Introduction to NASA Open Science and GeneLab

National Aeronautics and Space Administration

GL4U: Introduction 2024
Lecture 2 of 4

Amanda M. Saravia-Butler, Ph.D.

NASA GeneLab Science Lead

Contractor: KBR

Biological & **P**hysical **S**ciences

Acknowledgements

Open Science for Life in Space Teams

Open Science Analysis Working Group Members

Support

NASA Space Biology Program
NASA Science Mission Directorate
NASA Human Research Program
NASA Biological and Physical Sciences

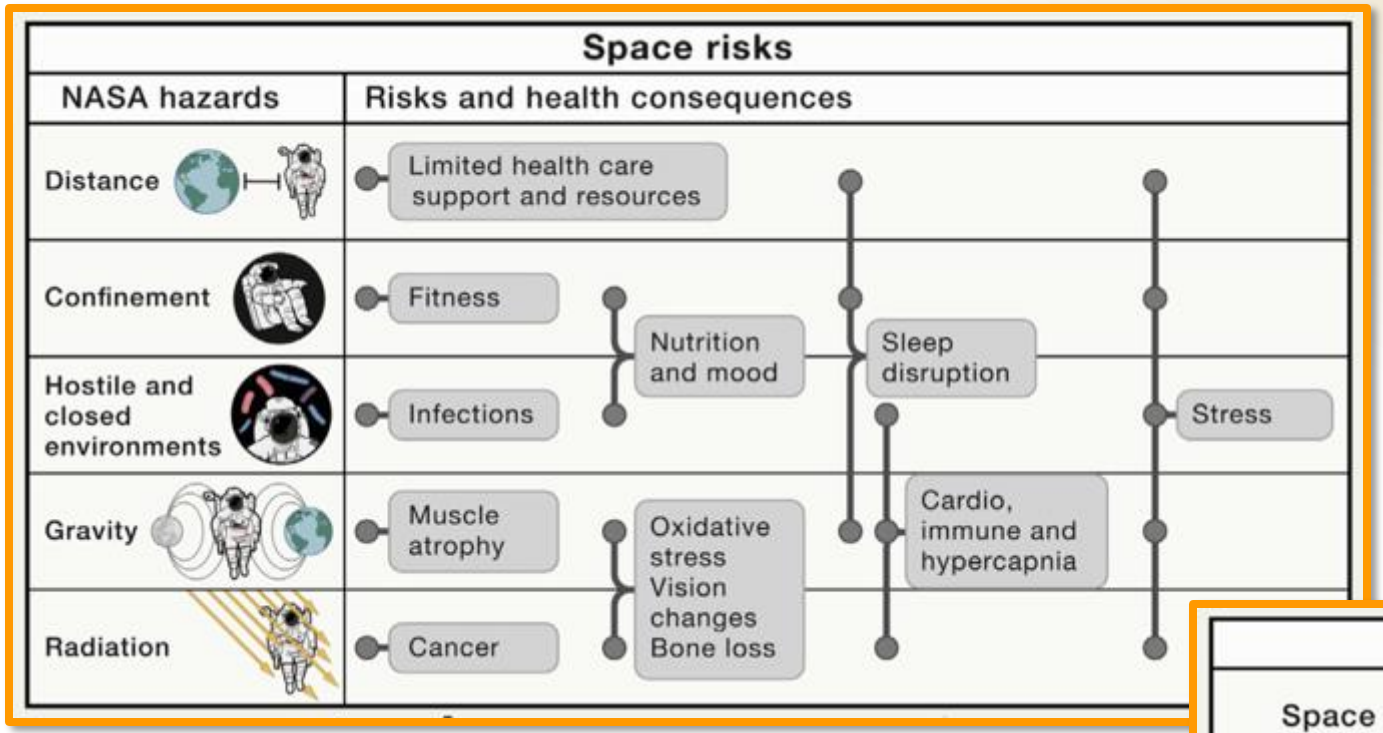
NASA Open Science Overview

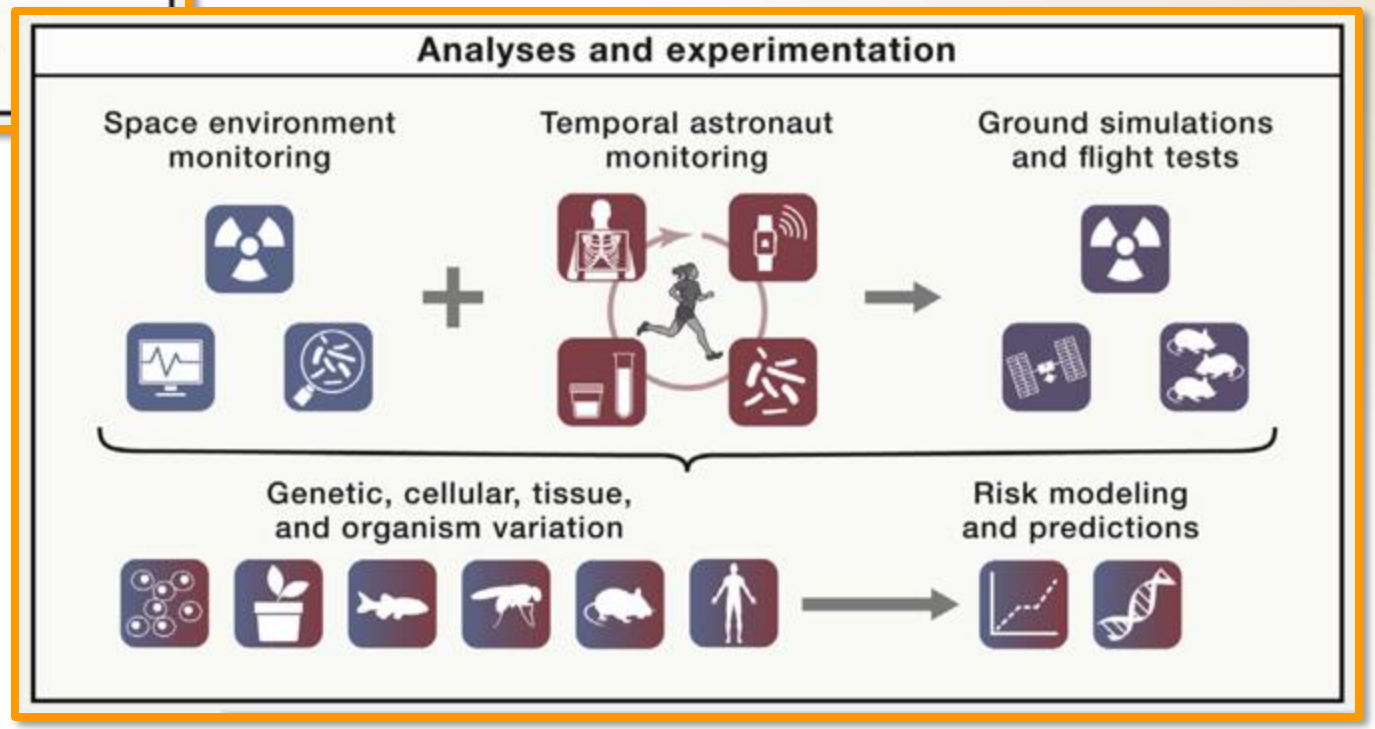
Open Science

"We define open science as a collaborative culture enabled by technology that empowers the **open sharing of data, information, and knowledge** within the scientific community and the wider public to accelerate scientific research and understanding."

Ramachandran, R., Bugbee, K., & Murphy, K. J. **Moving from Open Data to Open Science**. Earth and Space Science, Wiley Publication. https://doi.org/10.1029/2020EA001562

Space Travel Introduces Risks for Living Systems

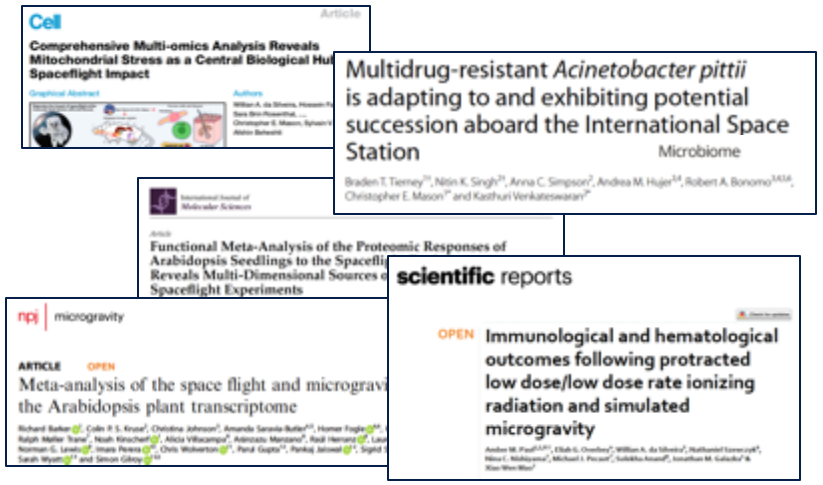




Afshinnekoo et al., *Cell*, 2020

Reusable Data is Essential for Life and Humans to Thrive in Deep Space

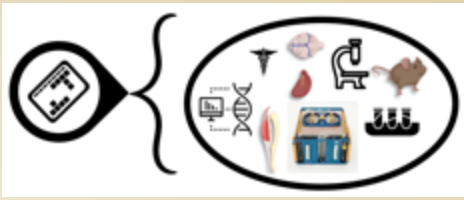
Knowledge Discovery



Advanced Computational Methods

Model development
Active and adaptive learning
Causal and network inference
Explainable AI
Federated learning
Few- or one-shot learning
Generative learning (synthetic datasets)
Knowledge graphs
Lifelong learning (living lab)
Out-of-distribution learning
Transfer learning

Telemetry for Biomedical Context

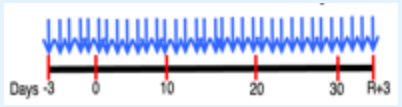


Analysis Across Missions

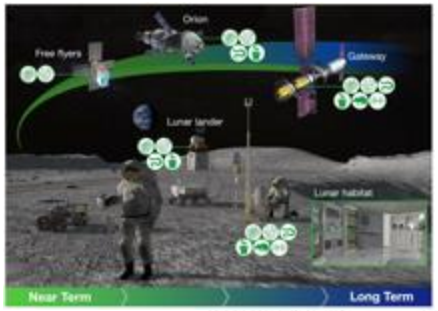
Across Hazards

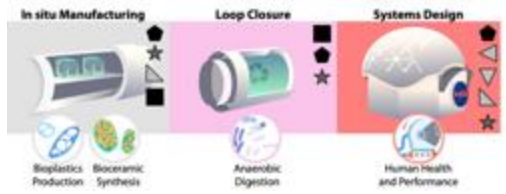
Across Species, Multimodal Data Types

Longitudinal, Biomonitoring, & Experimental

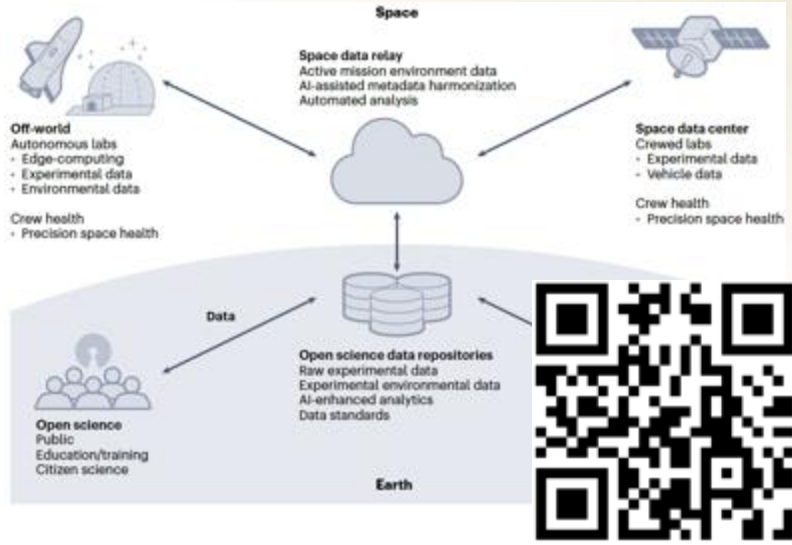


Novel Ecosystem, Platform, Countermeasure Development, Health Risk Quantification





Involve the World in Spaceflight



(Scott et al., 2023 https://rdcu.be/c8jSO;
Sanders et al., 2023 https://rdcu.be/c8jSS)

(Berliner et al., 2022 https://doi.org/10.1038/s44172-022-00012-9;
Everroad et al., 2021 https://ntrs.nasa.gov/citations/20210023324)

NASA Biological Open Science Resources

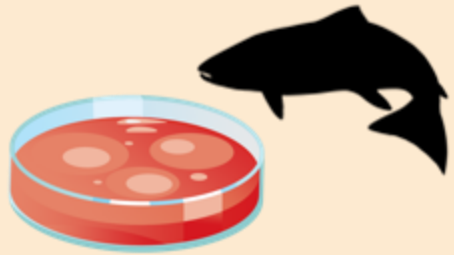
Biospecimen Sharing Program (BSP)

Dissection and preservation of rodent tissues from Flight and Ground investigations

Coordination of internal tissue sharing

NASA Biological Institutional Scientific Collection (NBISC)

Collection of non-human specimens and space microbial culture



NASA GeneLab (GL)

Collection and curation of omics data

Ames Life Sciences Data Archive (ALSDA)

Collection and curation of mission, project, and imaging data

NASA Internal Program

Open Source Science Programs — Available Globally

NASA Open Science Data Repository Overview

Physiological/Phenotypic/Imaging/ Environmental Telemetry Data

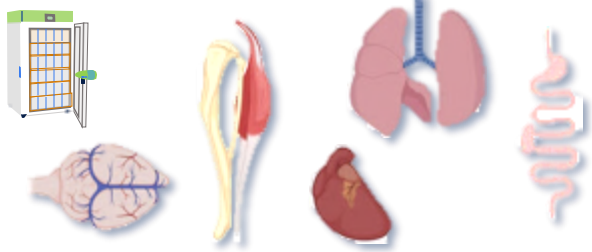
Tabular, text, imaging, video, code

ALSDA

Molecular/Omics Data

GeneLab

Biospecimens

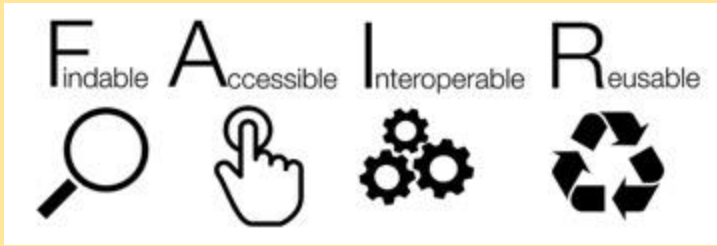


NBISC

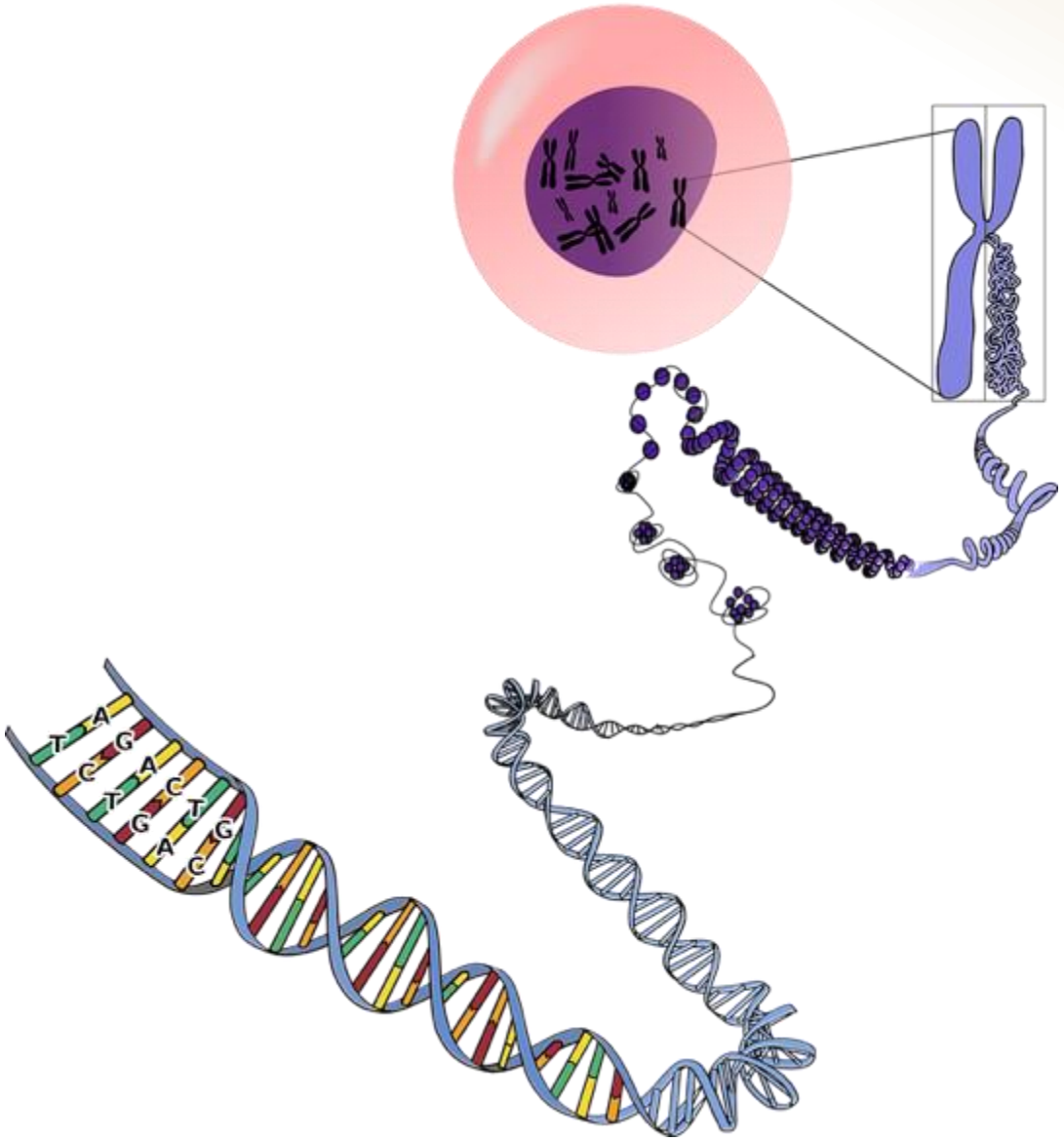
NASA Open Science Data Repository (OSDR)

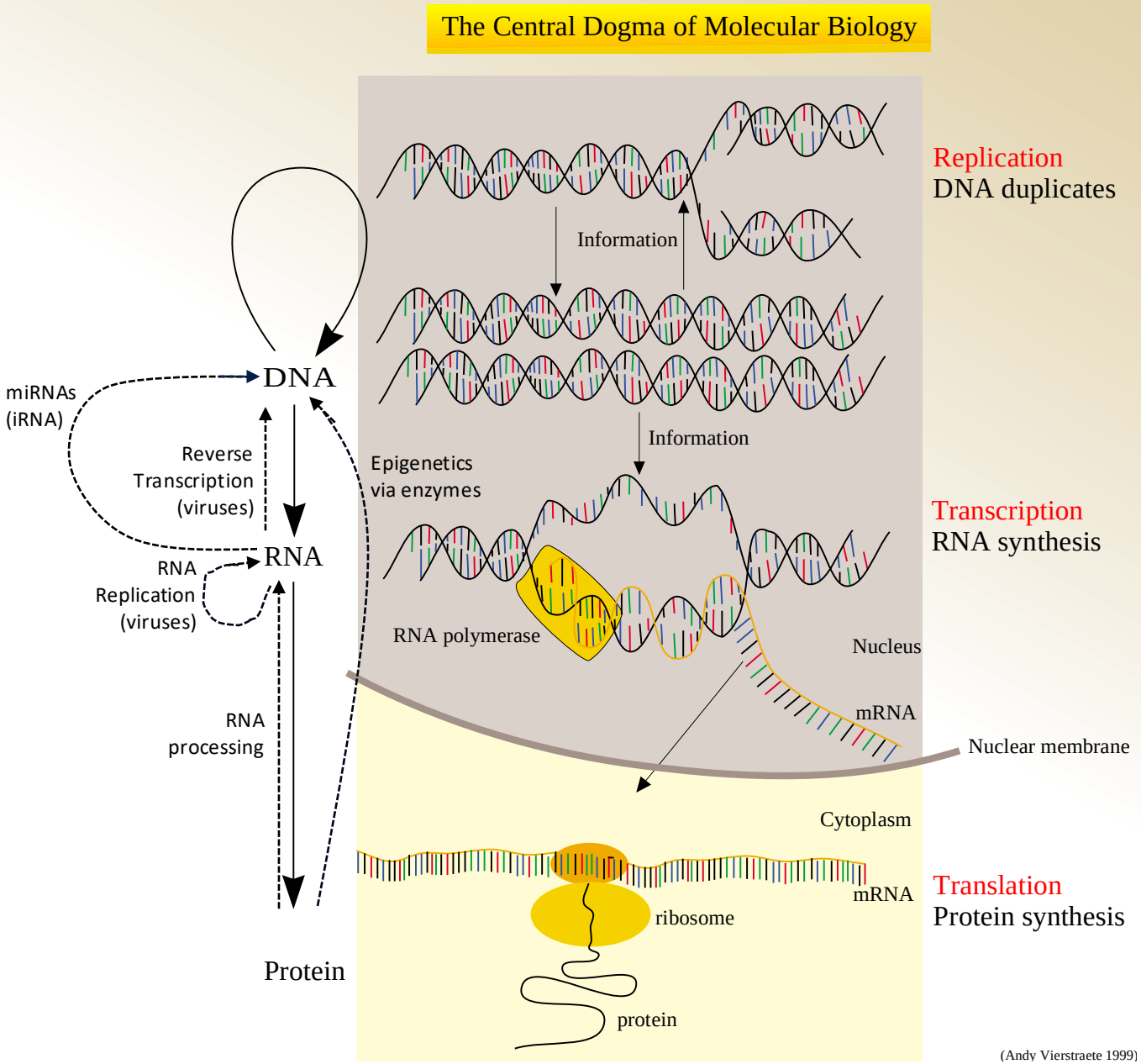
osdr.nasa.gov/bio

- Single Submission Portal (BDME)
- User Interface/Website Tool for RDSAs (Research Data Submission Agreements)
- Maximally Open Access with Necessary Controls for Sensitive Data
- Data Maximally FAIR



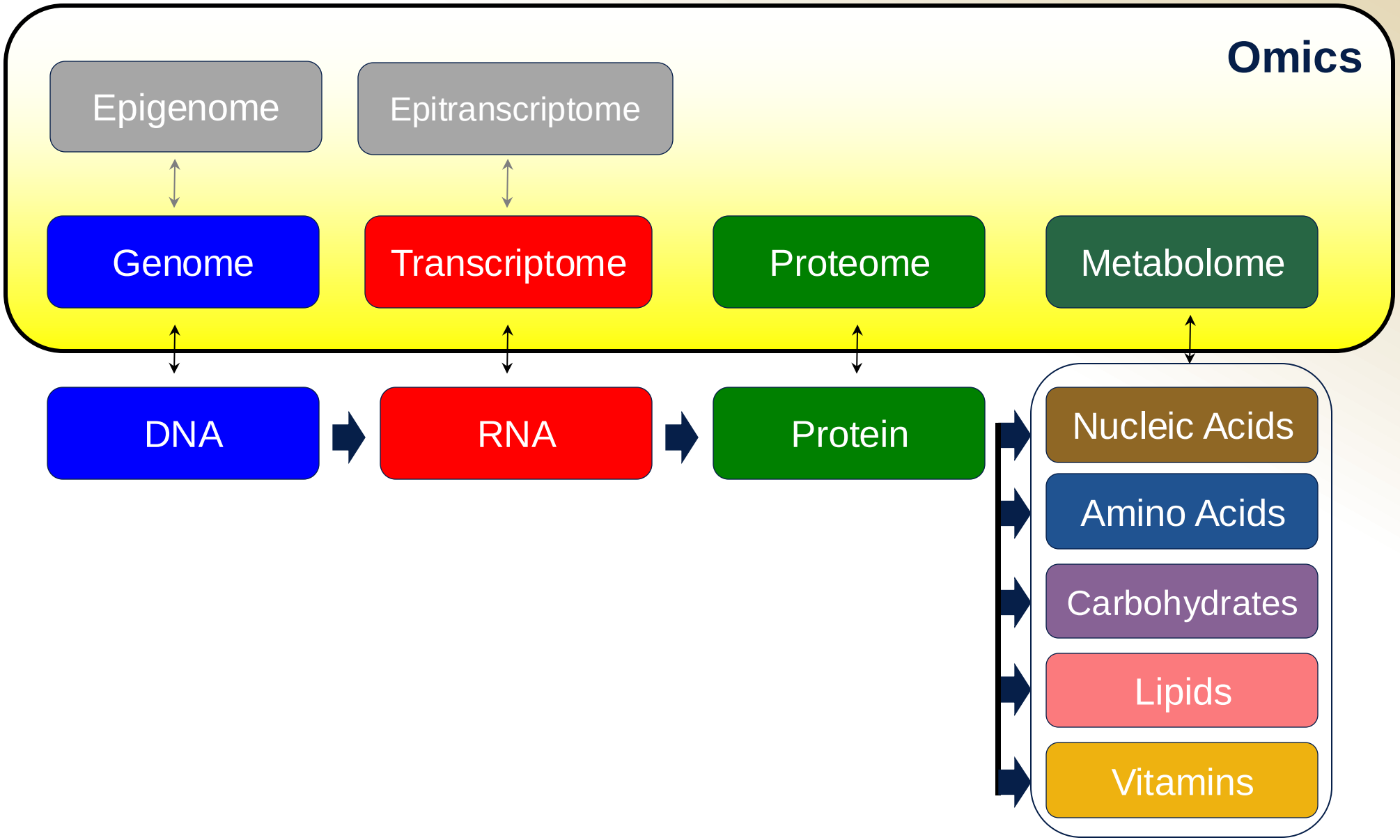
What are ‘omics?





(Andy Vierstraete 1999)

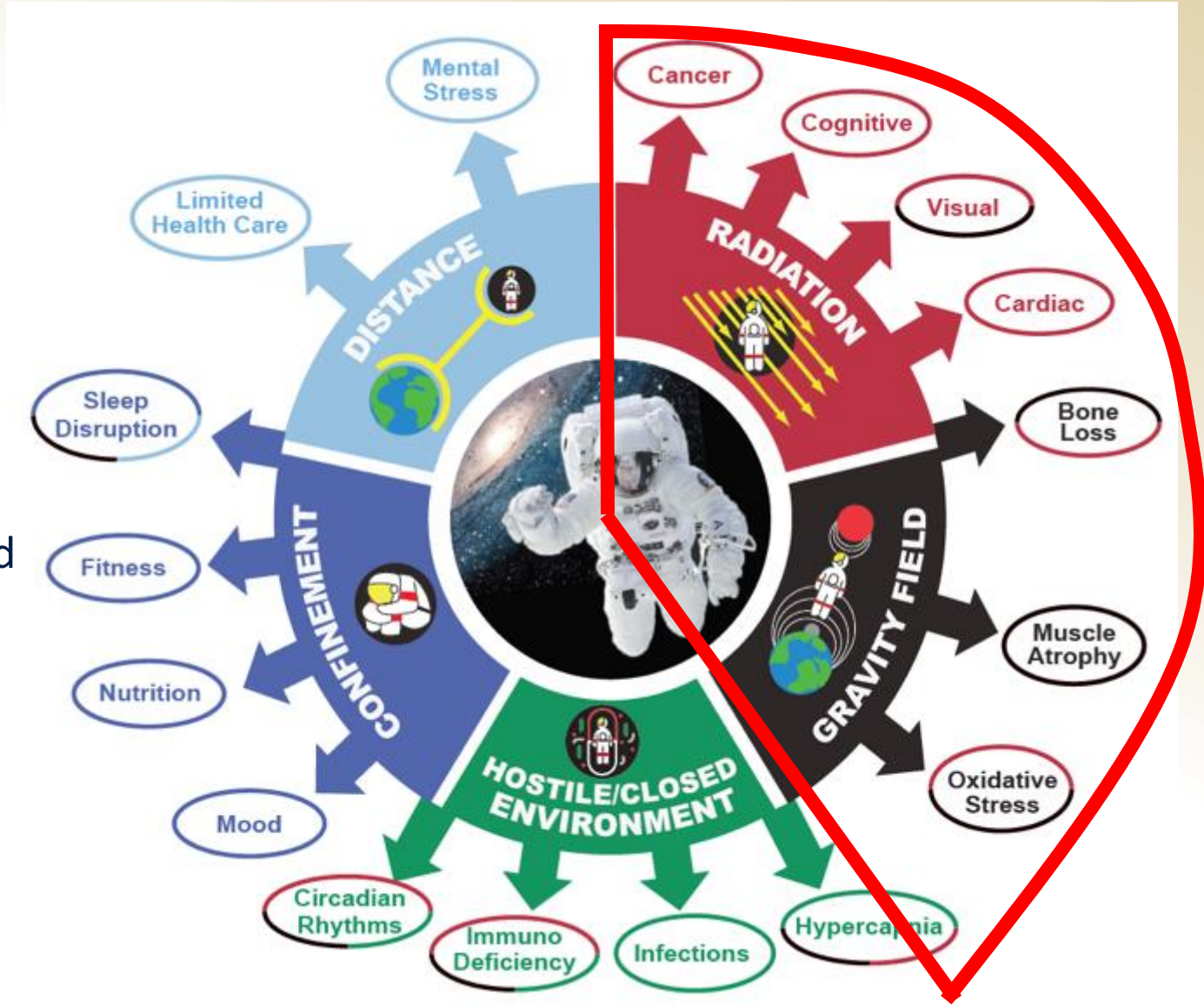
What are ‘omics?



Why is studying omics important for spaceflight?

- What, when, and where genes are expressed allow for cell type diversity and enable living organisms to respond and adapt to surroundings
- Gene expression is primarily regulated by environmental factors both micro (cell's micro-environment) and macro (organism's external stimuli or stressors)
- Spaceflight alters the transcriptional patterns and molecular signaling networks within our cells, which in turn causes physiological changes
- Understanding such changes will enable development of mitigation strategies to better withstand the rigors of long-duration spaceflight

Primary Stressors of Spaceflight



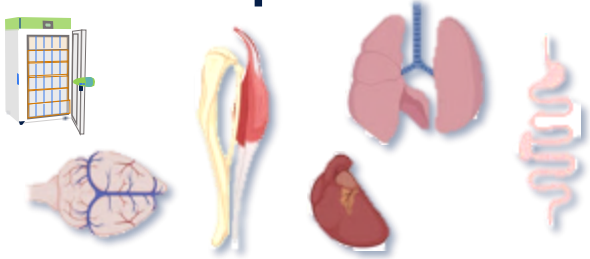
NASA Open Science Data Repository Overview

Physiological/Phenotypic/Imaging/ Environmental Telemetry Data

Tabular, text, imaging, video, code

Molecular/Omics Data

Biospecimens



ALSDA

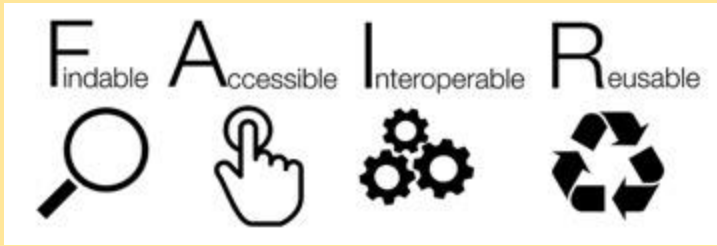
GeneLab

NBISC

NASA Open Science Data Repository (OSDR)

osdr.nasa.gov/bio

- **Single Submission Portal (BDME)**
- **User Interface/Website Tool for RDSAs (Research Data Submission Agreements)**
- **Maximally Open Access with Necessary Controls for Sensitive Data**
- **Data Maximally FAIR**



NASA OSDR – By The Numbers

505 Studies

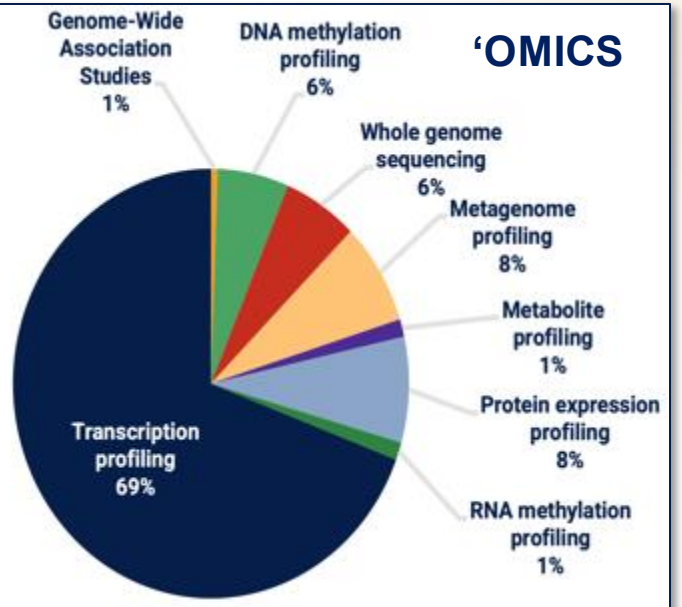
969 Datasets

45 Species

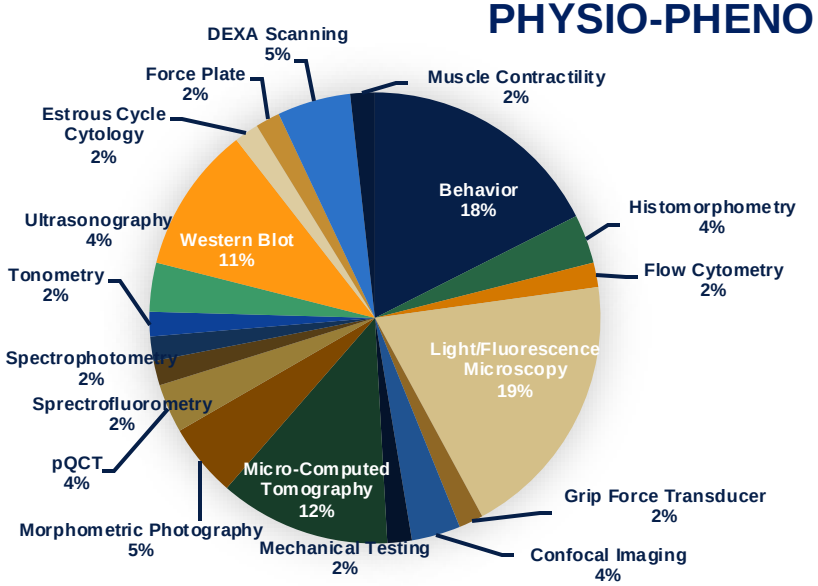
>80 Assays

>160TB Data

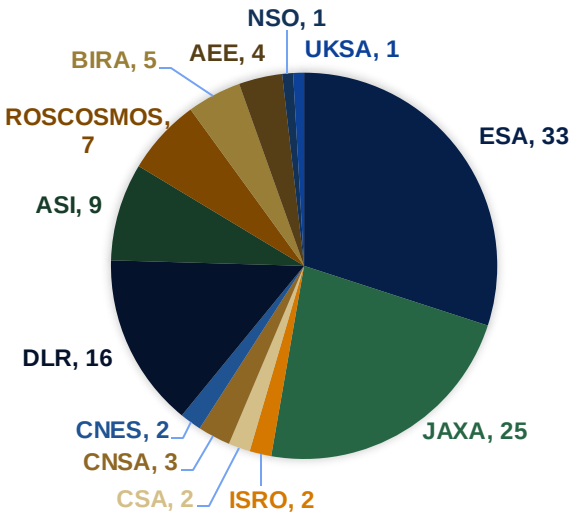
'OMICS



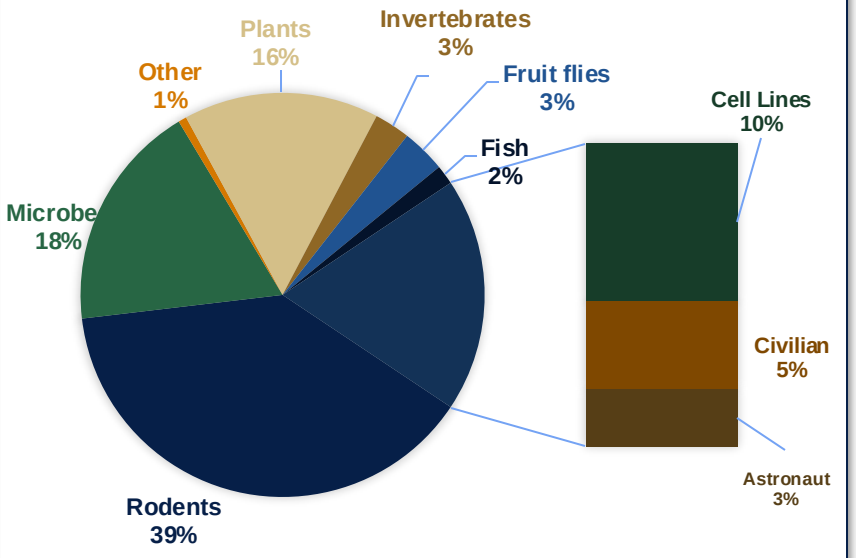
PHYSIO-PHENO



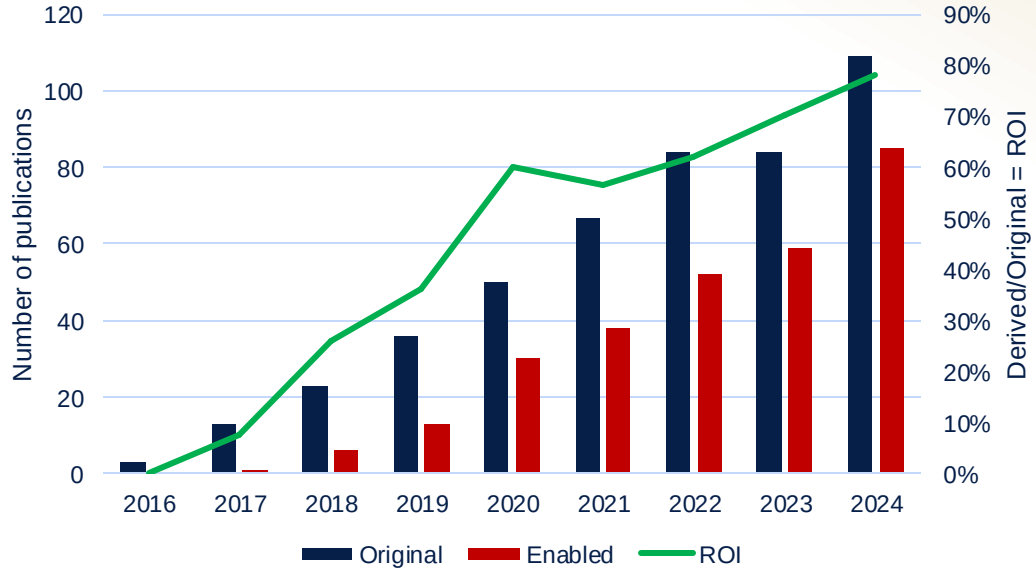
INTERNATIONAL STUDIES



ORGANISM



ROI grows faster than publications linked to original



85
Enabled Publications linked to OSDR

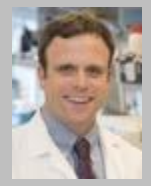
109
Original Publications linked to OSDR

150+
Datasets used in enabled publications

Civilian Human Spaceflight Data

INSPIRATION4

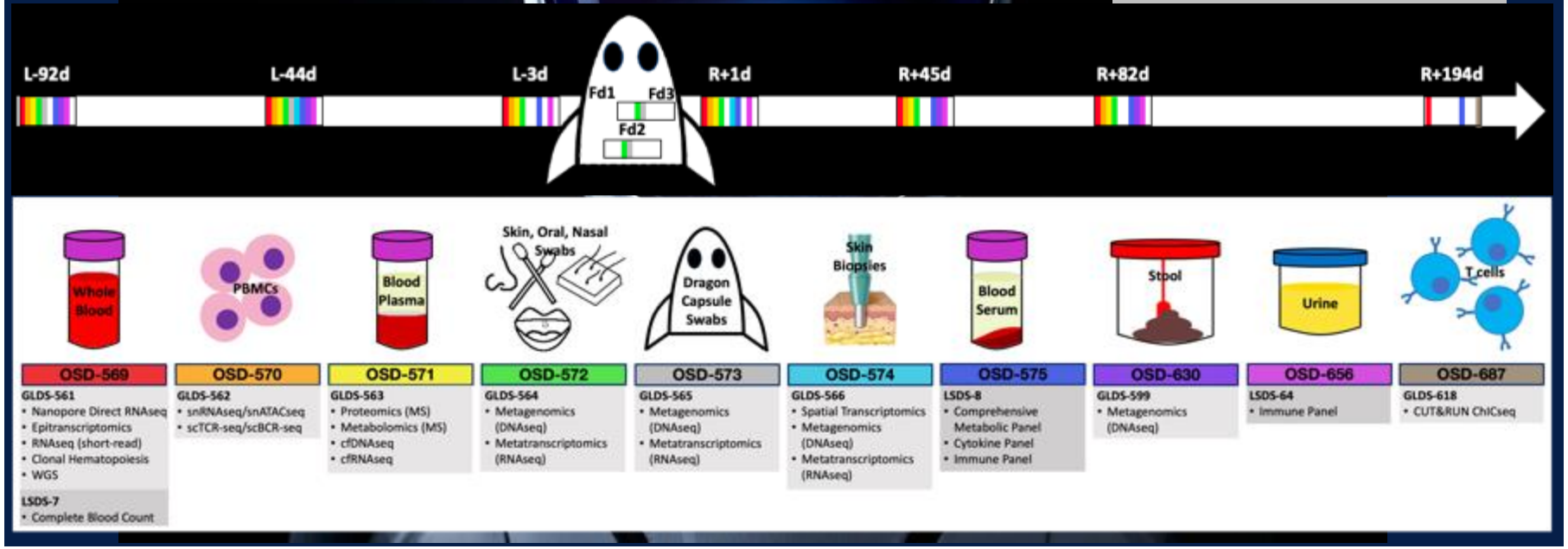
- Nature Portfolio Collection (2024): Space Omics and Medical Atlas (SOMA) across orbits
- I4 Image Credit: Chris Mason's Lab



Chris Mason

Eliah Overbey

Weil Cornell Medicine

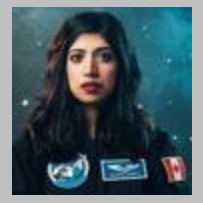


IIAS International Institute for Astronautical Sciences

- **Commercial Sub-Orbital**
- **Galactic-05 – Flew Nov. 2**
  - **Astroskin: ECG, systolic BP, heart rate (HR), breathing rate (BR), skin temperature, activity levels (provided by 3-axis accelerometer), and oxygen saturation in blood (SpO2)**
  - **Blood Glucose Monitoring Device**

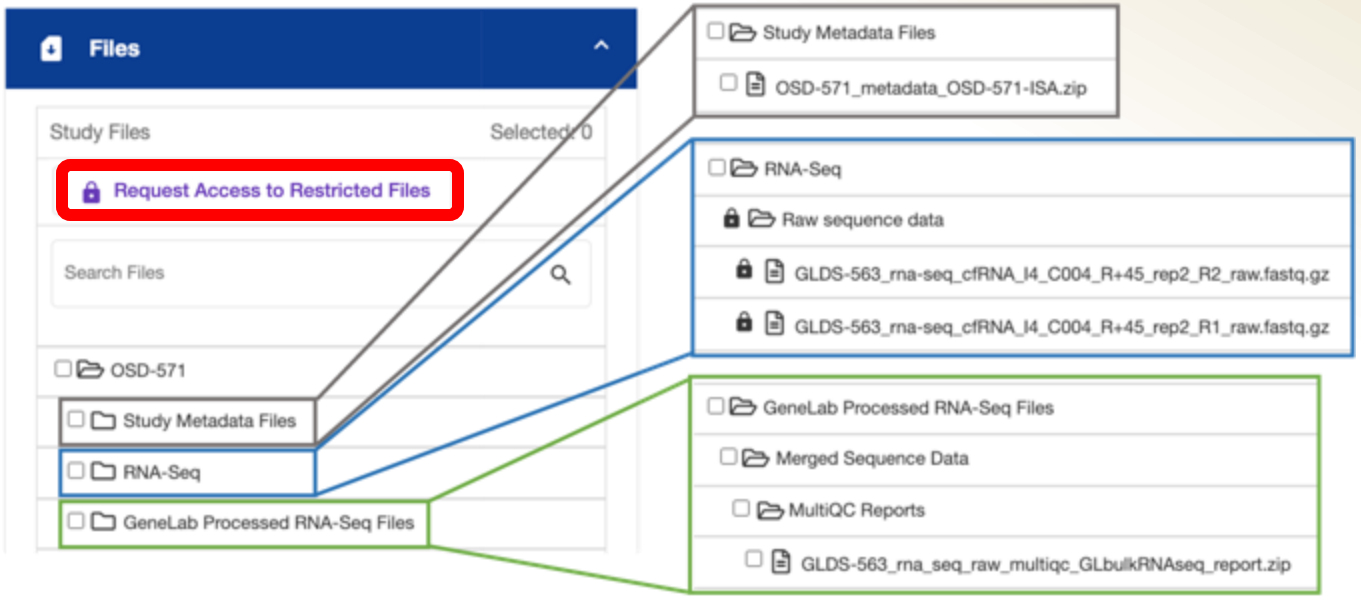
Soon!

- **Credit: Dr. Shawna Pandya**



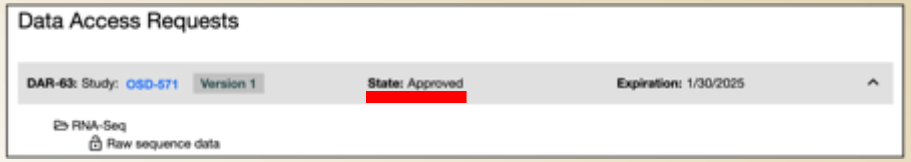
Shawna Pandya
IIAS Space Medicine

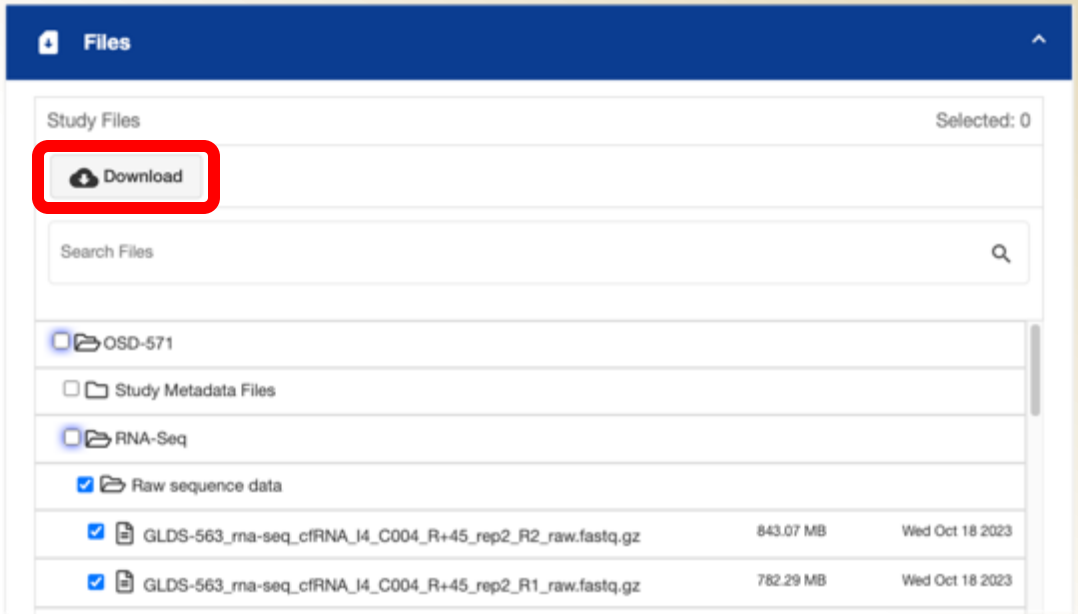
Access Request Process for Private Data



If Approved

Data Access Info Available On DAR Portal





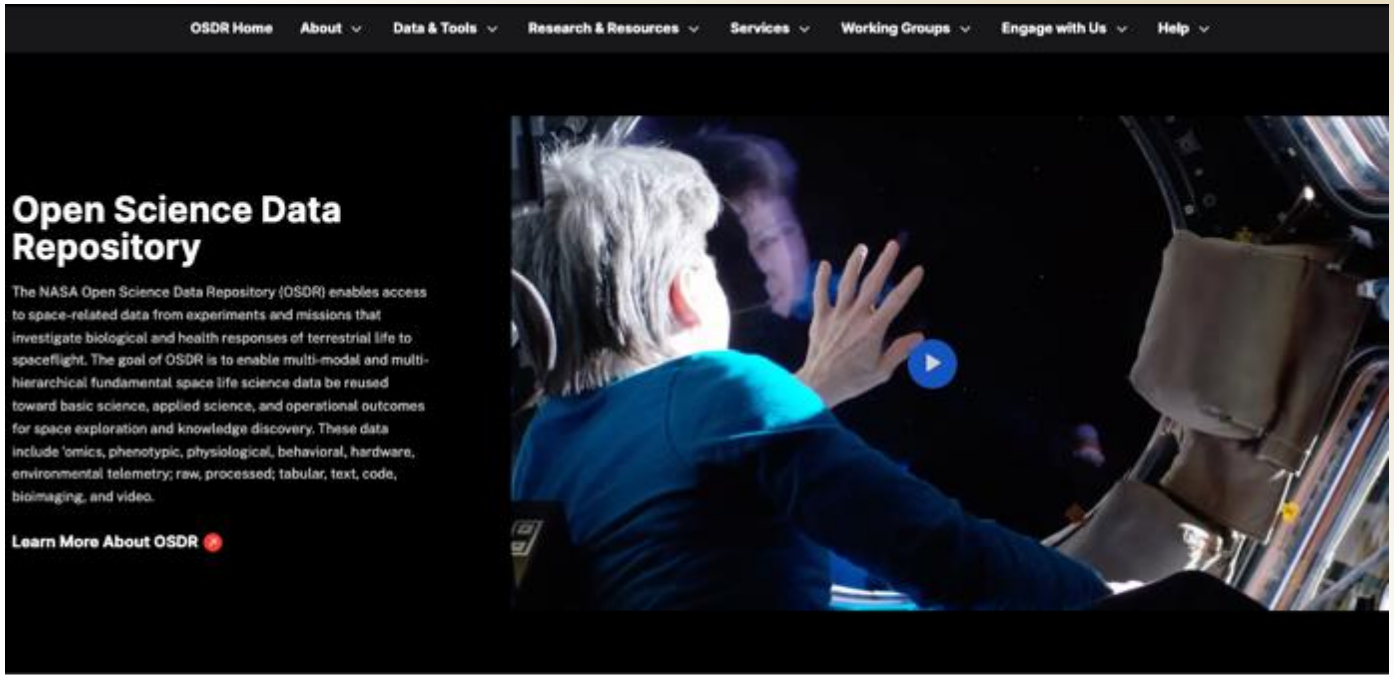
Access Request Process for Private Data

1	2	3	4	5	6	7
Pre-Submission	Investigator Submits	Signing Official Reviews	Data Access Committee Reviews	Access Data	Renew Project (Optional)	Close Out Project
Request Button clicked in OSD dataset	OSDR Data Request Package sent w/ Data Use Agreement	OSDR Follow-Up with your institution	Review Board with Specific Stakeholders (OSDR, PIs, etc)	Approved / Denied Secure Access	Approved / Denied	Access removed at end of IRB, upon participant request, or if conditions of Data Use Agreement are violated

Process inspired by NIH's Database of Genotypes and Phenotypes (dbGaP)

Navigating OSDR

Open Science Data Repository (nasa.gov/osdr)



Explore and Contribute

Open Science Data Repository

Search and upload spaceflight datasets

Explore the Data Repository

Submit Data to OSDR

Have space-relevant data to submit?

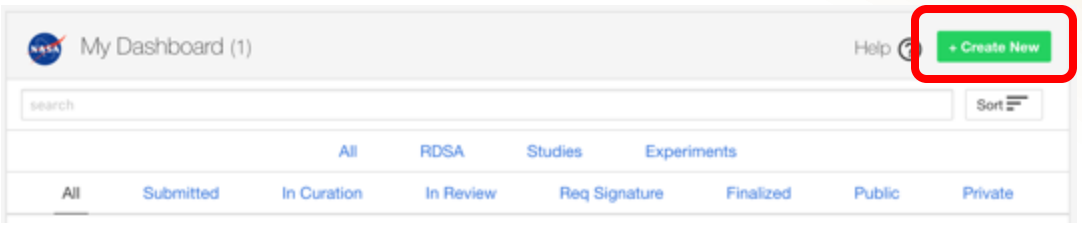
Contribute via the Submission Portal

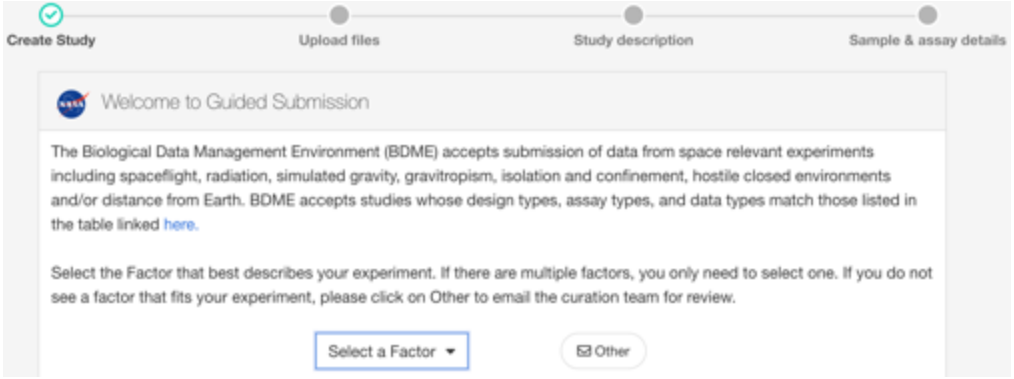
Join OSDR AWG's

We look forward to collaborating with you!

Interested in participating in an AWG?

Submission Portal – BDME (https://osdr.nasa.gov/bio/submission-sso-login.html)





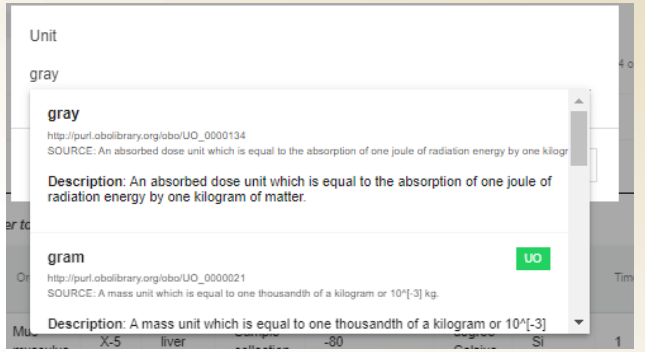
Sample level data

Sample Name	Organism	Strain	Material Type	Protocol REF	Parameter Value - Storage Temperature Setting	Unit	Ionizing Radiation	Time	Unit	Dose	Unit
Sample1	Mus musculus	X-5	liver	Sample collection	-80	degree Celsius	Silicon	1	hour	1	gray
Sample2	Mus musculus	X-5	liver	Sample collection	-80	degree Celsius	Sham-irradiated	0	hour	0	gray

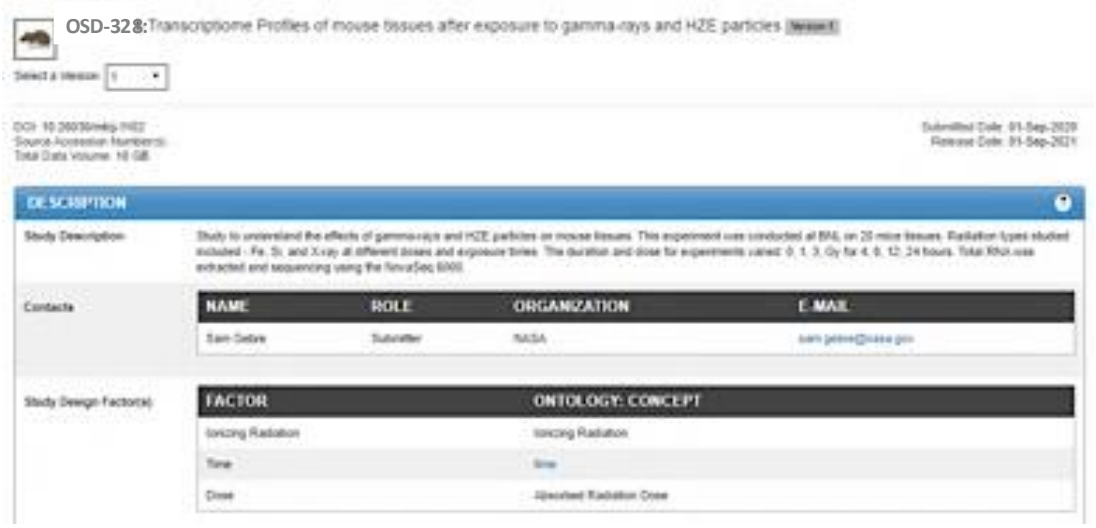
Assay level data with assay-specific fields

Sample Name	Protocol REF	Parameter Value - QA Instrument	Parameter Value - QA Assay	Parameter Value - QA Score	Unit	Extract Name	Protocol REF
Sample1	Nucleic Acid Extraction	BioAnalyzer	NanoChip	8	RIN	Extract 1	Library Construction
Sample2	Nucleic Acid Extraction	BioAnalyzer	NanoChip	9	RIN	Extract 2	Library Construction
Sample3	Nucleic Acid Extraction	BioAnalyzer	NanoChip	9.1	RIN	Extract 3	Library Construction

Terms are connected to a controlled ontology

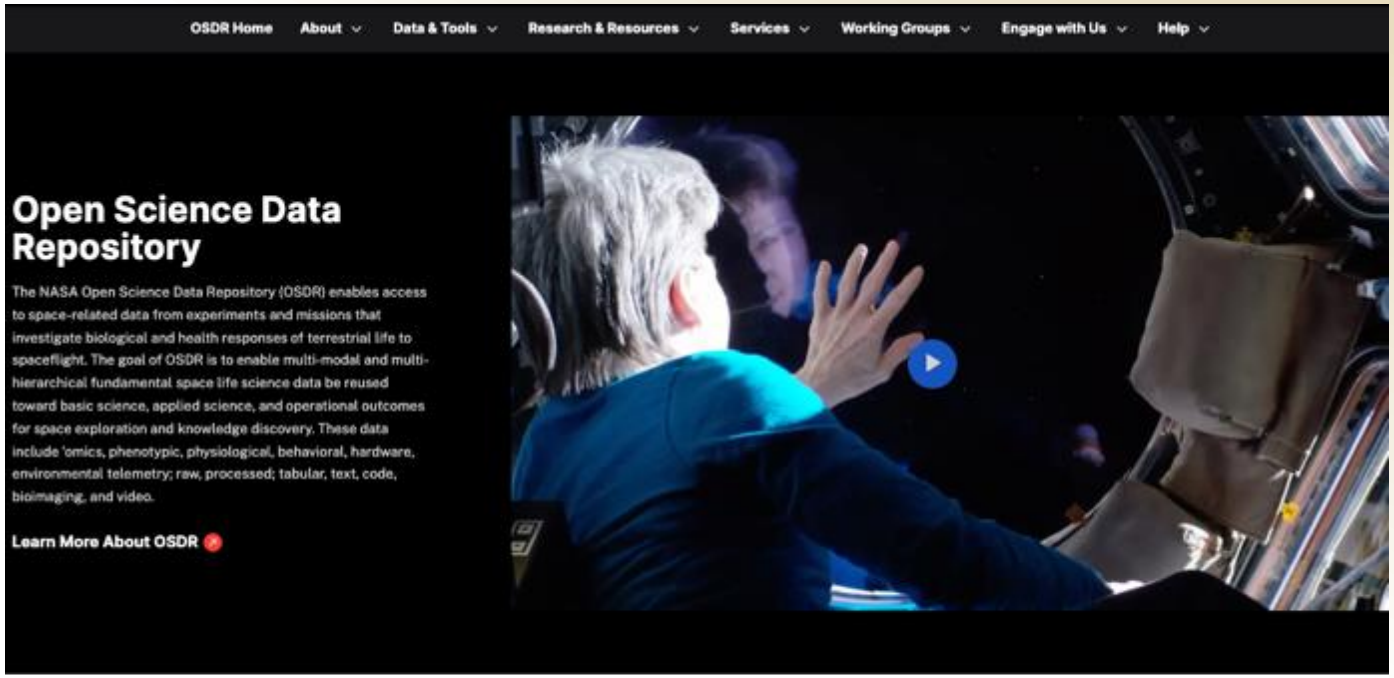


- PIs can now submit data through BDME.
- Allows efficient input of sample and assay level metadata alongside data files.
- Tailored to specialized needs of spaceflight samples.
- Used for both omics (GeneLab) and non-omics (ALSDA) data submissions.



PIs are presented with a study preview and a private link for sharing with reviewers and collaborators.

Open Science Data Repository (nasa.gov/osdr)



Explore and Contribute

Open Science Data Repository

Search and upload spaceflight datasets

Explore the Data Repository

Submit Data to OSDR

Have space-relevant data to submit?

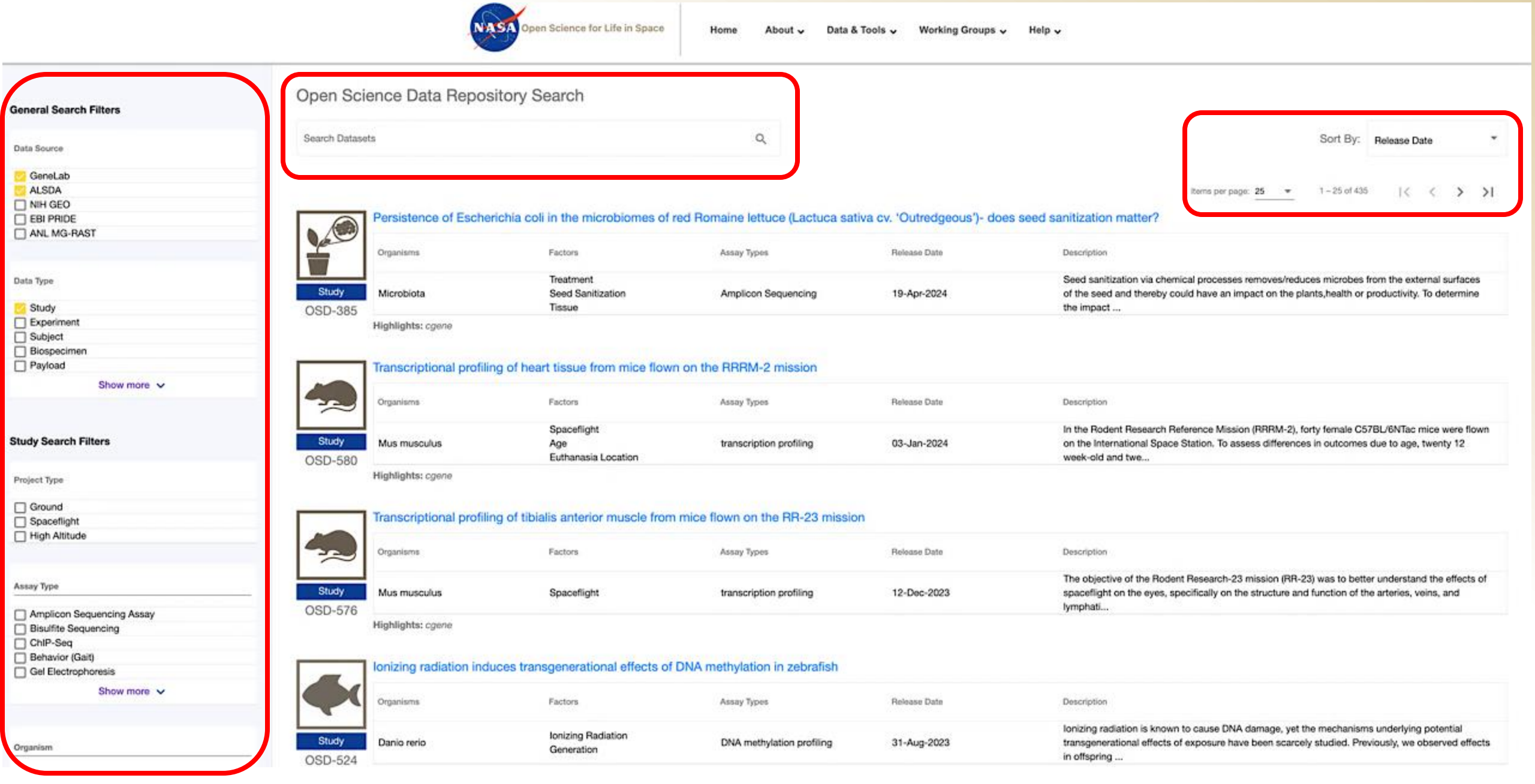
Contribute via the Submission Portal

Join OSDR AWG's

We look forward to collaborating with you!

Interested in participating in an AWG?

Open Science Data Repository (https://osdr.nasa.gov/bio/repo/)



OSDR: Study Page

- Description
- Experiments
- Payloads
- Missions
- Protocols
- Samples
- Assays
- Publications
- Files
- Version History
- Visualization

2.50 TB

OSD-137 Version 8

Rodent Research-3-CASIS: Mouse liver transcriptomic, proteomic, epigenomic and histology data

Study

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

GeneLab ID: GLDS-137
ALSDA ID: LSDS-28
DOI: 10.26030/9k6w-4c28
License: CC0-1.0

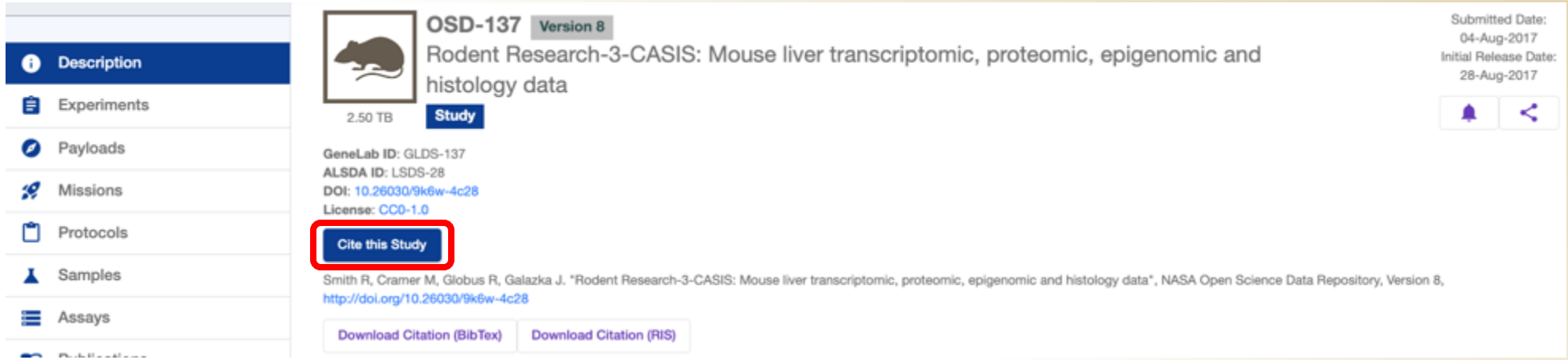
Cite this Study

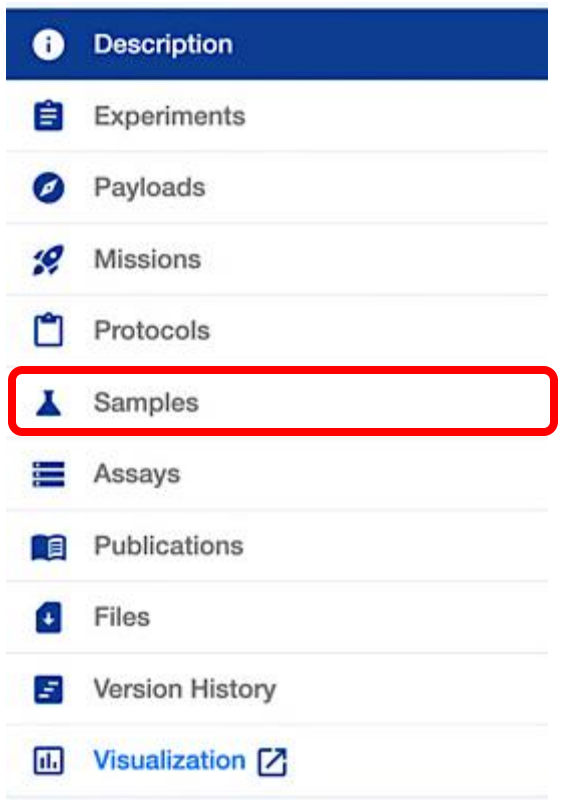
Description

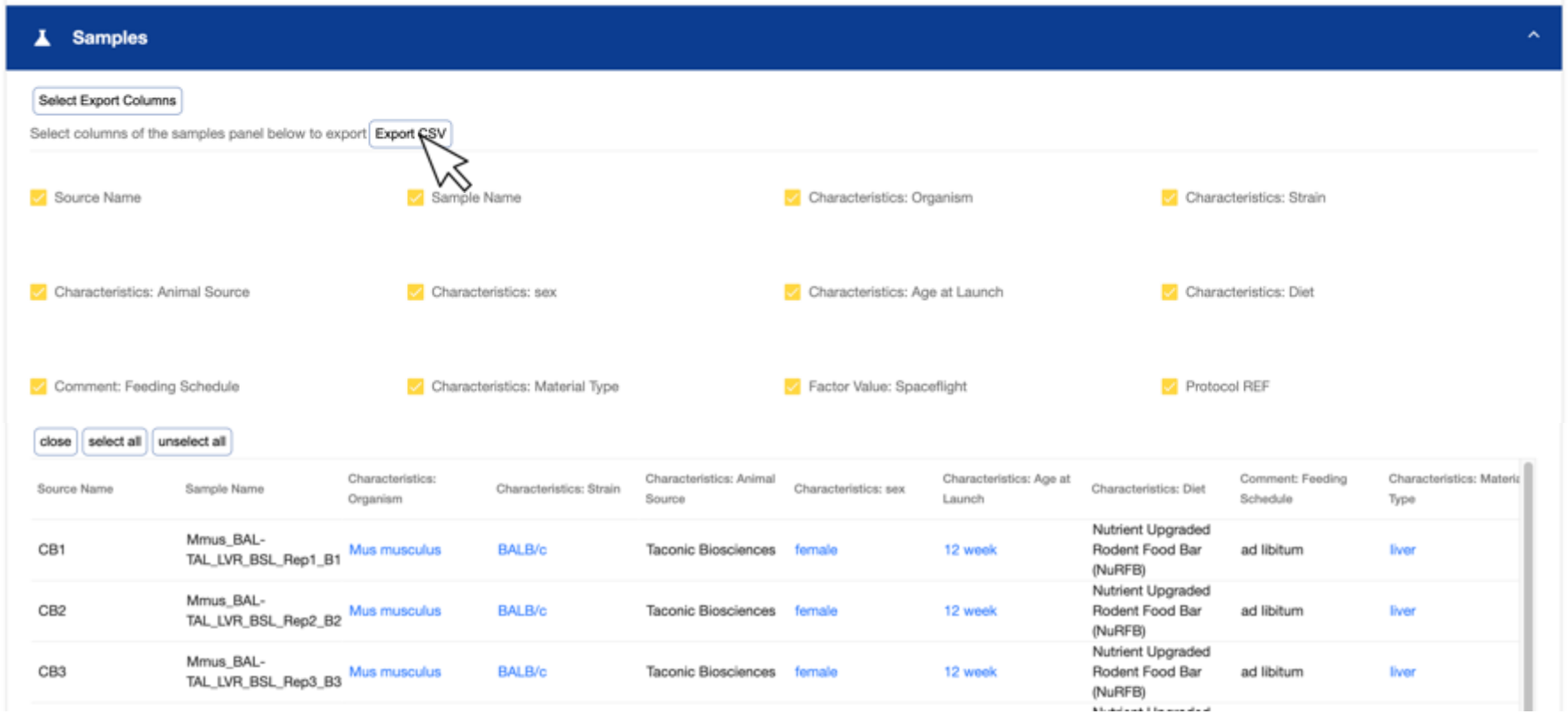
Description

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, 12-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 in identical hardware and matching ISS environmental conditions. Basal controls were housed in standard vivarium cages. Spaceflight, ground controls and basal groups had blood collected, then were euthanized, had one hind limb removed, and finally whole carcasses were stored at -80 C until dissection. All mice in this data set received only the control/sham injection.

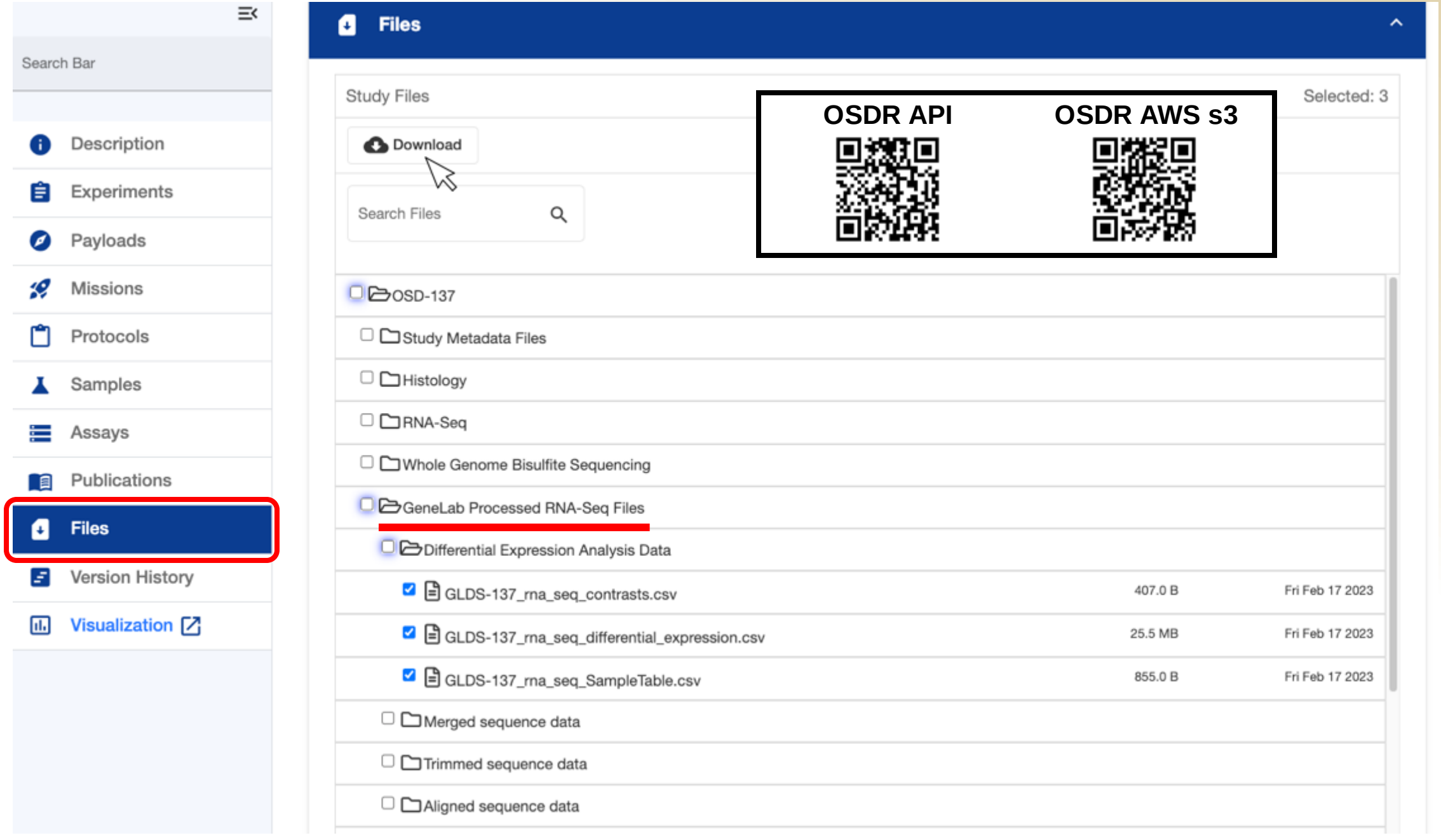
OSDR: Cite Study & Export Specific Metadata Table Columns







OSDR: Download Study Files



GeneLab Data Processing

Build consensus data processing pipelines with the scientific community

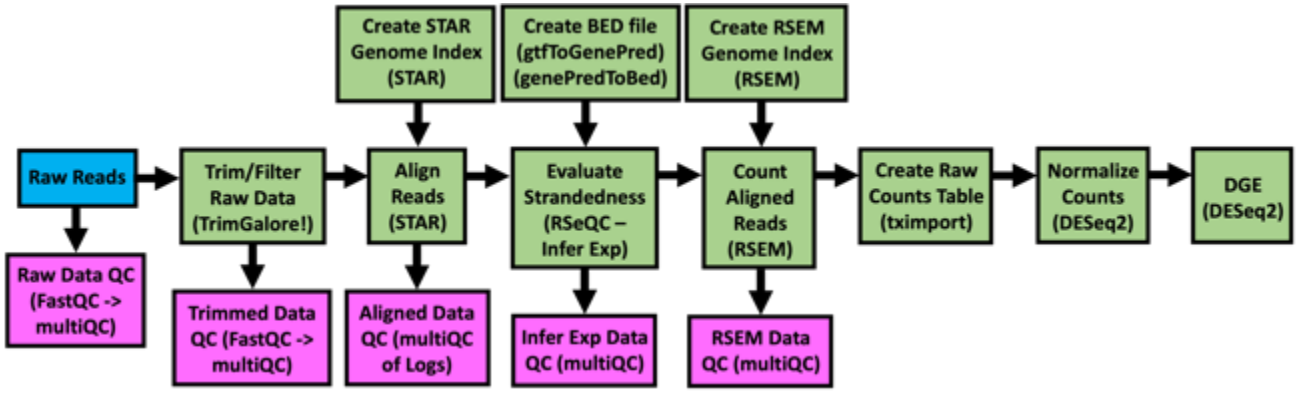
Raw Sequence Data

```
@A00654:12:HFY5YDSXX:1:1101:2121:1000_CTCGCGTTT 1:N:0:NGGAGAGT+TCCTACCT
AATATCAGTGATATTTAGAAACCACATAGTAAGCTAACAACTAATAATGGAATGGTTTTAATATCCTGTGACAAGTTAATGTGGATACTATGCGGTCTTCCTTAAAA
+
FFFFFFFFFFFFFFFFFFFF:FFFFFFFFFFFFFFF:FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF:FF
@A00654:12:HFY5YDSXX:1:1101:5466:1000_CCTCCCTCG 1:N:0:NGGAGAGT+TCCTACCT
GGCGGGTCTGAGCAAGAACACTCTGCCATCGCCGGCTCGGGACGGGAGGCCCACCCGTGGCCGAGCACGCAGACCTGGATGAACCGAGGCCTCGATGGGCCTGGACG
+
FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF
@A00654:12:HFY5YDSXX:1:1101:14109:1000_TGATCTACG 1:N:0:NGGAGAGT+TCCTACCT
TGTCACTACAACCAGCTGTGCCTGTGCTATTGCAGTTACACAGTGTCACTACAACCAACTGTGCCTGTGCTATTGCAGTTACACAGTGTCACTACAACCAGCAGATC
+
FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF:FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF,:FFFFFFF:FF:F:FF
```

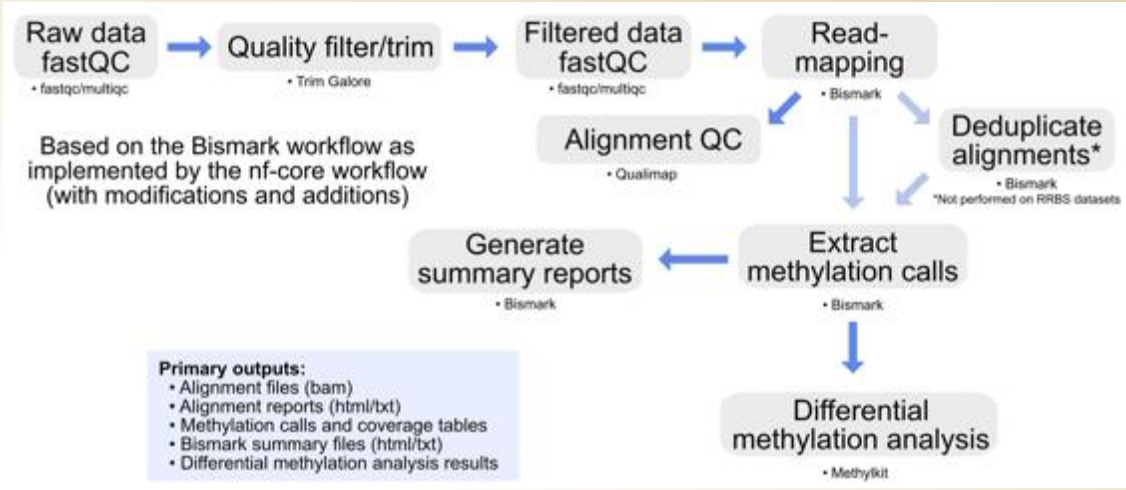
Processed RNA Sequence Data: Differential Gene Expression

SYMBOL	GENENAME	Group.Mean_(FLT)	Group.Mean_(GC)	Log2fc_(FLT)v(GC)	P.value_(FLT)v(GC)	Adj.p.value_(FLT)v(GC)
Timm17a	translocase of inner mitochondrial membrane 17a	3.91	15.11	-2.67	0.00	0.00
NA	NA	166.47	108.32	0.63	0.00	0.00
NA	NA	325.78	171.64	0.93	0.00	0.00
Dnajc7	DnaJ heat shock protein family (Hsp40) member C7	22.31	42.16	-0.95	0.00	0.00
Slc15a4	solute carrier family 15, member 4	2.75	9.38	-2.32	0.00	0.00
Ckap5	cytoskeleton associated protein 5	18.24	46.64	-1.39	0.00	0.00
NA	NA	104.49	67.33	0.64	0.00	0.00

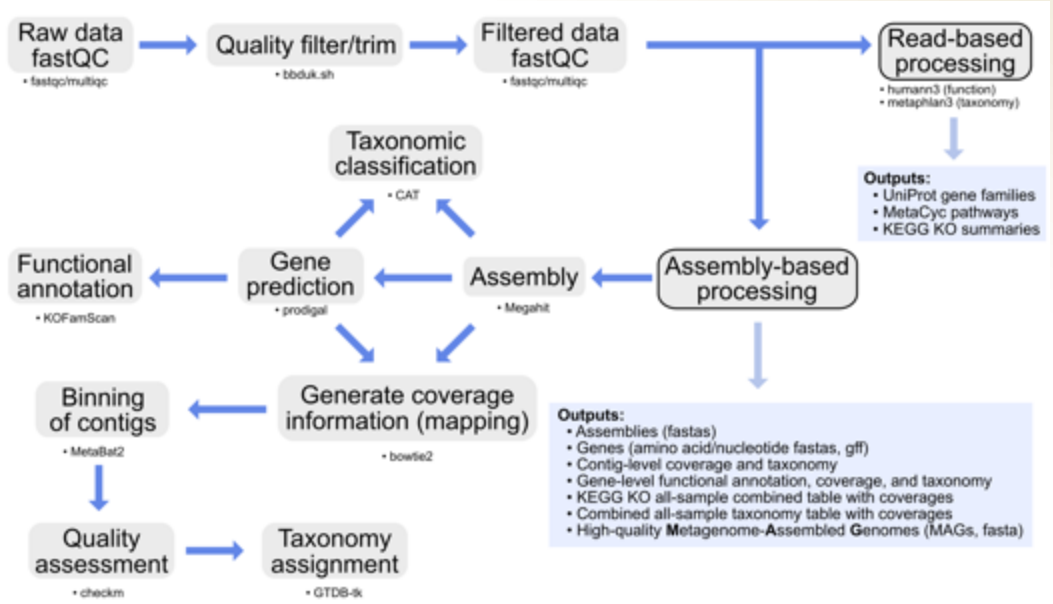
RNA Sequencing Data



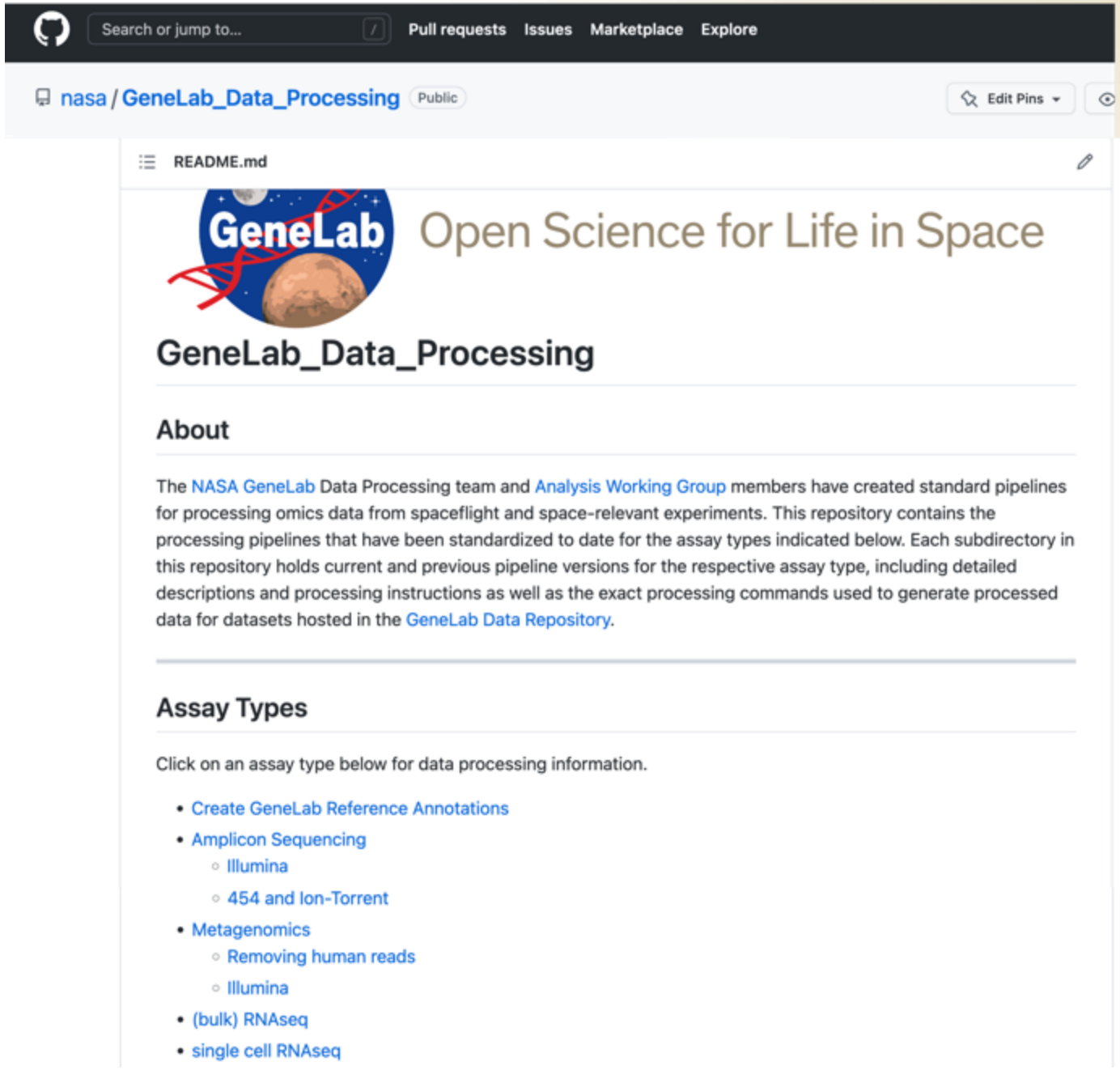
Methylation Sequencing Data



Metagenomics Data

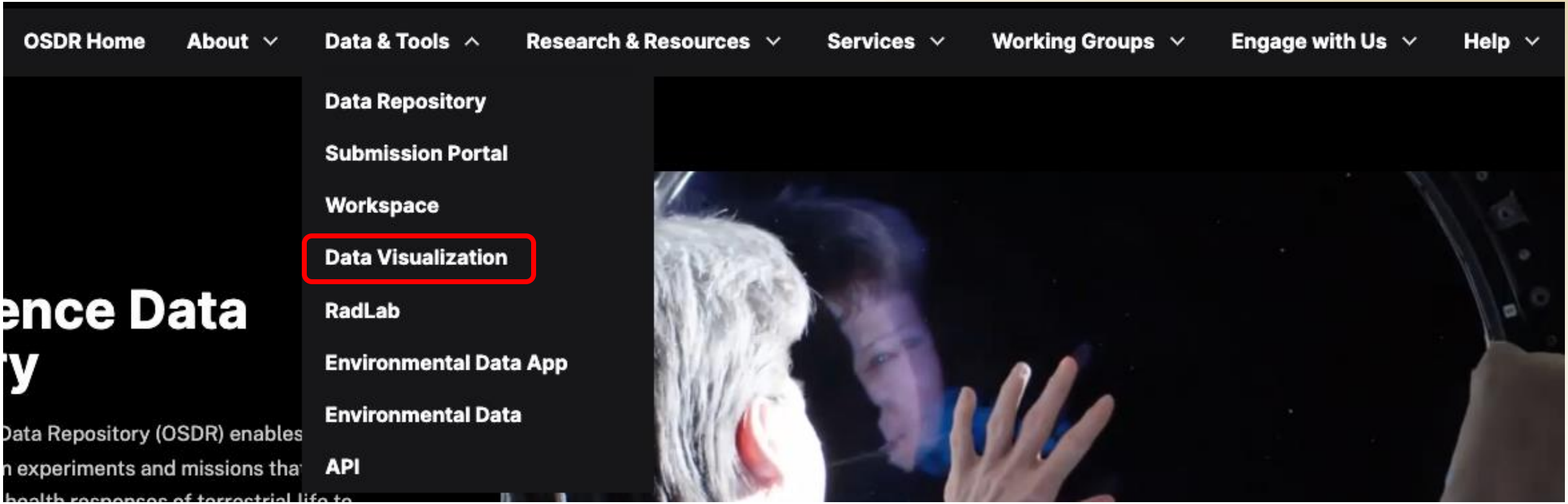


GeneLab Data Processing GitHub Repo

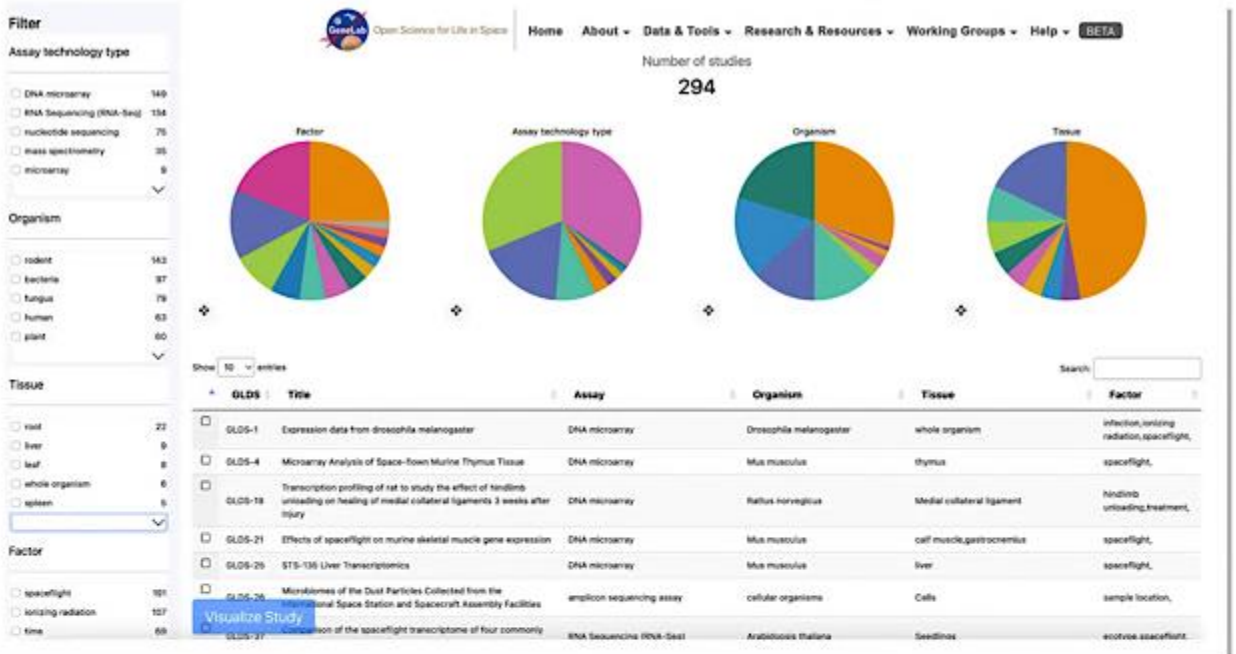


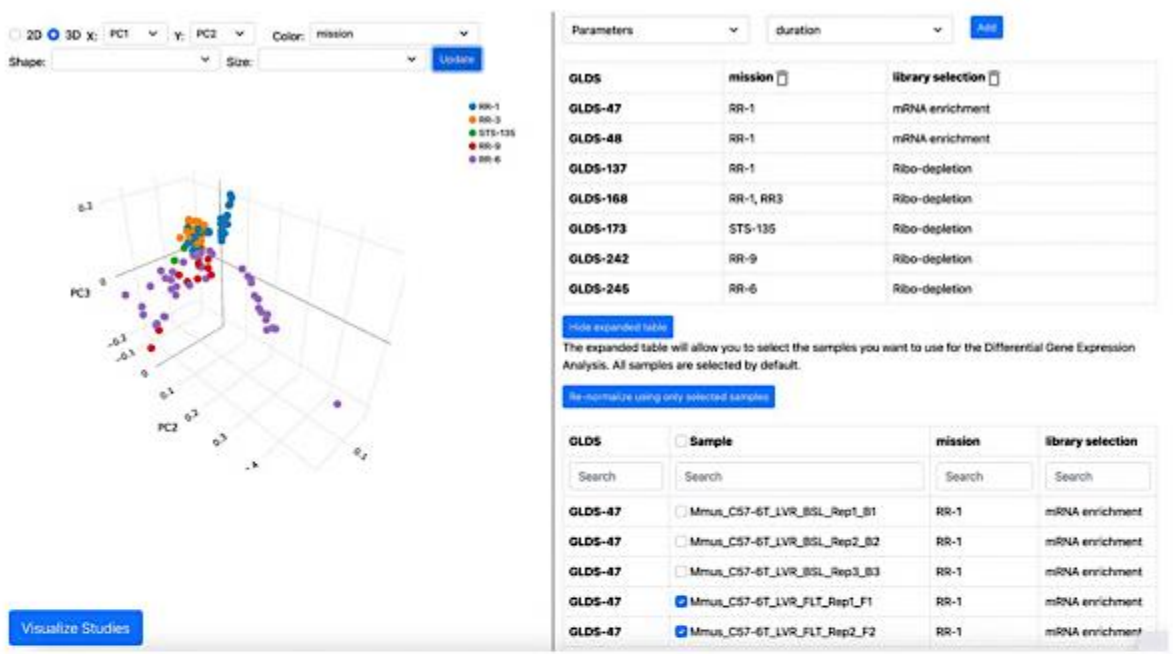
https://github.com/nasa/GeneLab_Data_Processing

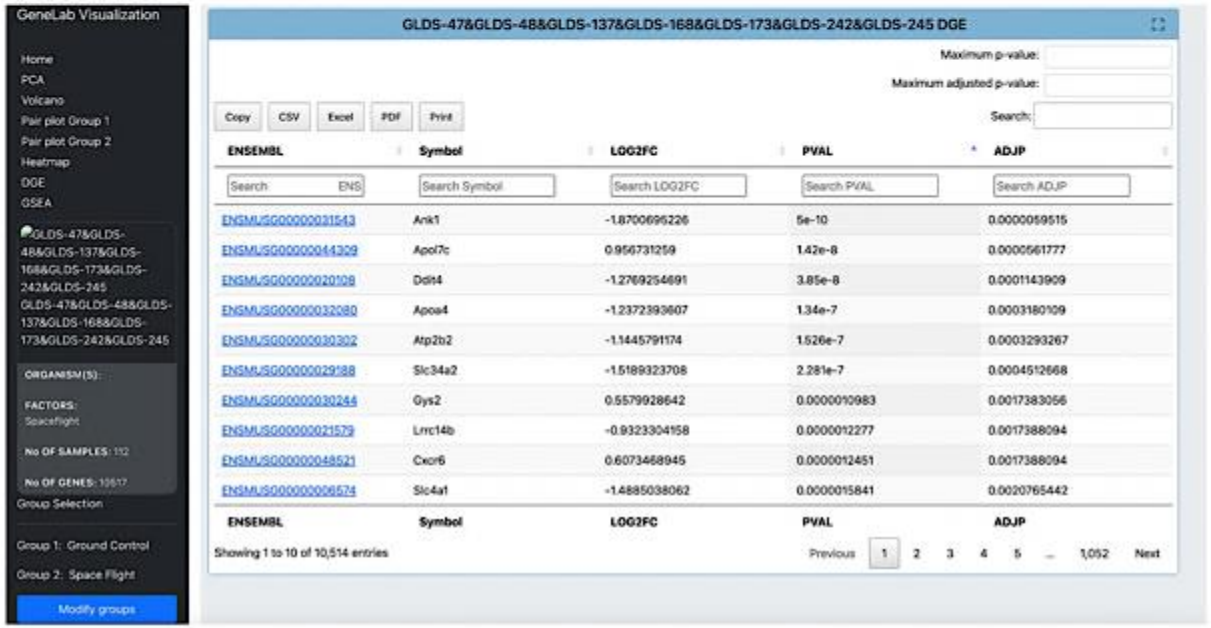
OSDR: Data Visualization



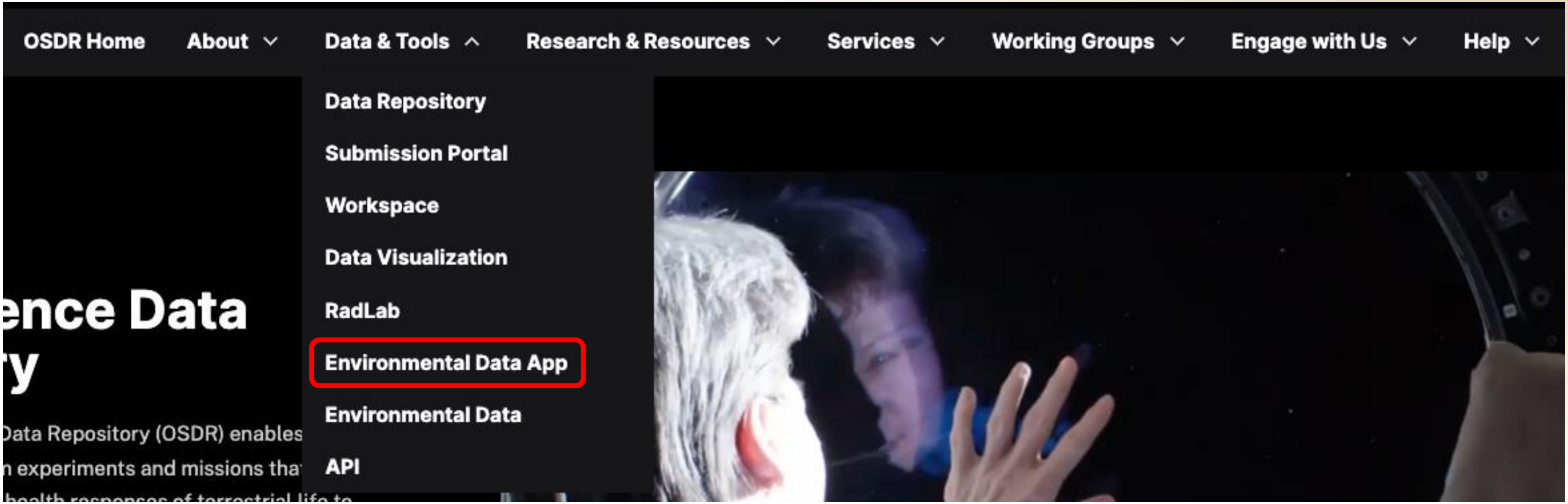
Data Visualization Portal (https://visualization.genelab.nasa.gov/data/)



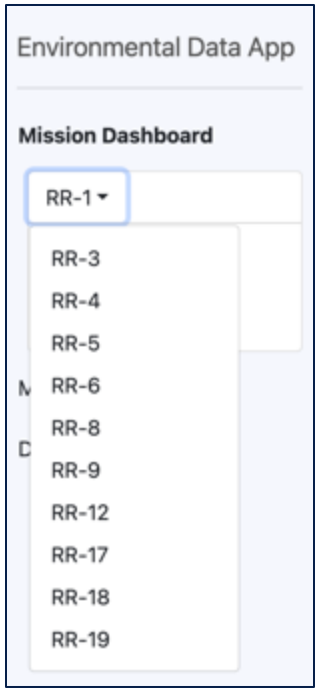


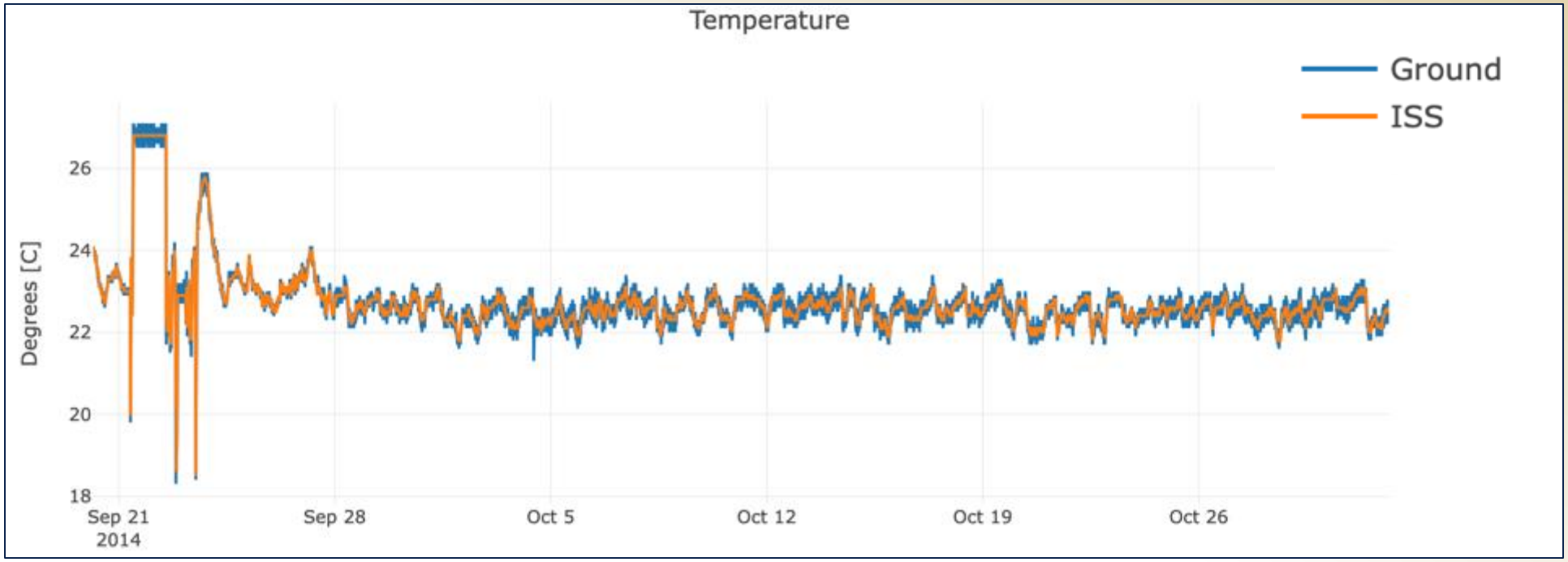


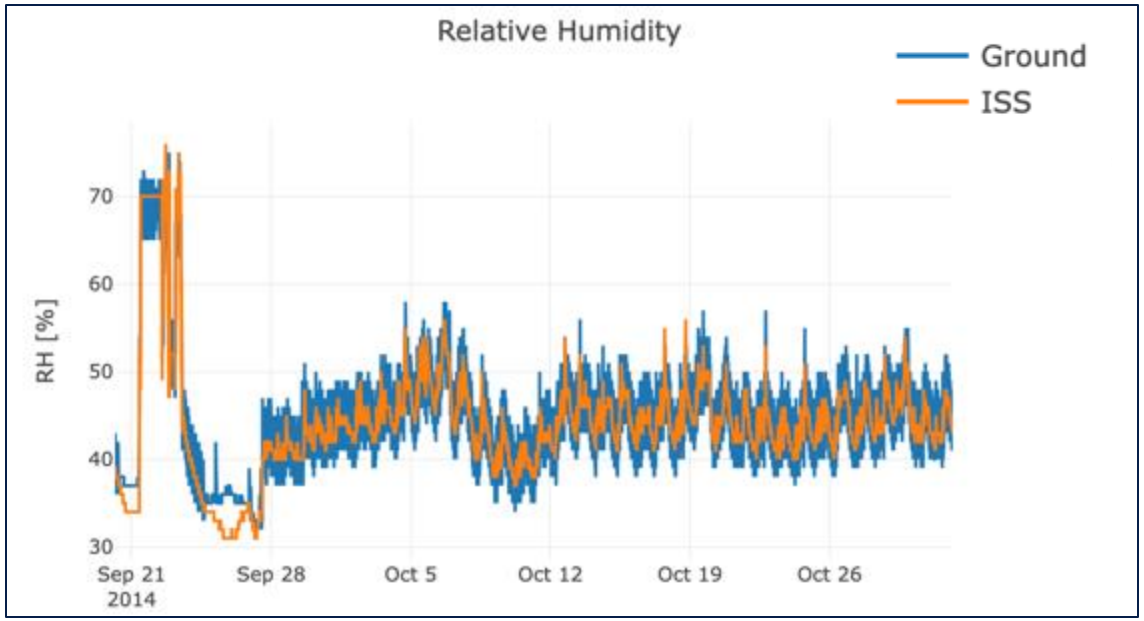
OSDR: Environmental Data Application (EDA)

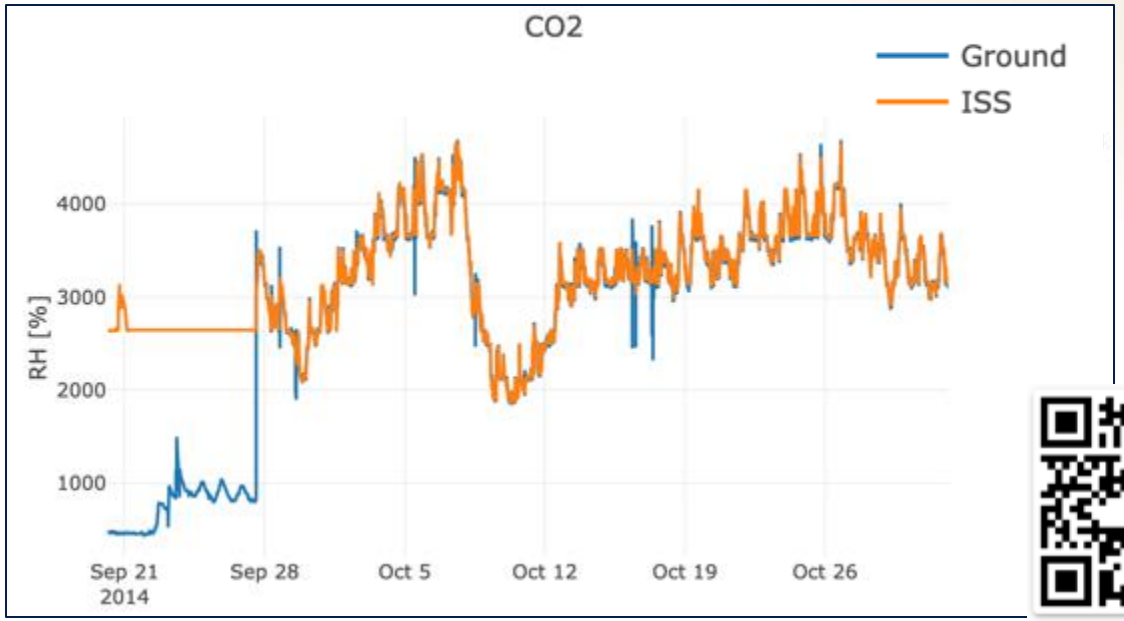


Environmental Data Application (EDA)









Environmental Data Application (EDA)

View Biologically Relevant Mission Milestones

Mission Milestone Dates

Display Mission milestones

Launch: 2014-09-21 09:49:00

Reached orbit: 2014-09-23 17:19:00

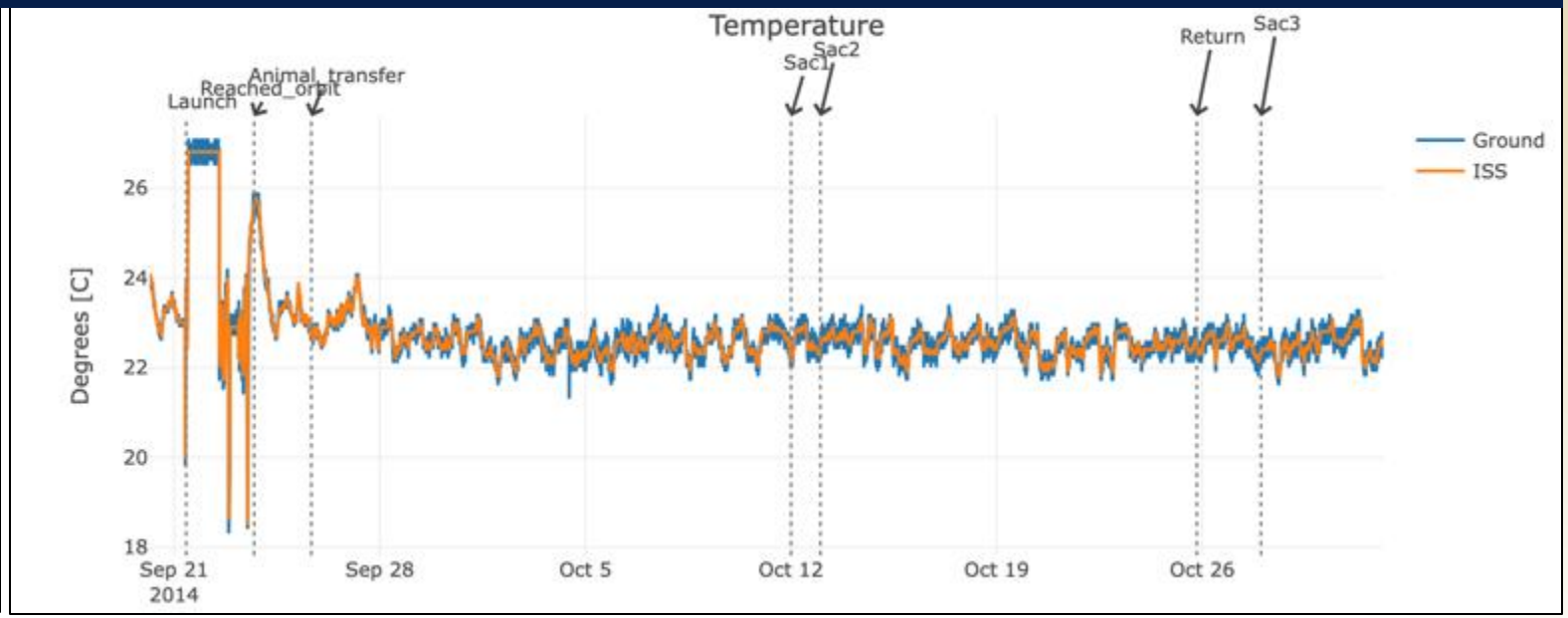
Animal transfer: 2014-09-25 16:05:00

Sac1: 2014-10-12

Sac2: 2014-10-13

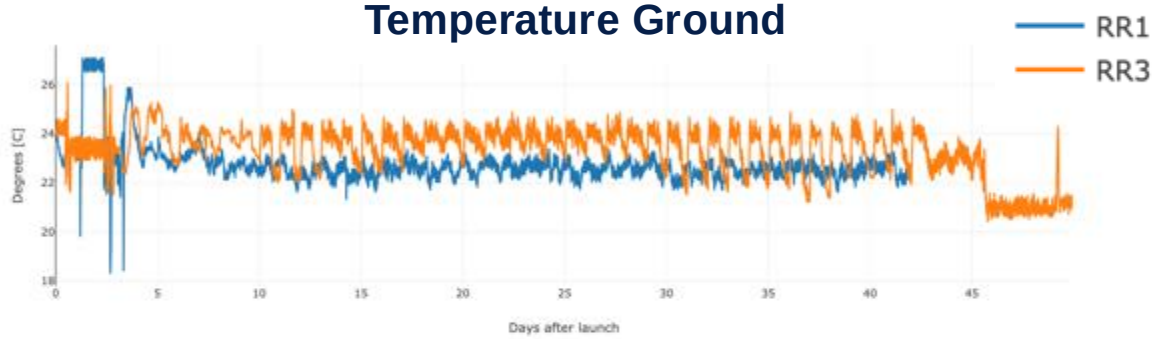
Return: 2014-10-25 19:36:00

Sac3: 2014-10-28 00:01:00



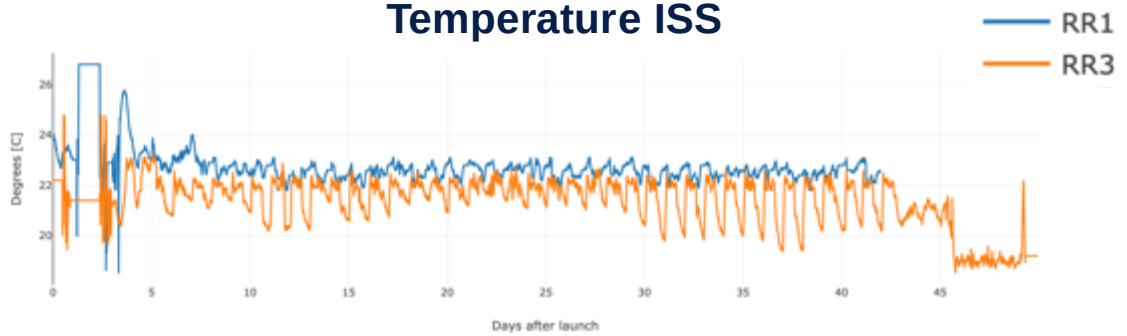
Compare Data From Multiple Missions

Temperature Ground



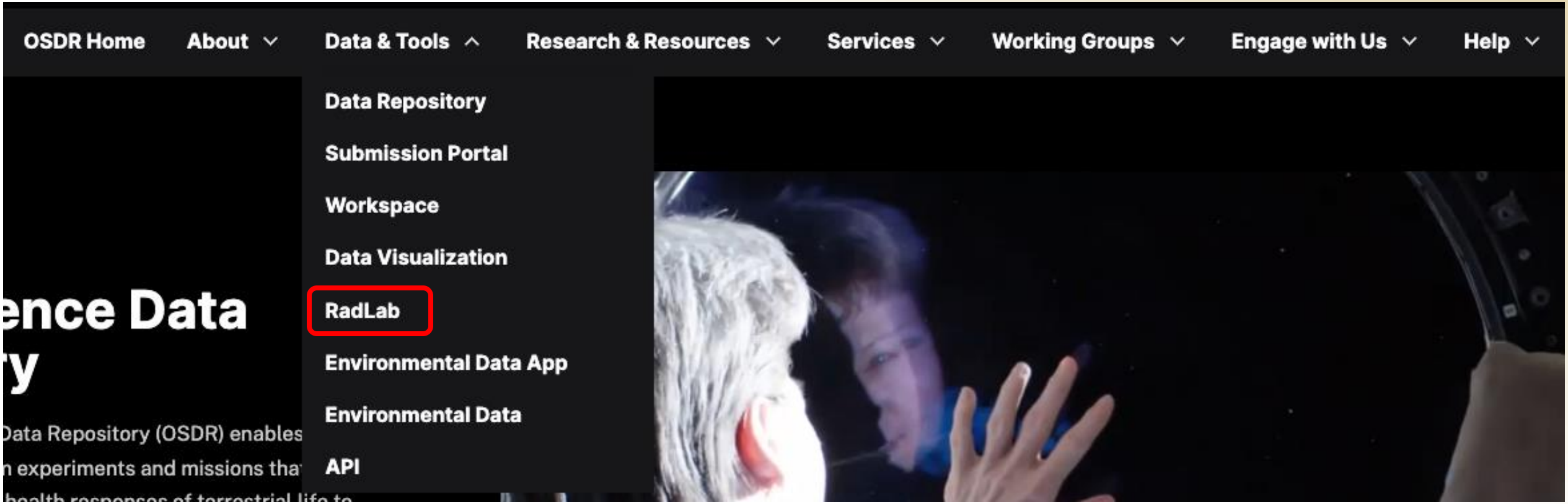
Mission	Min temperature Ground Degrees [C]	Max temperature Ground Degrees [C]	Mean temperature Ground Degrees [C]	Standard Deviation temperature Ground Degrees [C]	Median temperature Ground Degrees [C]
RR1	18.30	27.10	22.76	0.82	22.60
RR3	20.40	26.10	23.36	0.97	23.70
	Min	Max	Mean	Standard Deviation	Median

Temperature ISS



Mission	Min temperature ISS Degrees [C]	Max temperature ISS Degrees [C]	Mean temperature ISS Degrees [C]	Standard Deviation temperature ISS Degrees [C]	Median temperature ISS Degrees [C]
RR1	18.50	26.80	22.76	0.81	22.60
RR3	18.50	24.80	21.37	0.97	21.70
	Min	Max	Mean	Standard Deviation	Median

RadLab Platform

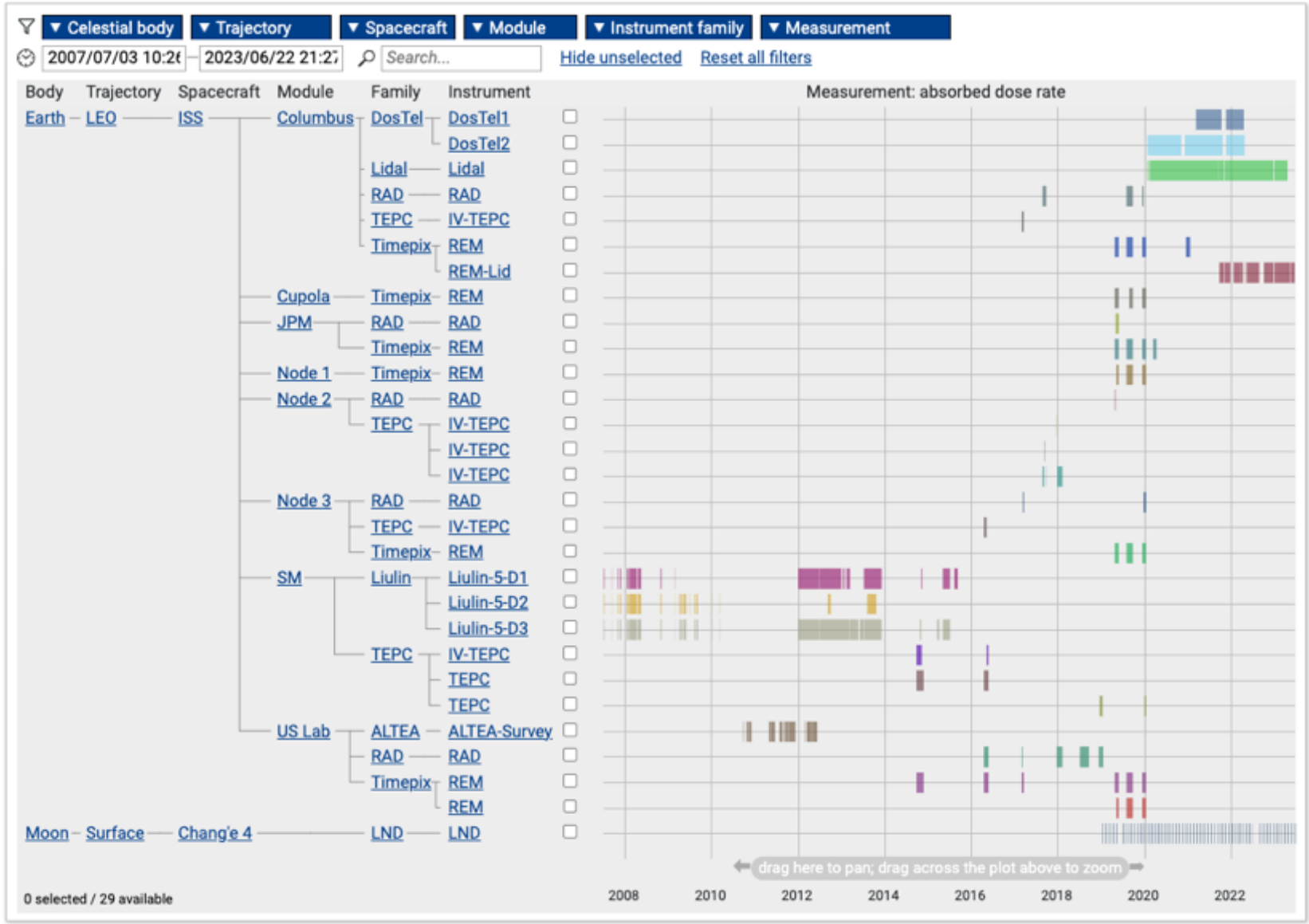


RadLab Portal and Data API

RadLab
- Overview
  - LEO
    - ISS
  - BLEO
- Time series plots
- Data comparison
- Geospatial plots
- Knowledgebase
- Data API
- Settings

Overview of available instrument readings

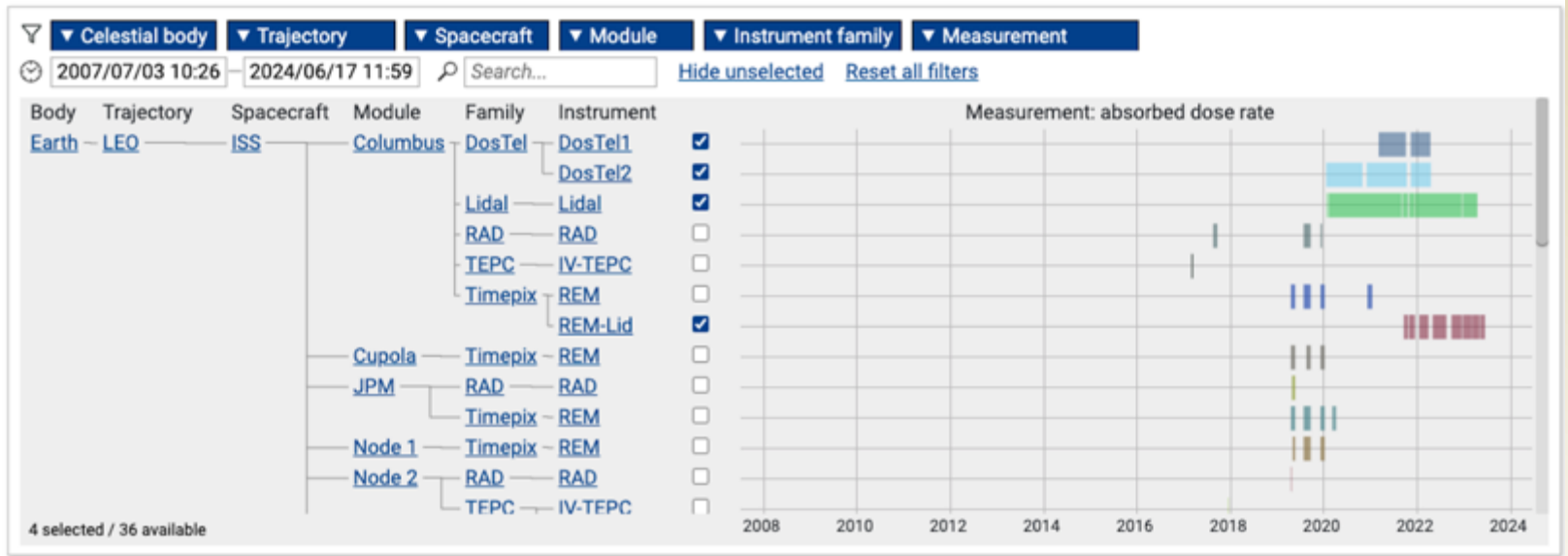
Explore the data available from the ISS or from instruments in BLEO, or directly explore the overview table below to filter for instruments (using the dropdowns above the table) and time spans of interest (by panning and zooming the interactive plot).
Tick the checkboxes next to the instruments of interest to proceed with more in-depth visualizations.



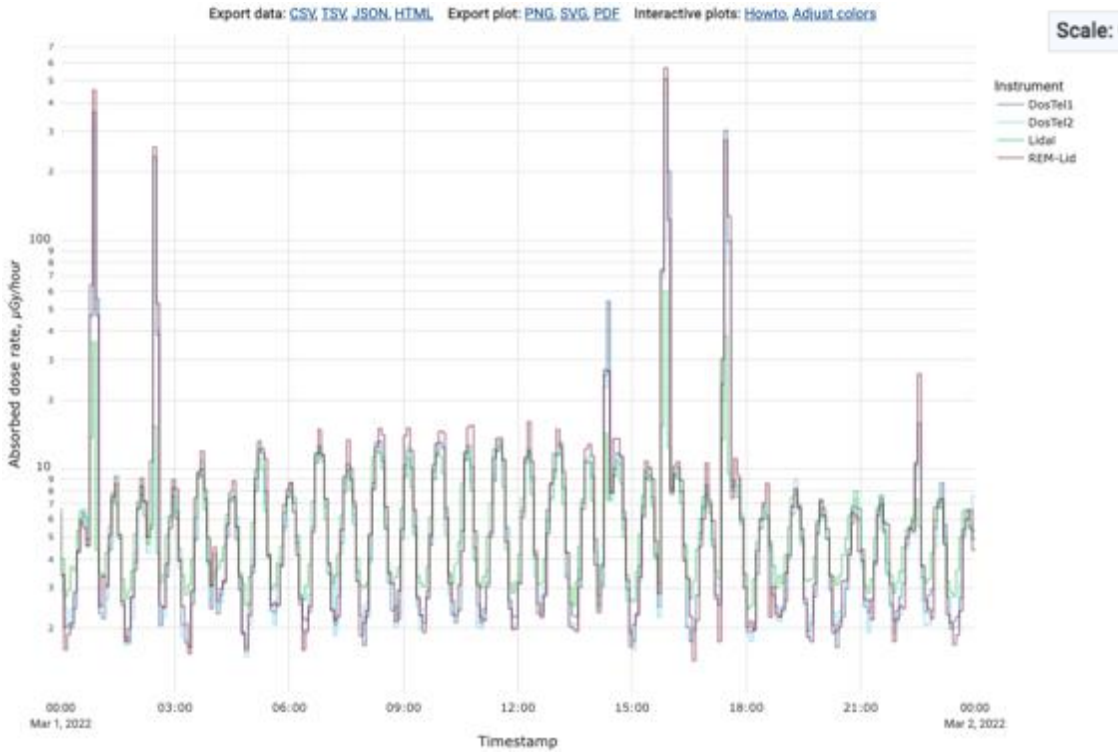
RadLab Portal and Data API

Time Series Plots

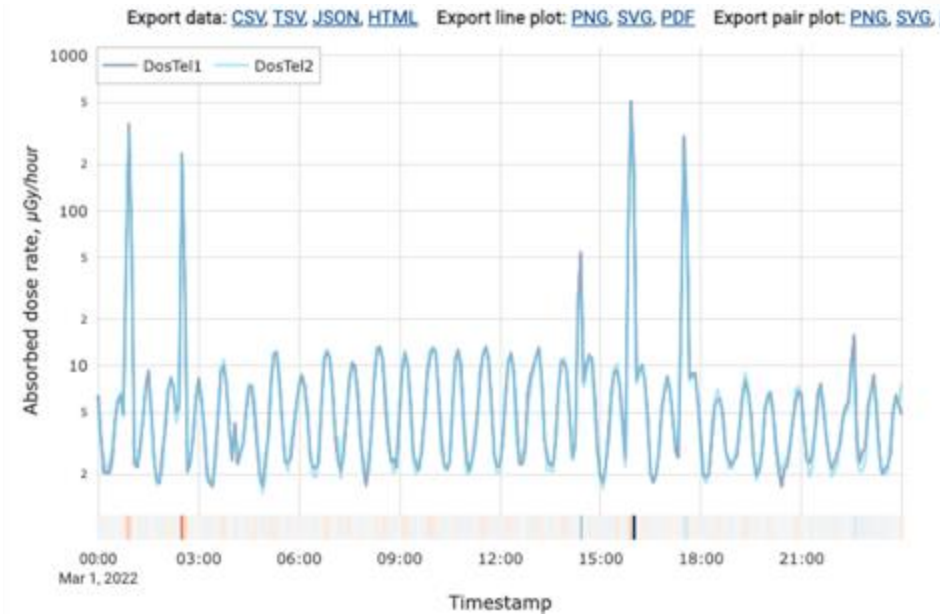
RadLab
- Overview
- LEO
- ISS
- BLEO
- **Time series plots**
- Data comparison
- Geospatial plots
- Knowledgebase
- Data API
- Settings

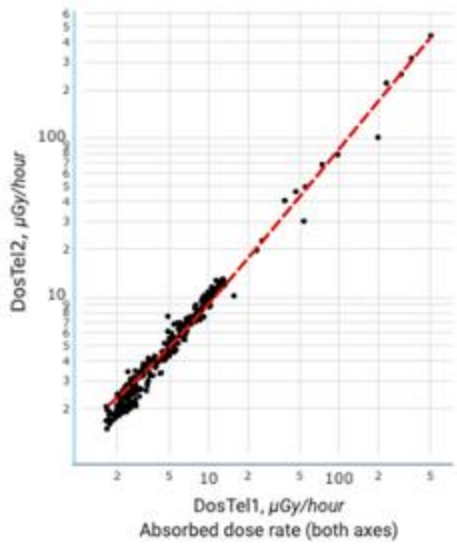


RadLab
- Overview
- LEO
- ISS
- BLEO
- Time series plots
- **Data comparison**
- Geospatial plots
- Knowledgebase
- Data API
- Settings



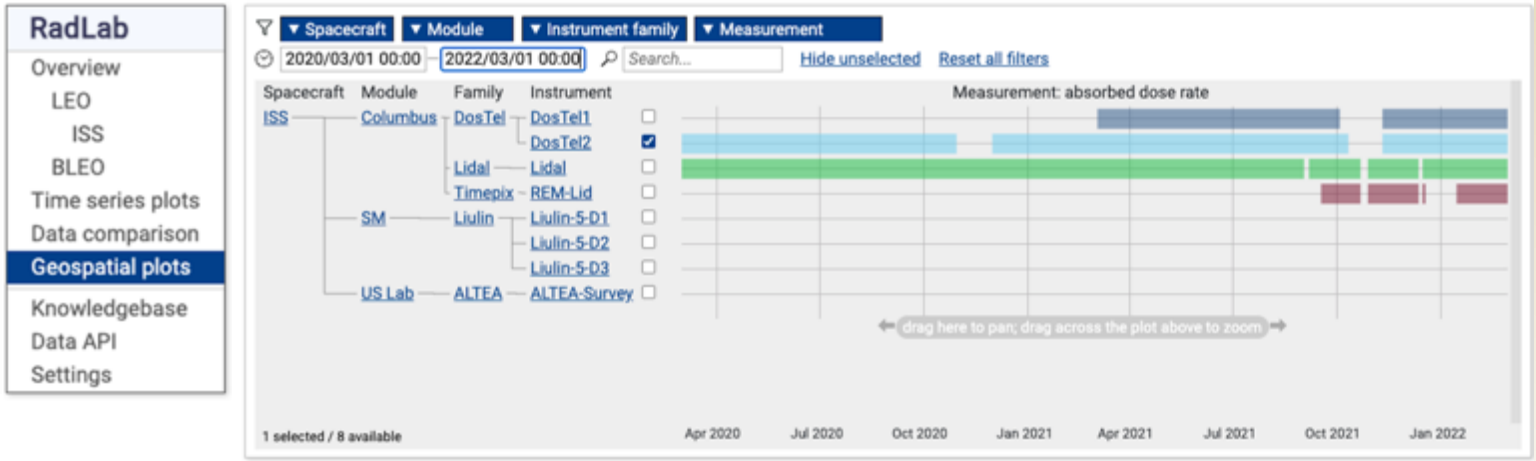
Compare Two Detectors

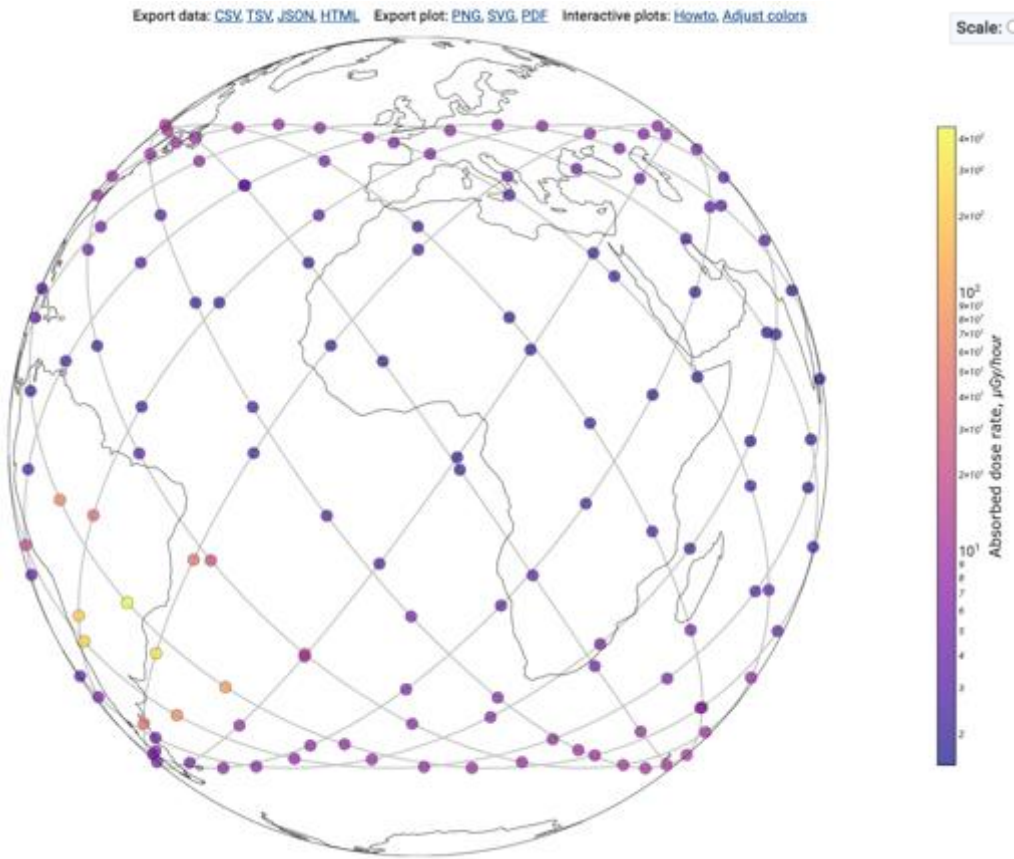


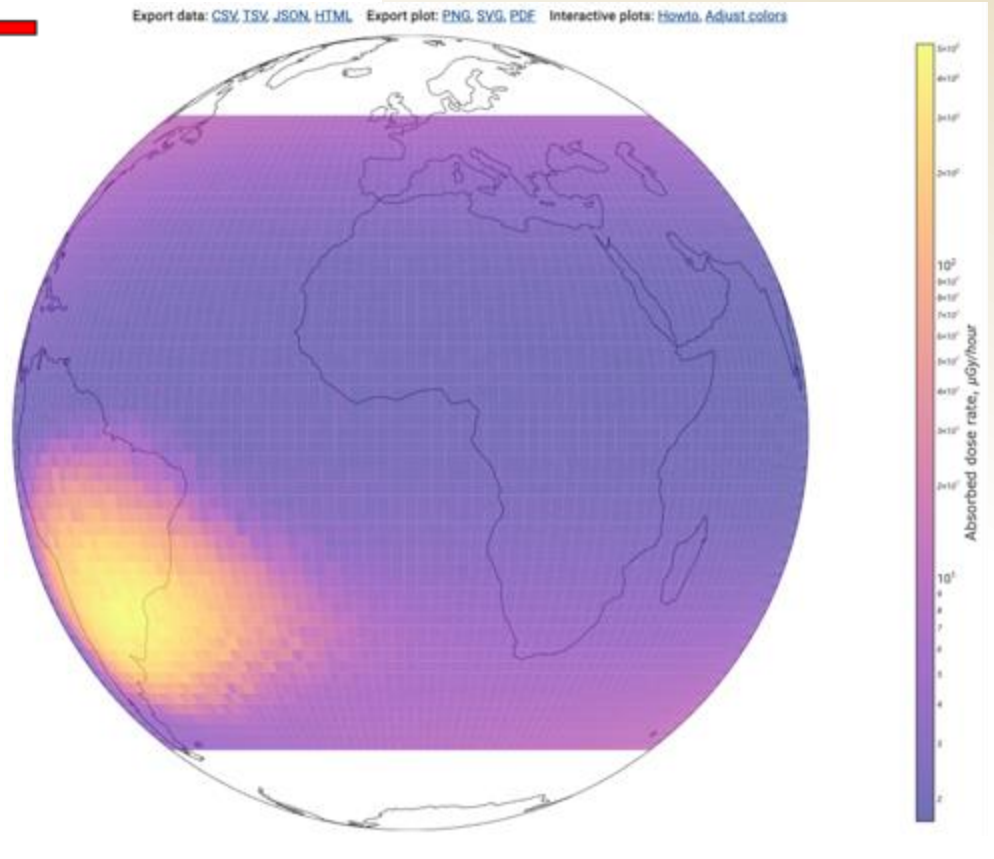


RadLab Portal and Data API

Geospatial Plots

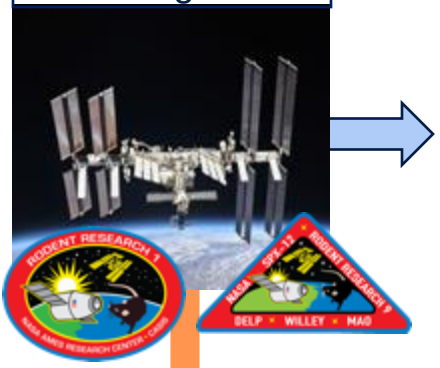






Maximizing Scientific Return Using Open Science Resources

Principal Investigation

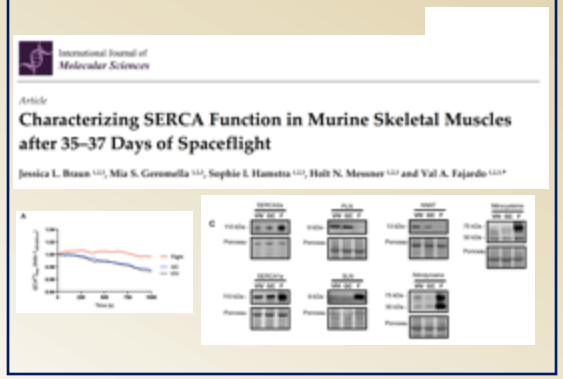


Dissections

Tissue Storage

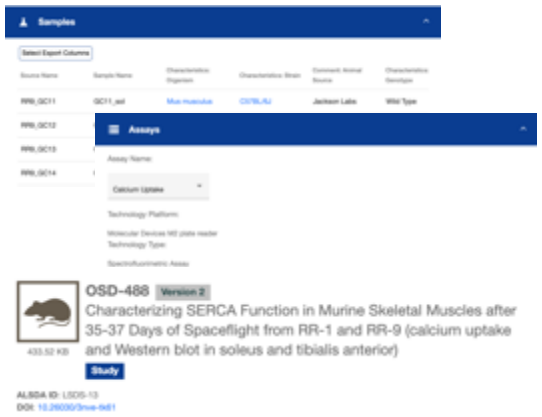
Tissue Request

Analysis and Publication



PI Results

Submitted & Curated Datasets



Data Reused

Soleus (SOL) Muscle Data:

- Proteomics (quantitative mass spectrometry)
- Transcriptomics (bulk RNA-seq)
- Epigenomics (DNA methylation)
- Calcium Reuptake

Tibialis Anterior (TA) Muscle Data:

- Transcriptomics (bulk RNA-seq)
- Epigenomics (DNA methylation)
- Calcium Reuptake

- Omics & phenotypic data: RR-9, RR-1
- Qlattice xAI (explainable AI) in AWG collaboration
- Li et al., 2023 npj Microgravity

Training Programs Using OSDR Data

GeneLab for High Schools (GL4HS)

Overview:

- **GeneLab for High Schools (GL4HS)** is a four-week space biology virtual training program for rising high school juniors and seniors in the US.
- For the **GL4HS Teacher Program**, educators learn alongside students for weeks 1-2 then spend weeks 3-4 developing personalized curriculum material using teacher template documents.
- The program has three components:
  - **LEARN** general cell biology, bioinformatics and omics, and applications for space biology both on Earth and in spaceflight conditions.
  - **NETWORK** with professionals including NASA researchers, and other guests from academia, industry, and private and non-profit sectors.
  - **RESEARCH** and present an analysis of a transcriptomic dataset from the Open Science Data Repository.

Objectives:

- To inspire the next generation of scientists to become involved in space-based research
- To introduce and teach high school students about 'omics data and how it relates to biological research, with a specific focus on space-based research
- To introduce and teach high school students basic bioinformatics skills enabling them to excel in both their college curriculum and academic research laboratories
- To provide students with networking opportunities with their peers, NASA research scientists, NASA Space Biology management and University professors
- To present students with an opportunity to conduct hypothesis driven research of a publishable standard in collaboration with a NASA mentor at a NASA research facility
- To provide tools for teachers to include GeneLab and bioinformatics analysis in their classrooms

GeneLab for Colleges and Universities (GL4U)

Background:

- NASA's GeneLab project empowers researchers with open access to space-relevant multi-omics data through the Open Science Data Repository (OSDR)
- GeneLab for Colleges and Universities (GL4U) offers **bioinformatics training for space biology** to:
  - ➢ Increase accessibility and interpretability of multi-omics Space Biology data
  - ➢ Train a diverse next generation of Space Biology researchers
  - ➢ Enhance awareness and understanding of Space Biology data

Objectives:

- **Educate and train** the next generation of scientists to process, analyze, and interpret space-relevant 'omics data using publicly available data and bioinformatics tools
- **Maximize** the number of scientists who understand and utilize NASA's open-source 'omics data and tools
- **Provide educators with** the **knowledge and resources** required to train and inspire their students using GeneLab bioinformatic analyses as an entree into Space Biology

GL4U Content Design

GL4U materials are organized into introduction and omics-specific modules

Introduction Module

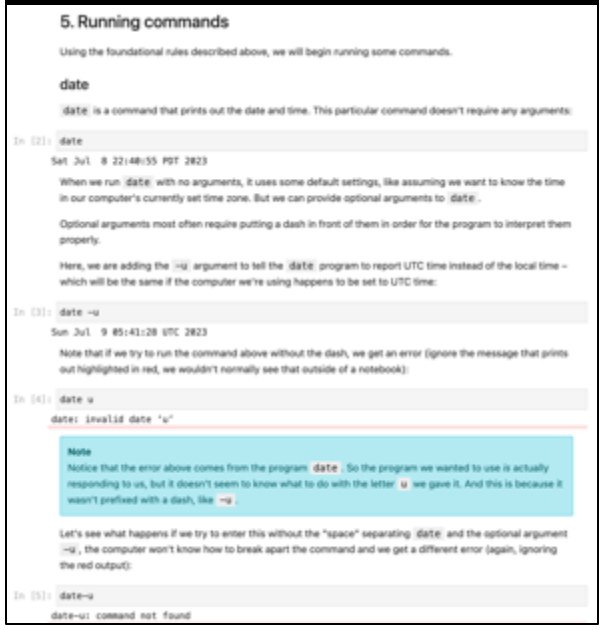
- Overview Lecture (NASA, SMD, BPS, OSDR, GeneLab)
- Jupyter Lab Tutorial
- Unix Jupyter Notebook (hands-on training)
- R Jupyter Notebook (hands-on training)

The Intro Module serves as a pre-requisite for any omics-specific module

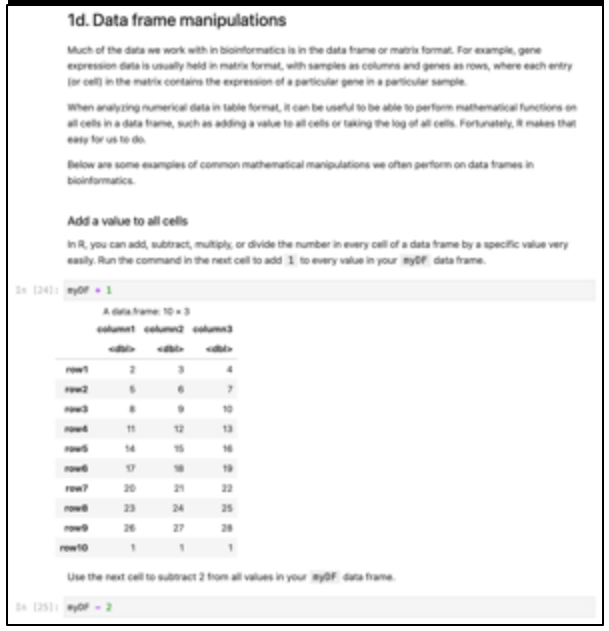
Omics-specific Modules (RNAseq, Amplicon Seq, etc.)

- Omics Data Lectures
  - Experimental design
  - Sample preparation and quality control
  - Data processing tools and visualizations
  - Results analysis and interpretation
- Hands-on data processing and analysis of an OSDR dataset via Jupyter Notebooks (JNs)

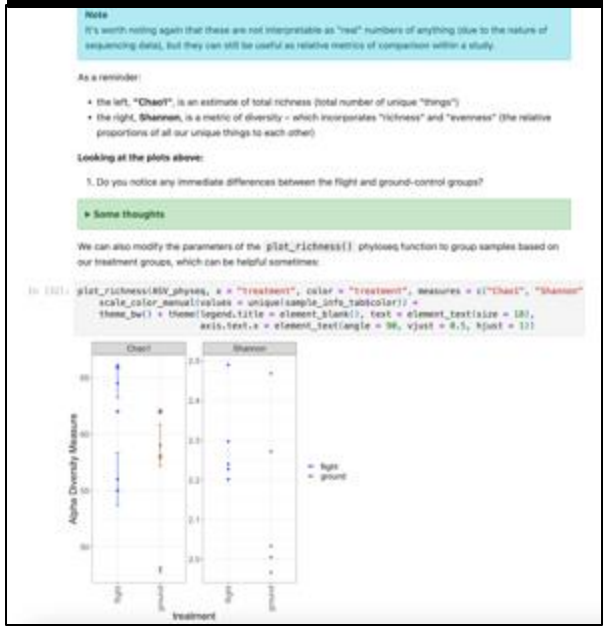
Unix Intro JN



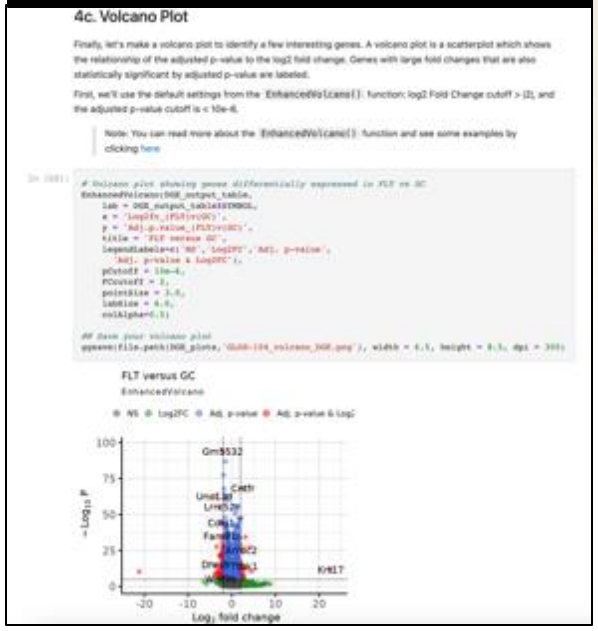
R Intro JN



AmpSeq JN



RNAseq JN



GL4U Approach

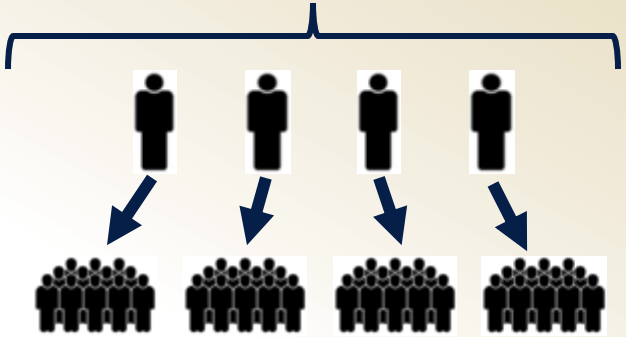
GL4U Direct Approach

Students

- Space biology-relevant training in bioinformatics
- Uses **direct** (training students) and **indirect** (training educators) approaches
- The GL4U Introduction module and one omics-specific module is taught to students or educators during a 1-2 week-long bootcamp
- Compute resources are provided to all participants
  - ➢ For indirect bootcamps: educators receive materials and training to enable them to run the bootcamp at their home institutions
- All GL4U modules are made publicly available on GitHub (https://github.com/nasa/GeneLab-Training/tree/main/GL4U) and virtual bootcamps are recorded to enable independent learning

GL4U Indirect Approach

Educators



GL4U RNAseq Certification

GL4U Introduction Module		GL4U RNAseq Module
☑ Pre-Intro Module Survey	+	☑ Pre-RNAseq Module Survey
☑ Intro Lectures		☑ RNAseq Lectures
☑ Intro JNs		☑ RNAseq JNs
☑ Post-Intro Module Survey		☑ Post-RNAseq Module Survey

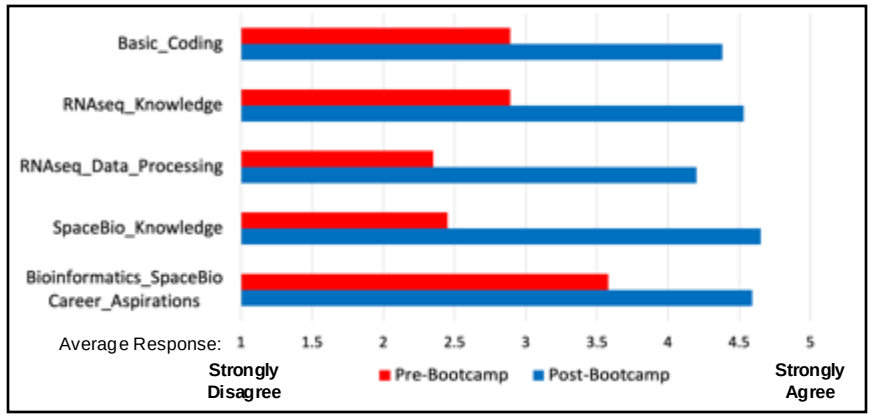
- Pre- and post- bootcamp surveys used to assess participant knowledge before and after training and to collect feedback
- GL4U certification offered for completing the GL4U Introduction module and one omics-specific module
  - ➢ *GL4U RNAseq certification shown as an example*

WINTER 2024: GL4U ON-DEMAND

GL4U Student & Educator Bootcamps To-Date

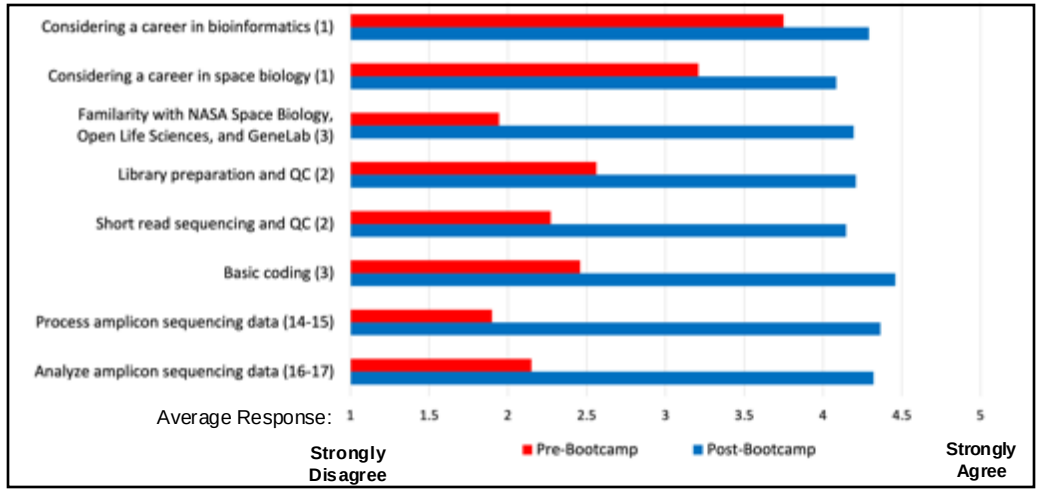
GL4U: RNAseq Student Bootcamp with SJSU, 06/2021

- Virtual 1-week long bootcamp with 17 **SJSU** students
- *Compute resources*: SJSU HPC
- Bootcamp covered GL4U Introduction and RNA Sequencing (RNAseq) modules



GL4U: AmpSeq Student Bootcamp at CSULA, 07/2023

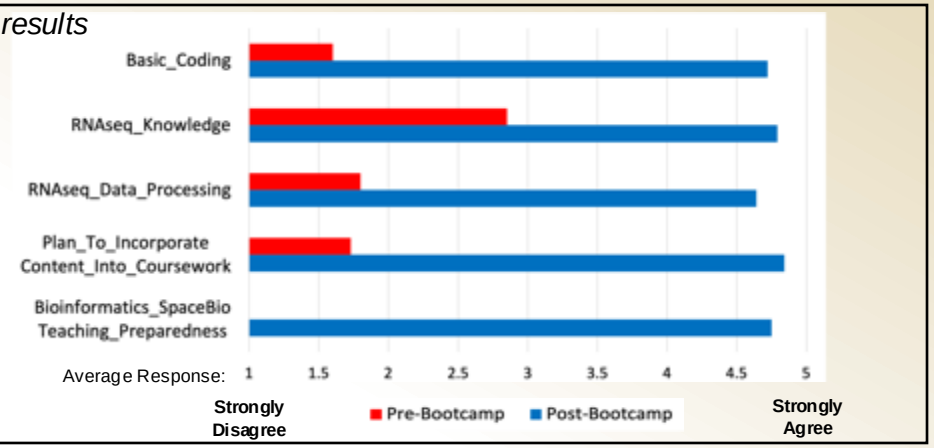
- In-person 4-day bootcamp with 24 **CSULA** students
- *Compute resources*: NSF ACCESS
- Bootcamp covered GL4U Introduction and Amplicon Sequencing (AmpSeq) modules



GL4U: RNAseq Educator Bootcamp, 06/2022

- Virtual 9-day long bootcamp with 6 professors and 4 graduate students from **4 institutions**:
  - CalState Northridge
  - New Mexico State University
  - Alabama A&M
  - CalPoly Pomona
- *Compute resources*: NASA NCCS SMCE
- Bootcamp covered GL4U Introduction and RNA Sequencing (RNAseq) modules

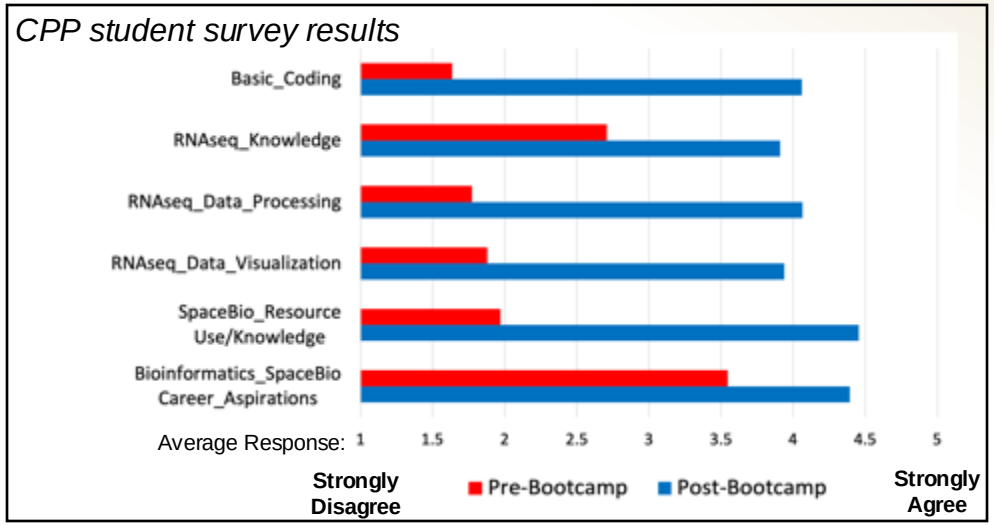
Educator survey results



CPP-GL4U RNAseq Independent Study Course, Fall 2022

- Dr. Lin taught GL4U RNAseq content for 11 junior & senior CalPoly Pomona (CPP) undergraduate students
- *Compute resources*: NASA NCCS SMCE

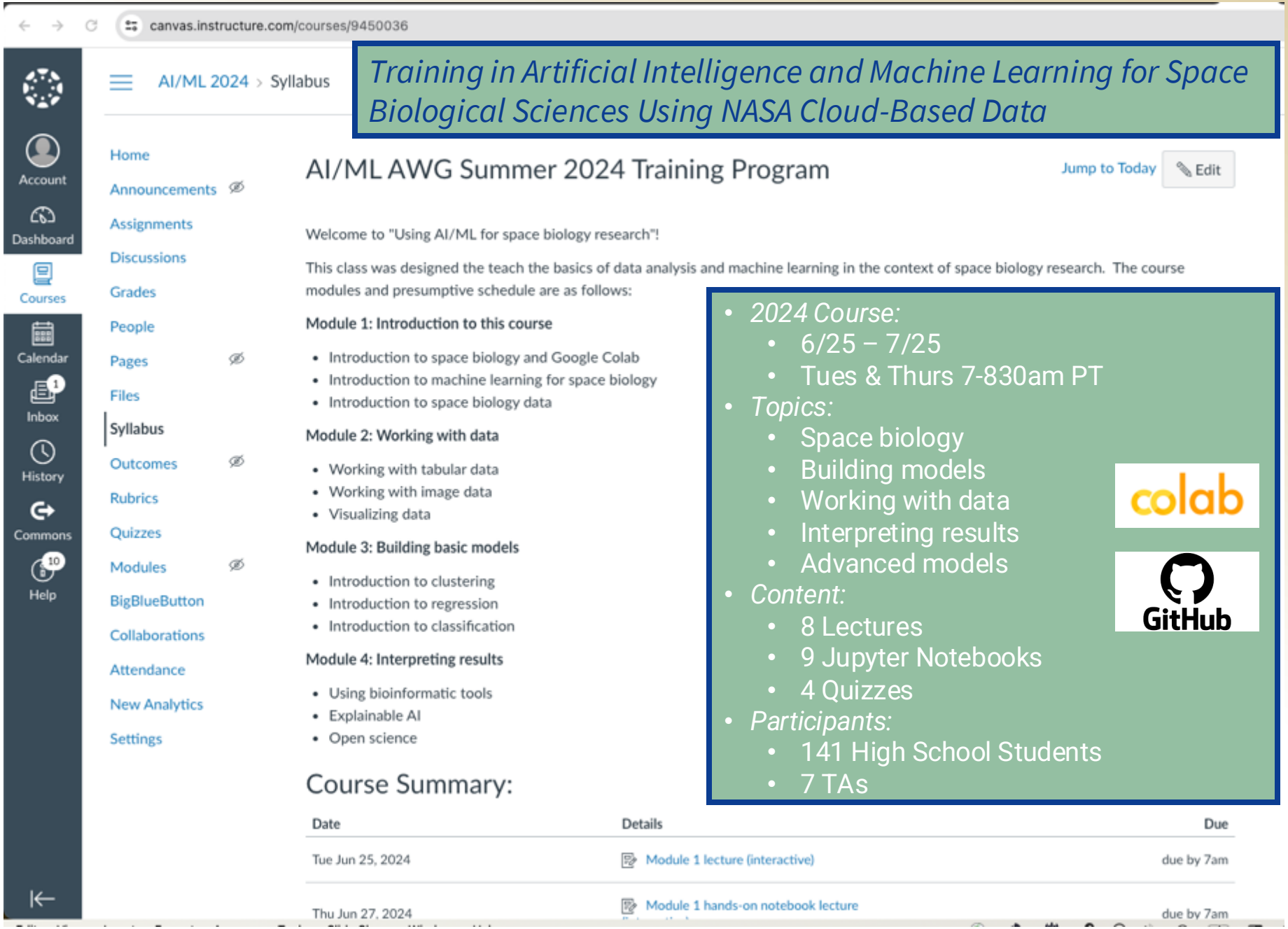
CPP student survey results



AI/ML Training – Synchronous

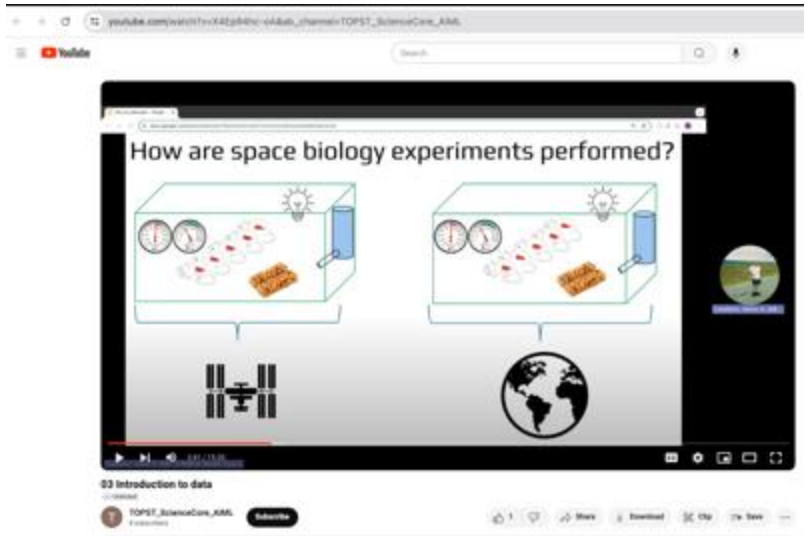
- *Training in Artificial Intelligence and Machine Learning for Space Biological Sciences Using NASA Cloud-Based Data*
  - AI/ML training with a focus on space biology
  - Real-world use cases with NASA cloud-based scientific datasets

Cloud based NASA biology data:
https://registry.opendata.aws/bps_rnaseq
https://registry.opendata.aws/bps_microscopy

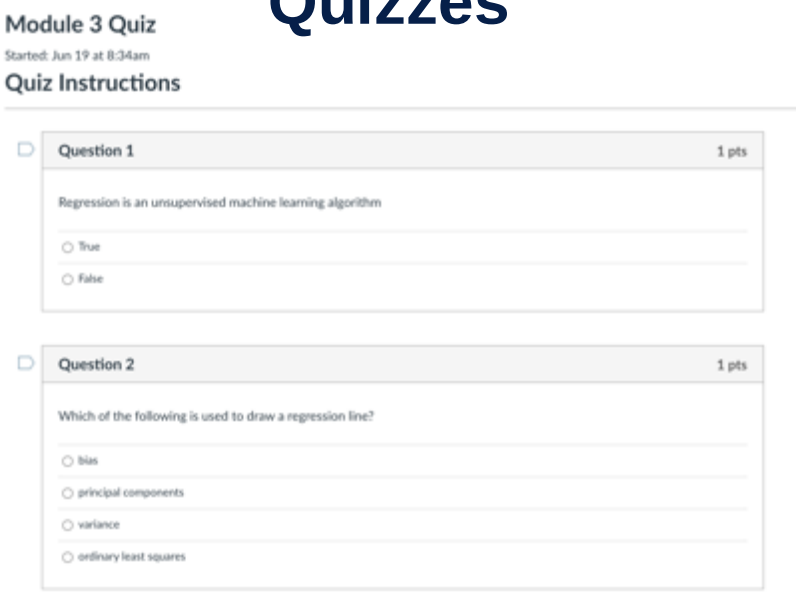


AI/ML Training – Asynchronous

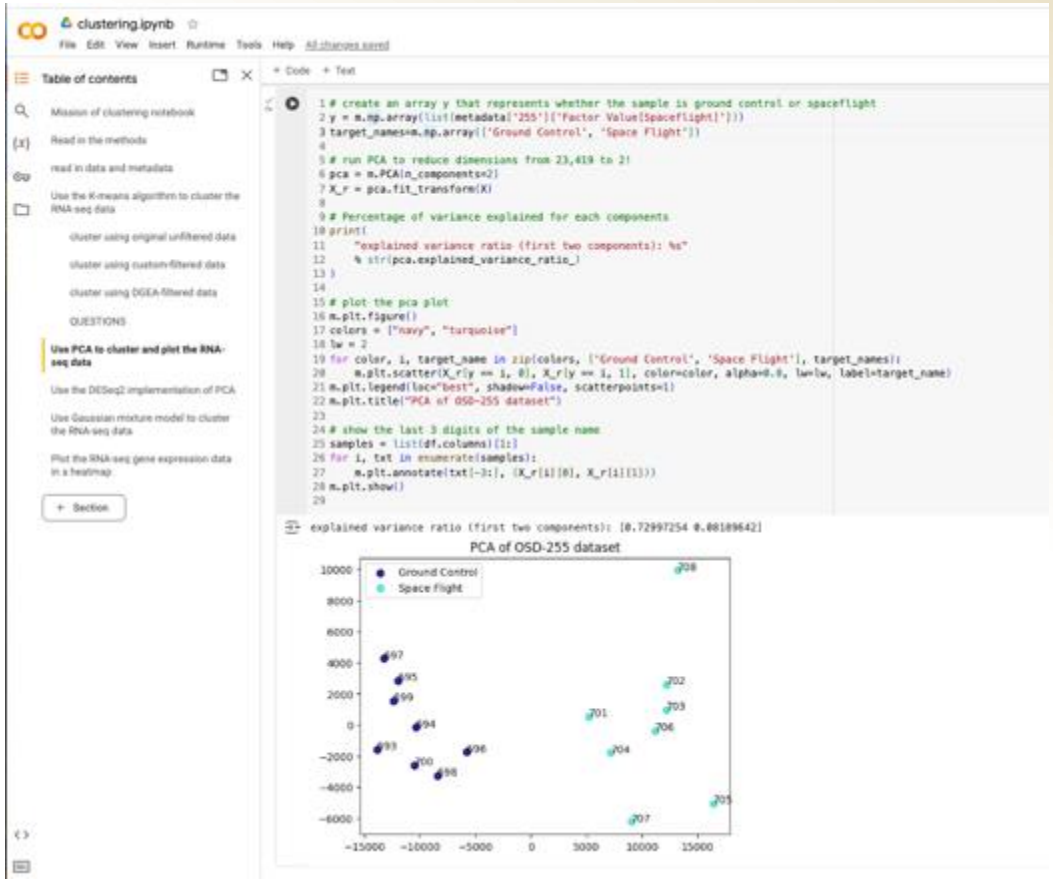
Pre-recorded lectures



Quizzes



Interactive Python notebooks



To enroll in this course, visit:

https://canvas.instructure.com/enroll/8JYKD7

Training in Artificial Intelligence and Machine Learning for Space Biological Sciences Using NASA Cloud-Based Data

- 2 week curriculum
- 2.5 hours / day
- *Topics:*
  - Space biology
  - Working with data
  - Building models
  - Interpreting results
  - Capstone projects
- Available Summer 2024 for independent learning

colab

GitHub

Get Involved!

Open Science Analysis Working Groups (AWGs)

Join an AWG!

ANIMAL

180 members

MULTI-OMICS

306 members

MICROBIAL

163 members

PLANTS

132 members

AI/ML

230 members

ALSDA

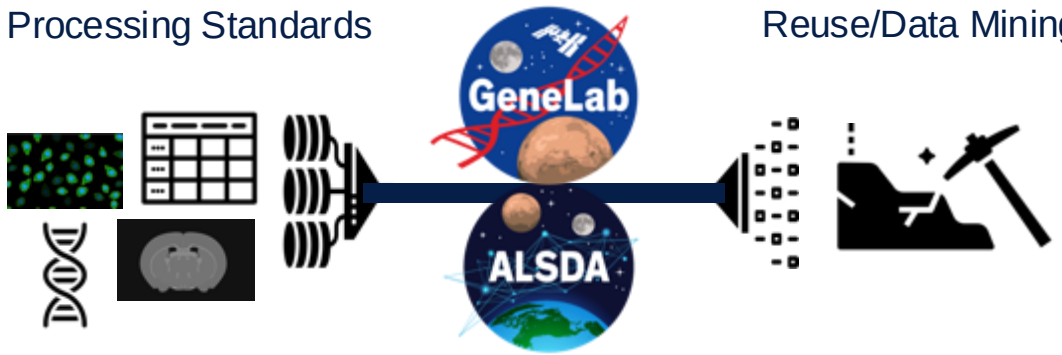
(Physiological/BioMedical)

189 members

Feedback on Ingress/Egress Standards

Ingestion, Curation, Processing Standards

Standards to Enable Reuse/Data Mining



Collaborate on Data Mining/Publications

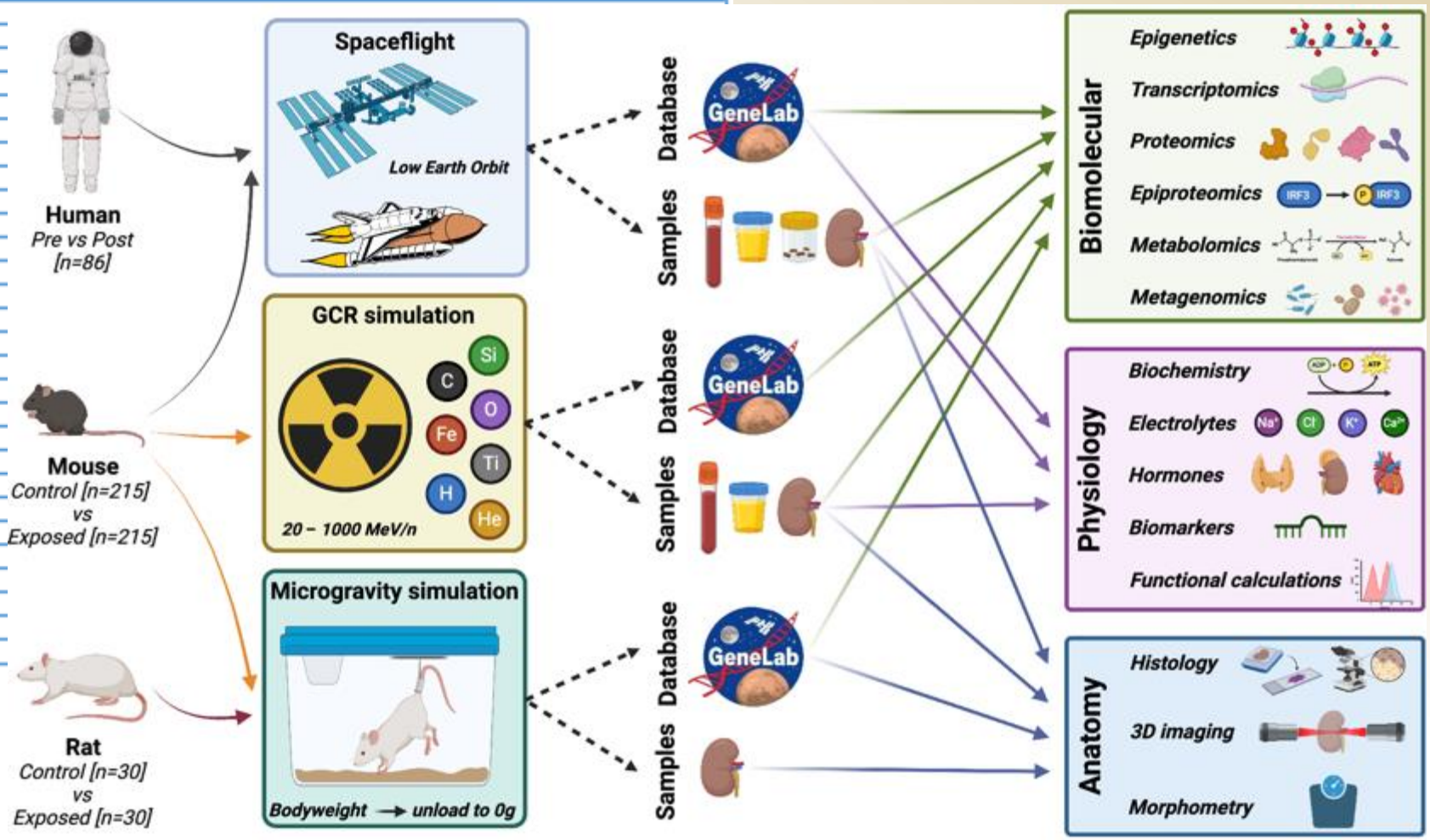
Cell Press Package 2020:
The Biology of Spaceflight
https://www.cell.com/c/the-biology-of-spaceflight

Nature Portfolio Collection 2024:
Space Omics and Medical Atlas across orbits (SOMA)
https://www.nature.com/collections/ebdbcahdgc

List of 74 publications enabled by OSDR
https://osdr.nasa.gov/bio/data/publications.html

Example of Open Science AWG Collaboration and Publication

Mission	Experimental Condition	Species	Sex	Background	GCR (mG)	Mission Duration (d)	Sample Collection	n	Control	Data Source	Tissue	DNA	RNA	Protein	PTMs	Metab	Microbes	Physiol	Anatomy
Inspiration4	Spaceflight	Human	Mix	USA	Unknown	3	R+0	4	Paired	Original	P, U			✓		✓		✓	
JAXA	Spaceflight	Human	♂	Japan	Unknown	120	L+120	6	Paired	Original	P		✓						
NASA	Spaceflight	Human	Mix	USA	Unknown	180	L+120	66	Paired	Database	P, U							✓	
Roscosmos	Spaceflight	Human	♂	Russia	Unknown	199	R+1	10											
BNL-1	simGCRsim	Mouse	♀	C57BL/6	500.0	~1.5yr sim	PE+1	10											
	MGsim				N/A	14	PE+1	10											
	simGCRsim + MGsim				500.0	~1.5yr sim	PE+1	10											
BNL-2	simGCRsim	Mouse	♀	C57BL/6	500.0	~1.5yr sim	PE+1	10											
	MGsim				N/A	14	PE+1	10											
	simGCRsim + MGsim				500.0	~1.5yr sim	PE+1	10											
BNL-3	simGCRsim	Mouse	♀	C57BL/6	500.0	~1.5yr sim	PE+1	10											
	MGsim				N/A	14	PE+1	10											
	simGCRsim + MGsim				500.0	~1.5yr sim	PE+1	10											
NSRL-22A	GCRsim	Mouse	♀	C57BL/6	750.0	~2.5yr sim	PE+180	12											
	GCRsim	Mouse	♂	C57BL/6	750.0	~2.5yr sim	PE+180	12											
STS-135	Spaceflight	Mouse	♀	C57BL/6CR	4.7	13	R+0	7											
RR-1	Spaceflight	Mouse	♀	C57BL/6J	5.0	38	L+38	6											
RR-3	Spaceflight	Mouse	♀	BALB/c	5.9	41	L+41	6											
RR-6	Spaceflight	Mouse	♀	C57BL/6NTac	8.2	29	R+4	9											
					16.1	57	L+75	9											
RR-7	Spaceflight	Mouse	♀	C3H/Hej	4.4	25	L+25	5											
					13.2	75	L+75	5											
				C57BL/6	4.4	25	L+25	5											
					13.2	75	L+75	5											
RR-9	Spaceflight	Mouse	♂	C57BL/6J	5.5	33	R+1	5											
RR-10	Spaceflight	Mouse	♀	B6129SF2/J	4.2	28	L+28	10											
RR-19	Spaceflight	Mouse	♀	C57BL/6N	5.2	35	R+0	8											
RR-23	Spaceflight	Mouse	♂	C57BL/6J	5.8	37	R+1	9											
MHU-1	Spaceflight	Mouse	♂	C57BL/6J	9.0	35	R+2	6											
MHU-3	Spaceflight	Mouse	♂	C57BL/6NCrl	9.0	31	R+0	6											
CNSA	MGsim	Rat	♂	Sprague-Dawley	N/A	28	R+0	30											



- **Renal/Kidney Research**
  - 20 spaceflight missions & ground simulations
  - 50+ datasets
  - 5 continents of collaborators

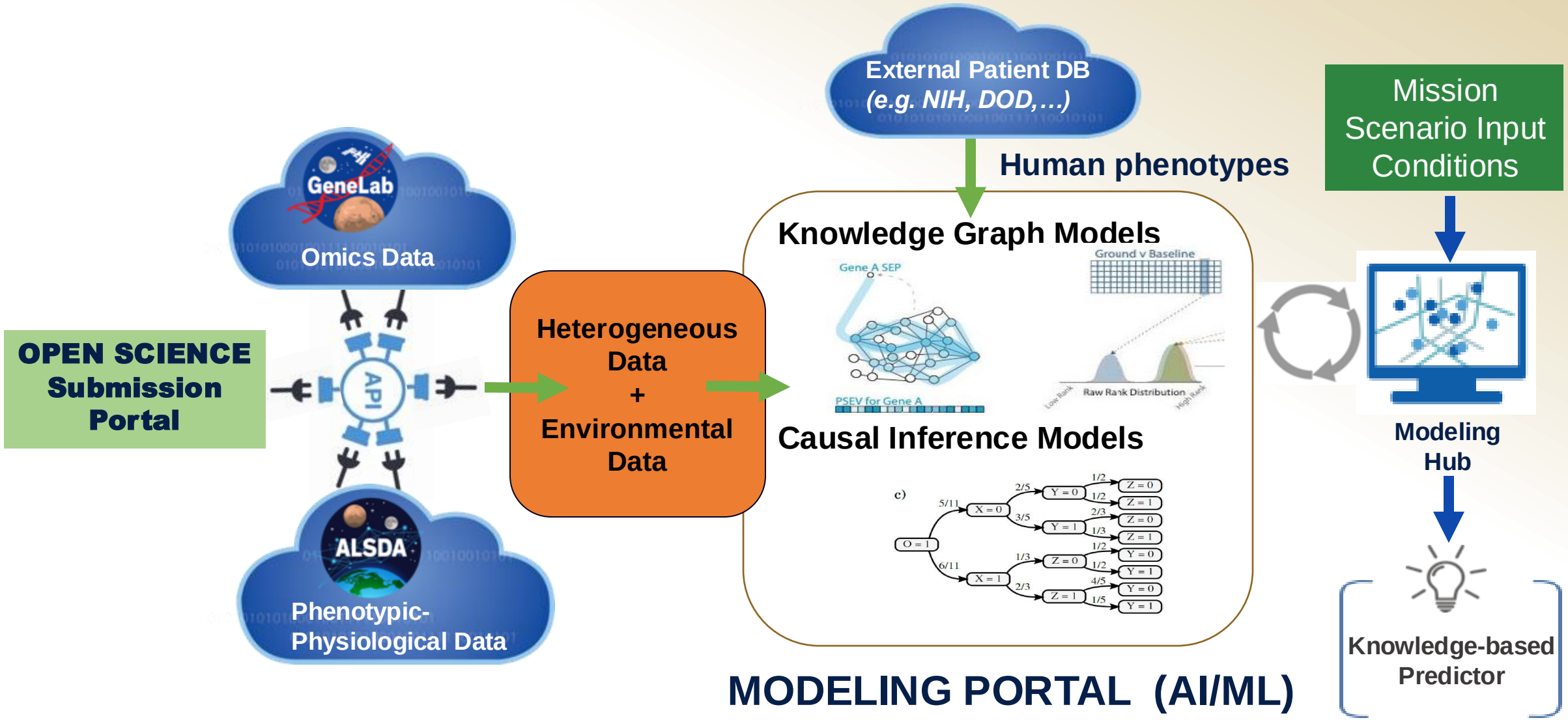
Cosmic Kidney Disease: The Effects of Spaceflight and Galactic Cosmic Radiation on Renal Structure and Function.

Siew, K., et al., Nat Communications 15, 4568 (June 2024). https://doi.org/10.1038/s41467-024-49212-1

What's Next?

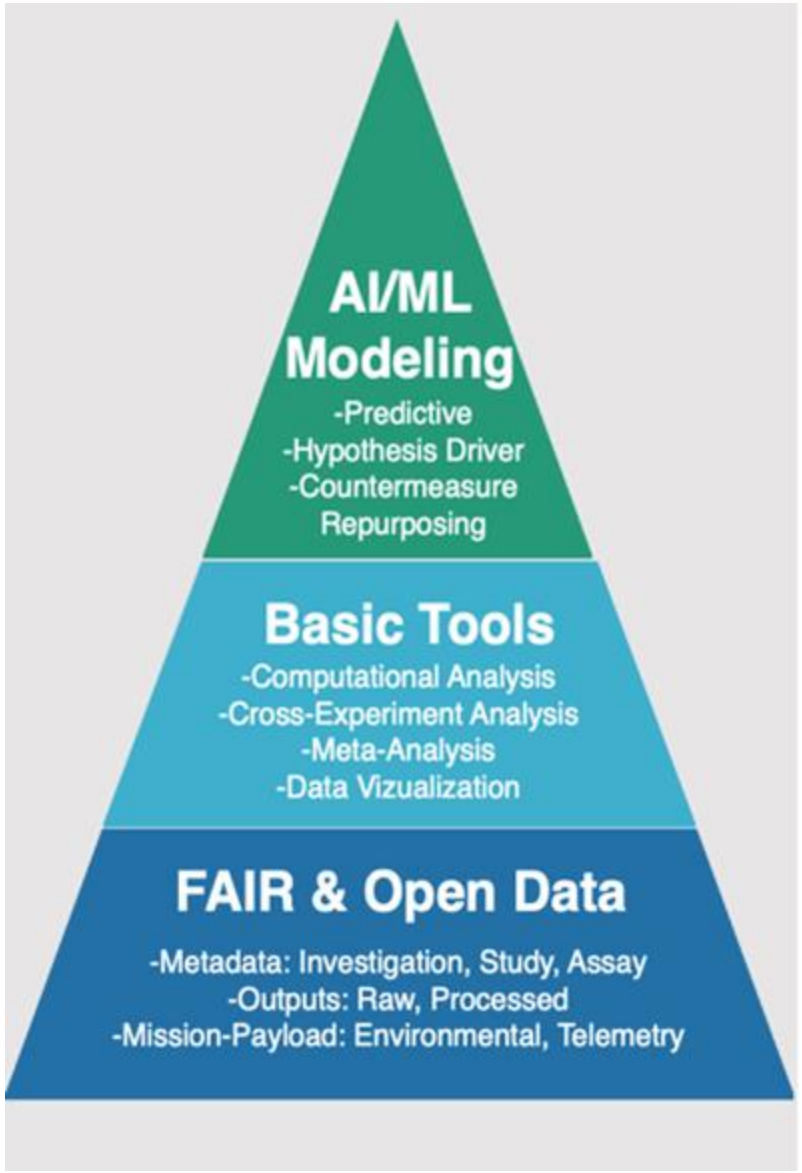
From Multiple Databases to a Knowledge-Based System

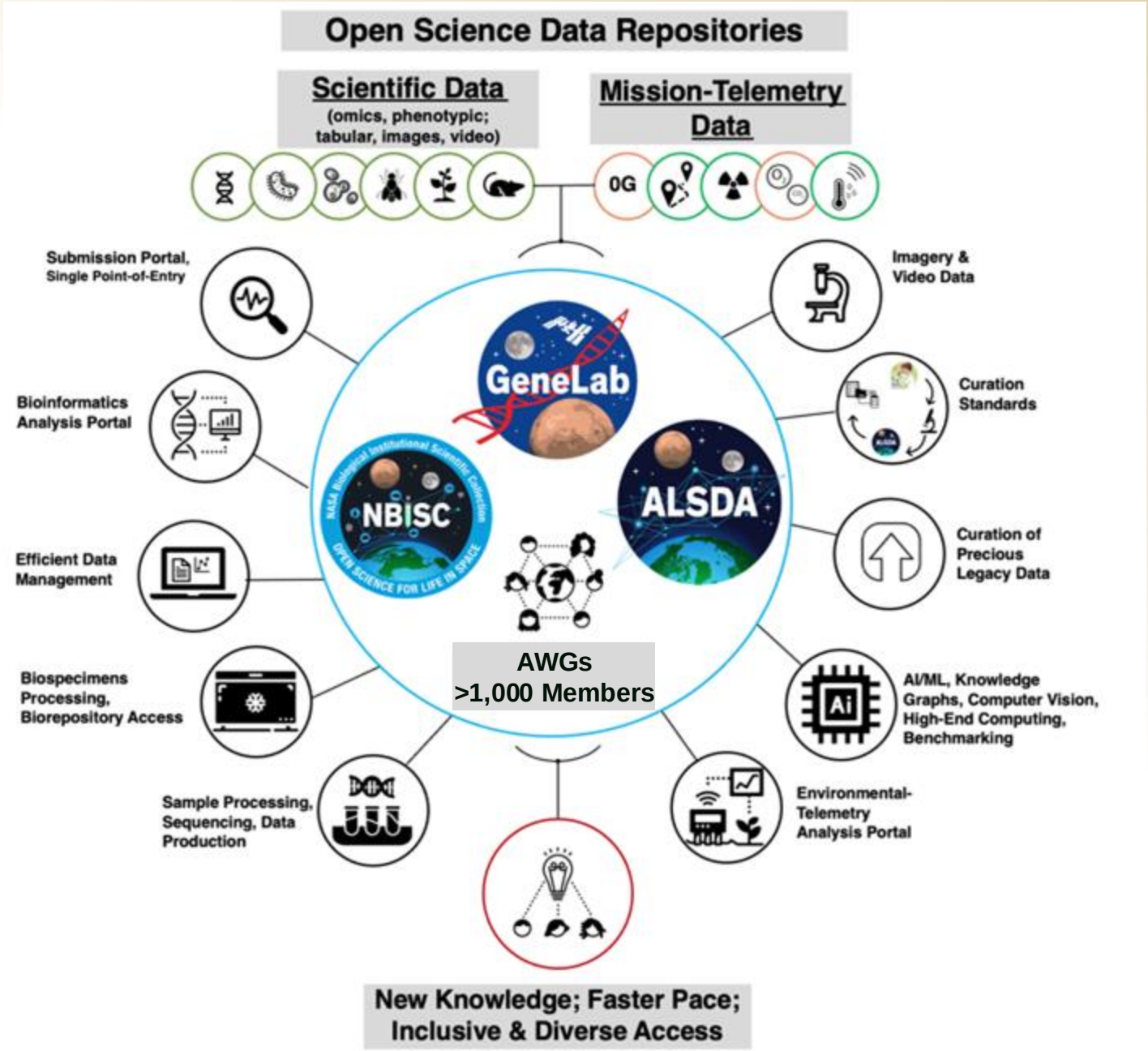
AI ready FAIR data for modeling and risk assessment



Credit: Nelson et al., 2021, https://doi.org/10.3390/life11010042; Genewein et al., 2020, arXiv:2010.12237v1 [cs.AI])

Big Picture for Open Science for Life in Space





Opportunities & Resources

Training and Other Opportunities at NASA

GeneLab for High Schools (GL4HS): A four-week intensive training summer program for rising high school juniors and seniors to learn bioinformatics and computational biology methods and techniques to analyze space omics data.

Learn more and apply at: https://www.nasa.gov/ames/genelab-for-high-schools

GeneLab for Colleges/Universities (GL4U): For educators and students to learn how analyze omics data using GeneLab standard pipelines and space-relevant data

Access to course materials: https://github.com/nasa/GeneLab-Training/tree/main/GL4U

Space Life Sciences Training Program (SLSTP): Provides undergraduate students entering their junior or senior years, and entering graduate students, with professional experience in space life science disciplines.

Learn more and apply at: https://www.nasa.gov/ames/research/space-life-sciences-training-program

---

NASA Postdoctoral Program: If you're an early-career or senior scientist, apply to the NASA Postdoctoral Program to help pursue NASA's mission and experience the world's most diverse technology and expertise.

Learn more and apply at : https://npp.orau.org/

NASA Pathways Program: Provides current students and recent grads with a dynamic career development program at a NASA center. Participants are given an opportunity for permanent employment with NASA.

Learn more and apply at : https://www.nasa.gov/careers/pathways-program

Spaceflight Technology, Applications, and Research (STAR): The annual course targets principal investigators (PIs), senior research scientists, and postdoctoral scholars and aims to facilitate their entry to space biology and preparation for conducting spaceflight experiments using NASA and commercial platforms.

Learn more and apply at : https://science.nasa.gov/biological-physical/programs/star

Acknowledgements

Open Science for Life in Space Teams

Open Science Analysis Working Group Members

Support

NASA Space Biology Program
NASA Science Mission Directorate
NASA Human Research Program
NASA Biological and Physical Sciences

END
GL4U: Introduction
Lecture 2 of 4