



GeneLab Analysis Working Group kick-off meeting

Sylvain Costes, Ph.D.
Project Manager





Agenda



- Goals to achieve for GeneLab AWG
 - GL vision
 - Review of GeneLab AWG charter
- Timeline and milestones for 2018
- Logistics
 - Monthly Meeting
 - Workshop
 - Internship
 - ASGSR
- Introduction of team leads and goals of each group
- Introduction of all members
- Q/A



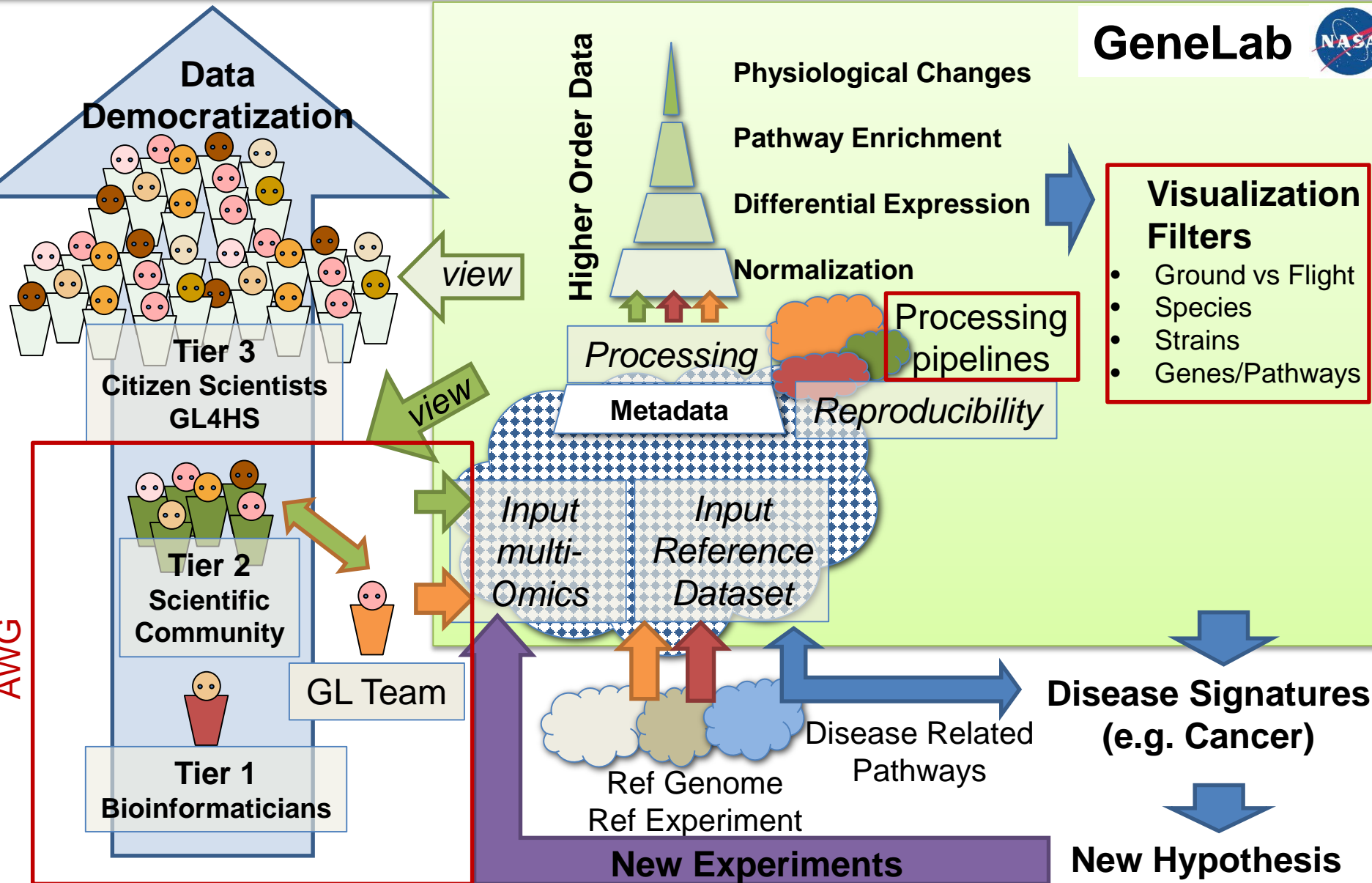
GeneLab

Open Science for Exploration

Three-tier Client Strategy to Democratize Data



GeneLab



Data federation/integration with heterogeneous bioinformatics external databases (GEO, PRIDE, MG-RAST)



Home Repository Data Data Mining Tools Submit Data Help Workspa

mouse myostatin x Q

All GeneLab NIH GEO EBI PRIDE ANL MG-RAST

Search results for: **mouse myostatin** using filter(s):

Sort by Relevance 25

Myostatin inactivation effects on myogenesis in vitro and in vivo
http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28986

Key words: dystrophin, mdx mouse, Duchenne, fibrosis, dystrophy ABSTRACT Stem (MDSC) into myogenic, as opposed to lipofibrogenic, lineages is a promising therapeutic counteracting myostatin, a negative regulator of muscle mass and a pro-lipofibrotic fibrogenic capacity of MDSC from wild...

Organism: Mus musculus Accession: GSE28986 PI/Contact: Robert Gelfand Re

The transcriptomic signature of myostatin inhibitory influence on the differentia
http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE59674

GDF8 (myostatin) is a unique cytokine strongly affecting the skeletal muscle phenoty molecular mechanism of myostatin influence on the differentiation of mouse C2C12 m technique. Treatment with exogenous GDF8 strongly affected the growth and devel proliferation and differentiatio...

Organism: Mus musculus Accession: GSE59674 PI/Contact: Zofia Wick Releas

Development of gene expression signature for defining the cell potency of mu
genotypes
http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE39765

In order to determine the cell potency, by identification of genes responsible for plur isolated from five week old male wild type(WT), C57B6J and another hypertrophied microarray analysis and compared this gene expression to that of a standard mouse and Mstn null mice using an esta...

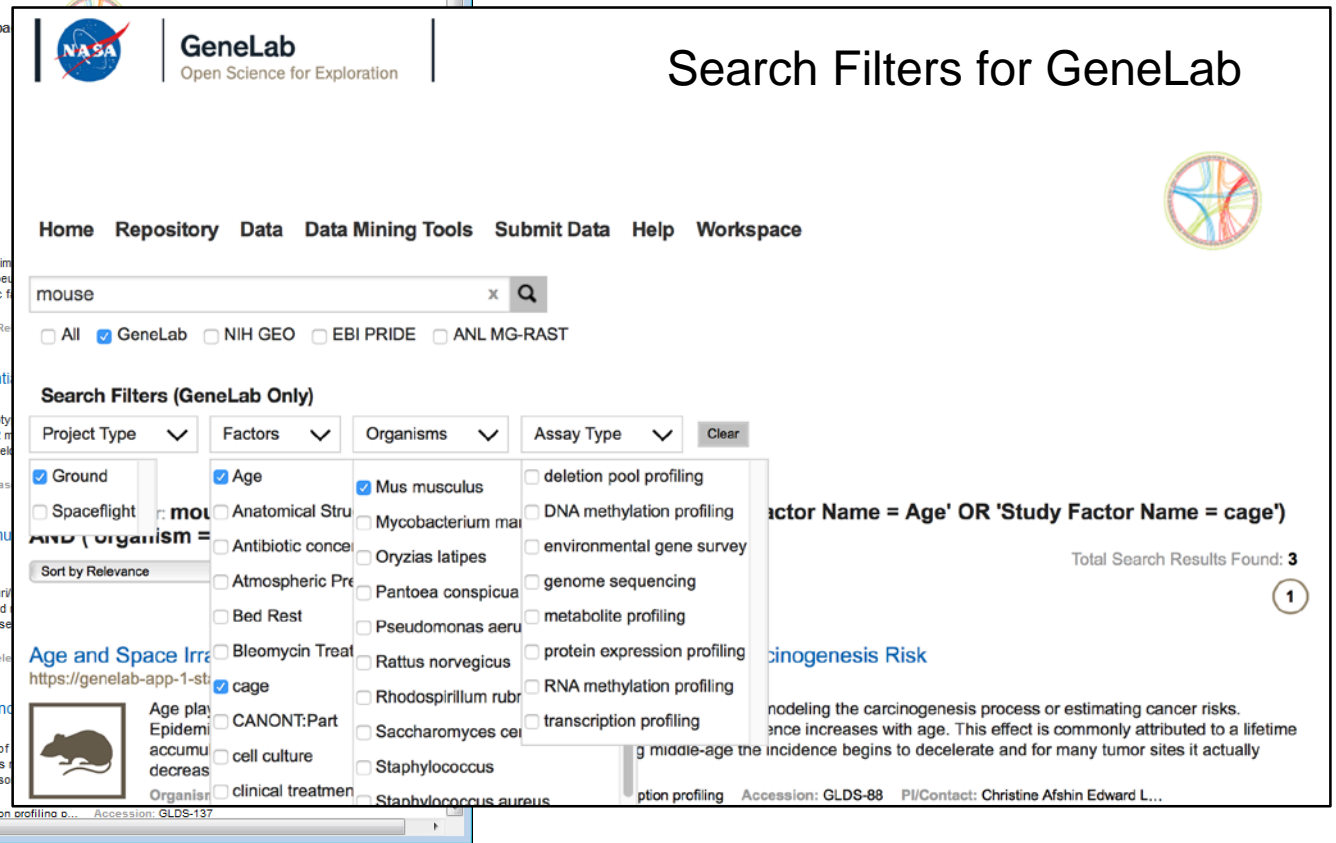
Organism: Mus musculus Accession: GSE39765 PI/Contact: Bipasha Bose Rele

Rodent Research-3-CASIS: Mouse liver transcriptomic proteomic and epigen
https://genelab-data.nsl.nasa.gov/genelab/accession/GLDS-137

The Rodent Research-3 (RR-3) mission was designed to study the effectiveness of occurs during spaceflight. Myostatin is a protein secreted by myoblasts that inhibits block myostatin cause increases in muscle mass. The RR-3 experiment was sponso Advancement of Science in Space and ass...

Organism: Mus musculus Factor: Microarray Treatment Assay Type: transcription profiling p... Accession: GLDS-137

Federated Search



Home Repository Data Data Mining Tools Submit Data Help Workspace

mouse x Q

All GeneLab NIH GEO EBI PRIDE ANL MG-RAST

Search Filters (GeneLab Only)

Project Type Factors Organisms Assay Type Clear

Ground Spaceflight mou

Age Anatomical Stru Antibiotic conce Atmospheric Pre Bed Rest Bleomycin Treat cage CANONT:Part cell culture clinical treatment

Mus musculus Mycobacterium ma Oryzias latipes Pantoea conspicua Pseudomonas aeru Rattus norvegicus Rhodospirillum rubr Saccharomyces cel Staphylococcus Stanhylococcus aureus

deletion pool profiling DNA methylation profiling environmental gene survey genome sequencing metabolite profiling protein expression profiling RNA methylation profiling transcription profiling

Factor Name = Age' OR 'Study Factor Name = cage'

Total Search Results Found: 3

1

carcinogenesis Risk
modeling the carcinogenesis process or estimating cancer risks. nce increases with age. This effect is commonly attributed to a lifetime g middle-age the incidence begins to decelerate and for many tumor sites it actually ption profiling Accession: GLDS-88 PI/Contact: Christine Afshin Edward L...



GLDS Phase 2 (Release 2.0) Customized NASA Collaborative Workspace



User Account Mgmt., Access Controls (e.g., Private, Shared, Public Folders)

The image displays three overlapping screenshots of the GeneLab web interface:

- Top Left:** The main GeneLab homepage. It features a search bar, navigation tabs (Home, Repository, Data, Data Mining Tools, Submit Data, Contact Us, Workspace), and a list of study results. One study is highlighted: "Dissecting Low Atmospheric Pressure Stress: Transcriptome Responses to the Components of Hypobaria in Arabidopsis [Experiment 2]".
- Top Right:** The login page titled "NASA GeneLab-GenomeSpace OpenID Login". It includes fields for "USERNAME:" and "PASSWORD:", "Sign In" and "Cancel" buttons, and a "Register new NASA GeneLab user" link.
- Bottom:** A file browser interface for the "genelab-data" repository. It shows a tree view on the left and a table of files on the right. The table columns are "Filename", "Tags", "Owner", "Size", and "Last Modified".

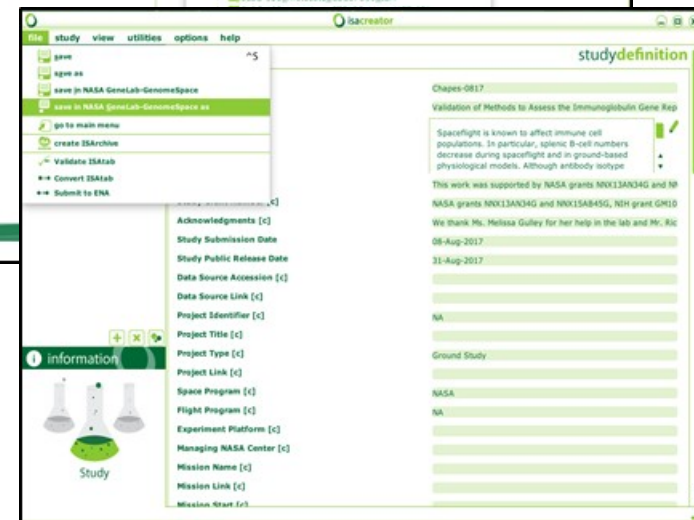
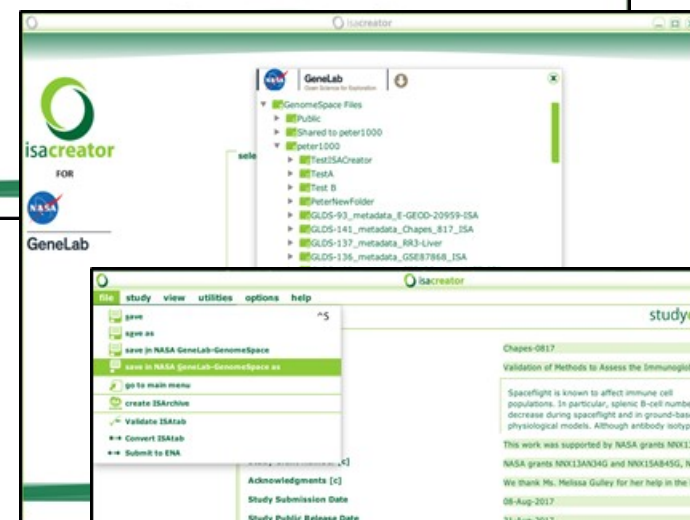
Filename	Tags	Owner	Size	Last Modified
GLDS-1		genelab		
GLDS-10		genelab		
GLDS-100		genelab		
GLDS-101		genelab		
GLDS-102		genelab		
GLDS-103		genelab		
GLDS-104		genelab		
GLDS-105		genelab		
GLDS-106		genelab		
GLDS-107		genelab		
GLDS-108		genelab		
GLDS-109		genelab		
GLDS-11		genelab		
GLDS-110		genelab		
GLDS-111		genelab		
GLDS-112		genelab		
GLDS-113		genelab		



GLDS Phase 2 (Release 2.0) Metadata Curation via ISACreator Tool



GeneLab-GenomeSpace Integration with ISACreator for Streamlining Data Processing Operations



Metadata Source Mappings



GeneLab **Analysis Working Groups (AWG)** will be tasked with analyzing all data across the GLDS with relevance to a specific domain to generate higher-order data.

Goals:

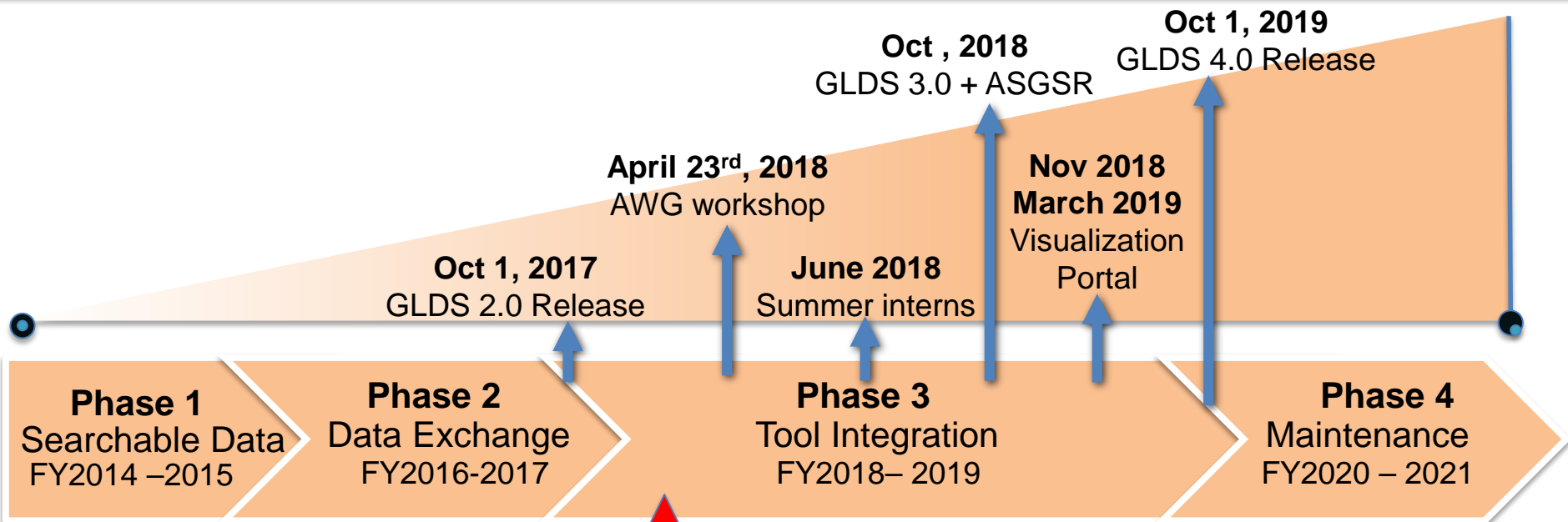
1. Peer-reviewed publications describing AWG's comprehensive analysis.
2. Consensus data analysis pipelines relevant to AWG domains to be used on the GLDS will help domains harmonize their data.
 - a) Summer interns will process all data based on AWG recommendation
 - b) Processed "higher-order" data relevant to domains will be posted on the GLDS.
 - c) Strategies needed to link metadata to processed data will be put in place for the visualization portal deployment
3. Critiques of the GLDS to be used for improving its utility; test driving passed along to scientific community via the AWG
 - a) Access to galaxy toolshed and Jupyterlab notebook within GeneLab provided with CPU and RAM AWS resources
 - b) Integration of GenomeSpace workspace with processing tools
 - c) GLDS 2.0 search query needs to be improved – What should we do different?

AWGs emphasis:

1. Mammalian
2. Invertebrate
3. Plant
4. Microbial
5. Multi-omics systems biology challenge



Phased Implementation



Data System

- ✓ Public Website
- ✓ Searchable Data Repository
- ✓ Top Level Requirements
- ✓ New Data and Legacy Data

Data System

- ✓ Link to Public Databases via Data Federation
- ✓ Integrated Search (e.g., data mashup)

Data System

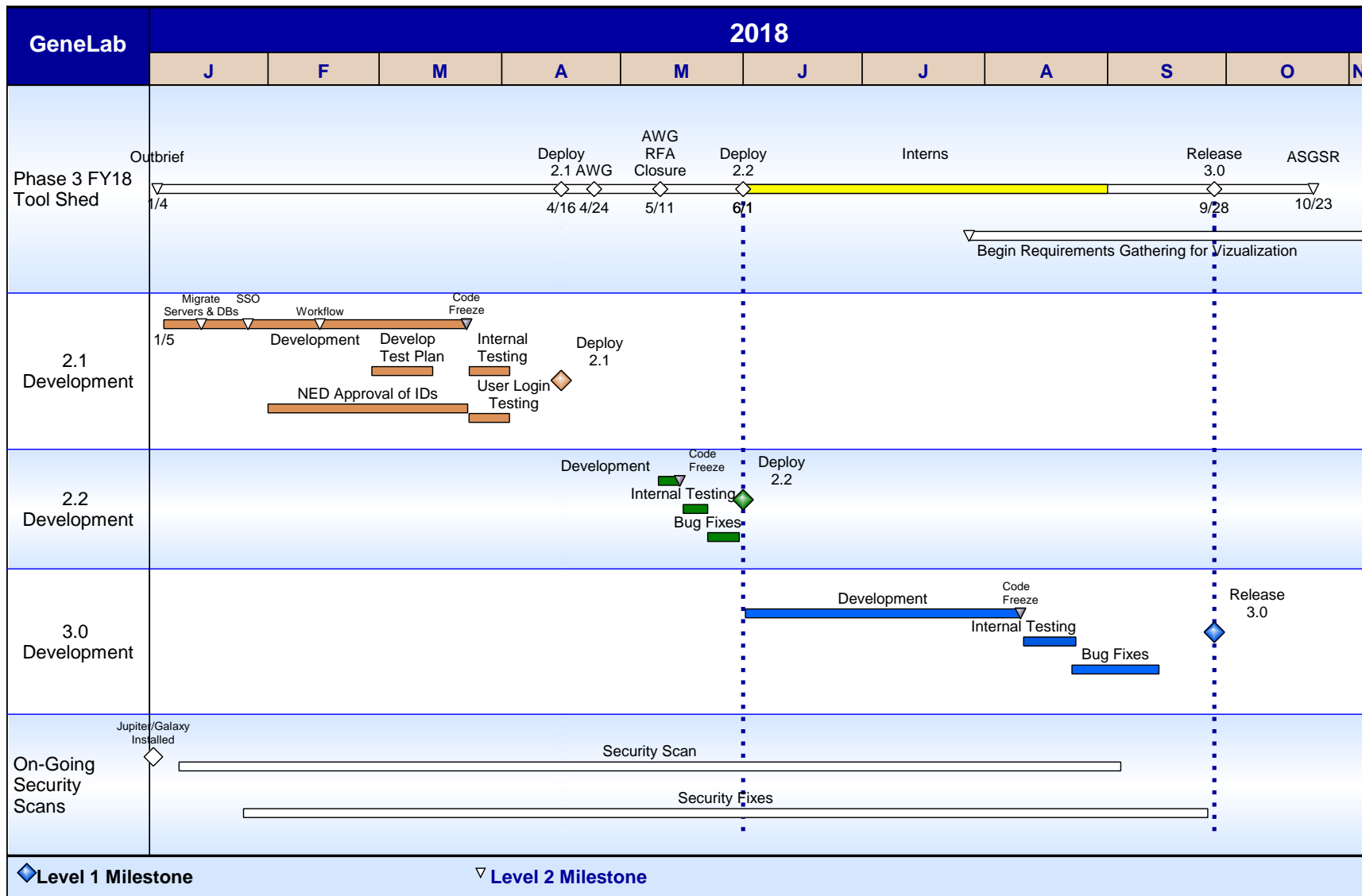
- Integrated Platform across model organisms
- Build Community via AWG
- Provide access to biocomputational tools for omics analysis
- Provide collaboration framework and tools

Open Source Maintenance

- User community becomes primary provider of new tools/knowledge
- Maintain integrity of data, and data system



Detailed Timeline for 2018





AWG Charter



- Please add...



- Monthly Meetings
 - Logistics details please...
 - Mention need for members to get their credential sorted out for Galaxy account
 - Mention need for members to get acquainted with the data
 - Make sure all data are available within the toolshed
 - Communication – mention ResearchGate page



GeneLab will host a two-day in-person workshop in April to finalize the establishment of processing pipelines. In this workshop, each group will demonstrate the pipeline that will be used to analyze the datasets in the GeneLab Data Repository.

When: April 23-24, 2014

Where: Orlando, FL

Workshop agenda- with times are expected arrival

GeneLab will cover travel and accommodations cost for members to attend this two day conference. Spaces are limited and available on first come first served basis. **An email will be sent out after this meeting for confirmation of attendance.**



GeneLab

Open Science for Exploration

Student Internship Program



GeneLab will host students for 10 weeks during the summer of 2018 at NASA Ames Research Center. Students will be working on processing the data using the pipelines established by the AWG.

We encourage AWG members to recommend their students to apply. GeneLab cannot guarantee acceptance due to AWG participation but will preference students with GeneLab data processing experience. Applications will be managed by the NASA OSSI system, which has [basic eligibility requirements](#).

When the posting is public, we will notify all AWG members.



- Please add slides for each group
 - Lead, support people, theme
 - List of confirmed members
 - Sam, can you take care of my group please, Thx!
 - Make sure final slide summarizes the key points (AWG workshop, internship, whatever logistics needed so it stays on while we have the Q/A)