GeneLab: Open Science for Life in Space

National Aeronautics and Space Administration









- GeneLab background and motivation
- GeneLab Data System (GLDS) and datasets
 - Data system architecture (AWS, search, federation,)
 - Datasets, dataset metadata, and curation
 - Collaborative workspace
 - Cloud analysis platform
- GeneLab community
 - Analysis Working Groups (AWGs)
- Increasing the scientific return on Space Biology investments
 - Leveraging biomedical research
 - Aggregating Space Biology research
 - Increasing data volume from Space Biology research





- GeneLab background and motivation
- GeneLab Data System (GLDS) and datasets
 - Data system architecture (AWS, search, federation,)
 - Datasets, dataset metadata, and curation
 - Collaborative workspace
 - Cloud analysis platform
- GeneLab community
 - Analysis Working Groups (AWGs)
- Increasing the scientific return on Space Biology investments
 - Leveraging biomedical research
 - Aggregating Space Biology research
 - Increasing data volume from Space Biology research



Challenges of Spaceflight









Humans body: 79 organs, 200 cell types, 20,000 protein coding genes, Many more mRNA splice isoforms, Many more protein isoforms, 16,000 endogenous metabolites (22,000 from food), Plus...DNA, RNA, and protein modifications, histone modifications, protein-protein interaction, etc., etc., etc., We're complicated!



SLPSRA 2011 NRC Decadal Survey and the Sequencing Paradigm Shift





"...**genomics, transcriptomics, proteomics, and metabolomics** offer an immense opportunity to understand the effects of spaceflight on biological systems..."

"...Such techniques generate considerable amounts of **data that can be mined and analyzed** for information by multiple researchers..."



Recapturing a Future for Space Exploration





Omics Acquisition in Space





Sample Preparation Module



Oxford Nanopore MinION Gene Sequencer



Cepheid Smart Cycler qRT-PCR



GeneLab



- Mission: To enable scientific discovery and space exploration through multi-omics data-driven research.
- Currently funded by SLPSRA Space Biology program. Previously received funding from the ISS program.





Overview: Database content









- GeneLab background and motivation
- GeneLab Data System (GLDS) and datasets
 - Data system architecture (AWS, search, federation)
 - Datasets, dataset metadata, and curation
 - Collaborative workspace
 - Cloud analysis platform
- GeneLab community
 - Analysis Working Groups (AWGs)
- Increasing the scientific return on Space Biology investments
 - Leveraging biomedical research
 - Aggregating Space Biology research
 - Increasing data volume from Space Biology research



GeneLab Data Democratization







GLDS Systems Architecture Phase I





11/21/2019



GLDS Systems Architecture Phase II





11/21/2019



GLDS Systems Architecture Phase III







Overview: Database content







69 Ground Data Sets: Radiation and simulated microgravity







Searching the Repository



mouse liver transcripto	omics x Q	
🔲 All 🗷 GeneLab 🛛	mouse liver transcriptomics x Q	
Search Filters (Gen	🖉 All 🕜 GeneLab 🖉 NIH GEO 🕜 EBI PRIDE 🖉 ANL MG-RAST	
Project Type 🗸 🗸	Search results for: mouse liver transcriptomics using filter(s): Total Search Results Found: 20	
	Sort by Relevance	
Search results for: mou	Rodent Research-3-CASIS: Mouse liver transcriptomic proteomic and epigenomic data https://genelab-data.ndc.nasa.gov/genelab/accession/GLDS-137	
Rodent Research-3.	The Rodent Research-3 (RR-3) mission was sponsored by the pharmaceutical company Eli Lilly and Co. and the Center for the Advancement of Science in Space to study the effectiveness of a potential countermeasure for the loss of muscle and bone mass that occurs during spaceflight. Twenty BALB/c 18-weeks old female mice (ten controls and ten treated) were flown to the ISS and housed in the Rodent Habitat for 39-42 days. Twenty mice of similar age sex and strain were used for ground controls housed	
The Rod Space tc weeks of sex and	Corganism: Mus musculus Factor: spacenight freatment Assay type: transcription profiling prot Accession: GLDS-137 Pi/Contact: Ruth Globus, NASA Gen STS-135: Mouse Liver Transcriptomics using RNA-Seq https://genelab-data.ndc.nasa.gov/genelab/accession/GLDS-173	
STS-135: Mouse Lin https://genelab-data.ndc.	Female C57BL/6 mice were flown onboard STS-135 for 13 days and returned to Earth for analysis. Livers were collected within 3-4 hours of landing and snap frozen in liquid nitrogen. Purified RNA samples that were used for microarray analysis for GLDS-25 were provided to GeneLab. GeneLab added ERCC control spike-in to the samples and performed RNA-Seq analysis. Organism: Mus musculus Factor: Space Flight Absorbed R Assay Type: transcription profiling Accession: GLDS-173 PI/Contact: Jonathan Galazka, NAS Release/Publication Date: 04-May-2018	
Female (snap fro; ERCC ci	Transcriptomics of single-cell and bulk sorted and micro-dissected mouse bone marrow and fetal liver cells http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE89379	
Organism	This SuperSeries is composed of the SubSeries listed below. Organism: Mus musculus Accession: GSE89379 PI/Contact: Jean-Charles Boisset Release/Publication Date: 01-Nov-2016	
STS-135 Liver Tran: https://genelab-data.ndc.		
Mice we	mouse liver SWATH-MS http://www.ebi.ac.uk/pride/archive/projects/PXD003266	
Organisn	Here, we have profiled 386 individuals in 80 cohorts of the BXD mouse genetic reference population across two environmental states through a metabolic phenotyping program including glucose response, exercise capacity, and cold response. To understand how the observed phenotypic differences are related to genetic variance, we generated a multilayered set of molecular phenotypes—genomics, transcriptomics, proteomics, and metabolomics across al cohorts, then modeled these molecular patterns with t	I
	Organism: Mus musculus (Mouse) Accession: PXD003266 PI/Contact: Yibo Wu Release/Publication Date: 13-Jun-2016	

DD-MM-YYY





GeneLab recognizes the need to include environmental data to spaceflight experiments. These include but not limited to:

- Radiation dosimetry (High and low LET)
- Temperature
- Humidity
- $-CO_2$ levels
- Payload specific sensors: temperature, light cycle, humidity, CO₂ levels
 - We recommend radiation specific sensors for payloads
- Events in space planned or unplanned*
- Launch and Return environment*

*Not currently in scope but will add if provided or requested by PI













Search Filters (GeneLab Only)

	Project Type	\checkmark	Factors	\checkmark	Organisms	\checkmark	Assay Type	\checkmark	Clear
--	--------------	--------------	---------	--------------	-----------	--------------	------------	--------------	-------

NASA GeneLab Data System

NASA GeneLab expands scientists' access to experiments onboard the International Space Station that explore the molecular response of terrestrial biology to spaceflight environments. Our mission is to maximize the utilization of the valuable biological research resources aboard the International Space Station by collecting genomic, transcriptomic, proteomic, and metabolomic data known as "omics".

Data Repository

The data repository hosts space biology and space-related datasets funded by multiple space agencies around the world. Use the search options above to filter through the GeneLab repository or to search across several databases.

Collaborative Workspace

GeneLab has customized a workspace for file sharing and access to data analysis tools. To access the workspace, users will need to create an account. Data analysis tool integration is ongoing and more tools will be available in future releases.







Environmental Metadata – currently only have radiation data available for STS and BION-M1 mission datasets

Radiation Metadata for GLDS Studies

DESCRIPTION	PROTOCOLS		SAMPLES	ASS	SAYS	PUBLICATIO	DNS	STUDY FI	LES
									Com
Source Name	Sample Name	:	Exposure Dura	tion :	Maxim Absori	um Total : bed Dose	Min Abs	imum Total orbed Dose	:
Mouse Tissue	AEM Control 18		12.76 day						•
Mouse Tissue	AEM Control 28		12.76 day						
Mouse Tissue	AEM Control 32		12.76 day						
Mouse Tissue	AEM Control 36		12.76 day						
Mouse Tissue	AEM Control 44		12.76 day						
Mouse Tissue	Flight 52		12.76 day		5.46 mGy		3.26 mGy	,	
Mouse Tissue	Flight 58		12.76 day		5.46 mGy		3.26 mGy	,	
Mouse Tissue	Flight 60		12.76 day		5.46 mGy		3.26 mGy	,	
Mouse Tissue	Flight 64		12.76 day		5.46 mGy		3.26 mGy	,	
Mouse Tissue	Flight 66		12 76 dav ∢		5 46 mGv		3 26 mGv	,	→ <u>/</u>



Radiation Dosimetry for STS samples (ISS to follow)









STS = Space Transportation System (Shuttle Program)





- GeneLab utilizes the Investigation Study Assay (ISA) model for organizing metadata
- GeneLab participates in ontology groups to help standardize spaceflight ontology





GeneLab Workspace



- Sequencing data can be massive
- Terabytes
- Moving data requires special considerations
- The GeneLab allows user to upload, download and share raw and processed data







space Tools $\mathbf{\Theta}$ search tools Data **Get Data** Analysis Send Data **Epigenomics** Tools **RNA-Seq Other Tools Picard Tool Suite SAMTools Suite Microarray** Sequencing General **Microarray Tools Metagenomics Tools** Workflows All workflows

Hello, NASA GeneLab Analysis Platform is running!

Help - User -

Take an interactive tour: Galaxy UI History

Workflow Shared Data -

Analyze Data

NASA GeneLab's Analysis Platform is an open platform for performing large-scale analysis of space-relevant omics data using a customized suite of data analysis tools, powered by <u>Galaxy</u>.

To access **GeneLab RNAseq processed data** go to 'Shared Data' \rightarrow `Data Libraries' \rightarrow 'GeneLab Data Repository' \rightarrow click on the GLDS dataset you are interested in.

To report any issues with this platform, contact the NASA GeneLab project at: <u>arc-dl-genelab-it@mail.nasa.gov</u>

Now used by universities to teach omics

History	<i>⊷</i> ¥
search datasets	
Unnamed history	
(empty)	
This history is empty. load your own data o from an external sou	. You can r get data rce
Jobs and output fi will be displaye here	d les d



GLDS Public Data Processing Platform



NATER GeneLab Using 1.7 GB

nole 主		_	History
against reference	This dataset is large and only the first megabyte is shown below.		History
genome	Show all Save		search datasets
TopHat Gapped-read		_	Unnamed history
mapper for RNA-seq data	@J00113:264:HGMATEBXX:8:1101:1560:1209 1:N:0:NAACATCG		2 shown, 2 <u>deleted</u>
Trim Galore! Quality and	NCGCCTCGTTCAGCGCTTTGGCGATTGGCGTCCAGCCAGCC		3.49 GB
adapter trimmer of reads	* * * * * * * * * * * * * * * * * * *		
Trimmomatic flexible	@J00113:264:HGMATEBXX:8:1101:1905:1209 1:N:0:NAACATCG		4: Trimmomatic on
read trimming tool for	NTGAGTTACTGCTCTGGTTTGTTTATCACCGTTCATTTGTTTTACAATGTATTCAGCATCACTTCCTACGCCTTGAAGCAACGCTGAAC		A2-RNA S64 L008 R1 00
Illumina NGS data	= +.		ta.gz
MPileup multi-way pileup	= TATEIGAJGGAGFGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG		782.5 MB
of variants	NTTCTTTGCTCTTTTTGCCGAAAATCGCAGAAATGATTCCGATGATCGCTGCAATAATAAGCGGATTGGTCAGTAAATCTTCCATCAGCG		format: fastqsanger.gz,
Reheader copy SAM/BAM	•		database: baciSubt
header between datasets	*AAPAJJJAJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJ		TrimmomaticSE: Started v
<u>Slice</u> BAM by genomic	godulasizeti komki bakan ti laguli zakoli izao 1 ini u ini kakakago NCATCCATAGTICGAGAGATTTI AGULI ZAKOLI KAGAGAGAGAGCTTTAAGAGCAGAACCTTTAACAACTGGTACATCATCA		arguments:
regions	*		-threads 1 -phred33
Stats generate statistics	\$ARF333333333333333333333333333333333333		fasto out.fastosanger.gz
for BAM dataset	@J00113:264:HGMNTBEXX:8:1101:24535:1209 1:N:0:NAACATCG		SLIDINGWINDOW:4:20
Split BAM dataset on			Input Reads: 19037657
readgroups	\$AFF33333333333333333333333333333333333		Surviving: 19019674 (99.9
Sort order of storing	@J00113:264:HGMNTBEXX:8:1101:24596:1209 1:N:0:NAACATCG		TrimmomaticSE: Complet
aligned sequences	NCSCATTTGTGAAAACAAAGAAATTATGGCCGAGCATATTCATTTGCAAGATCGCTTCTTCACTATCCATCGGCTTTAAATTAAAGCGTT		successfully
BedCov calculate read	**************************************		
depth for a set of	@J00113:264:HGMNTBEXX:8:1101:25185:1209 1:N:0:NAACATCG		
genomic intervals	NCAGAACGTTTTGCCGTTTCAACAATCTTCTCTGCAGATTGATCAAGGATTCTATGATCATATGCTTTCAAACGAATACGAATTTTTTGT		@J00113:264:HGW/T88XX:8:1101:1560:120
CalMD recalculate MD/NM	*		NCGCCTCGTTCAGCGCTTTGGCAATTGGCGTCCAGCC
tags	@J00113:264:HGMNTBEXX:8:1101:25246:1209 1:N:0:NAACATCG		•
Id×Stats reports stats of	NGCAGCTTTTCCTTCTCAGTGTAGATCCCATTTTTTATATCCTCTTTAATTCGTTCAATTACTTTCAAATACAAGTGCCGATTATCAGCT		#AAFFF33A333F33333333333333333333333333
the BAM index file	*		g100113:264:HGM/TEEXX:8:1101:1905:120
offread Filters and/or	TARE US USUBUSUBUSUBUSUBUSUBUSUBUSUBUSUBUSUB		
converts GFF3/GTF2	NCTGCCTCTTTAGCTCCTGCCATACCGCCGCCCGCAAGAAGATCCTCATAAGCAACTGTCTCAGCACGAATGAAGCCCCGGCTCAAAGTCA		3: FLT-BS-
records	+		A2-RNA 564 L008 R1 00
edgeR Perform	\$AAFAJJJJJJFJEJJJJJJJJJJJJJJJJJJJJJJJJJJ		tq.qz
differential expression of	gJUUI3:200:HUMNIBAX:0:1101:250/5:1209 1:N:U:MAACATCG NCArCeGTTCAACGGTTGACGATGCTGTTGTTCCGGCGTTTCATCACTGATAATGCCAGGCCGTCAGCTTGTTTCGCATCTTCTTCG		
count data	+		
DESeq2 Determines	\$AAPFJJJFJJJFJJJJJJJJJJJJJJJJJJJJJJJJJJJ		
differentially expressed	@J00113:264:HGMNTBEXX:8:1101:2595:1209 1:N:0:NAACATCG		
reatures from count	+		
	\$AFF533333333333333333333333333333333333		
reatureCounts Measure	@J00113:264:HGMNTBEXX:8:1101:26017:1209 1:N:0:NAACATCG		
gene expression in	NCGGCAAAATCAATTATTCGCCTTTATTTTTTTGATAATTCTTCTGCGACGCTCTTCGGACGCTCTTCGGCACGTCATGGATGCATGGCATAG		
		-	(III

Site Information -

Contacts -

Questions and Feedback

NASA

GeneLab

- **Galaxy Server** •
- Galaxy Toolshed (subset) ullet



GLDS-Galaxy Data Processing Workflows & Data Libraries









- GeneLab background and motivation
- GeneLab Data System (GLDS) and datasets
 - Data system architecture (AWS, search, federation,)
 - Datasets, dataset metadata, and curation
 - Collaborative workspace
 - Cloud analysis platform
- GeneLab community
 - Analysis Working Groups (AWGs)
- Increasing the scientific return on Space Biology investments
 - Leveraging biomedical research
 - Aggregating Space Biology research
 - Increasing data volume from Space Biology research
- Future efforts and conclusion



How can we increase user engagement and









GeneLab Analysis Working Groups (AWGs)



- Modeled on the NIH TCGA Analysis Working Groups
- Analyze a single type of cancer across all TCGA platforms and publish comprehensive analyses
- Have generated very high profile publications



Analysis Working Groups

The TCGA Analysis Working Groups (AWGs) are composed of scientific and clinical experts who study a particular type of cancer. With the support of the TCGA Genome Data Analysis Centers (GDACs) they work together to analyze a single type of cancer across all TCGA platforms and to publish a comprehensive analysis. These papers are authored by The Cancer Genome Atlas Research Network, and a complete list can be found on the TCGA Publications Page.

Coordinated by the TCGA program office, the AWGs are self-assembled teams of scientists from all over the world who bring a range of expertise to the analysis. AWG members come from a spectrum of specialties, including:

- Oncology
- Pathology
- Bioinformatics
- Systems Biology

Members of the AWG agree to a special set of rules called the AWG charter that binds them to keep data learned in the context of AWG discussions confidential until their analysis is published.

An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics Cell. Volume 173 Issue 2: p400-416 Read the full article 화

A Pan-Cancer Analysis of Enhancer Expression in Nearly 9000 Patient Samples Cell. Volume 173 Issue 2: p386-399 Read the full article 법

Comprehensive Characterization of Cancer Driver Genes and Mutations Cell. Volume 173 Issue 2: p371-385 Read the full article ଜ୍ୟ

Pathogenic Germline Variants in 10,389 Adult Cancers Cell. Volume 173 (2): p355-370 Read the full article ଜିମ

Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation Cell. Volume 173 Issue 2: p338-354 Read the full article &





The purpose of the GeneLab Analysis Working Groups (AWGs) is to optimize the processing of raw omics data from the GeneLab repository in order to maximize the gain of new knowledge from such complex datasets. The AWGs will also assess and improve the effectiveness of the GeneLab Data System (GLDS) through intensive utilization of the analytics to be deployed therein.

The AWGs aspire to scientific excellence, and participation in AWGs is strictly on a volunteer basis. The primary activity of each AWG is to establish analytical processes to generate higher order data from data housed in the GLDS with relevance to one or more specific application areas. For example, a mammalian AWG would analyze all data relevant to mammalian systems (human, mouse, rat, etc.). The GeneLab AWGs aspire to yield:

Canonical data analysis pipelines. The interpretation of analysis results is too often confounded by variability in data-processing protocols. This can make comparison of results between studies even more difficult. By developing canonical analysis pipelines, the AWGs will help speed the harmonization of results across space biology experiments

Processed ("higher-order") space biology data. The AWGs will work with GeneLab staff to generate processed data using the canonical analysis pipelines. This higher-order data will be made available to the broader scientific community through the GLDS. This data will serve as a reference to the community on use of the pipelines, and will be fully reproducible using the GLDS and referenced pipelines. Optimal strategies regarding the display of processed data to the scientific community will also be discussed.

Critiques of the GLDS. The value of the GLDS to the scientific community is best judged through extensive use of the systems. As the AWGs will be using the GLDS intensively (see 1 and 2 above), they will be poised to give valuable feedback to GeneLab on its strengths and weaknesses.

Scientific communication of available analyses and interpretations. The AWGs are expected to communicate the development and application of the canonical data analysis pipelines with the broader scientific community. Such communications could take the form of self-published white papers, newsletters, or peer-review publications and conference presentations. These communications will also serve to catalyze the generation of new hypotheses for further experimentation.

AWG Recruitment



 Reached out to members of Space Biology community, scientist at different NASA centers, outside domain experts

- Sent open invitation through GeneLab newsletter
- Posted open invitation on GeneLab website (still active)
- Continue to recruit and accept new members

NASA	Analysis Working Groups	
GeneLab Open Science for Life in Space	GeneLab is a science collaboration initiative and data system effort that provides a workspace with tools to analyze and visualize omics data. GeneLab has established four different Analysis Working Groups with the goal of maximizing the generation of new knowledge from these rare and complex spaceflight-relevant omics datasets within the <u>GeneLab Data System</u> . About & Charter How to join an AWG Internship opportunity AWG workshop	
Home About Access GeneLab Data Analysis Working Groups →	MEMBERS - Plants	(h)
Intern Resources Publications Newsletters FAQ	Multi-Omics/System Biology	•
Contact Us	Microbes Microbes Animals	•

GeneLab is recruiting members for Analysis Working Groups to analyze NASA omics data!

NASA's GeneLab Project is recruiting investigators, bioinformaticians, graduate students, and postdocs to participate in Analysis Working Groups (AWGs) that will investigate specific subsets of omics data from experiments conducted onboard the International Space Station, the Space Shuttle, as well as ground-based research with relevance to spaceflight (e.g. radiation or unloading/weightlessness).





- >100 individuals
- 4 Groups: Plants, Microbes, Animals, Multi-omics
- Monthly meetings
- Deliverables:
 - Consensus pipelines for primary analysis of data (Microarray, RNASeq, Bisulfite sequencing, Proteomics, 16S metagenomics, Whole genome metagenomics)
 - Recommendations for visualization of data













- GeneLab background and motivation
- GeneLab Data System (GLDS) and datasets
 - Data system architecture (AWS, search, federation,)
 - Datasets, dataset metadata, and curation
 - Collaborative workspace
 - Cloud analysis platform
- GeneLab community
 - Analysis Working Groups (AWGs)
- Increasing the scientific return on Space Biology investments
 - Leveraging biomedical research
 - Aggregating Space Biology research
 - Increasing data volume from Space Biology research





	Release						
GLDS	date	Flight	Flight date	Title	Organism	Platform	Publication
GLDS-21	2008	STS-108	2001	Effects of spaceflight on murine skeletal muscle gene expression	Mouse	Affymetrix	Effects of spaceflight on murine skeltal muscle gene expression. Allen et al. 2008. J Appl Physiol.
GLDS-25	2015	STS-135	2011	STS-135 Liver Transcriptomics	Mouse	Affymetrix	Spaceflight activates lipotoxic pathways in mouse liver. Jonscher et al. 2016. Plos One.
GLDS-87	2016	STS-135	2011	Spaceflight effects on the mouse retina: Histological, gene expression and epigenetic changes after flight on STS-135	Mouse	Affymetrix	Molecular effects of spaceflight in the mouse eye after space shuttle missions STS-135. Theriot & Zanello. 2014. Gravitational and Space Research.
GLDS-4	2010	STS-118	2007	Microarray Analysis of Space-flown Murine Thymus Tissue	Mouse	Affymetrix	Microarray analysis of spaceflown murine thymus tissue reveals changes in gene expression regulating stress and glucocorticoid receptors. Lebsack et al. 2010. J Cell Biochem.

- Data has been analyzed and published.
- What's the value in reanalyzing?



Value of Data Reanalysis



Impact of knowledge accumulation on pathway enrichment analysis. Wadi L, Meyer M, Weiser J, Stein L, Reimand J. Nature Methods. 2016













- GeneLab background and motivation
- GeneLab Data System (GLDS) and datasets
 - Data system architecture (AWS, search, federation,)
 - Datasets, dataset metadata, and curation
 - Collaborative workspace
 - Cloud analysis platform
- GeneLab community
 - Analysis Working Groups (AWGs)
- Increasing the scientific return on Space Biology investments
 - Leveraging biomedical research
 - Aggregating Space Biology research
 - Increasing data volume from Space Biology research





- Expertise:
 - DNA/RNA/protein extraction
 - Animal work
- Develop standards for sample processing (species dependent)
- Responsible for ~50% of GeneLab data by volume







Since 2015, GeneLab has helped generate 70 published datasets equaling ~14 TB of data













GeneLab suggests the following sequencing parameters to best capture important signals in RNAsequencing data.

		Transcriptome		
	Representative genome	complexity	Ribodepleted RNA-seq	Poly-A enriched RNA-seq
Human	H. sapiens GRCh38.p12	High	60 M clusters, 150 bp PE	40 M clusters, 150 bp PE
Rodents (mouse and rat)	M. musculus GRCm38.p6	High	60 M clusters, 150 bp PE	40 M clusters, 150 bp PE
	<i>D. melanogaster</i> Release 6 plus			
Fruit fly	ISO1 MT	Medium	40 M clusters, 150 bp PE	30 M clusters, 150 bp PE
Worms	C. elegans WBcel235	Medium	40 M clusters, 150 bp PE	30 M clusters, 150 bp PE
Plants	A. thaliana TAIR10.1	High	60 M clusters, 150 bp PE	40 M clusters, 150 bp PE
Fungi	S. cerevisiae S288C R64	Low	20 M clusters, 150 bp PE	13 M clusters, 150 bp PE
Bacteria	<i>E. coli</i> str. K-12 substr. MG1655	Low	20 M clusters, 150 bp PE	13 M clusters, 150 bp PE

Spike-in Controls. At a minimum, GeneLab recommends that all sequencing samples include ERCC Spike-in Mix 1 in all samples. This will allow assessment of the dynamic range within each sample. In addition, in experiments focused upon comparing 2 levels of a given factor (spaceflight vs. ground control), GeneLab recommends including Mix 1 in the level A samples (spaceflight) and Mix 2 in the level B samples (ground control). This allows direct assessment of the available power in the data. In all cases, these Mixes should be added to samples before library preparation and sequencing. These Spike-in standards are available from Thermo Fisher: Mix 1 (https://www.thermofisher.com/order/catalog/product/4456740?SID=srch-srp-4456740), Mix 1 and 2 (https://www.thermofisher.com/order/catalog/product/4456739?SID=srch-srp-4456739).

Library preparation and sequencing standard. GeneLab recommends that all sample pools to be sequenced include at least one replicate of the SEQC Universal Reference RNA available from Agilent (https://www.agilent.com/en/product/gene-expression-microarray-platform/gene-expression-microarray-kits-reagents/gene-expression-universal-reference-rnas-228491). This RNA should be included as a sample during each library preparation batch.





Ames Life Sciences Data Archive (ALSDA)

- Data and Descriptions from 837 Experiments
 - Flight and Ground-based Experiments (non-human life sciences)
 - Space Biology and Human Research Program funded
 - Data Submission Agreements
- Information is curated and made available (when approved) on public LSDA website

Data Archived

Biospecimen metadata Payload Mission Hardware Personnel Research Subject Images Publication citations PI and Project Data Sets



Collection Management

Secure database

Metadata: Conditions, Tissue type, Descriptions, Species, Fixations, Chain of Custody Record, Telemetry Storage: -80°C, -20°C, +4°C, Ambient Features: Power backup, Alarm system for power failure (24-7 staff on-call)

NASA ARC Institutional Scientific Collection (ISC)

- 32,037 Non-Human Biospecimens
 - Spaceflight investigations and related ground controls
 - Tissue request process managed by ALSDA
 - Tissue request proposal reviewed by appropriate funding organization (SB or HRP)
 - Housed in ARC Biospecimen Storage Facility (BSF)

000

000

• Curated by ALSDA









Biospecimens Available for Research

		2016	2014-17	2017	2018	2017	2017	2017	2017	2016	2014	2011	2010	2007	2001	1996	1995	1993	1993	1991	1985	1974
		ISS Biological Research in Canister23	Hindimb Unloaded	ISS/JAXA Mouse Habitat Unit-2	ISS Rodent Research 7	ISS Rodent Research 9	ISS Rodent Research 6	ISS Rodent Research 5	ISS Rodent Research 4	ISS Rodent Research 3	ISS Rodent Research 1	STS-133 Immune2	STS-131 Immune	STS-118 CBTM2	STS-108 CBTM	STS-72 NIH.R3	STS-70 NIH.R2	STS-56 PARE.03	STS-58 SLS-2	STS-40 SLS-1	STS- 51B SL-3	Cosmos 1129
System	System Specimen Types		Rat	Mouse	Mouse	Mouse	Mouse	Mouse	Mouse	Mouse	Mouse	Mouse	Mouse	Mouse	Quail	Rat	Rat	Rat	Rat	Rat	Rat	Rat
Circulatory	Blood, Heart, Lymph nodes, Spleen		х			х					х			Х	х			х	Х	х		
Digestive	Digestive Cecum, Colon, Duodenum, GI Tract, Ileum, Intestine, Jejunum, Pancrea, Stomach			х	x	x	x		х	х	х	x		x	х	х		х	х	x		
Endocrine	Adrenal glands, Liver, Salivary glands, Thymus, Thyroid		х		х	х	х			х	х	x	x	x		х	x	х	х	x		
Excretory	Bladder, Fecal pellets, Kidney		х			х	х					x	x	х		х		х	Х	х		
Integumentary	Adipose, Skin, Whiskers		Х	Х		Х	X	X	Х	Х	Х			Х	Х	Х		Х				
Muscular	Adductor longus, Diaphragm, EDL, Gastroc, Plantaris, Soleus		x		x	x				x	x					x		x	х	x	x	
Neurosensory	Brain, Eyes, Hypothalamus, Pituitary									х	х			х		Х		х	x		х	
Reproductive	Ovary, Uterus, Testes		Х	Х		X	X			Х				Х		Х		Х	Х	Х		
Respiratory	Esophagus, Lung, Trachea		Х	Х	Х	Х	X			Х			X	Х		Х		Х	Х	Х		
Skeletal	Calvaria, Femur, Humerus, Jawbone, Parietal, Pelvis, Ribs, Tail, Tibia, Vertebrae			Х	х	х		x	х	х				x	х	x			x			x
DNA and Protein extracts	Adrenal, Kidney, Liver, TA, EDL, Eye, Gastroc, Quadriceps, Soleus					x	x				x											
DNA and Protein extracts	Bacillus subtilis, Staphylococcus aureus	x																				

Next page For details



GeneLab Team



