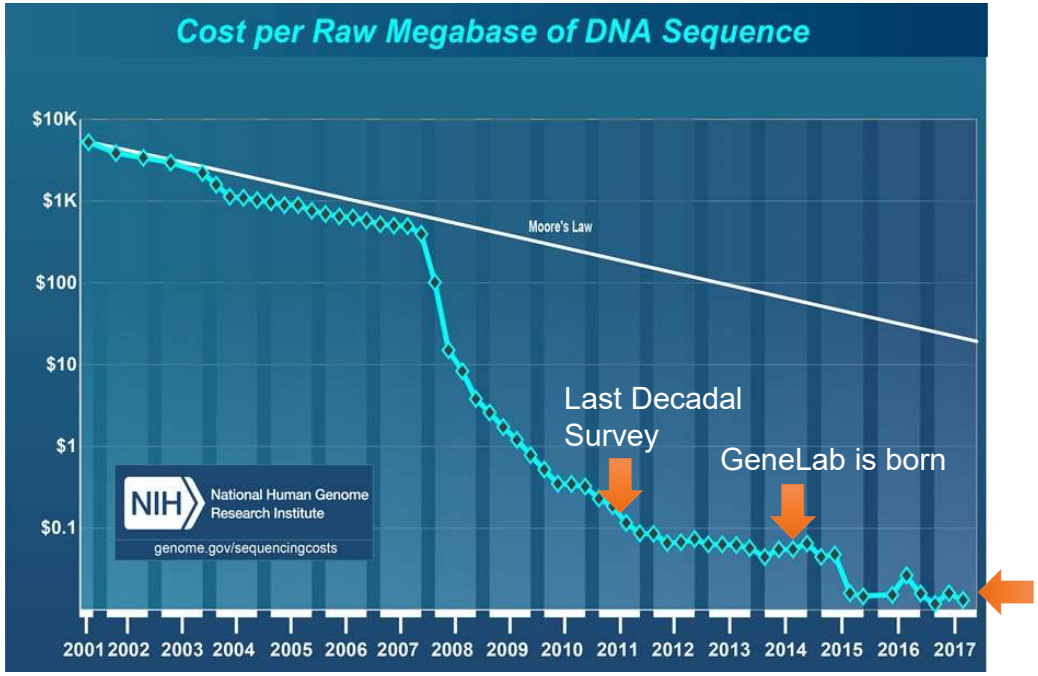
La science ouverte pour les « omiques » et données spatiales biologiques à la NASA

National Aeronautics and Space Administration

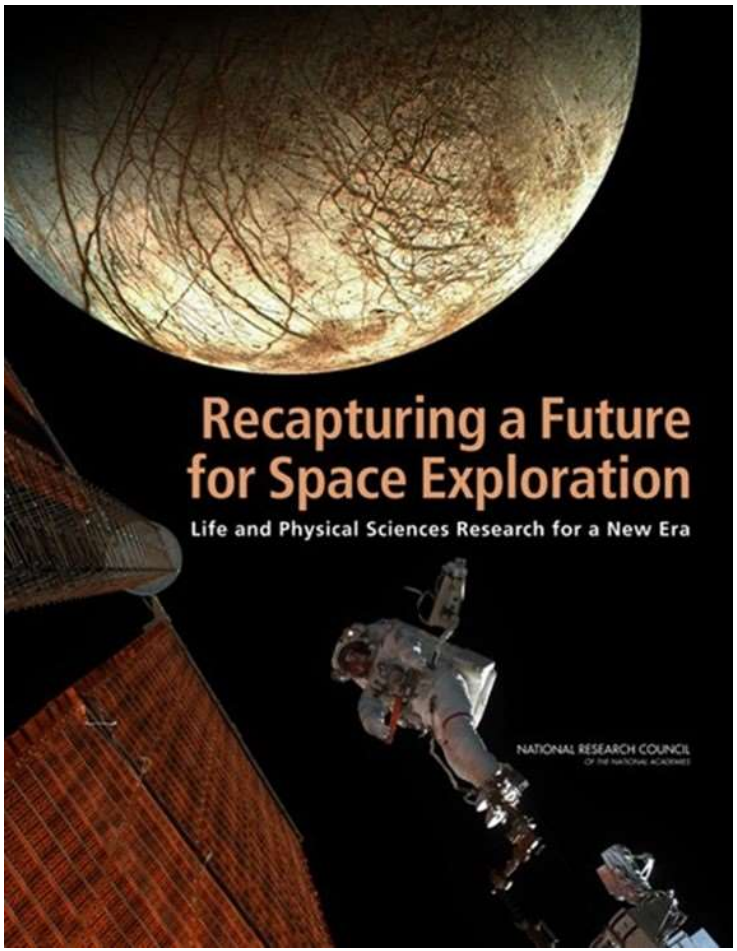
Sylvain Costes, Ph.D.
GeneLab Project Manager
NASA Ames Research Center

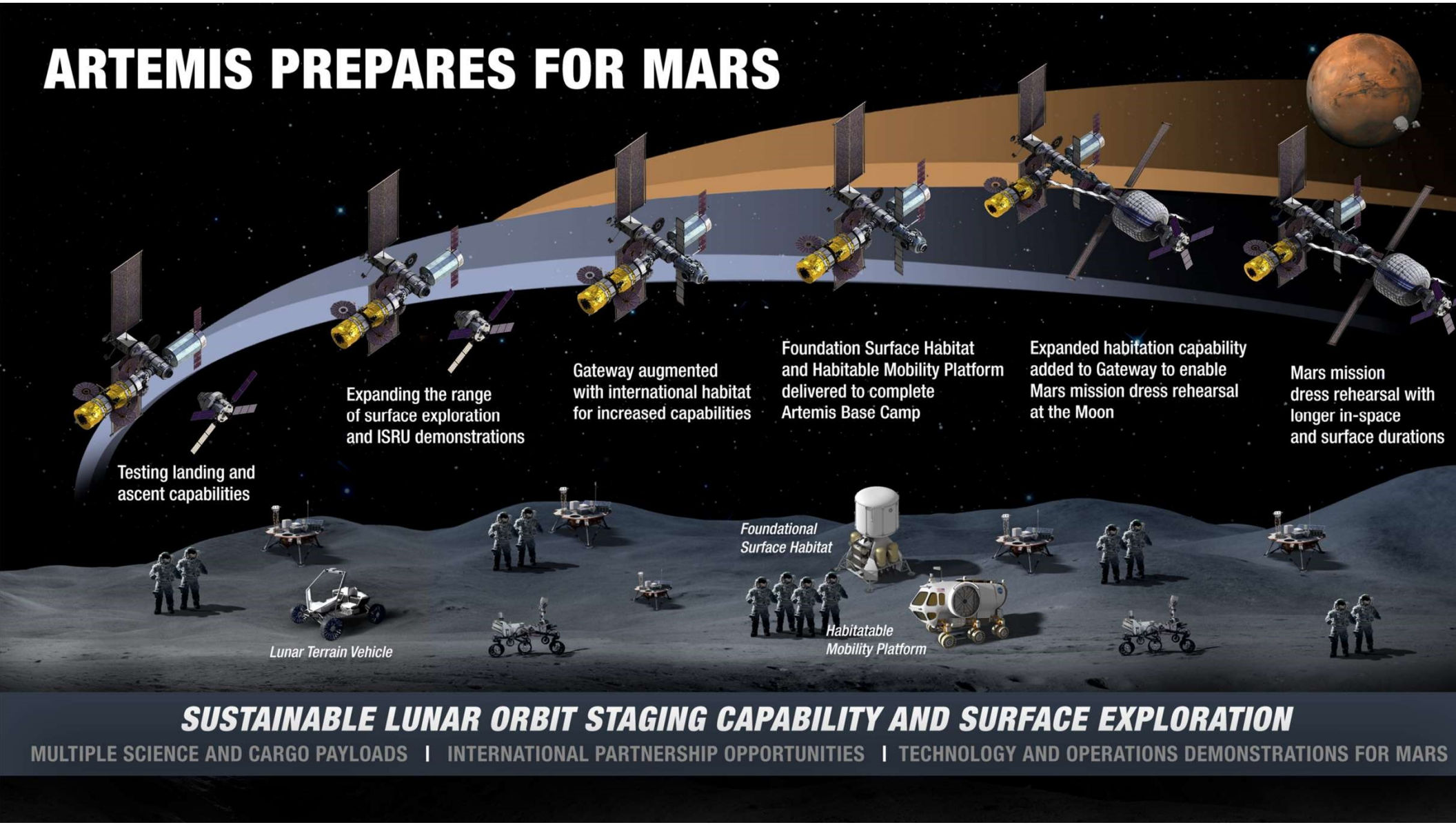
2011 NRC Decadal Survey and the Sequencing Paradigm Shift



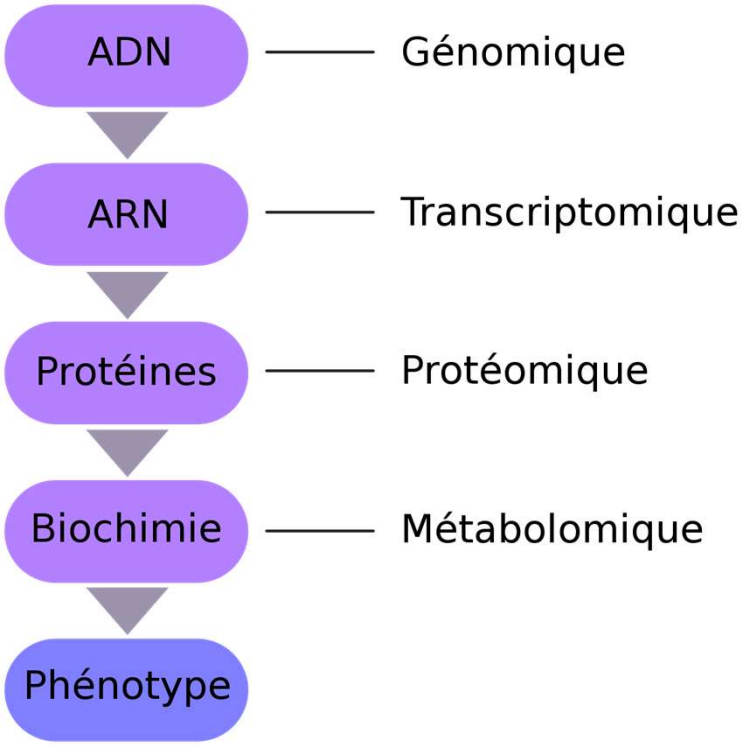
*"…**genomics, transcriptomics, proteomics, and metabolomics** offer an immense opportunity to understand the effects of spaceflight on biological systems…"*

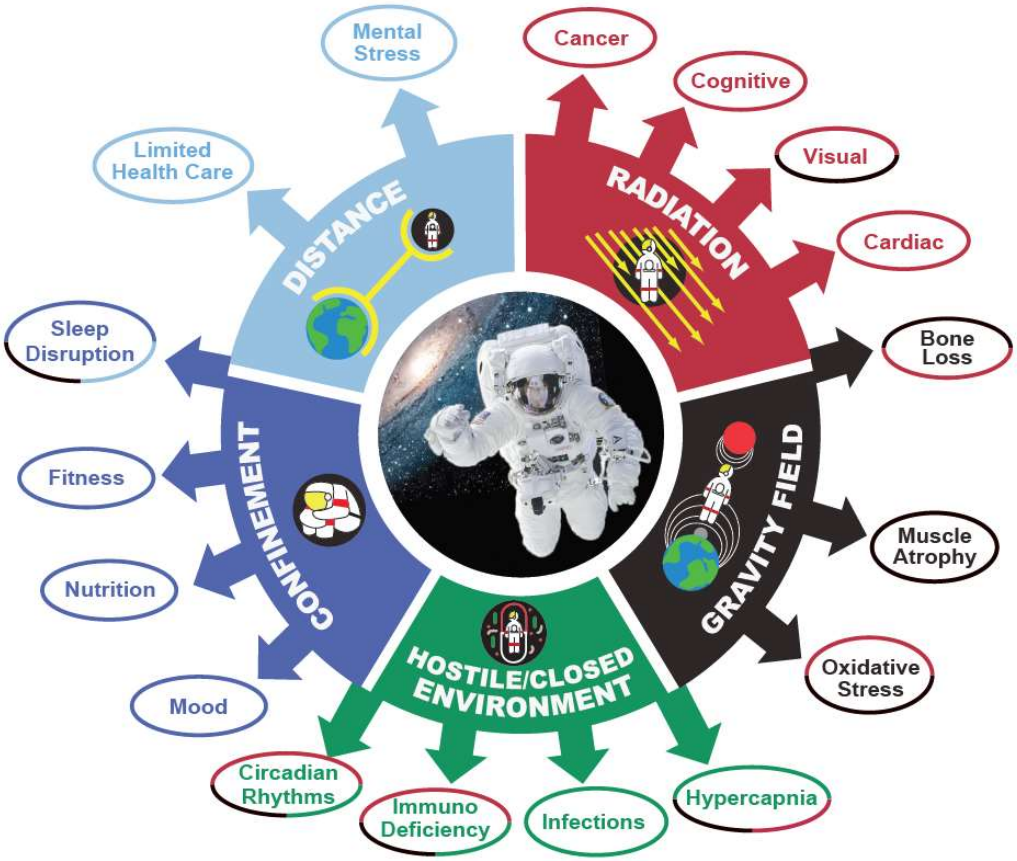
*"…Such techniques generate considerable amounts of **data that can be mined and analyzed** for information by multiple researchers…"*




ARTEMIS PREPARES FOR MARS
Testing landing and ascent capabilities
Expanding the range of surface exploration and ISRU demonstrations
Gateway augmented with international habitat for increased capabilities
Foundation Surface Habitat and Habitable Mobility Platform delivered to complete Artemis Base Camp
Expanded habitation capability added to Gateway to enable Mars mission dress rehearsal at the Moon
Mars mission dress rehearsal with longer in-space and surface durations
Foundational Surface Habitat
Lunar Terrain Vehicle
Habitatable Mobility Platform
SUSTAINABLE LUNAR ORBIT STAGING CAPABILITY AND SURFACE EXPLORATION
MULTIPLE SCIENCE AND CARGO PAYLOADS | INTERNATIONAL PARTNERSHIP OPPORTUNITIES | TECHNOLOGY AND OPERATIONS DEMONSTRATIONS FOR MARS

Stresses on Space Guardians

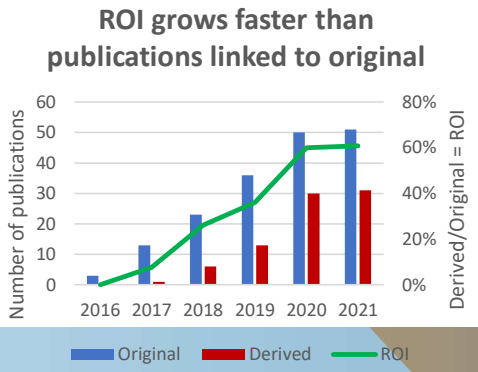


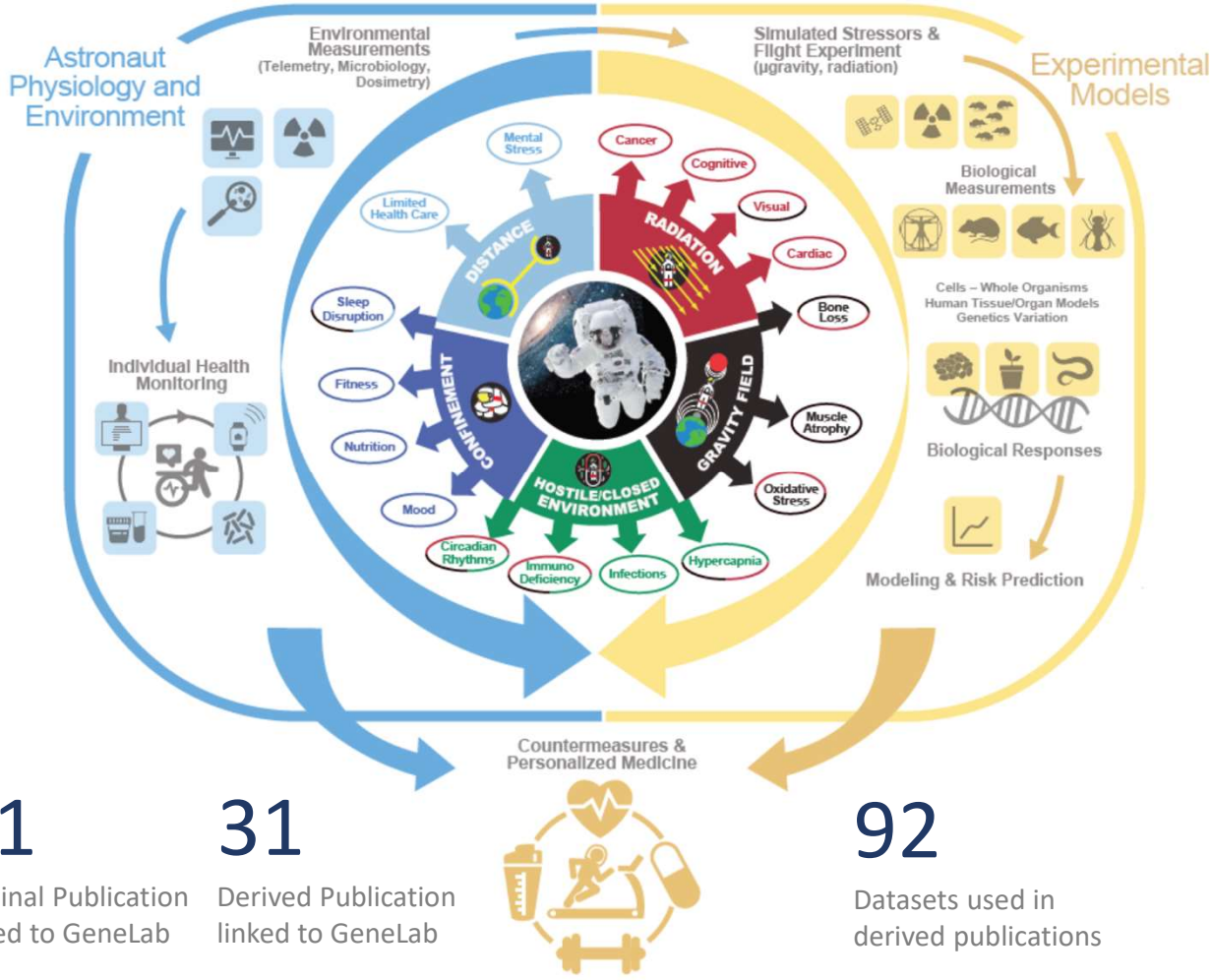


Application to Space Guardians

GeneLab established the following biological outcome using Omics and databases:

- Radiation Gene Signatures
- CO2 effects on mice
- Impact on mitochondria (Cell Package)
- Impact on liver
- Potential contermeasures





51
Original Publication linked to GeneLab

31
Derived Publication linked to GeneLab

92
Datasets used in derived publications

BPS Open Science

Physical Sciences Informatics (PSI)

- *Collects and Curates and Disseminates data*
- *Physical science experiments performed on the ISS*

GeneLab (GL)

- *Collects and Curates and Disseminates* data
- Molecular omics-level data
- Viz Portal - Enviro Portal
- Analysis Toolshed

Ames Life Sciences Data Archive (ALSDA)

- *Collects and Curates and Disseminates* data
- Physiological-level data
- Mission and Project data
- Imaging data

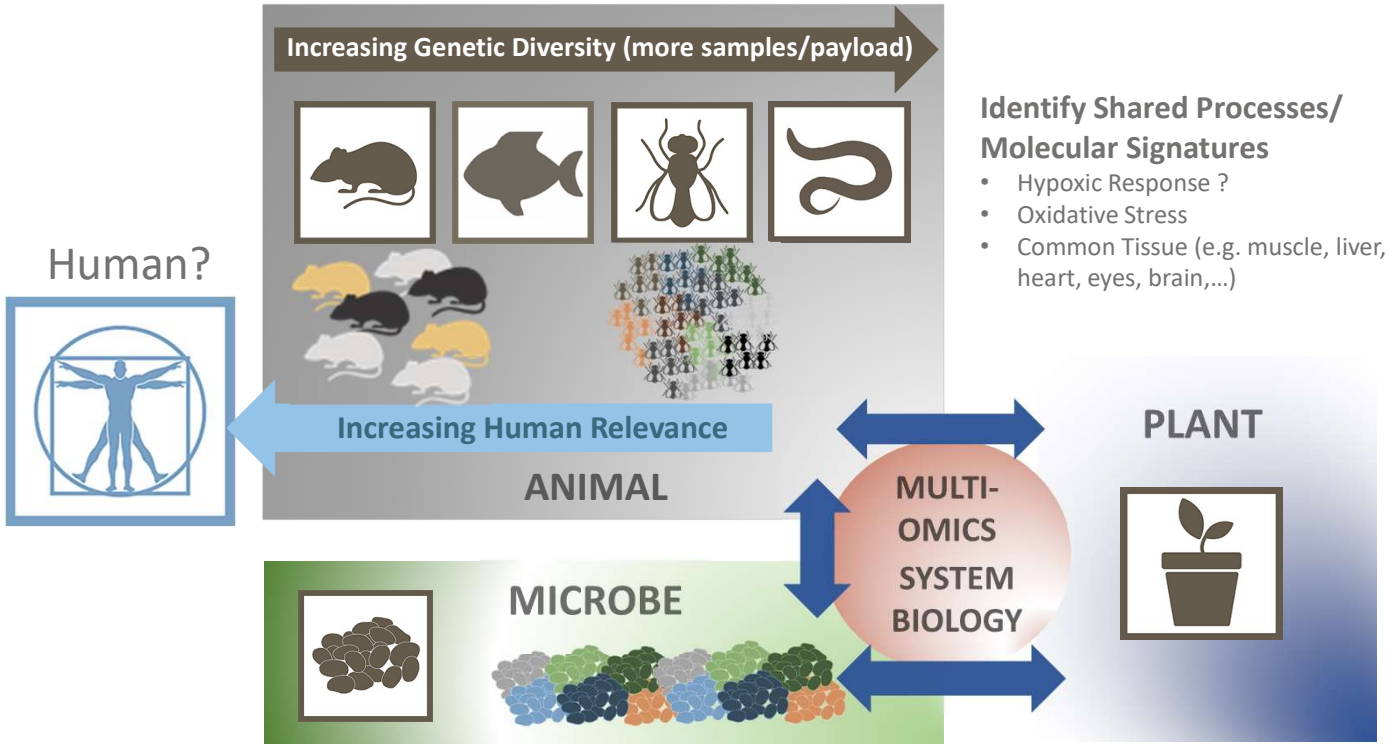
Biospecimen Sharing Program (BSP)

- *Dissects and preserves* rodent tissues from Flight and Ground investigations
- *Coordinates* internal tissue sharing

NASA Biological Institutional Scientific Collection (NBISC)

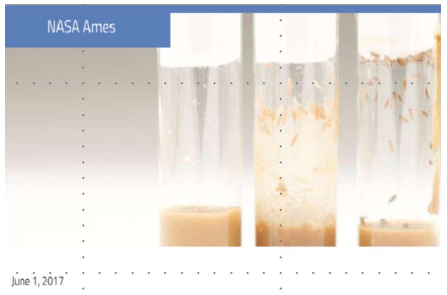
- *Collection* of non-human biospecimens
- ARC Biospecimen Storage Facility (BSF)
- Space Microbial Culture Collection

GeneLab ecosystem: maximizing knowledge by bringing experiments together as a system



Identify Shared Processes/ Molecular Signatures

- Hypoxic Response ?
- Oxidative Stress
- Common Tissue (e.g. muscle, liver, heart, eyes, brain,…)

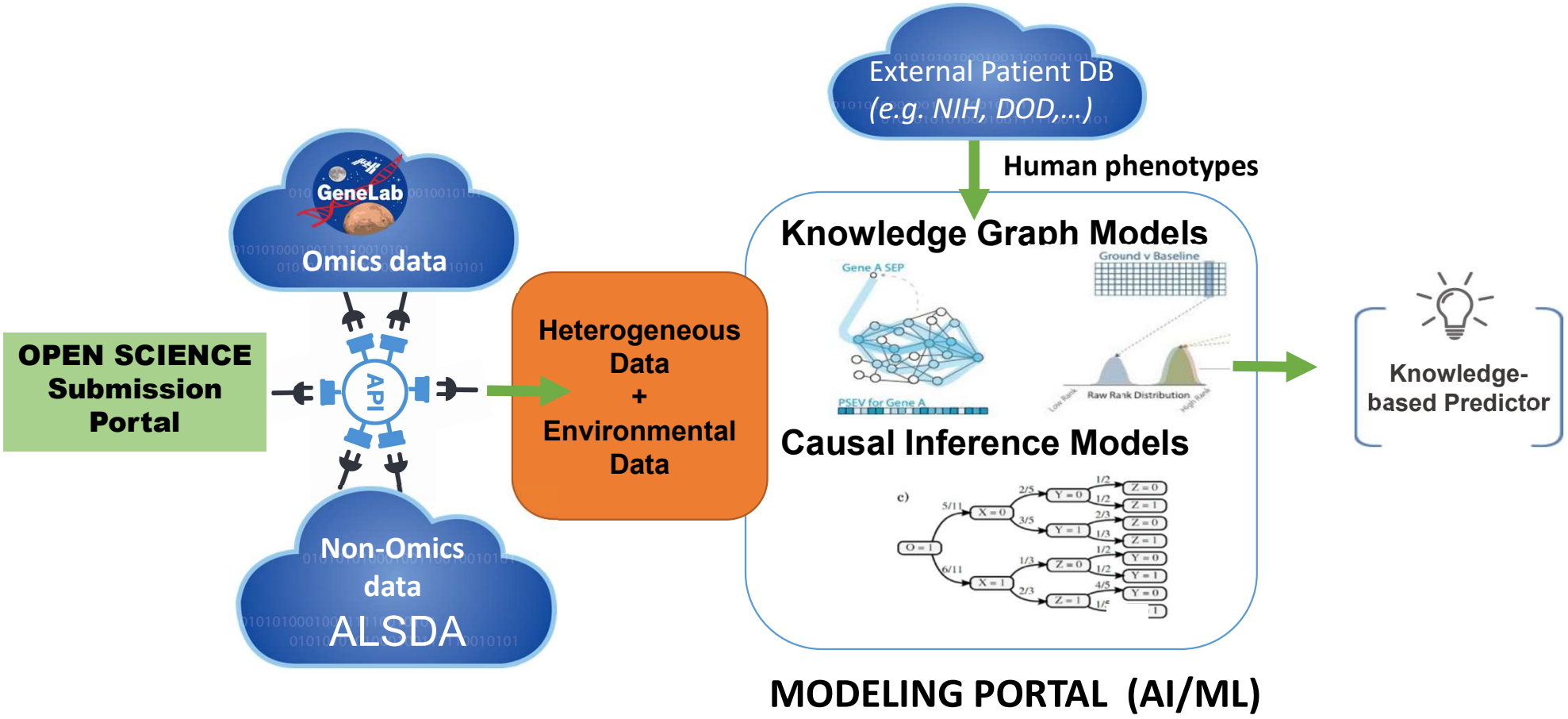


Fruit Fly Lab (FFL-02) Scientist's Blog

For Spaceflight

- High "n" number – statistically significant data
- Genetically identical animals
- Low resource requirements
- Short life cycle - multiple generations
- Measure response of a whole multicellular animal
- **Flies used as a model for humans for innate immunity, circadian rhythm, oxidative stress, neurobehavior, development, genetics, GWAS, "omics" studies etc**

From multiple databases to a knowledge-based system

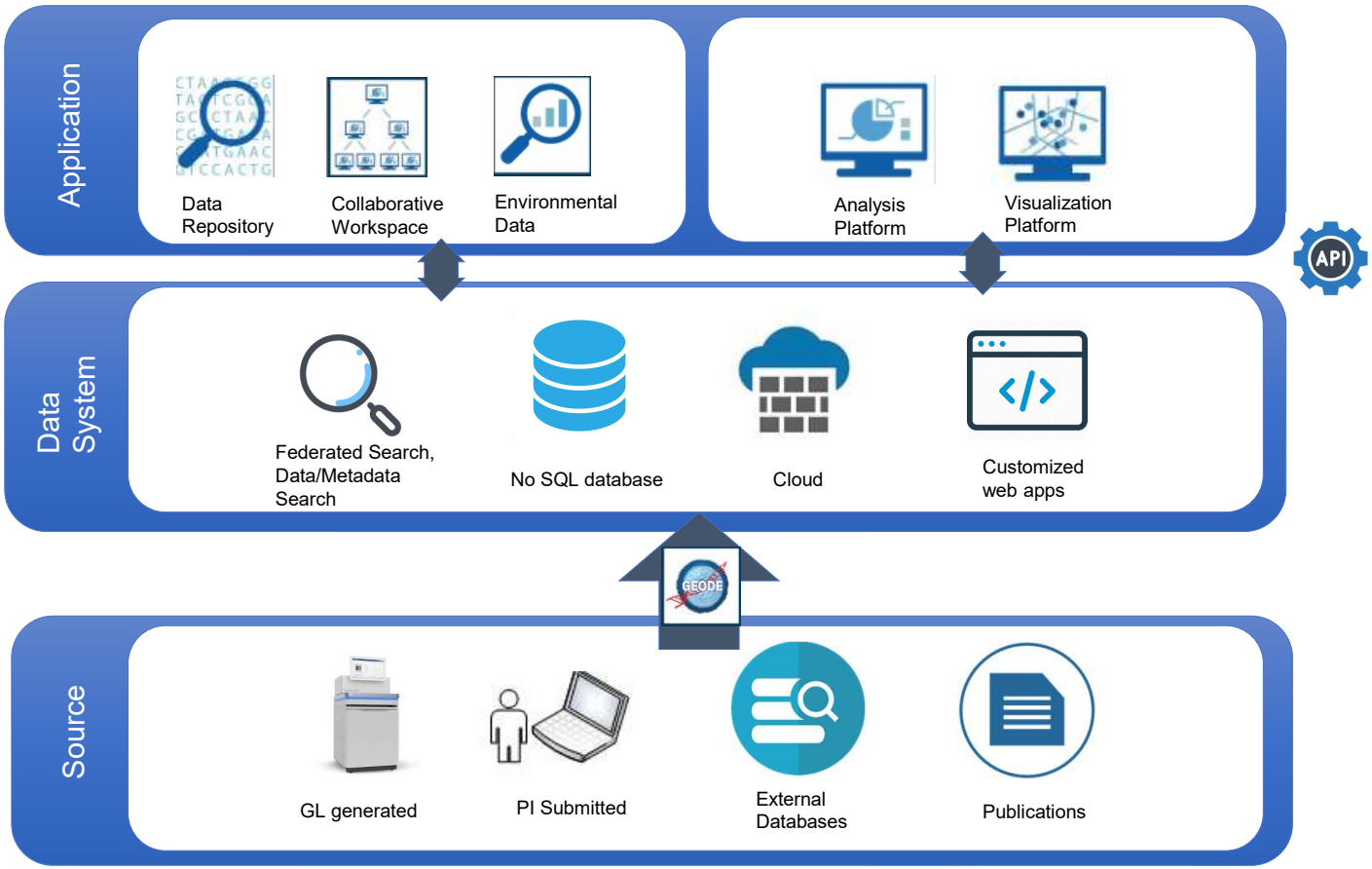


NASA BPS Open Science Enterprise Solution: genelab.nasa.gov

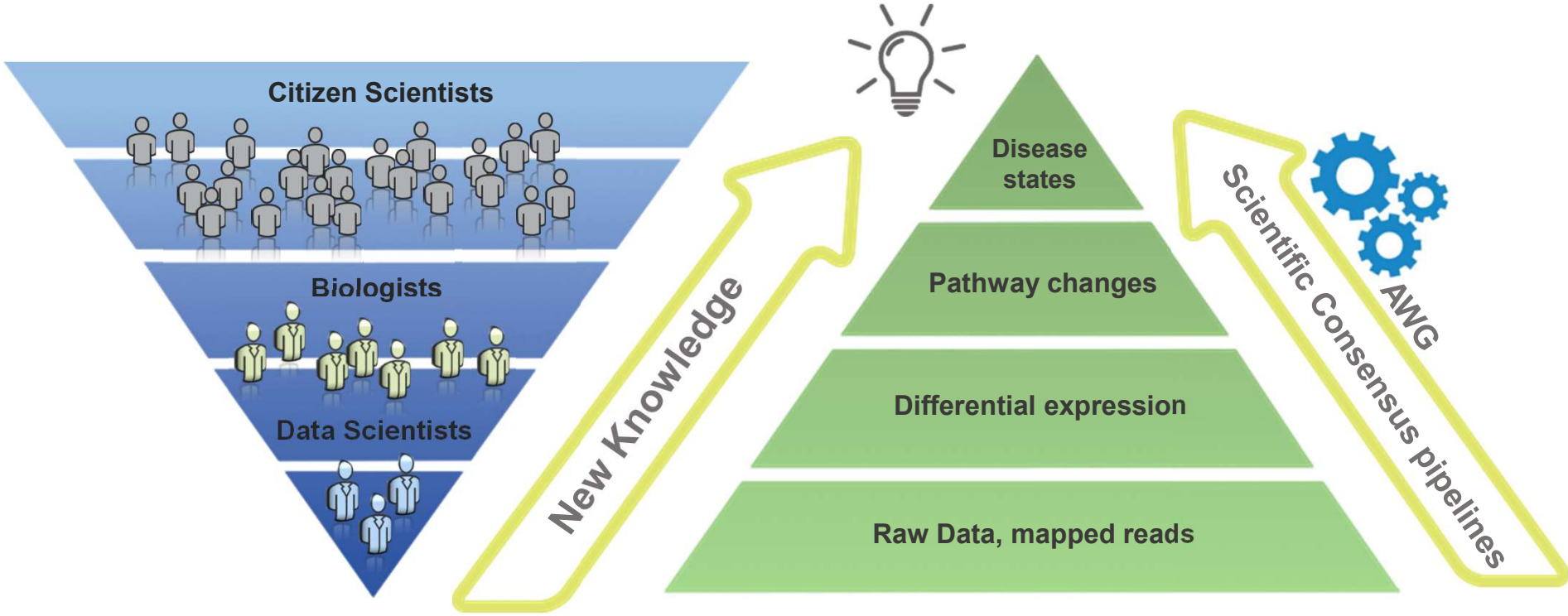
- Open access data
- **FAIR (Findable, Accessible, Interoperable, Reusable)**
- Controlled access tools
- API - internal and external
- User Friendly Interface
- Tutorials
- Self-service Submission Portal

- Federated search – GEO, PRIDE, MG-RAST, **ALSDA**
- Database & Cloud – Scalable, easy access, fast
- Web apps
  - Data Access & Management
  - Security
  - Operation
  - Governance and Integration
- Open Source software – *no maintenance cost for software*

- Multiple data sources
  - Standard metadata organization
  - Open file formats



GeneLab Omics Data Democratization



GeneLab Download Metrics

Dec 2018 - May 2020

**represents downloads per GLDS Study, not per file*

264
Avg Monthly Downloads

8702
Total Study Downloads

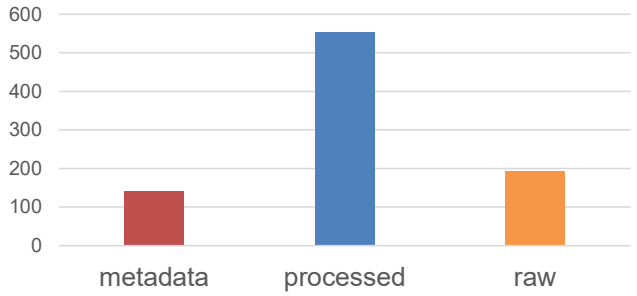
Introduction of Processed Data and Visualization

Downloads by Category

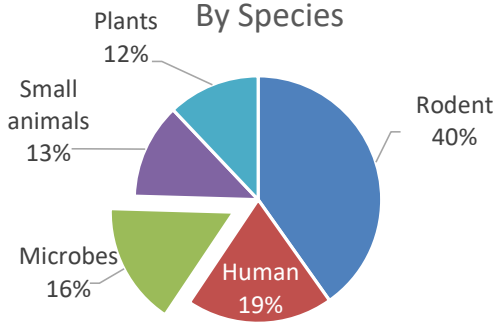
March 2020

**represents downloads per GLDS Study, not per file*

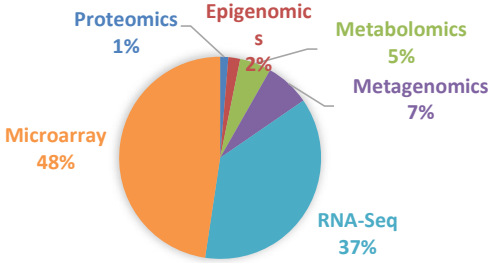
Data File Downloads by Type



By Species



By Assay



Ames Life Sciences Data Archive (ALSDA)

The Ames Life Sciences Data Archive (ALSDA) is the official repository of non-human science data generated by NASA's Space Biology Program and Human Research Program. The ALSDA captures, curates, preserves, and makes available NASA funded flight- and ground-based research.

Science Disciplines

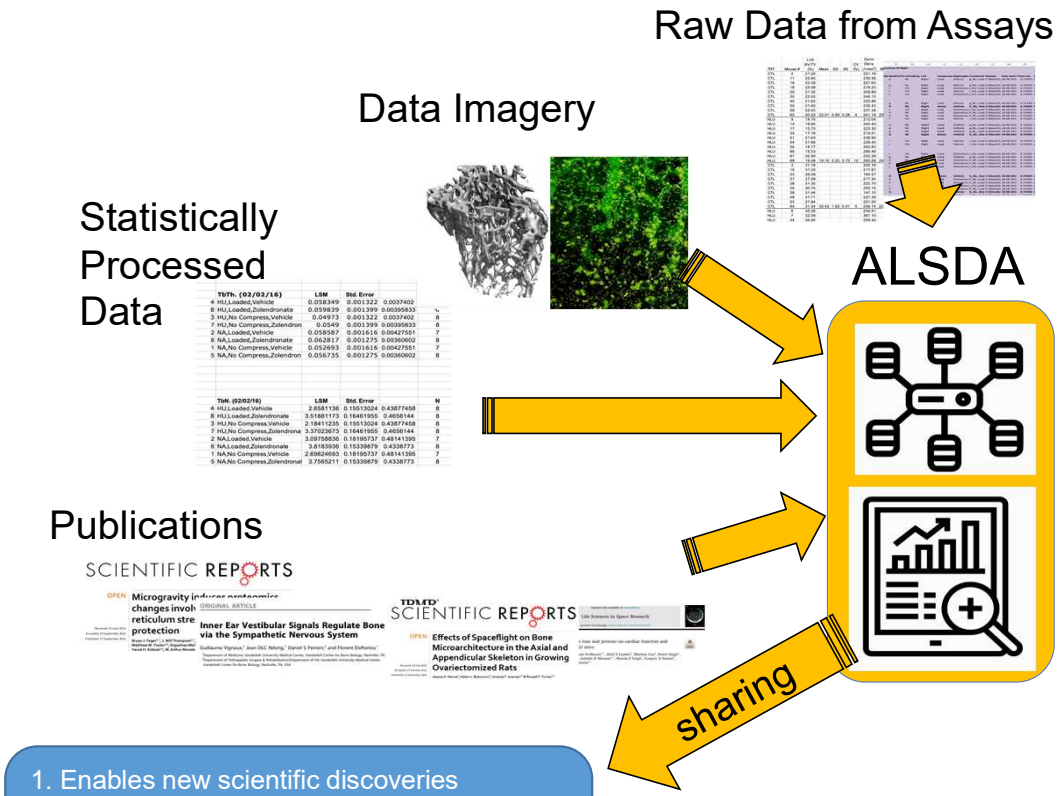
- Bone Physiology
- Cardiovascular/ Cardiopulmonary
- Developmental Biology
- Immunology
- Microbial Growth and Virulence
- Muscle Physiology
- Neurophysiology
- Chronobiology
- Endocrinology
- Hematology
- Metabolism and Nutrition

Subjects

- **Mouse, Rat**, Quail, Chicken, Gecko, Rhesus Monkey, Bacteria, Fungus, and more

Missions

- ISS, Shuttle (STS)
- COSMOS, Bion, Ground analogs



1. Enables new scientific discoveries
2. Enables new spaceflight hardware/platform development
3. Ensures long-term preservation
4. Ensures reproducibility and transparency

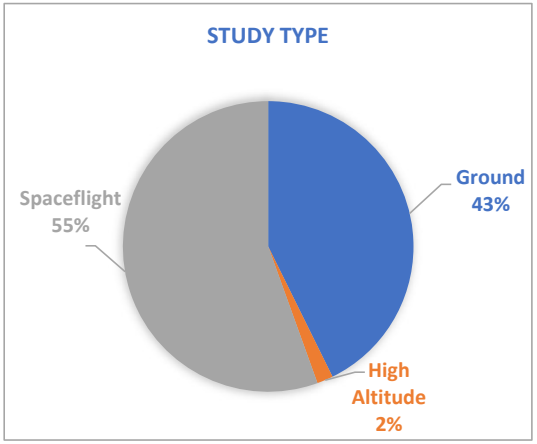
335
Studies

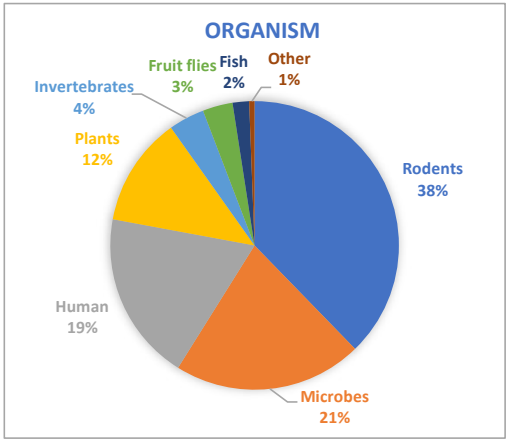
371
Datasets

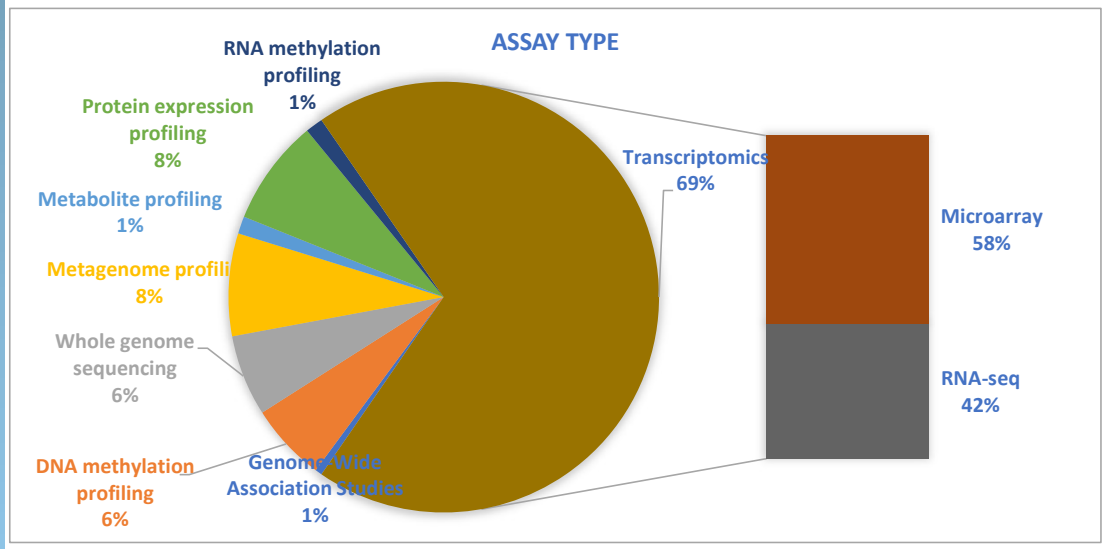
45
Species

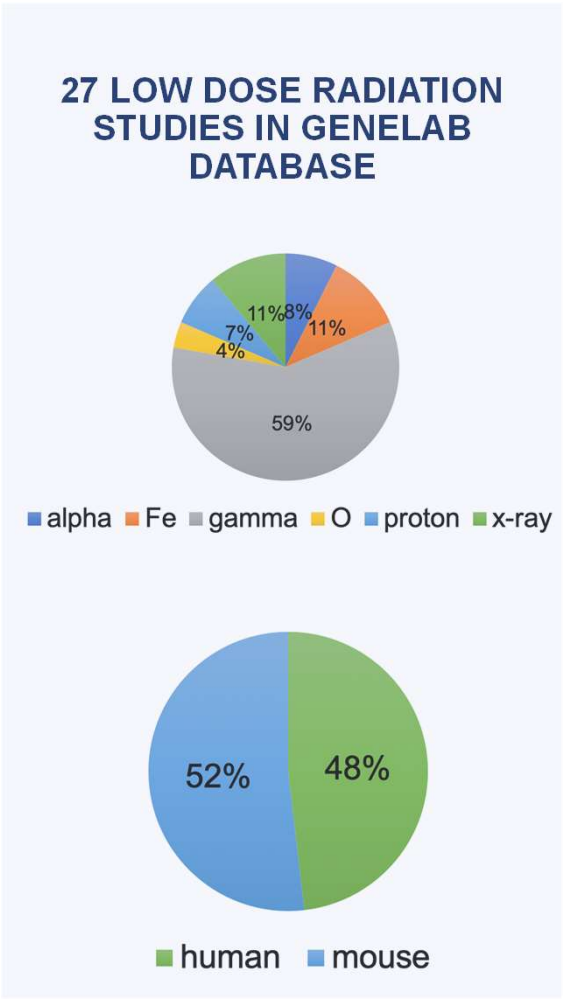
>10
Assays

>135TB
Data







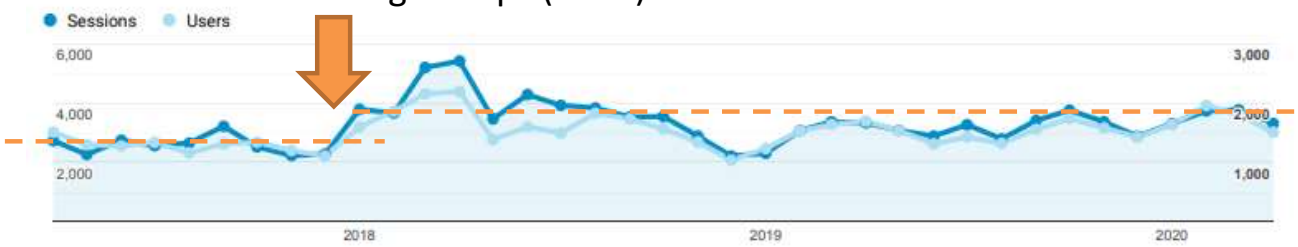


GeneLab Usage Metrics
April 2017 – April 2020

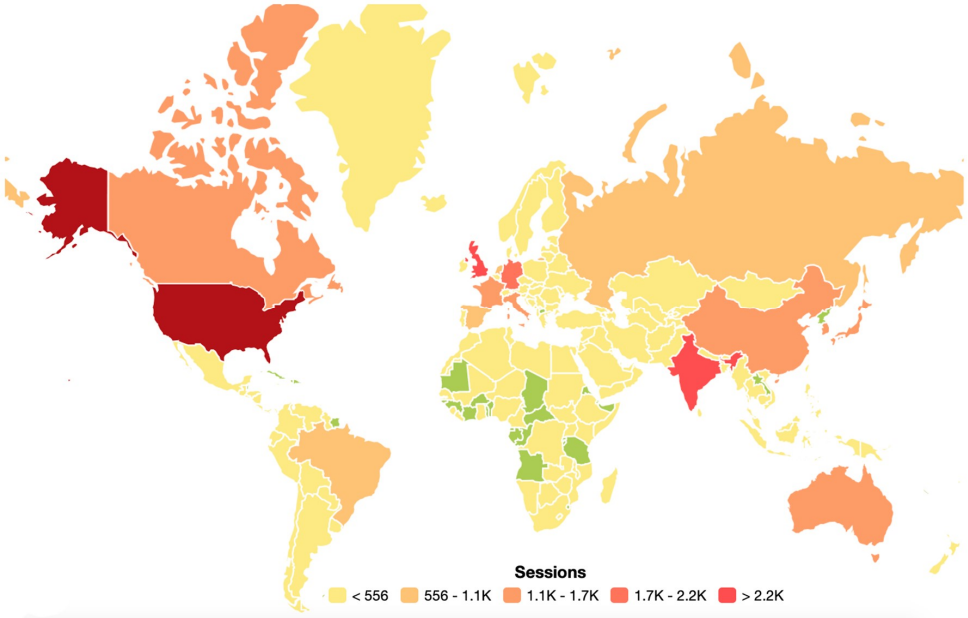
3266
Avg Monthly Sessions

1522
Avg Monthly Users

Introduction of GeneLab Analysis Working Groups (AWG)



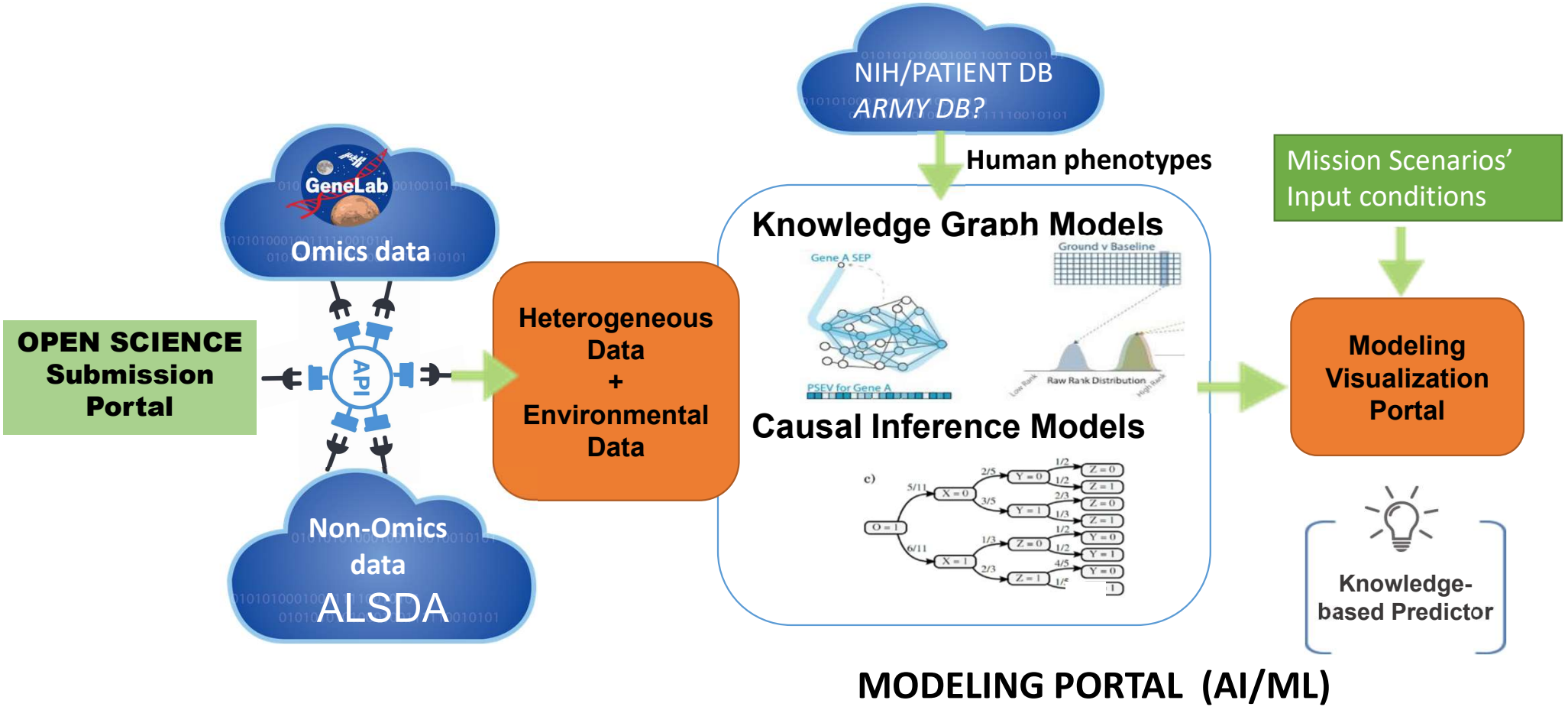
Sessions by Country



Top 10 Countries – Sessions

1. United States
2. India
3. United Kingdom
4. Germany
5. Canada
6. France
7. South Korea
8. Japan
9. China
10. Australia

From multiple databases to a knowledge-based system



Open Science for Life in Space

genelab.nasa.gov

Data Portal

- First multi-omics space related database
- 335 publicly available studies
- High-order data for transcriptomics and metagenomics studies
- Spaceflight **environmental data** associated with sample metadata (radiation dosimetry, temperature, humidity)
- High-order Data **visualization Portal**

Tools

- User friendly submission portal to submit and publish data
- Analysis platform to tools and workflows to analyze your own data or data from the GeneLab repository.
- Tutorials and online resources to learn how to analyze RNA-Seq data
- Workspace to store, share, and organize data files

Sequencing

- State-of-the-art sequencing facility to process spaceflight samples
- Optimized SOPs with standard processing workflow
- Provide sequencing service for NASA funded PI without any university overhead
- Generate high quality data from shared tissues for open science access

Community

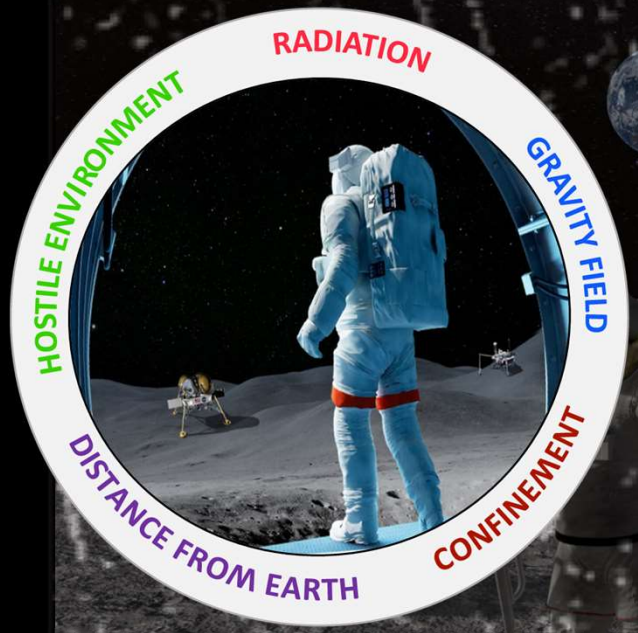
- Analysis Working Groups comprised of over 100+ scientists worldwide collaborating and analyze space omics data.
- Education Working Group focused on providing resources to educators and students to learn about space biology and bioinformatics

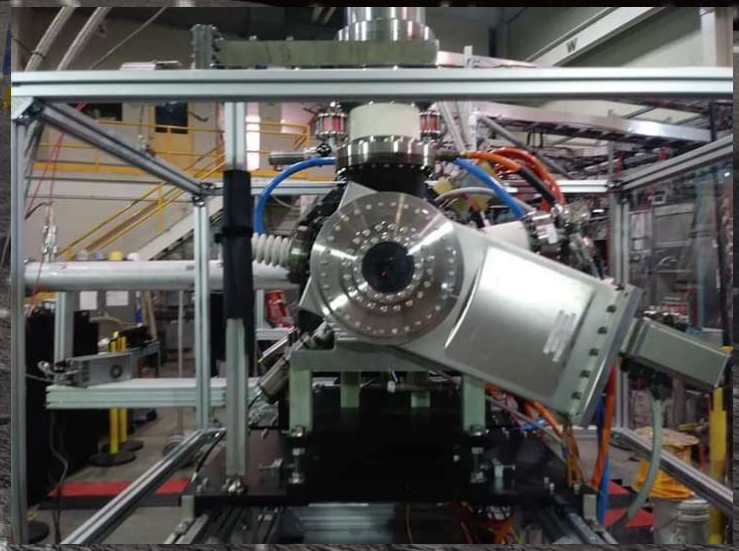
Environmental Portal

Radiation Quality	GCR
Proton	87%
^{4}He	12%
HZE	1%

ABSORBED DOSES	ISS (LEO)	MOON	MARS
TOTAL (mGy/day)	0.2-0.4	0.5-0.7	~0.5

- **LEO – 50% GCR, 50% trapped protons** (South Atlantic Anomaly), secondary neutrons
- **Moon, Mars – mainly GCR**, secondary neutrons, albedo neutrons





Simulated at NASA Space Radiation Laboratory in Brookhaven National Lab

Low dose data are highly relevant for NASA space missions

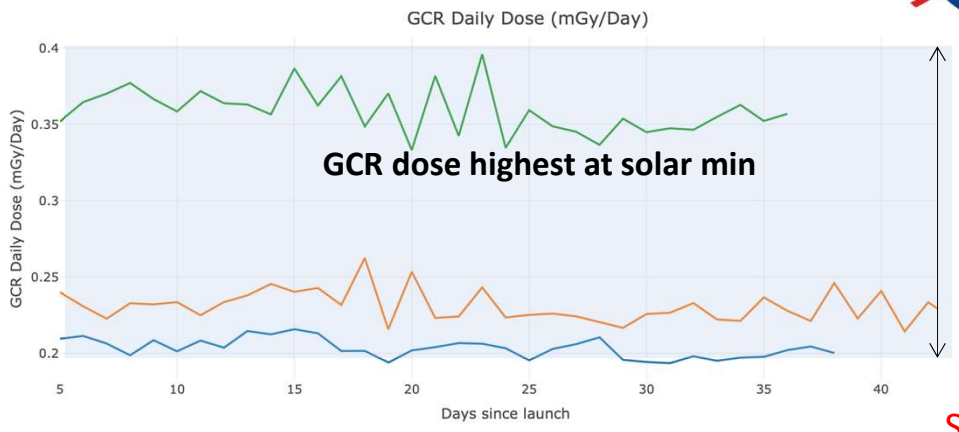
Application

Environmental Data

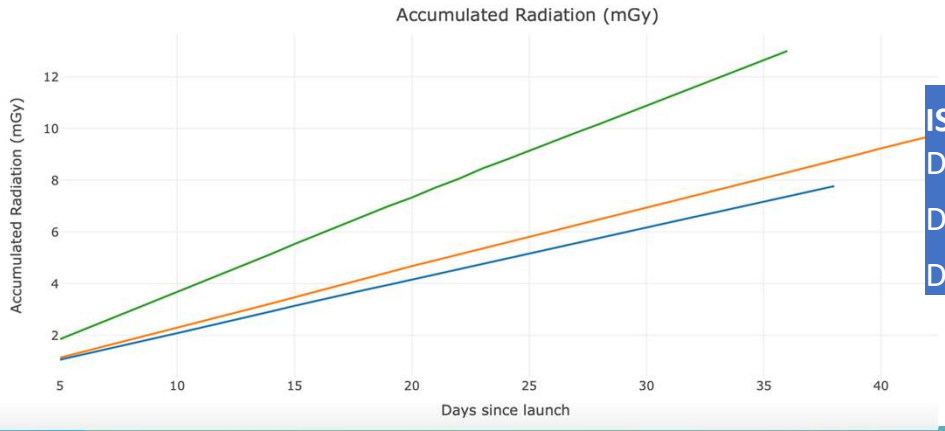
Mission	Date
RR1	2014-09-20 03:59:00
RR3	2016-04-08 04:00:00
RR5	2017-06-01 05:00:00
RR6	2017-12-06 15:06:00
RR7	2018-07-01 04:49:00
RR12	2019-04-17 19:19:00
RR19	0019-12-03 10:24:00

Select plots to display
- [x] Timeplot
- [ ] Boxplot
- [ ] Bar Chart
- [ ] Show landmarks

RR1
RR3
RR12



0.2-0.4 cGy/day



Space flight data:
Low dose rate: 0.2-0.4 cGy/day
& Low dose: <10 cGy

ISS	Mouse	Astronauts
Duration	30 days	180 days
Dose solar min	**1.2 cGy**	**7.2 cGy**
Dose solar max	**0.6 cGy**	**3.6 cGy**

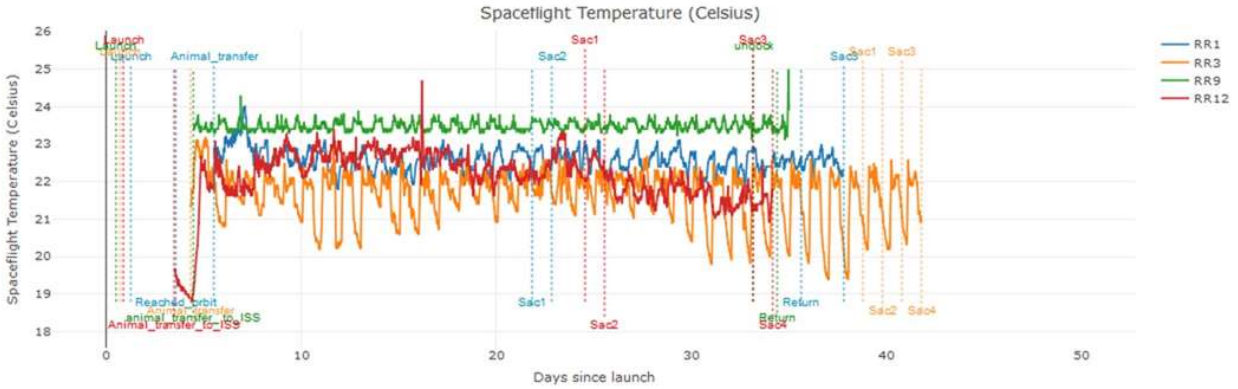
Rodent Research Environmental Data

NASA GeneLab | Individual missions | Compare

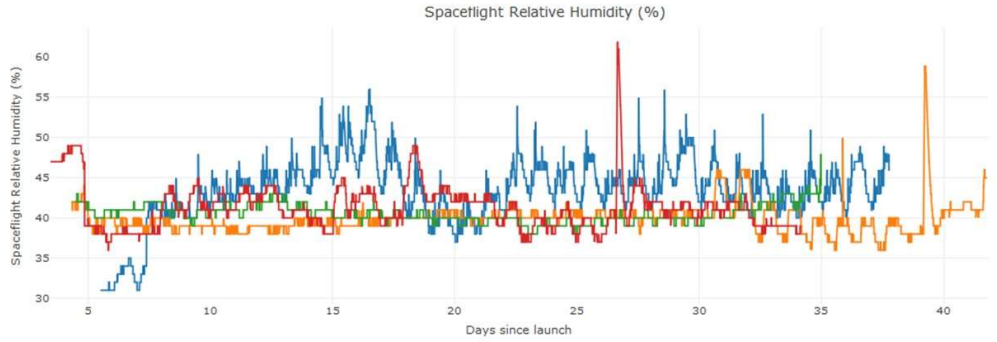
Compare Environmental Data Among Datasets

Mission	Date
RR1	14-09-20 03:59:00
RR3	16-04-08 04:00:00
RR4	
RR5	17-06-01 05:00:00
RR6	17-12-06 15:06:00
RR7	18-07-01 04:49:00
RR8	
RR9	18-08-14 04:53:00
RR12	19-04-17 19:19:00
RR17	19-07-24 23:05:00

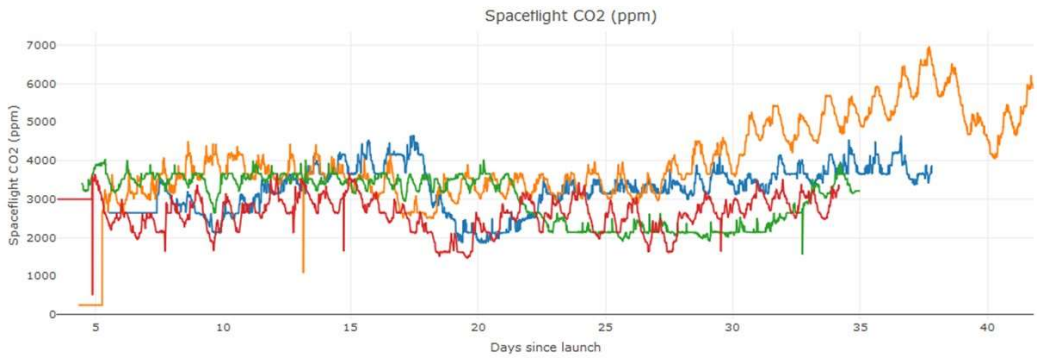
Radiation | Temperature | Relative Humidity | CO2



Radiation | Temperature | Relative Humidity | CO2

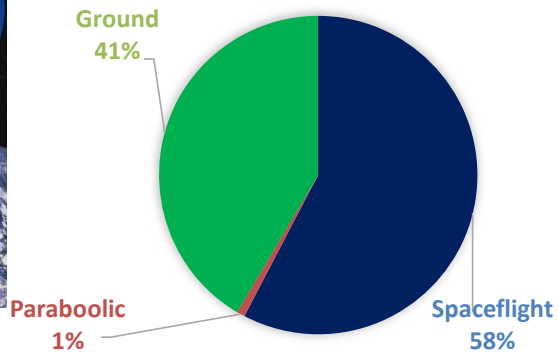


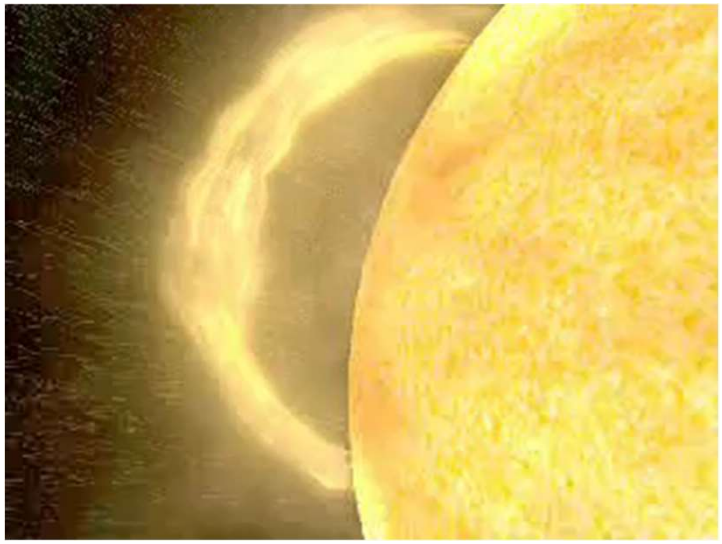
Radiation | Temperature | Relative Humidity | CO2

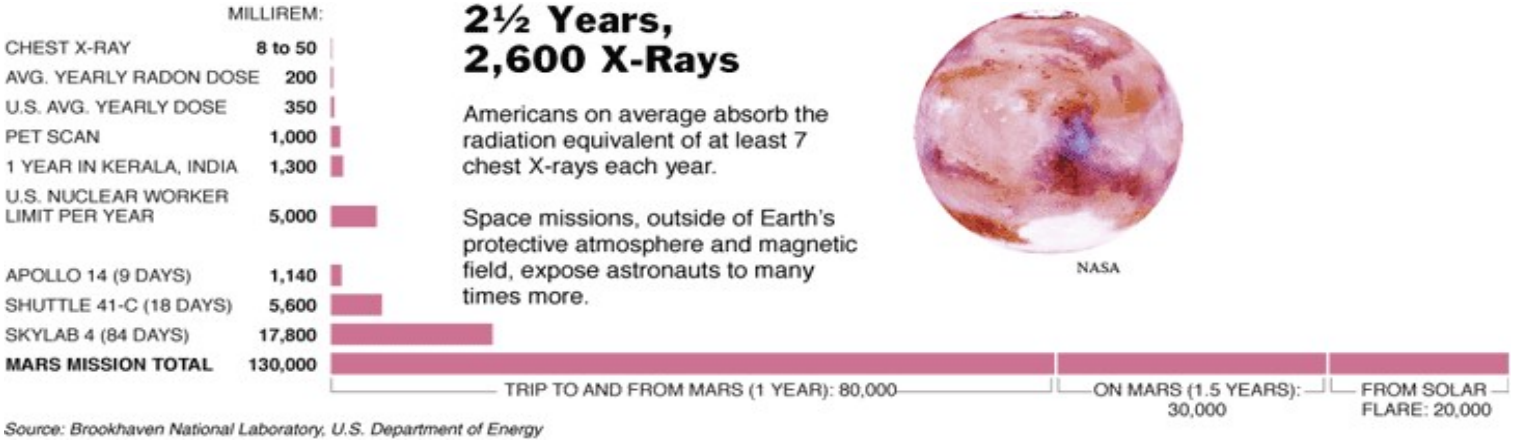


Radiation Stressor

GENELAB STUDY TYPE







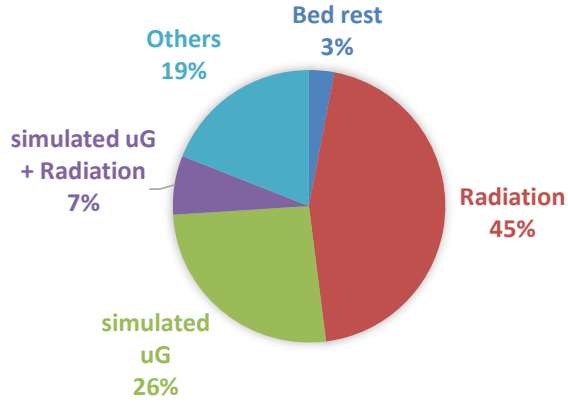
2½ Years, 2,600 X-Rays

Americans on average absorb the radiation equivalent of at least 7 chest X-rays each year.

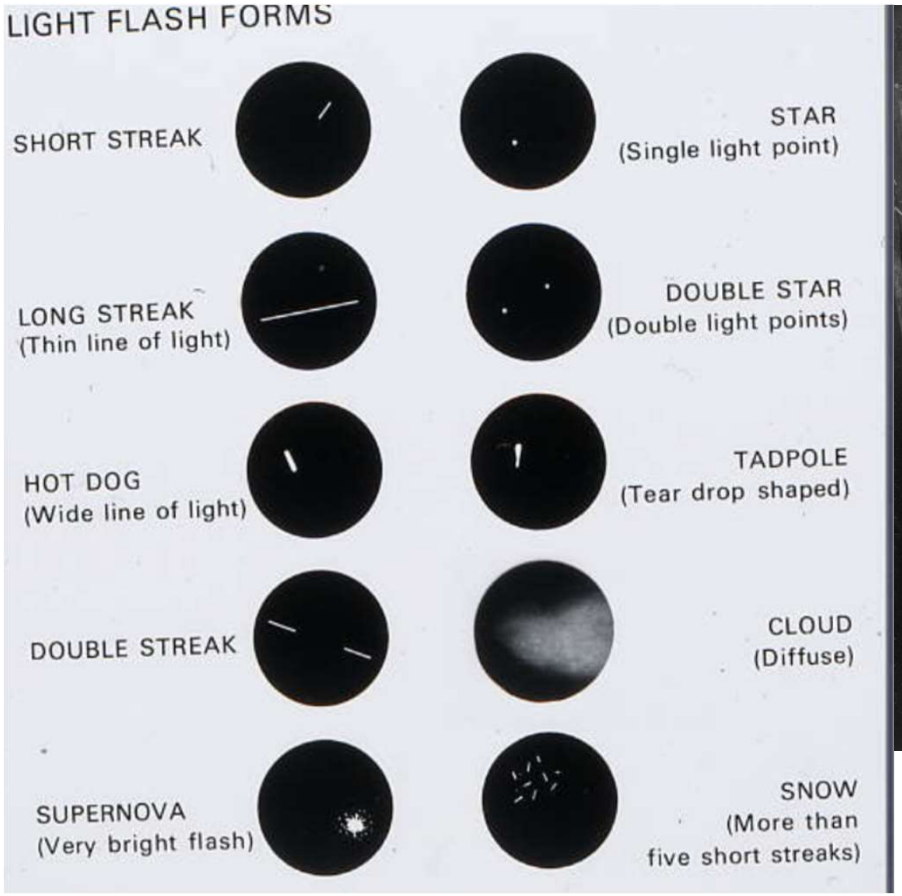
Space missions, outside of Earth's protective atmosphere and magnetic field, expose astronauts to many times more.

NASA

GROUND STUDY



Radiation effects

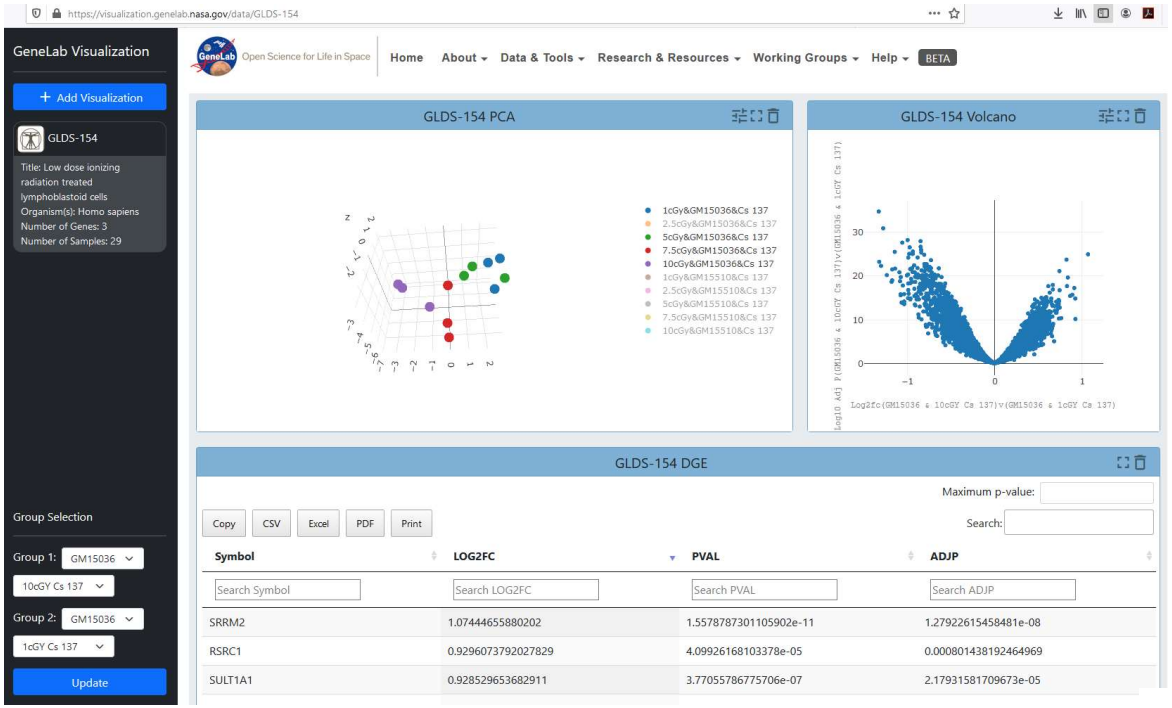


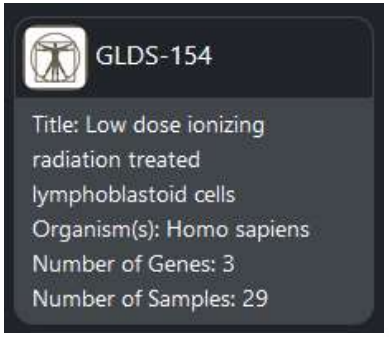
Budinger et al., Nature, 1972

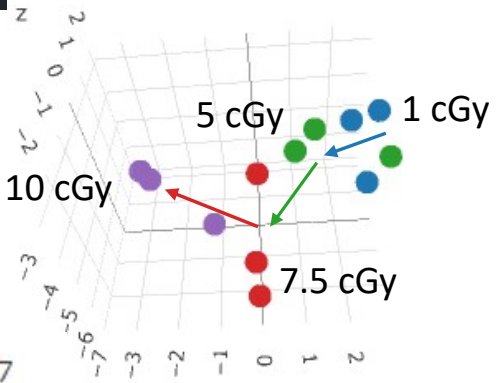
Bevatron Heavy-Ions Beam Group with Dr. Cornelius Tobias, Dr. Ed McMillan, and Dr. Thomas Budinger studying light flashes in nitrogen beam. (Photo courtesy of Lawrence Berkeley National Laboratory.)


CAUTION
480 VOLTS
NSRL
DIGITAL BEAM IMAGER

Low dose radiation transcriptomic data can be visualized







Application

Visualization Platform

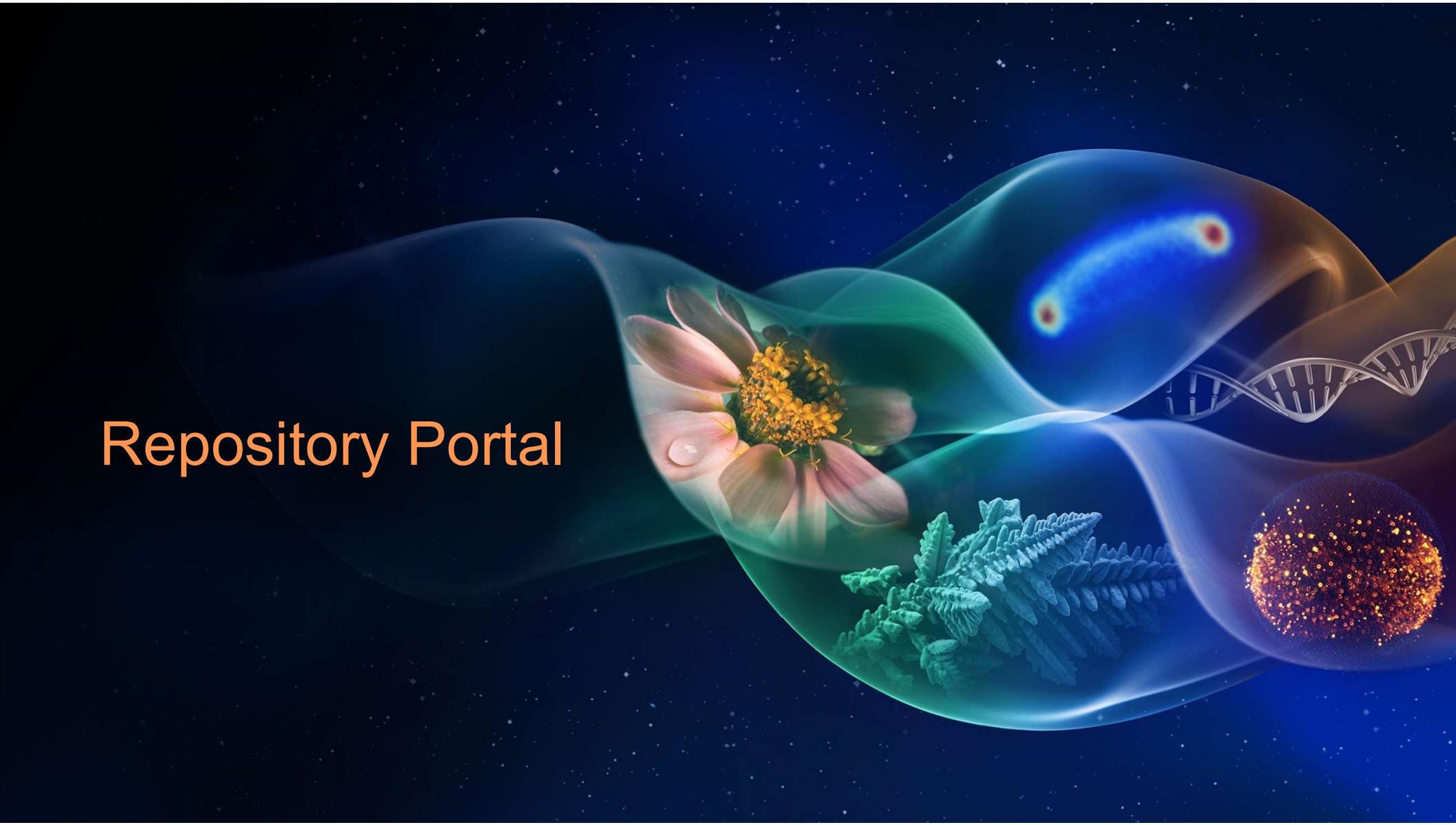
Wu P, Coleman M, Wyrobek AJ. "Low dose ionizing radiation treated lymphoblastoid cells", GeneLab, Version 3, http://doi.org/10.26030/hs0p-6w85

Video of mice in space

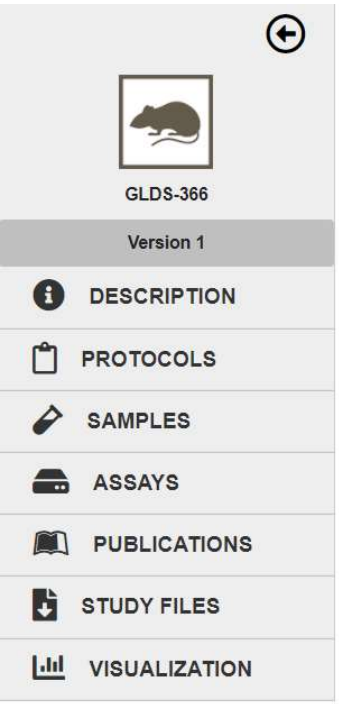
Application

Visualization Platform

Repository Portal



HRP and DOE Low dose funded research – All dataset are assigned a DOI (FAIR database)



Search Data

GLDS-366: Coalescence of DNA double strand breaks induced by galactic cosmic radiation is modulated by genetics in 15 inbred strains of mice

Version 1

Select a Version: 1

DOI: 10.26030/v8w4-rg83
Source Accession Number(s)
Total Data Volume: 64.2 MB

Submitted Date: 29-Sep-2020
Release Date: 19-Aug-2021

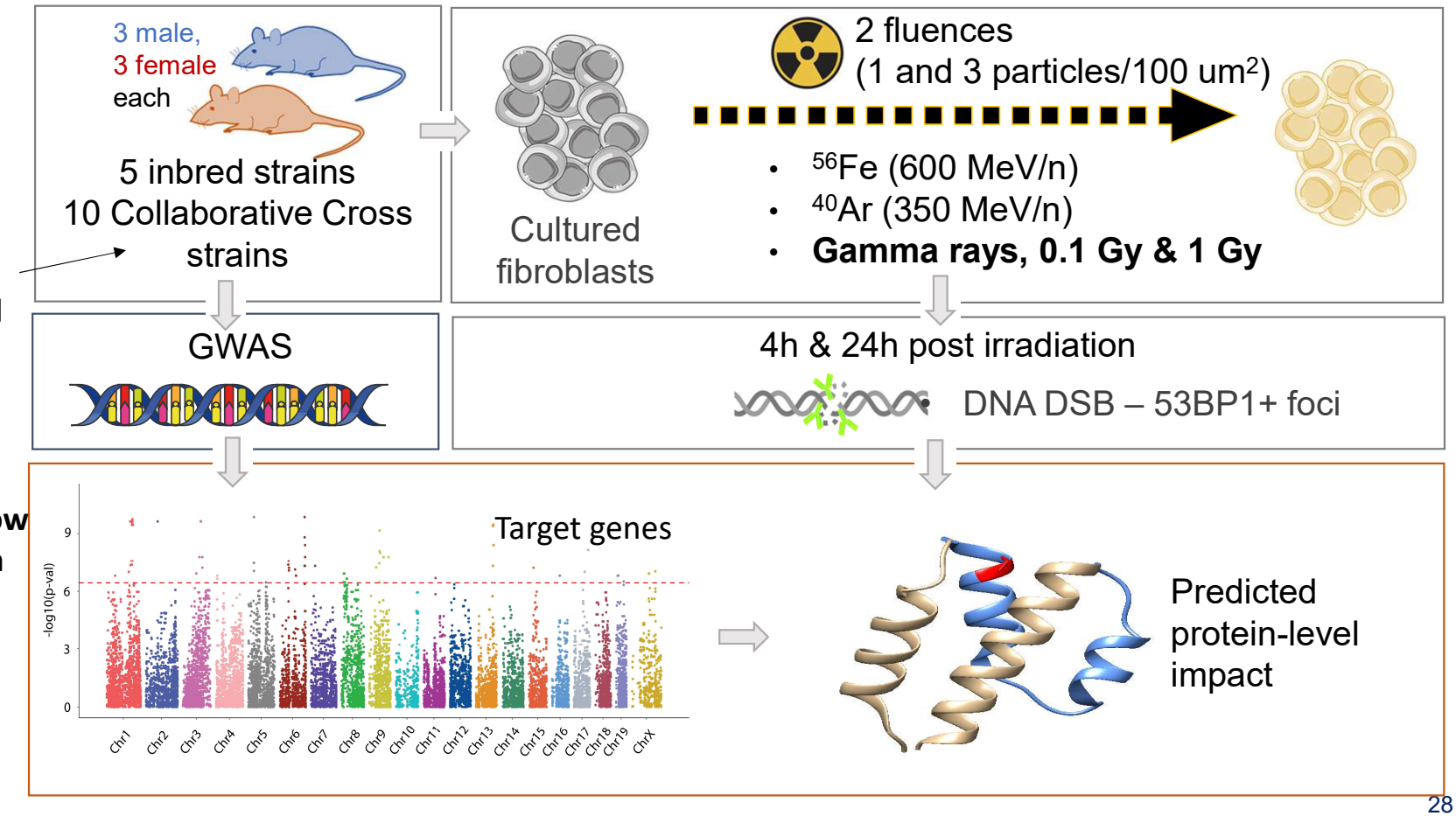
DESCRIPTION

Study Description

This study analyzes the variability of responses to simulated deep space radiation among 15 commonly used mouse strains. Ex vivo primary skin fibroblast responses to high mass-high charge particles and X-rays were analyzed by quantifying DNA damage-sensing protein 53BP1 positive radiation-induced foci (RIF) as a surrogate biomarker of DNA double strand breaks (DSBs). Primary skin fibroblasts were isolated from 10 collaborative cross strains and five reference inbred mice (C57Bl/6, BALB/CByJ, B6C3, C3H and CBA/CaJ) and exposed to 350 MeV/n Ar and 600 MeV/n Fe particles as well as X-rays. Our results indicate that nearby DSBs coalesced into repair units characterized by large RIFs. Such model has the advantage of being much more efficient molecularly, but is poorly suited to deal with cosmic radiation, where energy is concentrated along the particle trajectory. Thus, we observed a large density of DSBs along each particle track and the percentage of unrepaired DSBs that increased with linear energy transfer of the particle over 48 hours post irradiation. Furthermore, persistent RIF levels ex vivo were well correlated with T and B lymphocyte survival in vivo in 10 collaborative cross strains 24 hours after 0.1 Gy whole-body dose of X-rays, suggesting that persistent RIFs might serve as an ex vivo biomarker for in vivo radiation toxicity. Finally, we performed genome-wide association study to identify the genomic associations with dose responses to ionizing radiation, marked as Foci per Gray (FPG), as well as with background DNA repair levels (BGD). This dataset includes GWAS data as well as FPG and BGD values for each sample and condition.

Mouse variability and genomic associations with radiation responses

Low dose animal model developed by Lawrence Berkeley National Laboratory Low dose program



Different strains for GWAS are easily visible in sample tables

Application

Data Repository

GLDS-366

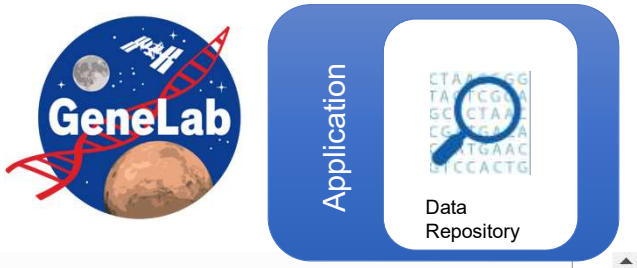
Version 1

- DESCRIPTION
- PROTOCOLS
- SAMPLES
- ASSAYS
- PUBLICATIONS
- STUDY FILES
- VISUALIZATION

SAMPLES

Source Name	Sample Name	Characteristics: Organism	Factor Value: Strain	Characteristics: sex	Characteristics: Material Type	P
B6C3F1	B6C3F1_P235_F4_B6C3...	Mus musculus	B6C3 Mouse	female	primary cultured fibroblast cell	growt
BALBCF1	BALBCF1_P235_A8_BAL...	Mus musculus	BALB/cByJ	female	primary cultured fibroblast cell	growt
C3HF1	C3HF1_P235_B4_C3H_F...	Mus musculus	C3H/HeMsNsrf	female	primary cultured fibroblast cell	growt
C57BLF1	C57BLF1_P235_A6_C57...	Mus musculus	C57BL/6J	female	primary cultured fibroblast cell	growt
CBAF1	CBAF1_P235_E2_CBA_F...	Mus musculus	CBA/CaJ	female	primary cultured fibroblast cell	growt
B6C3F1	B6C3F1_P234_F4_B6C3...	Mus musculus	B6C3 Mouse	female	primary cultured fibroblast cell	growt

Assays are selectable



GLDS-366

Version 1

- DESCRIPTION
- PROTOCOLS
- SAMPLES
- ASSAYS
- PUBLICATIONS
- STUDY FILES
- VISUALIZATION

Select Export Columns

ASSAYS/MEASUREMENTS

Assay Name: Genome-wide Association Study

Technology Platform: MegaMUGA

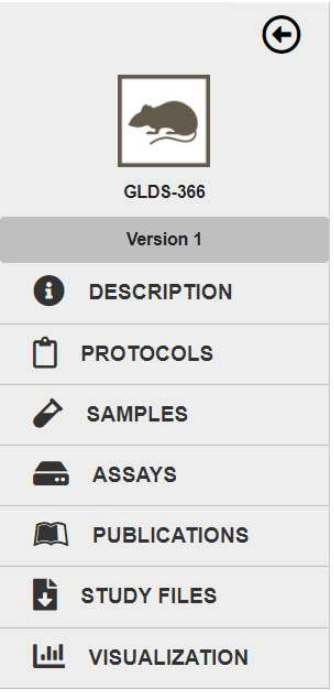
Technology Ty

Genome-wide Association Study

histology

Sample Name	Protocol REF	Parameter Value: QA Instrument	Parameter Value: Platform	Extract Name	Array Data File	Pr
B6C3	Genotyping	NanoDrop 2000 UV-vis spectrophotometer	MegaMouse Universal Genotyping Array (MegaMUGA platform)	B6C3	GLDS-366_SNP_Merged_GENO…	data t
BALBC	Genotyping	NanoDrop 2000 UV-vis spectrophotometer	MegaMouse Universal Genotyping Array (MegaMUGA platform)	BALBC	GLDS-366_SNP_Merged_GENO…	data t
C3H	Genotyping	NanoDrop 2000 UV-vis spectrophotometer	MegaMouse Universal Genotyping Array (MegaMUGA platform)	C3H	GLDS-366_SNP_Merged_GENO…	data t
C57	Genotyping	NanoDrop 2000 UV-vis spectrophotometer	MegaMouse Universal Genotyping Array (MegaMUGA platform)	C57	GLDS-366_SNP_Merged_GENO…	data t
CBA		NanoDrop 2000 UV-vis	MegaMouse Universal	CBA	GLDS-	

ALSDA non-omics histology data are linked to GeneLab via the assay selection



Select Export Columns

ASSAYS/MEASUREMENTS

Assay Name: histology

Technology Platform:

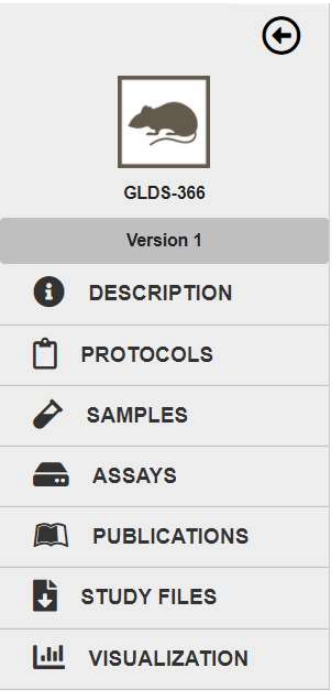
Technology Ty… py assay

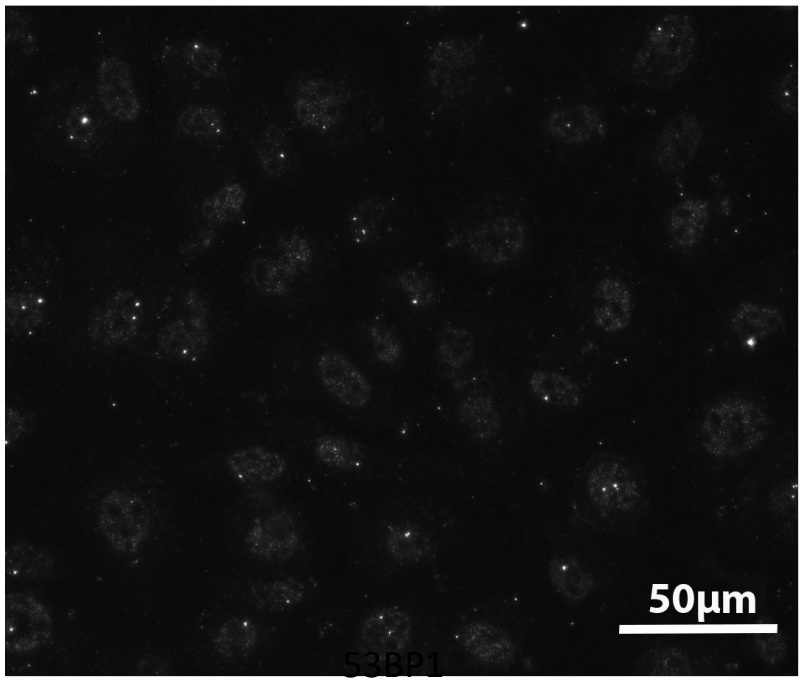
Genome-wide Association Study

histology

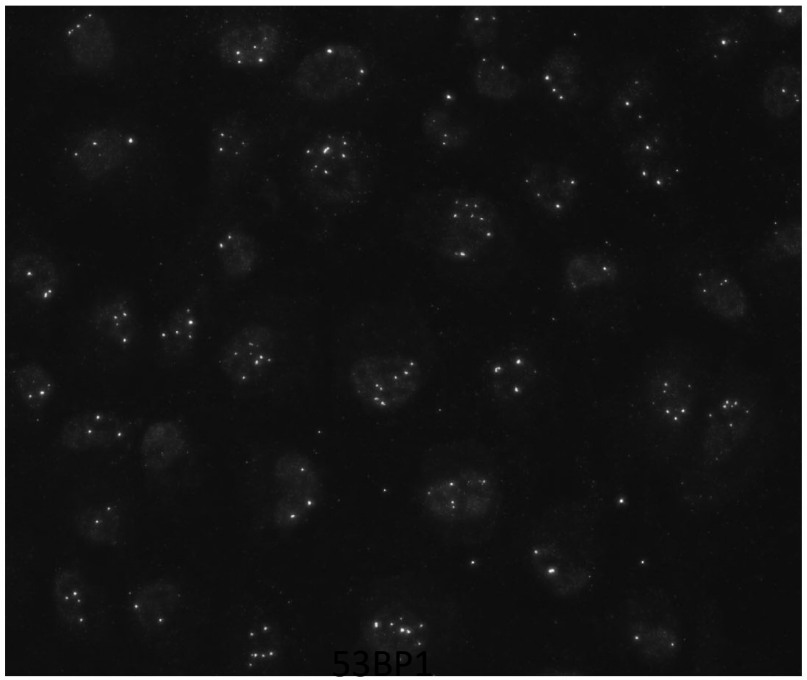
Sample Name	Protocol REF	Parameter Value: primary antibody	Comment: Dilution factor	Comment: Incubation time	Parameter Value: secondary antibody	Di
B6C3F1_P235_F4_B6C3…	immunostaining and imaging	rabbit polyclonal anti-53BP1	1 to 400	1 hour	Alexa Fluor 488 goat anti-rabbit	1 to 4
B6C3F2_P235_E3_B6C3…	immunostaining and imaging	rabbit polyclonal anti-53BP1	1 to 400	1 hour	Alexa Fluor 488 goat anti-rabbit	1 to 4
B6C3F3_P235_F3_B6C3…	immunostaining and imaging	rabbit polyclonal anti-53BP1	1 to 400	1 hour	Alexa Fluor 488 goat anti-rabbit	1 to 4
B6C3M1_P235_C3_B6C3…	immunostaining and imaging	rabbit polyclonal anti-53BP1	1 to 400	1 hour	Alexa Fluor 488 goat anti-rabbit	1 to 4
B6C3M2_P235_D4_B6C3…	immunostaining and imaging	rabbit polyclonal anti-53BP1	1 to 400	1 hour	Alexa Fluor 488 goat anti-rabbit	1 to 4
B6C3M3_P235_D3_B6C3…	immunostaining and imaging	rabbit polyclonal anti-53BP1	1 to 400	1 hour	Alexa Fluor 488 goat anti-rabbit	1 to 4

Images can be accessed via ALSDA (work in progress)





0 cGy, 4 hours post-IR, 360 KV Xray

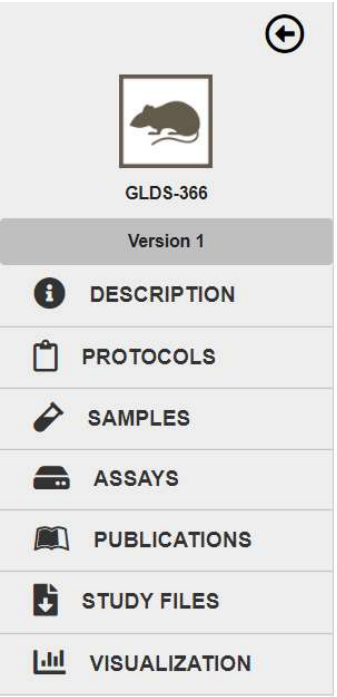


10 cGy, 4 hours post-IR, 360 KV Xray

Application

Data Repository

All peer-reviewed publications associated to dataset are listed



PUBLICATIONS

Considering Cell Proliferation to Optimize Detection of Radiation-Induced 53BP1-Positive Foci in 15 Mouse Strains Ex Vivo
Authors: Sebastien Penninckx, Eloise Pariset, Ana Uriarte Acuna, Stephane Lucas, Sylvain V. Costes
PubMed ID: 33181852
DOI: 10.1667/RADE-20-00165.1

Dose, LET and Strain Dependence of Radiation-Induced 53BP1 Foci in 15 Mouse Strains Ex Vivo Introducing Novel DNA Damage Metrics
Authors: Sebastien Penninckx, Egle Cekanaviciute, Charlotte Degorre, Elodie Guiet, Louise Viger, Stephane Lucas, Sylvain V. Costes
PubMed ID: 31081741
DOI: 10.1667/RR15338.1

53BP1 Repair Kinetics for Prediction of In Vivo Radiation Susceptibility in 15 Mouse Strains
Authors: Eloise Pariset, Sebastien Penninckx, Charlotte Degorre Kerbaul, Elodie Guiet, Alejandra Lopez Macha, Egle Cekanaviciute, Antoine M Snijders, Jian-Hua Mao, Francois Paris, Sylvain V Costes
PubMed ID: 32991727
DOI: 10.1667/RADE-20-00122.1

STUDY FILES

To view files, click on the folder of interest.

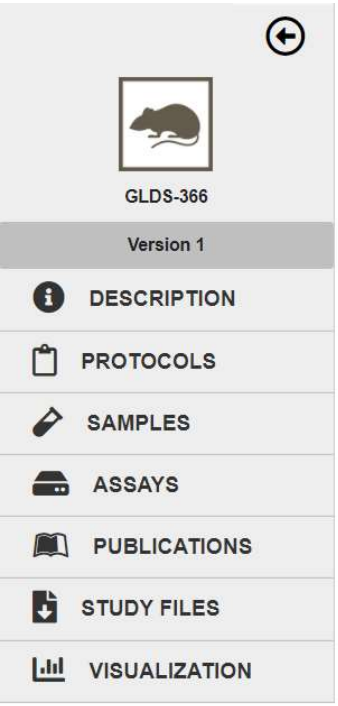
0 files selected

Download Selected Files

- All Files
  - Genome Wide Association Study Data Files
  - Histology Data Files
  - SNP array
  - Study Metadata Files

FILES	FILE SIZE	RESOURCE CATEGORY	RESOURCE DESCRIPTION
GLDS-366_GWAS_processed_associations.csv	53.82 MB	Processed data	Compressed collection of processed data files and quality report associated with this study. Formats are platform specific.

Data file manager allows access to all levels of data (from raw to processed)



STUDY FILES

To view files, click on the folder of interest.

- All Files
  - Genome Wide Association Study Data Files
    - Processed data
  - Histology Data Files
    - Raw Data Files
    - Processed Data Files
    - Raw Data Files/Supplemental Materials
  - SNP array
    - Raw Data Files
  - Study Metadata Files

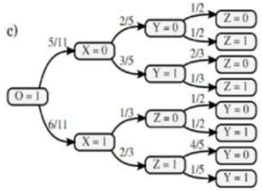
0 files selected

Download Selected Files

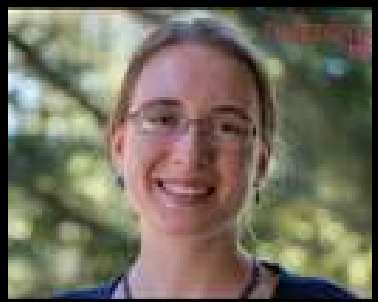
FILES	FILE SIZE	RESOURCE CATEGORY	RESOURCE DESCRIPTION
GLDS-366_GWAS_processed_associations.csv	53.82 MB	Processed data	Compressed collection of processed data files and quality report associated with this study. Formats are platform specific.
GLDS-366_Histology_raw_pheno_V3.csv	1.02 MB	Raw Data Files	raw or processed data files and quality report associated with this study. Formats are platform specific.
GLDS-366_Histology_processed_pheno_V2.csv	251.08 KB	Processed Data Files	raw or processed data files and quality report associated with this study. Formats are platform specific.
GLDS-366_Histology_Phenotypes_description.txt	6.51 KB	Raw Data Files/Supplemental Materials	raw or processed data files and quality report associated with this study. Formats are platform specific.
GLDS-366_SNP_Merged_GENO.txt	8.8 MB	Raw Data Files	An oligonucleotide DNA microarray used to detect polymorphisms in DNA samples.

FDL-2021
SPACE MEDIC

CRISP Causal Inference Model

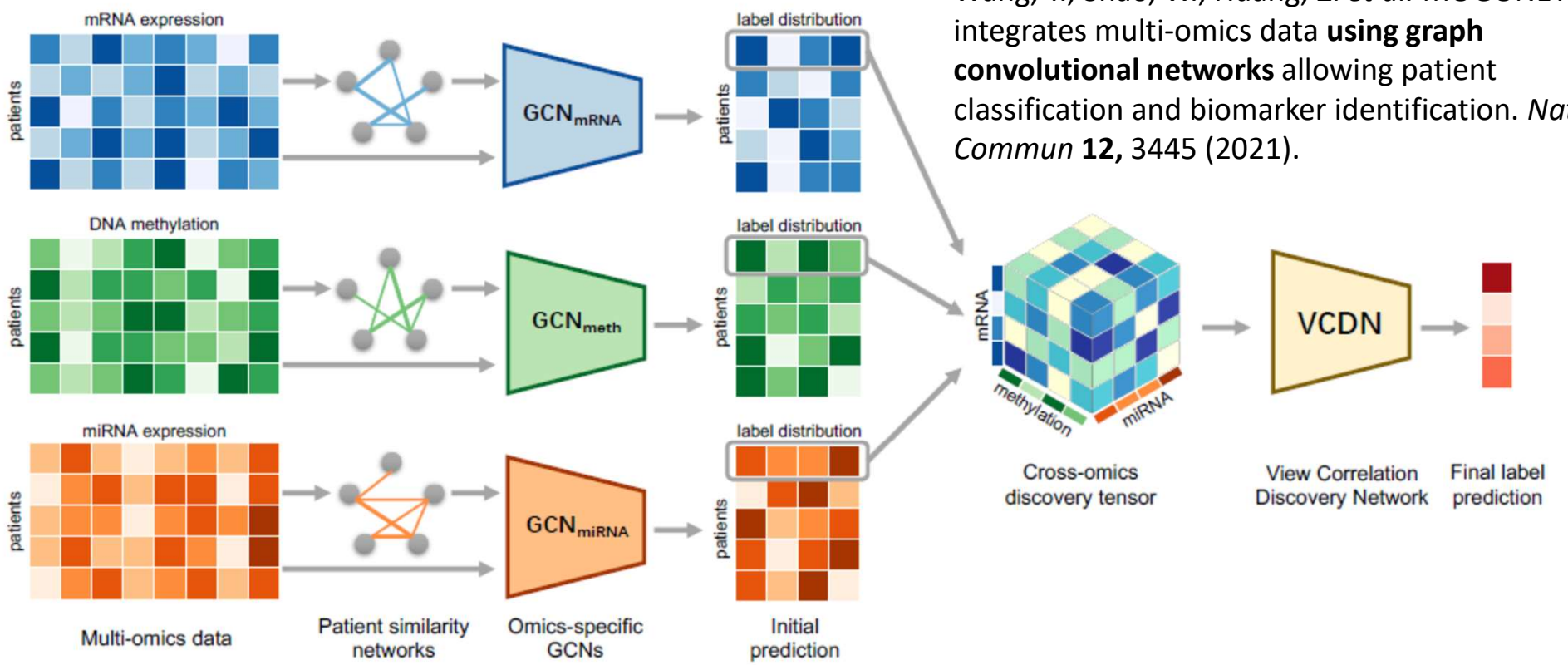


MODELING PORTAL (AI/ML)



Lauren Sanders, Ph.D.
AI/ML Scientist
GeneLab

The rapid advancement in high-throughput technologies have enabled the collection of various types of "omics" data at an unprecedented detailed level.



Wang, T., Shao, W., Huang, Z. *et al.* MOGONET integrates multi-omics data **using graph convolutional networks** allowing patient classification and biomarker identification. *Nat Commun* **12,** 3445 (2021).

GeneLab radiation data: enabling AI/ML meta-analysis

6 individual radiation exposure gene expression datasets:

GLDS-71: Immediate Transcriptional Changes in Response to High Dose Radiation Exposure Version 4

0.3 Gy – Cs-137

GLDS-152: Transcription profiling of human peripheral blood to development gene expression signatures for practical radiation biodosimetry Version 2

0.5 Gy – gamma ray

GLDS-156: Identifying radiation exposure biomarkers from mouse blood transcriptome Version 1

2 Gy – Cs-137

GLDS-157: Gene expression in human peripheral blood 48 hours after exposure to ionizing radiation Version 2

0.5 Gy – gamma ray

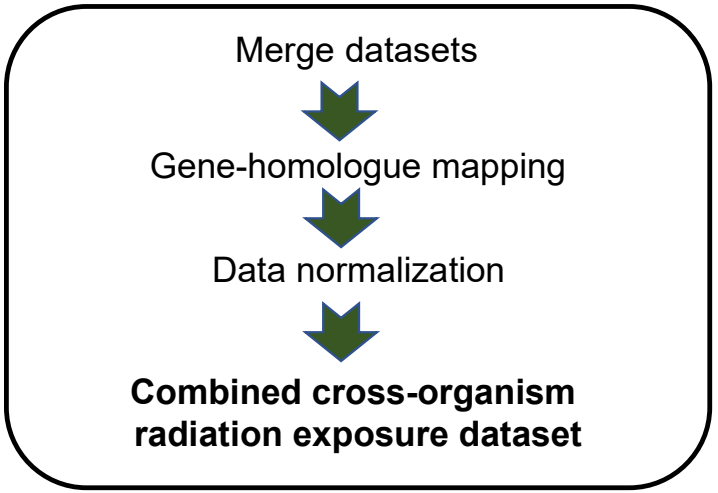
GSE124612: Transcriptomic responses in mouse blood during the first week after in vivo gamma irradiation

1.5 Gy – gamma ray

GSE62623: Gene expression in mouse blood following low dose-rate or acute x-ray exposure

1.1 Gy – X ray

Data Pipeline:



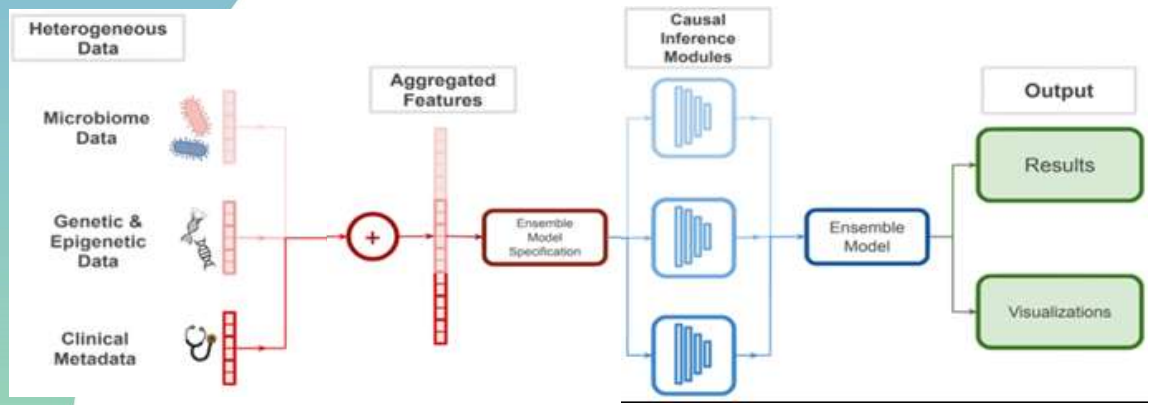
Combined Dataset:

- ~25,000 human-mouse gene homologues
- 455 total samples

- Gamma irradiation or non-irradiated controls:
  - *ex vivo* irradiation of human blood samples
  - *in vivo* irradiated mouse whole blood

Causal inference in complex biological data

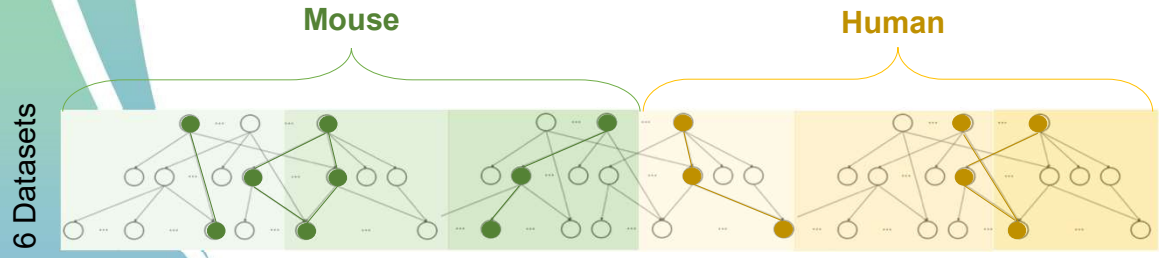
CRISP Platform Overview



Generalizable to different data types with the ability to learn causal drivers of any feature within a dataset

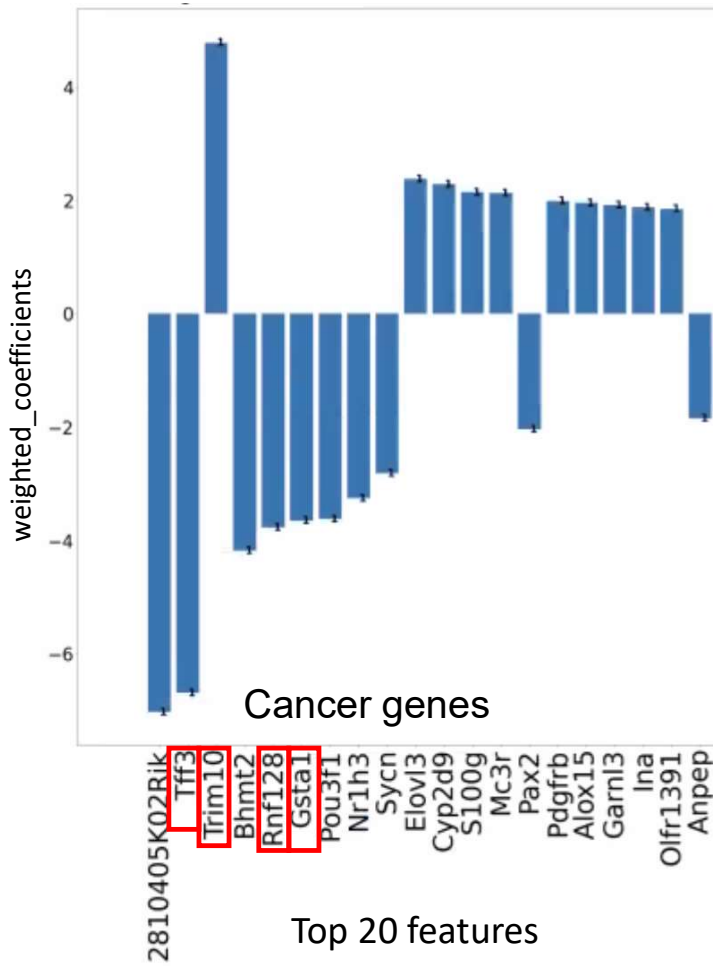
Automated dimensionality reduction using machine learning selects the most important input variables for causal analysis

An ensemble of causal inference models "votes" on the selection of causal features



TARGET: radiation state (irradiated or control)

CRISP identifies features that are *causal* of a target variable



Developed by FDL Astronaut Health 2020 Team

GeneLab – Open Science for Life in Space

genelab.nasa.gov

Data Portal

- First multi-omics space related database
- 301 publicly available studies
- High-order data for transcriptomics and metagenomics studies
- Spaceflight environmental data associated with sample metadata (radiation dosimetry, temperature, humidity)

Tools

- User friendly submission portal to submit and publish data
- Analysis platform to tools and workflows to analyze your own data or data from the GeneLab repository.
- Tutorials and online resources to learn how to analyze RNA-Seq data
- Workspace to store, share, and organize data files

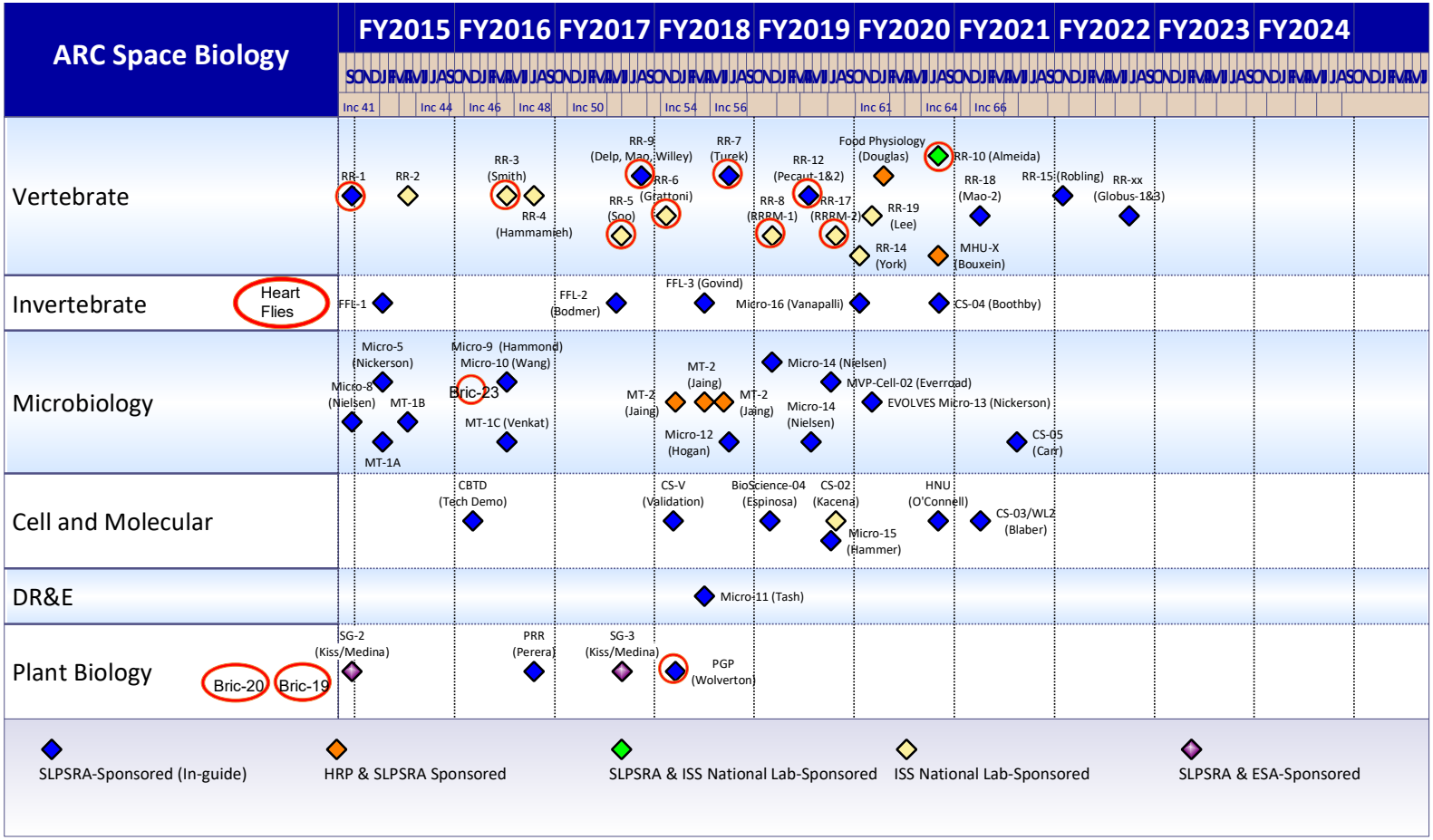
Sequencing

- State-of-the-art sequencing facility to process spaceflight samples
- Optimized SOPs with standard processing workflow
- Provide sequencing service for NASA funded PI without any university overhead
- Generate high quality data from shared tissues for open science access

Community

- Analysis Working Groups comprised of over 100+ scientists worldwide collaborating and analyze space omics data.
- Education Working Group focused on providing resources to educators and students to learn about space biology and bioinformatics

Adding value: Data generation



◯ = GeneLab data generation (actual and planned)

NASA Omics In-house Data Generation

Ames Research state-of-the-art sequencing facility, GeneLab Sequencing Group processes and sequences samples from spaceflight experiments. Optimized for spaceflight experiments, GeneLab uses standard protocols and automation to maximize the scientific return of each experiment.

GeneLab – Open Science for Life in Space

genelab.nasa.gov

Data Portal

- First multi-omics space related database
- 301 publicly available studies
- High-order data for transcriptomics and metagenomics studies
- Spaceflight environmental data associated with sample metadata (radiation dosimetry, temperature, humidity)

Tools

- User friendly submission portal to submit and publish data
- Analysis platform to tools and workflows to analyze your own data or data from the GeneLab repository.
- Tutorials and online resources to learn how to analyze RNA-Seq data
- Workspace to store, share, and organize data files

Sequencing

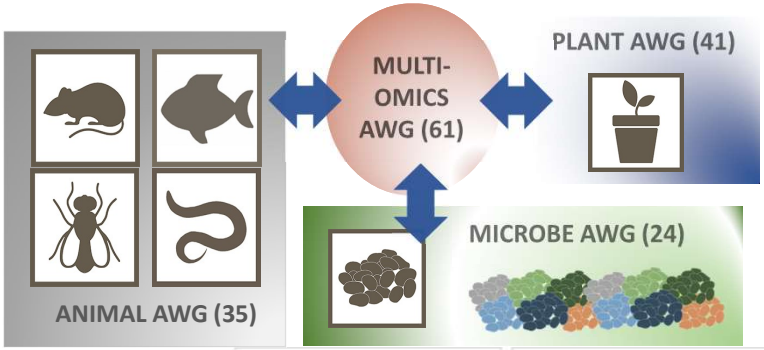
- State-of-the-art sequencing facility to process spaceflight samples
- Optimized SOPs with standard processing workflow
- Provide sequencing service for NASA funded PI without any university overhead
- Generate high quality data from shared tissues for open science access

Community

- Analysis Working Groups comprised of over 100+ scientists worldwide collaborating and analyze space omics data.
- Education Working Group focused on providing resources to educators and students to learn about space biology and bioinformatics

GeneLab Power Users

- ***GeneLab Analysis Working Groups (AWGs)*** *consist of* ***130+ scientists*** *from multiple space agencies, international institutions, and industry. Scientists meet monthly with each group to analyze data in the GeneLab repository. Majority of members are non-NASA PI's – many have applied for NASA funding following AWG interactions.*
- ***https://genelab.nasa.gov/awg/join***



AWG Members represent:

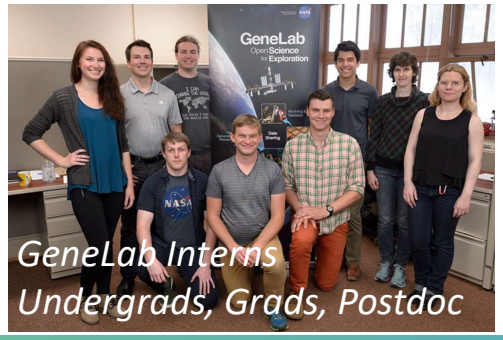
- 48 US Universities
- 4 NASA Centers
- 4 Other Government-funded Organizations
- 3 Institutes or Private Industry
- 3 International Universities

- ***Educational Working Work (EWG)*** *consists of educators and scientists spanning across high school and college, focusing on developing content to learn about spaceflight OMICS, promoting the use of GeneLab's database via the repository, and the visualization and analytical portals.*
  - *GeneLab for High Schools (GL4HS)*
  - *GeneLab for College/University (GL4U – Pilot with SJSU)*

SJSU SAN JOSÉ STATE UNIVERSITY

GeneLab included in Bioinformatics curriculum of degree granting university

GL4HS (Tutorial + Tools)



GeneLab Interns
Undergrads, Grads, Postdoc

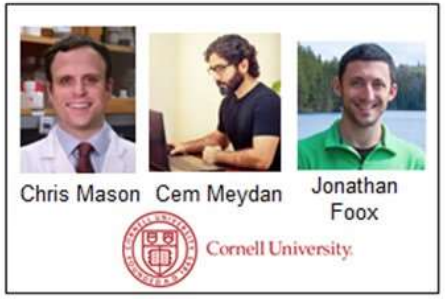
Annual AWG Workshop

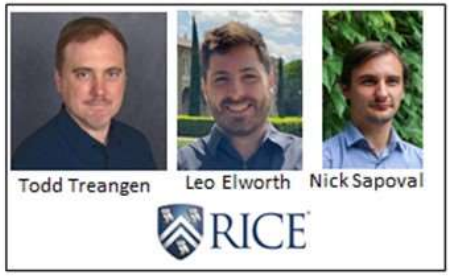
58	37	70+	100+
Original Publication linked to GeneLab	Derived Publication linked to GeneLab	Presentations linked to GeneLab	Datasets used in derived publications

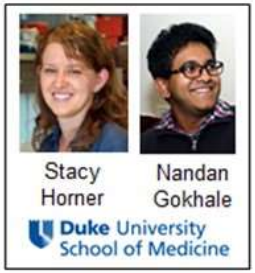
Cell Press- The biology of spaceflight

- A coordinated package of 29 scientific papers published in five *Cell Press* journals
  - including *Cell*, *Cell Reports*, *iScience*, *Cell Systems*, and *Patterns*
- These manuscripts span >200 investigators from dozens of academic, government, aerospace, and industry groups
- Representing the largest set of astronaut data and space biology data ever produced
- Including longitudinal multi-omic profiling, single-cell immune and epitope mapping, novel radiation countermeasures, and detailed biochemical profiles of 59 astronauts
  - Represents >10% of all humans that have ever been in space
- 9 papers utilize data or resources in GeneLab

Analysis Working Groups (AWGs) are driving data reuse



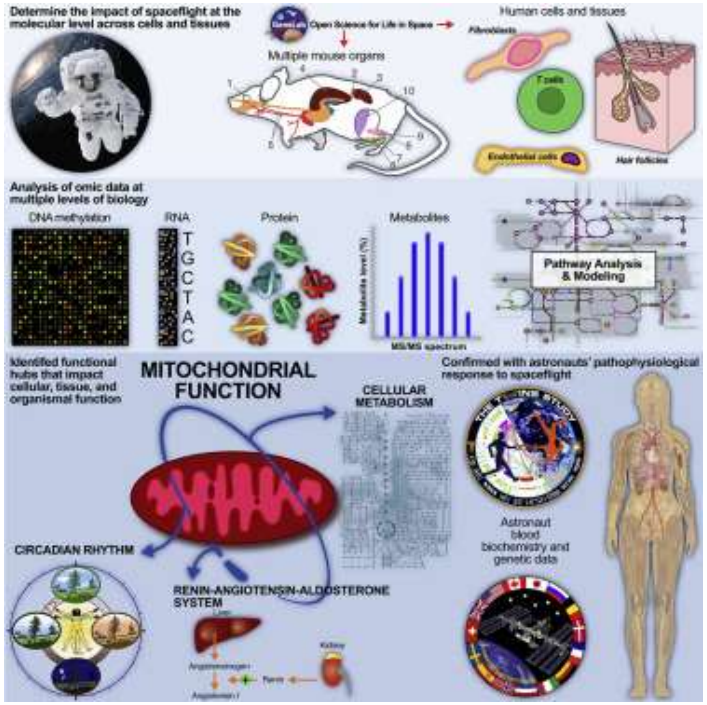




Comprehensive Multi-omics Analysis Reveals Mitochondrial Stress as a Central Biological Hub for Spaceflight Impact

Authors: Willian A. da Silveira, Hossein Fazelinia, Sara Brin Rosenthal, Evagelia C. Laiakis, Man S Kim, Cem Meydan, Yared Kidane, Komal S. Rathi, Scott M. Smith, Benjamin Stear, Yue Ying, Yuanchao Zhang, Jonathan Foox, Susana Zanello, Brian Crucian, Dong Wang, Adrienne Nugent, Helio A. Costa, Sara R. Zwart, Sonja Schrepfer, R. A. Leo Elworth, Nicolae Sapoval, Todd Treangen, Matthew MacKay, Nandan S. Gokhale, Stacy M. Horner, Larry N. Singh, Douglas C. Wallace, Jeffrey S. Willey, Jonathan C. Schisler, Robert Meller, J. Tyson McDonald, Kathleen M. Fisch, Gary Hardiman, Deanne Taylor, Christopher E. Mason, Sylvain V. Costes, Afshin Beheshti

Journal: Cell



Highlights:

- Multi-omics analysis and techniques with NASA's GeneLab platform.
- The largest cohort of astronaut data to date utilized for analysis.
- Mitochondrial dysregulation driving spaceflight health risks.
- NASA Twins Study data validates mitochondrial dysfunction during space missions.

Relevance and Impact:

- Uncovered insights into fundamental biological mechanisms affected by spaceflight.
- Power of comparing and integrating multiple omics and data types to understand further how life adapts to spaceflight conditions.
- This concept can guide new nutritional and pharmaceutical interventions and studies that will increase the viability of long-term human-crewed space missions.

RadBio Acknowledgments

Radiation Biophysics Laboratory
NASA Ames Research Center
Egle Cekanaviciute
Sonali Verma
Sherina Malkani
Alejandra Lopez Macha
Eloise Pariset

University of Bologna
Giulia Babbi

Brookhaven National Laboratory
Adam Rusek
Peter Guida
NSRL Support Staff

Funding
NASA HRP
DOE Low Dose Program

Former Collaborators

LBNL: Gary Karpen, Jian-Hua Mao, Antoine Snijders

NASA: Steve Blattnig, Artem Ponomarev, Ianik Plante

CSU: Mike Weil

UCSF: Mary Helen Barcellos-Hoff

INSERM, France: François Paris

Université de Namur, Belgium: S. Penninckx, S. Lucas

Exogen Biotechnology Inc.: Jonathan Tang, Antony S. Tin

Thank you!

Twitter: @NASAGeneLab
https://genelab.nasa.gov

2017-2018

2018-2019

GeneLab is funded by the NASA Space Biology program within the NASA Science Mission Directorate's (SMD) Biological and Physical Sciences (BPS) Division