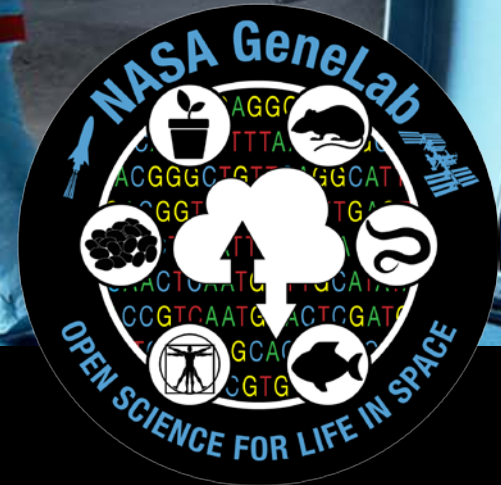




NASA's GeneLab: An Integrated Omics Data Commons and Workbench

Daniel C. Berrios, MD MPH PhD
Sylvain V. Costes PhD
Peter B. Tran, MS

USRA & NASA Ames Research Center
Moffett Field, CA USA





Agenda



- Motivation for GeneLab
- GeneLab Data Systems
- Demonstration
 - Data Repository
 - Collaborative Workspace
 - Biocomputation
- Steps Forward



Motivation

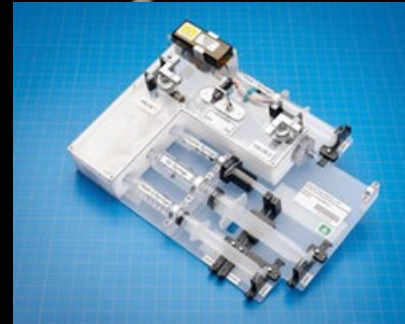
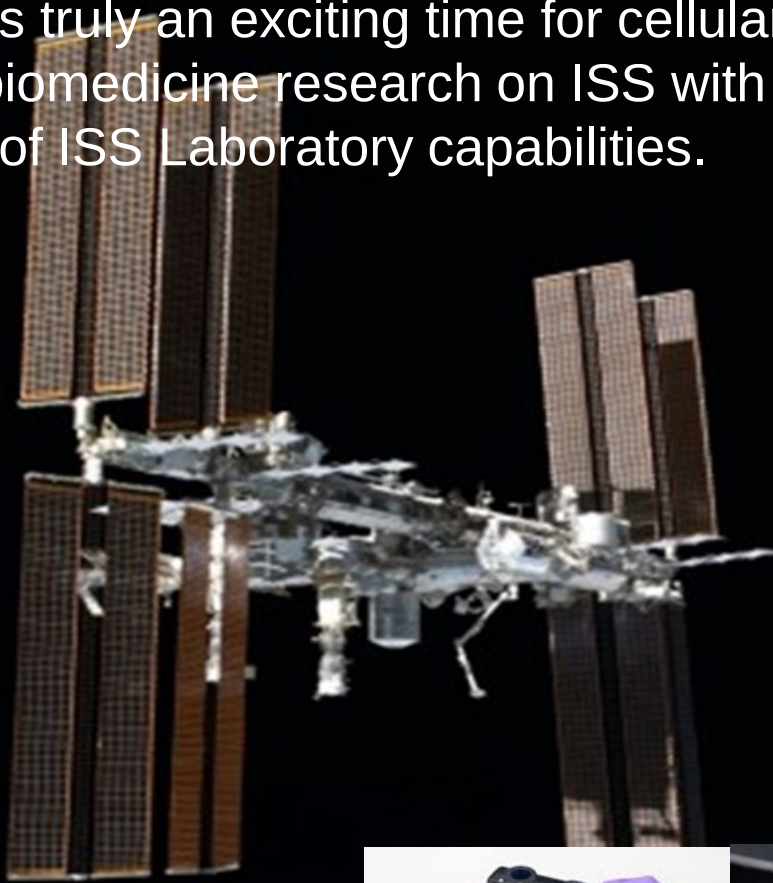


- Primary Goal: Help investigators
 - corroborate findings from “omics” (genomics, transcriptomics, proteomics, and metabolomics) assays
 - translate them into systems biology knowledge (and eventually therapeutics, including countermeasures to support life in space)
- Increased Biological Sample and Data Sharing
 - Extremely high costs of experimentation in space
 - More data from fewer samples
- Synergy of Scientific Investigation
 - Help ground-based investigations analyze and correlate findings with spaceflight-relevant investigations



Omics Acquisition in Space is Now a Reality

This is truly an exciting time for cellular and molecular biology, omics and biomedicine research on ISS with these amazing additions to the suite of ISS Laboratory capabilities.

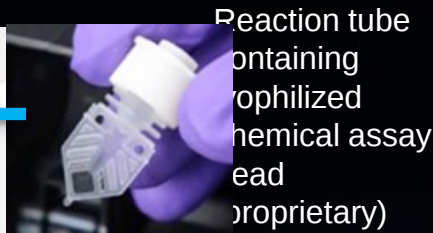


Sample Preparation Module

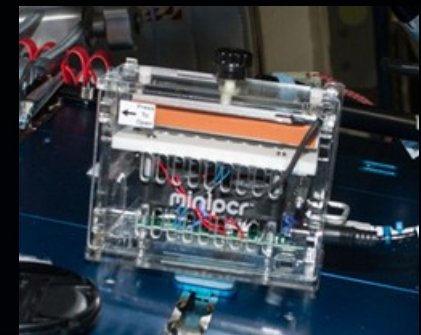


Oxford Nanopore MinION Gene Sequencer

Cepheid Smart Cycler qRT-PCR



Reaction tube containing lyophilized chemical assay bead (proprietary)

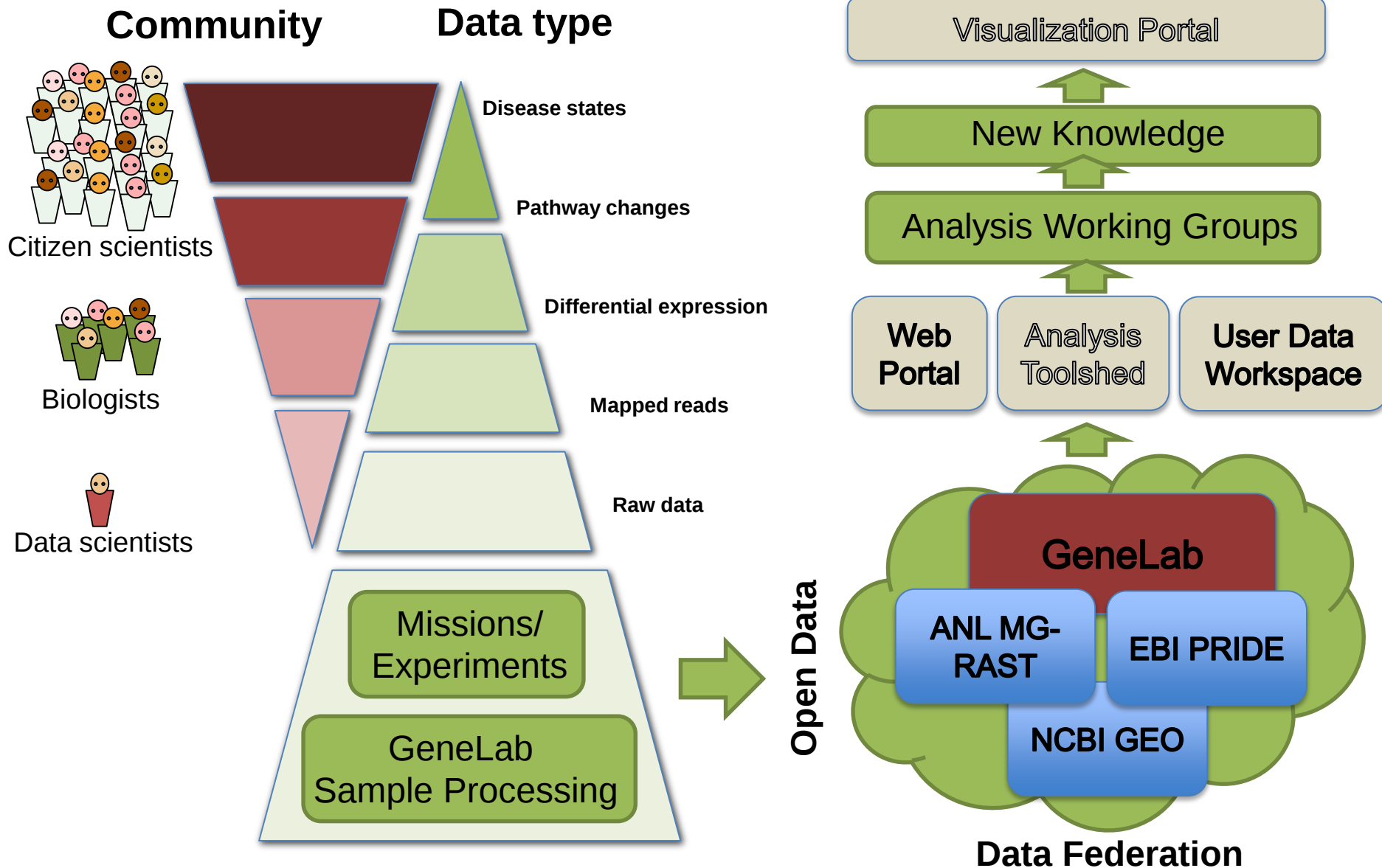


Mini-PCR

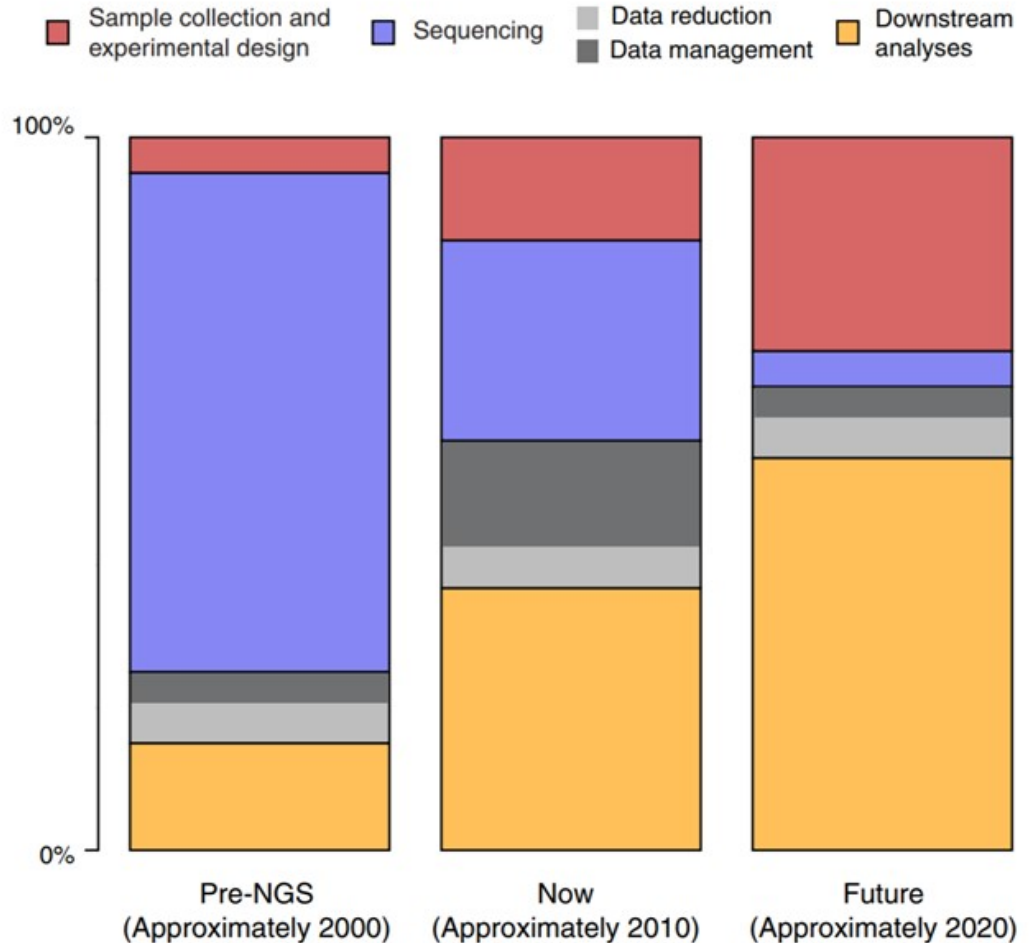
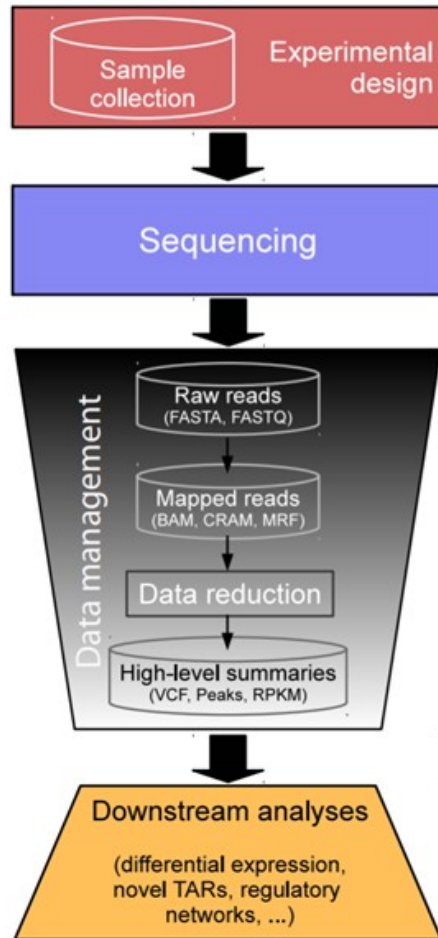


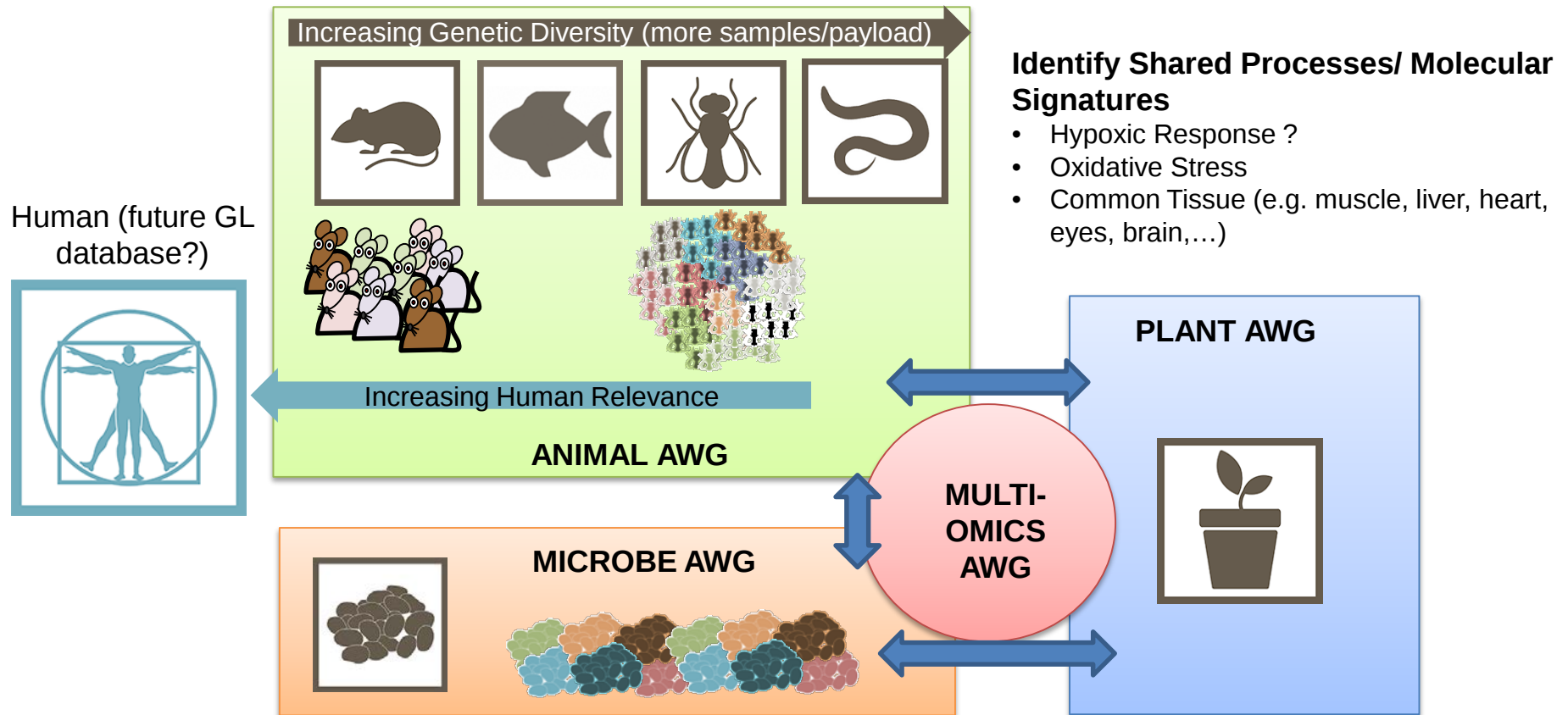
GeneLab Overview video



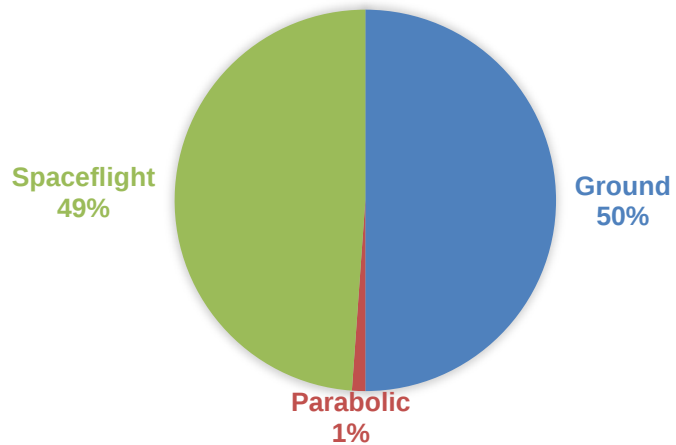


The real cost of OMICS

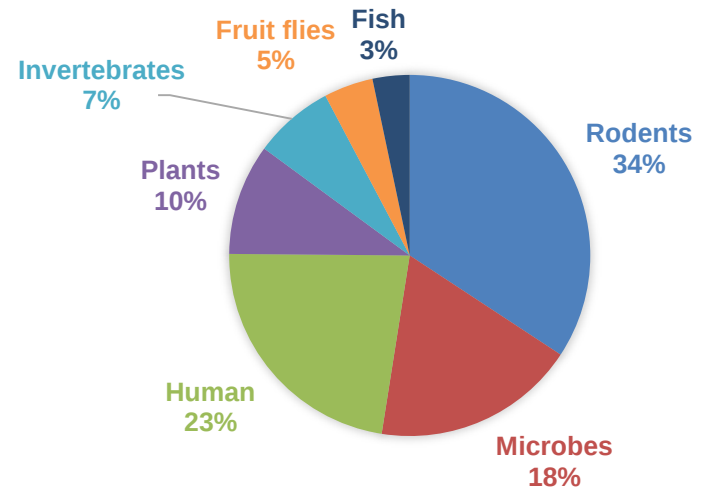




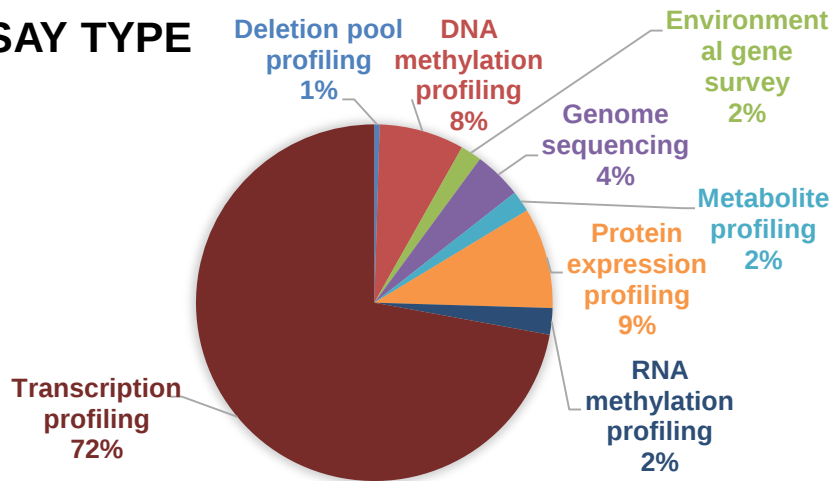
STUDY TYPE



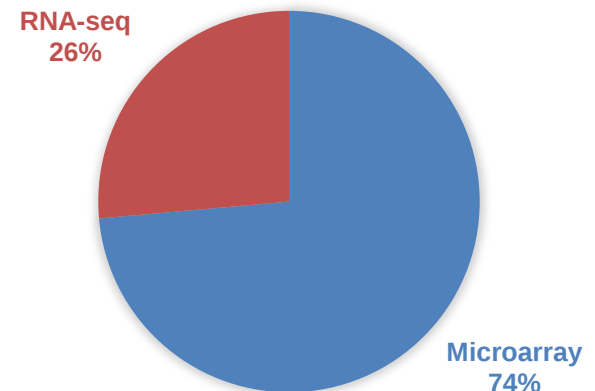
ORGANISM



ASSAY TYPE

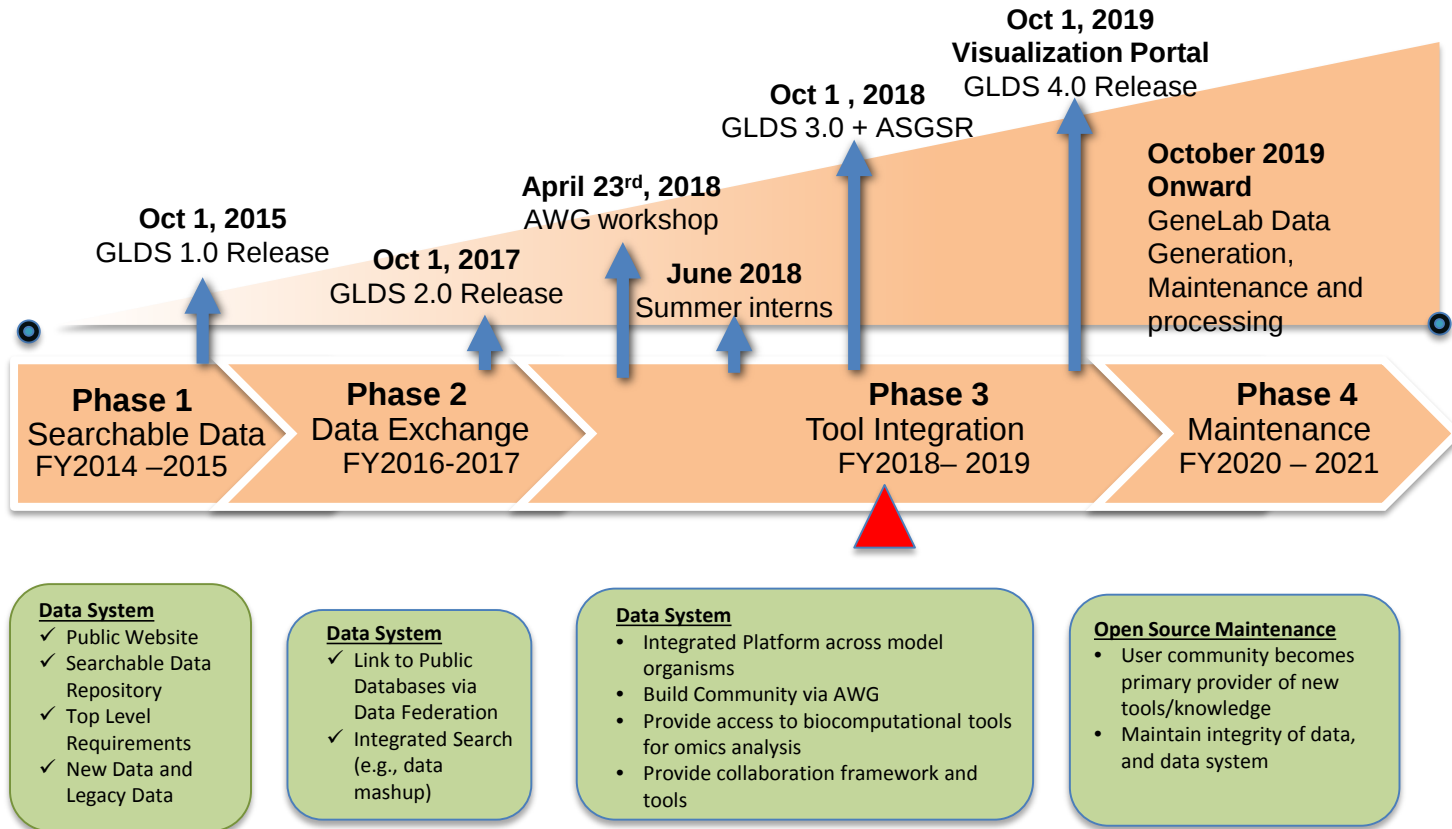


TRANSCRIPTION PROFILING

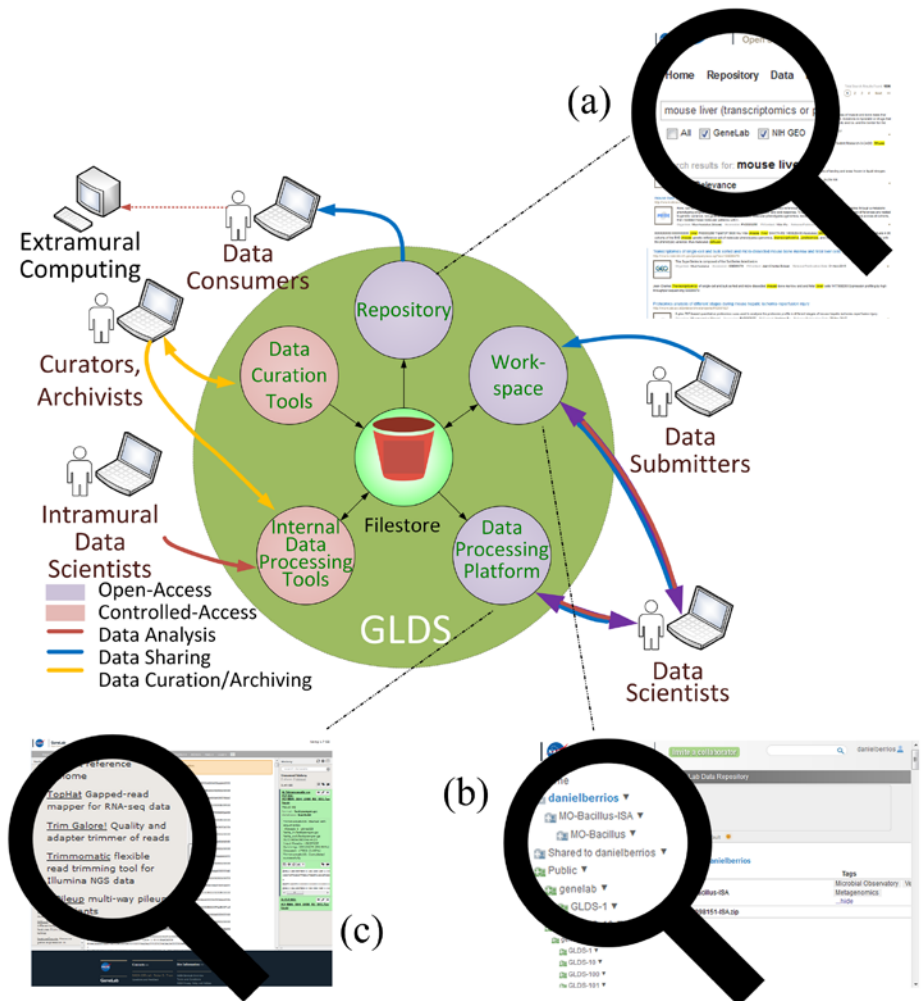


Total # of studies: 178

Phased Implementation



- Single platform for Omics data archiving, sharing, and analysis
 - Diversity of users: citizen scientists, students, educators, NASA-funded PIs, other investigators
 - Diversity of usage: data browsing to data analysis, data submission, data set curation
 - Open access, with private data option (for pre-publication)





GLDS Repository Phase I



GeneLab

Open Science
for Exploration

Home Studies Submit Data Help Search:



All Studies

page 1 of 1



Root transcriptome remodeling of Arabidopsis in response to high levels of magnesium sulfate

Organisms	Factors	Assay Types	Release Date	Description
Arabidopsis thaliana	treatment time genotype treatment	transcription profiling	Sep-17-2010	Martian regolith (unconsolidated surface material) is a potential medium for plant growth in bioregenerative life support systems during manned missions on Mars. However, hydrated magnesium sulfate mineral levels in the regolith of Mars can...



Transcription profiling of activated human T cells induced by microgravity to identify apoptotic genes and other immune response genes

Organisms	Factors	Assay Types	Release Date	Description
Homo sapiens	simulated microgravity	transcription profiling	Feb-09-2006	The purpose of this study was to search for microgravity-sensitive genes, specifically for apoptotic genes influenced by the microgravity environment and other genes related to immune response. Experiment Overall Design: Two-group design wi...



Spaceflight adaptation requires organ specific alterations in the proteomes of Arabidopsis

Organisms	Factors	Assay Types	Release Date	Description
Arabidopsis thaliana	spaceflight	protein expression profiling	Mar-24-2015	Life in spaceflight demonstrates remarkable adaptive processes within the specialized environments of space vehicles which are subject to the myriad of attending and unique environmental issues associated with orbital trajectories. To exami...

Root transcriptome remodeling of Arabidopsis in response to high levels of magnesium sulfate



2 Datasets available:

ISA-TAB Metadata file
Microarray Data Files

GeneLab Accession Number	GLDS-779		
Source Accession Number	E-GEOD-21019		
Contacts	Name	Role	Organization
	Anne Visscher	submitter	
	Anne Visscher		
	Anna-Lisa Paul		
	Robert Ferl		
	Matias Kirst		
	Charles Guy		
Submission Date	Jan-29-2010		
Public Release Date	Sep-17-2010		
Study Description	Martian regolith (unconsolidated surface material) is a potential medium for plant growth in bioregenerative life support systems during manned missions on Mars. However, hydrated magnesium sulfate mineral levels in the regolith of Mars can reach as high as 10 wt%, and would be expected to be highly inhibitory to plant growth. A global approach was used to identify novel genes with potential to enhance tolerance to high MgSO ₄ stress. The early Arabidopsis root transcriptome response to elevated concentrations of magnesium sulfate was characterized in col-0, and also between col-0 and the mutant line cax1-1 a mutant relatively tolerant of high levels of MgSO ₄ -7H ₂ O in soil solution. After 3 weeks of growth under hydroponic conditions, Arabidopsis thaliana col-0 roots were exposed to a basic nutrient solution (0.25 g/L MES, 1164 M ₂ , pH 5.7) with an additional 2.05 mM magnesium sulfate (total Ca:Mg ratio = 1:15) for 45 min., 90 min., or 180 min., while a col-0 control set was exposed to the basic nutrient solution without additional magnesium sulfate for 45 minutes. Arabidopsis thaliana cax1-1 roots were exposed to the basic nutrient solution with additional magnesium sulfate for 180 min. Only four replicate containers were harvested for the control and each of the treatment sets, resulting in a total of 20 samples. Gene expression of the col-0 sets exposed to magnesium sulfate treatment for 45 min., 90 min., or 180 min. was compared to gene expression of the col-0 control set. Gene expression of the cax1-1 set exposed to magnesium sulfate treatment for 180 min. was compared to gene expression of the col-0 set exposed to magnesium sulfate treatment for 180 minutes.		
Organisms	Arabidopsis thaliana		
Study Design Factor(s)	Factor	Ontology: Concept	
	treatment time	time	
	genotype	genotype	
	treatment	treatment	
Assay(s)	Assay Type	Device Type	Device Platform
	transcription profiling	DNA microarray	
Project	Project Identifier	NA	
	Project Link	http://fsls.jsc.nasa.gov/docs/forResearchers/for_researchers_home.aspx	
	Project Type	Ground Study	
	Flight Program	NA	
	Experiment Platform	NA	
	Space Program	NA	
	Managing NASA Center	NA	
	Mission Name	Start Date	End Date
	NA	NA	NA
	Mission Link		
	http://fsls.jsc.nasa.gov/cf/scripts/mis/mis_search_start_adv.cfm		
Publications	Visscher AM, Paul AL, Kirst M, Guy CL, Schuerger AC, Ferl RJ. Growth performance and root transcriptome remodeling of Arabidopsis in response to Mars-like levels of magnesium sulfate. PubMed ID 20808807, DOI 20808807		



GLDS Repository Phase II



Multiple Tabs User Interfaces for Viewing Detailed GeneLab MetaData

**GLDS-137: Rodent Research-3-CASIS: Mouse liver transcriptomic, proteomic, and epigenomic data**

Source Accession Number
Total Data Volume: 1.7 TB

Submitted Date: 04-Aug-2017
Release Date: 28-Aug-2017

DESCRIPTION		PROTOCOLS		SAMPLES		ASSAYS		PUBLICATIONS		STUDY FILES											
DESCRIPTION																					
Study Description		The Rodent Research-3 (RR-3) mission was designed to study bone mass that occurs during spaceflight. Myostatin is a protein that inhibits muscle growth. Mutations in myostatin, or drugs that block myostatin, have been used by pharmaceutical company Eli Lilly and Co. and the Center for Space and Human Health to prevent skeletal muscle atrophy and weakness in mice flown to the ISS and housed in the AEM-X habitat for 10 days. The mice were housed in identical hardware and matching ISS environment. The mice of the experiment, spaceflight, ground controls and baseline were sacrificed and carcasses were stored at -80 C until BSP (Biospecimen Processing) analysis from flight, ground and basal (pre-flight) mice were processed exclusively by Eli-Lilly.																			
Contacts		<table><thead><tr><th>NAME</th><th>ROLE</th></tr></thead><tbody><tr><td>Ruth Globus</td><td>Project Scientist</td></tr><tr><td>NASA GeneLab</td><td>Sample Processing</td></tr><tr><td>Rosamund Smith</td><td>Principal Investigator</td></tr><tr><td>Marty Cramer</td><td>Co-Investigator</td></tr></tbody></table>										NAME	ROLE	Ruth Globus	Project Scientist	NASA GeneLab	Sample Processing	Rosamund Smith	Principal Investigator	Marty Cramer	Co-Investigator
NAME	ROLE																				
Ruth Globus	Project Scientist																				
NASA GeneLab	Sample Processing																				
Rosamund Smith	Principal Investigator																				
Marty Cramer	Co-Investigator																				
Study Design Factor(s)		<table><thead><tr><th>FACTOR</th><th>ORGANISM</th></tr></thead><tbody><tr><td>Microgravity</td><td>Weightlessness</td></tr><tr><td>Treatment</td><td>Treatments</td></tr></tbody></table>										FACTOR	ORGANISM	Microgravity	Weightlessness	Treatment	Treatments				
FACTOR	ORGANISM																				
Microgravity	Weightlessness																				
Treatment	Treatments																				
Organisms		Mus musculus																			
Assay(s)		<table><thead><tr><th>ASSAY TYPE</th><th>DEVICE</th></tr></thead><tbody><tr><td>transcription profiling</td><td>nucleotide sequencing</td></tr><tr><td>protein expression profiling</td><td>mass spectrometry</td></tr><tr><td>DNA methylation profiling</td><td>nucleotide sequencing</td></tr></tbody></table>										ASSAY TYPE	DEVICE	transcription profiling	nucleotide sequencing	protein expression profiling	mass spectrometry	DNA methylation profiling	nucleotide sequencing		
ASSAY TYPE	DEVICE																				
transcription profiling	nucleotide sequencing																				
protein expression profiling	mass spectrometry																				
DNA methylation profiling	nucleotide sequencing																				



GLDS-137: Rodent Research-3-CASIS: Mouse liver transcriptomics

Source Accession Number
Total Data Volume: 1.7 TB

DESCRIPTION		PROTOCOLS		SAMPLES			
FILES		:	FILE SIZE	:	RESOURCE CATEGORY	:	
GLDS-137_metadata_RR3-Liver.zip		:	9.43 KB	:	metadata	:	S
GLDS-137_transcriptomics_RR3-BSL-B1.tar.gz		:	13.86 GB	:	transcriptomics	:	R
GLDS-137_transcriptomics_RR3-BSL-B2.tar.gz		:	14.07 GB	:	transcriptomics	:	R
GLDS-137_transcriptomics_RR3-BSL-B3.tar.gz		:	12.54 GB	:	transcriptomics	:	R
GLDS-137_transcriptomics_RR3-BSL-B4.tar.gz		:	13.7 GB	:	transcriptomics	:	R
GLDS-137_transcriptomics_RR3-BSL-B6.tar.gz		:	12.38 GB	:	transcriptomics	:	R
GLDS-137_transcriptomics_RR3-BSL-B7.tar.gz		:	12.6 GB	:	transcriptomics	:	R
GLDS-137_transcriptomics_RR3-FLT-F1.tar.gz		:	11.69 GB	:	transcriptomics	:	R

**GLDS-137: Rodent Research-3-CASIS: Mouse liver transcriptomic, proteomic, and epigenomic data**

Source Accession Number
Total Data Volume: 1.7 TB

Submitted Date: 04-Aug-2017
Release Date: 28-Aug-2017

DESCRIPTION	PROTOCOLS	SAMPLES	ASSAYS	PUBLICATIONS	STUDY FILES
FILES	FILE SIZE	RESOURCE CATEGORY	RESOURCE NAME	RESOURCE DESCRIPTION	
GLDS-137_metadata_RR3-Liver.zip	9.43 KB	metadata	Study Metadata Files	ISA-Tab format file. Tab delimited Investigation, Study, and Assay level metadata. See ISA-tools.org for format details and software.	
GLDS-137_transcriptomics_RR3-BSL-B1.tar.gz	13.86 GB	transcriptomics	RNA-Seq Data	Raw data files in FASTQ format associated with this study. Please view the associated ISA-TAB metadata files for additional details.	
GLDS-137_transcriptomics_RR3-BSL-B2.tar.gz	14.07 GB	transcriptomics	RNA-Seq Data	Raw data files in FASTQ format associated with this study. Please view the associated ISA-TAB metadata files for additional details.	
GLDS-137_transcriptomics_RR3-BSL-B3.tar.gz	12.54 GB	transcriptomics	RNA-Seq Data	Raw data files in FASTQ format associated with this study. Please view the associated ISA-TAB metadata files for additional details.	
GLDS-137_transcriptomics_RR3-BSL-B4.tar.gz	13.7 GB	transcriptomics	RNA-Seq Data	Raw data files in FASTQ format associated with this study. Please view the associated ISA-TAB metadata files for additional details.	
GLDS-137_transcriptomics_RR3-BSL-B6.tar.gz	12.38 GB	transcriptomics	RNA-Seq Data	Raw data files in FASTQ format associated with this study. Please view the associated ISA-TAB metadata files for additional details.	
GLDS-137_transcriptomics_RR3-BSL-B7.tar.gz	12.6 GB	transcriptomics	RNA-Seq Data	Raw data files in FASTQ format associated with this study. Please view the associated ISA-TAB metadata files for additional details.	
GLDS-137_transcriptomics_RR3-FLT-F1.tar.gz	11.69 GB	transcriptomics	RNA-Seq Data	Raw data files in FASTQ format associated with this study. Please view the associated ISA-TAB metadata files for additional details.	
GLDS-137_transcriptomics_RR3-FLT-F2.tar.gz	13.34 GB	transcriptomics	RNA-Seq Data	Raw data files in FASTQ format associated with this study. Please view the associated ISA-TAB metadata files for additional details.	



GLDS Workspace Phase II



GeneLab
Open Science for Exploration

Invite a collaborator

danielberrios

File Launch View Manage Help GeneLab Data Repository

GeneLab_ISACreator_v1.7.10

Default

Up to: Home ▸ danielberrios

Filename	Tags
MO-Bacillus-ISA	Microbial Observatory Ve Metagenomics ...hide
SRP098151-ISA.zip	

- User files
- Public repository files

- Privately Shared files
- Files submitted for Public Sharing (Repository)

Up to: Home ▸ Shared to danielberrios ▸ chapeslab

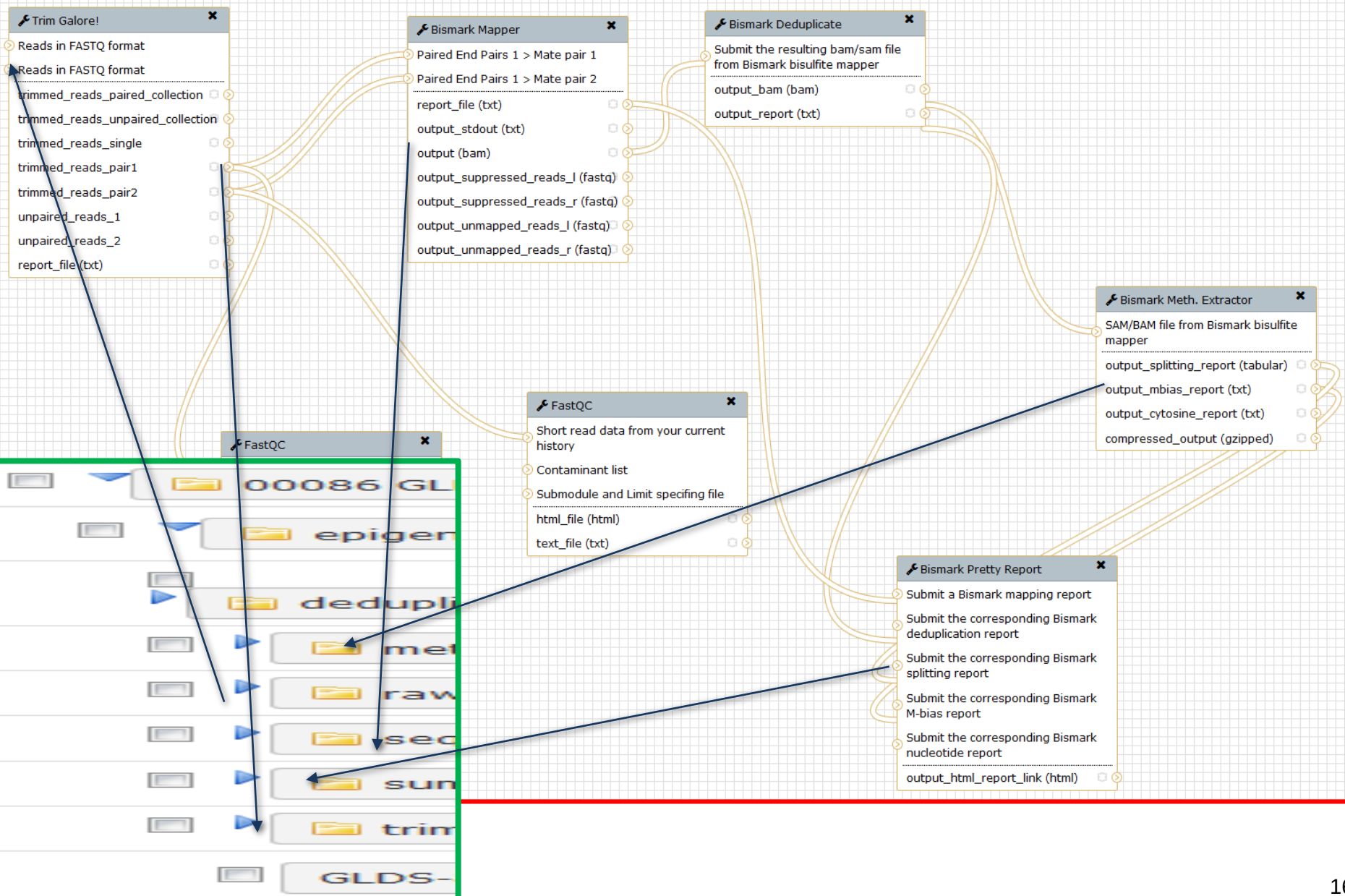
Filename	Tags	Owner	Size	Last Modified
Normal Mouse Pool 3 - March 16.zip		chapeslab	5.62 GB	3 days ago

- Privately Shared files



- Galaxy Server
- Galaxy Toolshed (subset)
- +Specialized Tools
 - (e.g., developed by GeneLab interns)
- Cluster of 5 Dedicated Compute Nodes
 - @8 vCPUs
 - @64 GiB RAM

GLDS-Galaxy Data Processing Workflows & Data Libraries





GLDS Live Demo



- <https://genelab-data.ndc.nasa.gov/genelab/>

GeneLab

D. Pham	D. Heher
J. Costa	S. Caldwell
C. Barreras	K. Staats
S. Dueck	T. Le
M. Geniza	J. Lin
H. Fogle	J. Galazka
A. Saravia-Butler	S. Reinsch
S. Gebre	V. Boyko
S. Ray	Y. Chen
J. Miller	M. Dinh
S. Lai	A. Beheshti
S. Reghnanandan	C. Tahimic
N. Paremski	S. Li
J. Moh Hashim	S. Richards
A. Williamson	



GeneLab is funded through NASA's
Space Life and Physical Sciences
Research Applications Division & the
International Space Station Program.

2018 Interns:

D. Mattox, N. Bense, C. Poling, L. Rutter, C. Erikson,
J. Rubin, A. Uriarte-Acuna, T. Matheny, M. Zajko