutorial explains new feature of Galaxy Interface - dataset collections. Dataset collections allow you to easily manipulate hundreds of samples in just a few clicks.
- **Diploid variant calling** - this tutorial demonstrates the use of Galaxy for finding sequence variants in diploid genomes.

Additional resources

- **Freiburg group tutorials** - an ever-growing collection of Galaxy tutorials that serves as a constant inspiration to this page.
- **Galaxy NGS101** - a collection of video tutorials detailing various stages of NGS analysis.
HPRC Maroon Galaxy

Welcome to the HPRC Maroon Galaxy

Notice that Maroon Galaxy will be offline on Tuesday, Nov 7 from 8:00 am to 5:00 pm for maintenance. All jobs running at that time will be stopped and you will need to restart your jobs after the maintenance is complete.

Please contact the HPRC helpdesk to request a new tool or indexed genome, report errors or if you just have questions about using Galaxy.

Newest tools added to Maroon Galaxy
- MyHPC SU balance
- Fastq-dirty-sort
- SamSan 0.72
- StringTo 1.33
- PICARO 2.03
- deepTools 2.5.1
- edgeR 3.14.0
- Tidy 2.4.6
- MACS2 2.1.0.20160308
- Gatk 3.8
- EMDR 2.4.0
- BURSCO 3.0.2b
- IntProScan 5.25.64.8

hprcgalaxy.tamu.edu/maroon
HPRC Maroon Galaxy

- Try Galaxy at usegalaxy.org to see if it appropriate for your project
- Getting Access to HPRC Maroon Galaxy
  - Available to Texas A&M students, staff and faculty with a NetID and Ada account
  - Apply for an Ada account first
    - https://hprc.tamu.edu/apply
  - Then send an email request for a Maroon Galaxy account
    - help@hprc.tamu.edu
- Read the Galaxy Usage Notes
  - hprc.tamu.edu/wiki/SW:Galaxy
- There are no backups of users’ Galaxy files
  - You can export a Galaxy history to a single file which can be uploaded to the same or a different Galaxy instance
    - Some custom tools may not be available in all Galaxies
### Repositories by Category

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Repositories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assembly</td>
<td>Tools for working with assemblies</td>
<td>88</td>
</tr>
<tr>
<td>ChIP-seq</td>
<td>Tools for analyzing and manipulating ChIP-seq data.</td>
<td>47</td>
</tr>
<tr>
<td>Combinatorial Selections</td>
<td>Tools for combinatorial selection</td>
<td>8</td>
</tr>
<tr>
<td>Computational chemistry</td>
<td>Tools for use in computational chemistry</td>
<td>30</td>
</tr>
<tr>
<td>Constructive Solid Geometry</td>
<td>Tools for constructing and analyzing 3-dimensional shapes and their properties</td>
<td>12</td>
</tr>
<tr>
<td>Convert Formats</td>
<td>Tools for converting data formats</td>
<td>79</td>
</tr>
<tr>
<td>Data Export</td>
<td>Tools for exporting data to various destinations</td>
<td>2</td>
</tr>
</tbody>
</table>
Get Data into Your Galaxy History

- Upload files < 2GB using Galaxy web interface
  - select local file to upload or paste URL
- Upload large files > 2GB
  - use sftp protocol instead of upload button (next slide)
- Can retrieve data from external websites directly into your Galaxy history with ‘Get Data’ tools
  - UCSC, BioMart, Ratmine, ...
Uploading a 2GB+ file to Galaxy

- Files larger than 2GB should be copied using ftp instead of the upload button.
- There are three options for uploading files:
  a. Use the sftp command in a Unix terminal on your Mac or Linux desktop
  b. Copy files from Ada $SCRATCH directory to ftp directory using sftp on Ada
  c. Use MobaXTerm on your Windows desktop
- After transferring file to the Galaxy ftp directory, go to Galaxy 'upload file' interface in Galaxy to see your ftp transferred file (next slide)

https://hprc.tamu.edu/wiki/SW:Galaxy#Uploading_Files_.3E_2GB_via_FTP_to_Maroon_Galaxy
Add Your FTP Uploaded 2GB+ File to Your History
HTTP URL Upload < 2GB File or Direct Paste
FTP URL Upload 2GB+ File via URL

You can URL upload a 2GB+ sized file if the URL is an ftp site

ftp://
Files can be added to a ‘Data Library’ which you can share with your colleagues. Send a request to help@hprc.tamu if you would like to create a Data Library.
Galaxy File Formats

- Many Galaxy tools require fastqsanger format
- FastQC tool will check fastq format of a fastq file
  - MiSeq, HiSeq, NextSeq, NovaSeq all use 1.8+

```
@ERR504787.5.1 M00368:15:000000000-A0HKH:1:2:16161:12630-1 length=100
TATTTTAAGGAATGACCAAGGAATGACTCCCCAATCATGGCTGTATCAACTCCAAAATTTTCTGCAACAGTCGCTGAAATATCTGCAAAATGCCCTTGTGGAA
+ERR504787.5.1 M00368:15:000000000-A0HKH:1:2:16161:12630-1 length=100
CCCFFFFHHHHHHJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJ JJ

```

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filename</td>
<td>ecoli_transcriptome_rna_seq_SRR933983_1_fastq</td>
</tr>
<tr>
<td>File type</td>
<td>Conventional base calls</td>
</tr>
<tr>
<td>Encoding</td>
<td>Sanger / Illumina 1.9</td>
</tr>
<tr>
<td>Total Sequences</td>
<td>128399</td>
</tr>
<tr>
<td>Sequences flagged as poor quality</td>
<td>0</td>
</tr>
<tr>
<td>Sequence length</td>
<td>35-51</td>
</tr>
<tr>
<td>%GC</td>
<td>51</td>
</tr>
</tbody>
</table>
FASTQ Format Encoding

<table>
<thead>
<tr>
<th>ASCII character</th>
<th>Decimal value</th>
<th>Sanger value</th>
<th>Illumina 1.8+ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>!#$%&amp;'()+-,./;&lt;=?@ABCDEFGHIJKLMNOPQRSTUVWXYZ{}^~`</td>
<td>33, 59, 64, 73, 104, 126</td>
<td>0, 26, 31, 40</td>
<td>0.2, 26, 31, 40</td>
</tr>
</tbody>
</table>

- **S** - Sanger: Phred+33, raw reads typically (0, 40)
- **X** - Solexa: Solexa+64, raw reads typically (-5, 40)
- **I** - Illumina 1.3+: Phred+64, raw reads typically (0, 40)
- **J** - Illumina 1.5+: Phred+64, raw reads typically (3, 41) with 0=unused, 1=unused, 2=Read Segment Quality Control Indicator (bold) (Note: See discussion above).
- **L** - Illumina 1.8+: Phred+33, raw reads typically (0, 41)

In Galaxy, Sanger (fastqsanger) is equivalent to Illumina 1.8+
Galaxy File Formats

- Fastq Groomer tool will convert fastq to fastqsanger
  - Will convert quality scores to fastqsanger encoding
    - Solexa to fastqsanger
    - Illumina 1.3 - 1.7 to fastqsanger
  - converted file will be an exact copy if the fastq file is already Illumina 1.8+
- If your fastq file is already in fastqsanger format, you can set format in file attributes instead of making a copy which will save you time and also save disk space
Setting Format to Fastqsanger on Upload
Setting Format to Fastqsanger in File Attributes

1. Select the dataset.
2. Change data type.
3. Choose "Fastqsanger".
Check Your HPRC SUs Balance

The image shows a webpage with a search bar and a list of options on the left side. Below the search bar, there is a section titled "My HPRC SU balance" with a dropdown menu labeled "I want to." The options include "Show my current SU balance," "Set or change my default account," and more. Below, there is a table labeled "List of users' Project Accounts" with columns for "Account," "Default," "Allocation," "Used & Pending SUs," and "Balance." The table contains entries with account numbers and corresponding SU balances. There are links or icons on the right side of the page, possibly for additional functionalities or support. The website's purpose is to allow users to check their SU balance and manage their accounts.
<table>
<thead>
<tr>
<th>Jobs Duration</th>
<th>BLAST+ Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Day Jobs</td>
<td>NCBI BLAST+ blast 480 SUs. Search nucleotide database with nucleotide query sequence(s) (max runtime 1 day, 480 SUs required)</td>
</tr>
<tr>
<td></td>
<td>NCBI BLAST+ blast 480 SUs. Search protein database with protein query sequence(s) (max runtime 1 day, 480 SUs required)</td>
</tr>
<tr>
<td></td>
<td>NCBI BLAST+ blast 480 SUs. Search protein database with translated nucleotide query sequence(s) (max runtime 1 day, 480 SUs required)</td>
</tr>
<tr>
<td></td>
<td>NCBI BLAST+ blast 480 SUs. Search translated nucleotide database with translated nucleotide query sequence(s) (max runtime 1 day, 480 SUs required)</td>
</tr>
<tr>
<td></td>
<td>BLAST reciprocal from two FASTA file 1 day, 480 SUs req</td>
</tr>
<tr>
<td>3 Day Jobs</td>
<td>NCBI BLAST+ blast 1440 SUs. Search nucleotide database with nucleotide query sequence(s) (max runtime 3 days, 1440 SUs required)</td>
</tr>
<tr>
<td></td>
<td>NCBI BLAST+ blast 1440 SUs. Search protein database with protein query sequence(s) (max runtime 3 days, 1440 SUs required)</td>
</tr>
<tr>
<td></td>
<td>NCBI BLAST+ blast 1440 SUs. Search protein database with translated nucleotide query sequence(s) (max runtime 3 days, 1440 SUs required)</td>
</tr>
<tr>
<td></td>
<td>NCBI BLAST+ blast 1440 SUs. Search translated nucleotide database with translated nucleotide query sequence(s) (max runtime 3 days, 1440 SUs required)</td>
</tr>
<tr>
<td>7 Day Jobs</td>
<td>NCBI BLAST+ blast 3360 SUs. Search nucleotide database with nucleotide query sequence(s) (max runtime 7 days, 3360 SUs required)</td>
</tr>
<tr>
<td></td>
<td>NCBI BLAST+ blast 3360 SUs. Search protein database with protein query sequence(s) (max runtime 7 days, 3360 SUs required)</td>
</tr>
<tr>
<td></td>
<td>NCBI BLAST+ blast 3360 SUs. Search protein database with translated nucleotide query sequence(s) (max runtime 7 days, 3360 SUs required)</td>
</tr>
<tr>
<td></td>
<td>NCBI BLAST+ blast 3360 SUs. Search translated nucleotide database with translated nucleotide query sequence(s) (max runtime 7 days, 3360 SUs required)</td>
</tr>
<tr>
<td>7 Day 5 Node Jobs</td>
<td>NCBI BLAST+ blast 16,800 SUs. Search protein database with protein query sequence(s) (max runtime 7 days on 5 nodes, 16,800 SUs required)</td>
</tr>
<tr>
<td></td>
<td>NCBI BLAST+ blast 16,800 SUs. Search nucleotide database with nucleotide query sequence(s) (max runtime 7 days, 5 nodes, 16,800 SUs required)</td>
</tr>
<tr>
<td></td>
<td>NCBI BLAST+ blast 16,800 SUs. Search protein database with translated nucleotide query sequence(s) (max runtime 7 days, 5 nodes, 16,800 SUs required)</td>
</tr>
<tr>
<td></td>
<td>NCBI BLAST+ blast 16,800 SUs. Search translated nucleotide database with translated nucleotide query sequence(s) (max runtime 7 days, 5 nodes, 16,800 SUs required)</td>
</tr>
</tbody>
</table>

**Galaxy jobs cannot be restarted and checkpoints are not supported**
Permanently Delete Unwanted Files

- delete file
- show deleted files
- permanently delete file
- hide deleted files

Example history:

This dataset has been deleted.
Undelete it
Permanently remove it from disk.
History Divider

Used to add a spacer between distinct job files so you can see which files were created with each job

Run the History Divider tool between jobs
cluttered history items

What it does

This tool just adds a spacer between distinct jobs in your Galaxy history panel.

divided history items

What it does

This tool just adds a spacer between distinct jobs in your Galaxy history panel.
Queued Jobs

- Make sure you have enough SUs to run the job
  - My HPRC SU Balance
- Make sure your account is renewed for current fiscal year
  - My HPRC SU Balance
- Check to see if there is an Ada maintenance scheduled
- Check hprc.tamu.edu to see if Ada node usage is high >95%
Failed Jobs

- Check the stderr file link
- Check error log
- Read the email you received to see if job requires more memory or more time

TERM_MEMLIMIT: job killed after reaching LSF memory usage limit.
Galaxy Notes

## Maroon Galaxy Accounts

Maroon Galaxy is available to students, faculty and staff for research use.

Before you request an account on Maroon Galaxy, you must do the following:

- Go to usegalaxy.org and get familiar with Galaxy. You can start with a free account and learn about Galaxy tools.
- Request a Maroon Galaxy account only if you have data to analyze, otherwise use FishCamp Galaxy for training and practice.
- If you decide that Galaxy is a good choice for your research project, then doing the following:
  - Establish an account on Ada by sending a request. See the NewUser page for details on how to request an account.
  - After you have your Ada account approved, send an email to help@hprc.tamu.edu requesting an account on Galaxy
  - Send us information about what type of data you will be analyzing and which tools you expect to use for your research project.

If you are off campus then you will have to install and run the TAMU VPN to connect to Maroon Galaxy.

Maroon Galaxy can be accessed using your favorite web browser such as Firefox, Chrome or IE.

https://hprcgalaxy.tamu.edu/maroon/

## Set Your Default Account

If you would like to change your default service unit (SU) account:

- Go to our portal website at: portal.hprc.tamu.edu
- After you login with your netid and password, click Clusters -> Ada Shell Access
  - Then enter your password when prompted.
- Then type in the following to see your accounts:
  - myproject
- Then use the following command and replace the ### with your project account from the myproject command output:
  - myproject -d ###
- Then verify that there is a Y in the 'Default' column of your account by typing:
  - myproject

[link: hprc.tamu.edu/wiki/SW:Galaxy]
usegalaxy.org Exercise

1. Add fasta file for S. cerevisiae chromosome I from genome.ucsc.edu to your current history and set the “Database/Build” attribute to the proper genome assembly
2. Run the “Compute Sequence Length” tool to get length of chromosome I
3. Add fasta file for chromosome M from genome.ucsc.edu to your history
4. Concatenate the two files into one file
5. Run the “Compute Sequence Length” tool on the concatenated file
6. Permanently delete the file you created in step 2.
Thank you.

Any question?