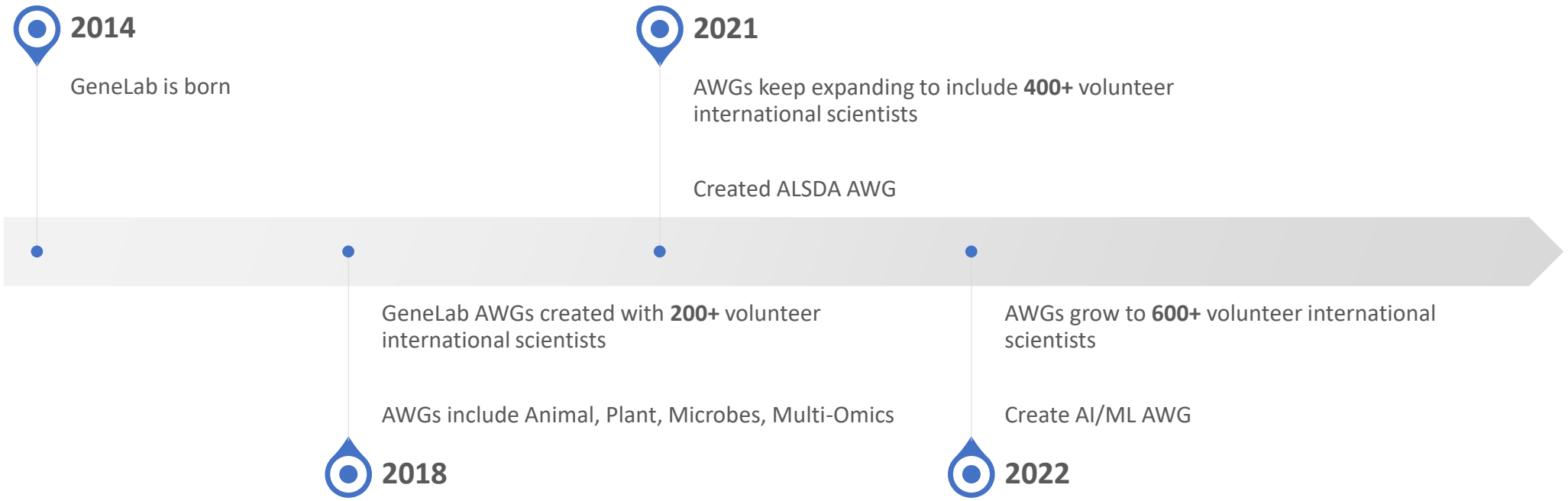
National Aeronautics and Space Administration

Fostering Open Science Through Analysis Working Groups (AWGs)

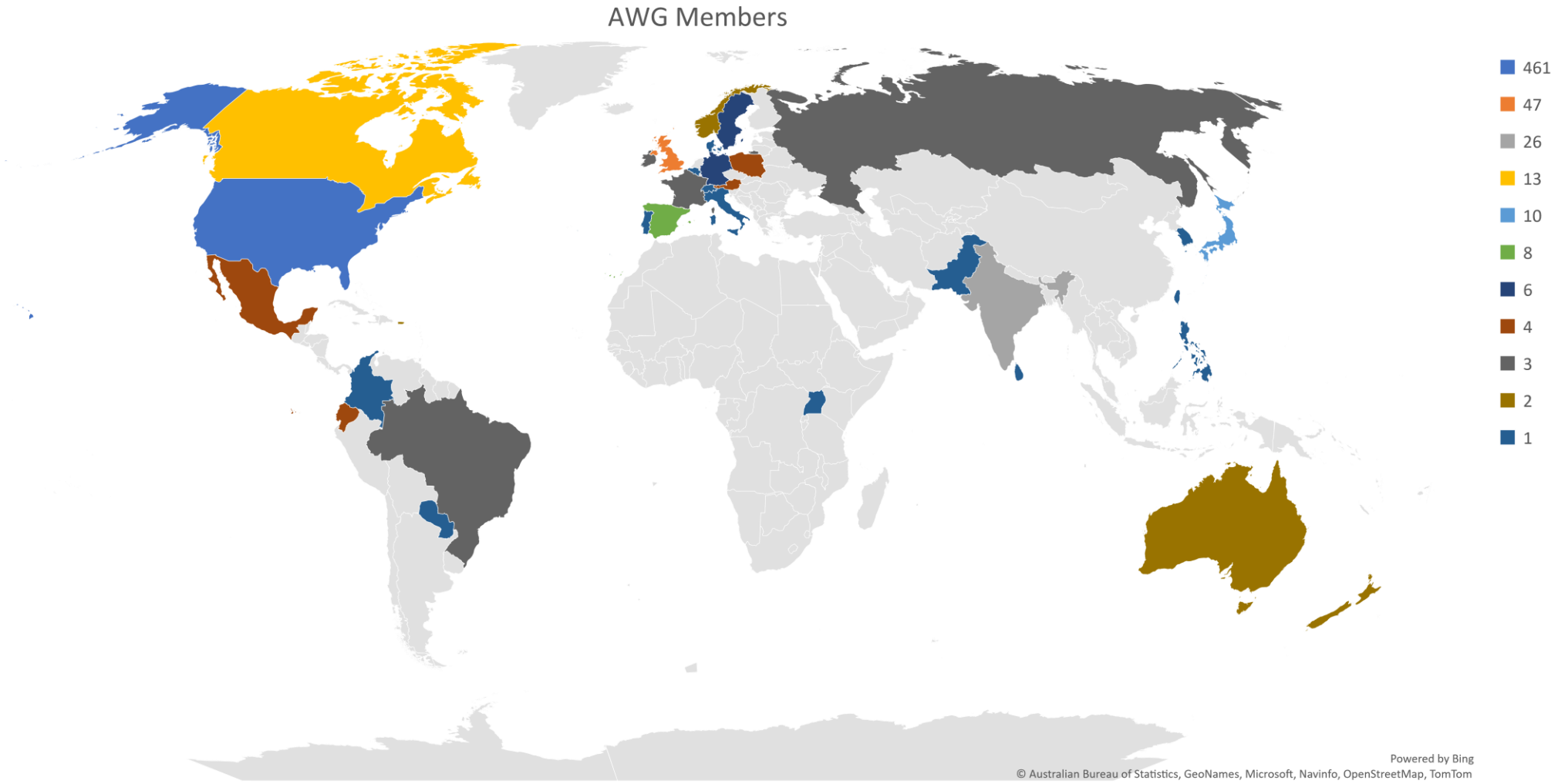
Lovorka Degoricija, PhD
GeneLab Science Communications Lead
KBR
NASA Ames Research Center

Connecting the World Through Science



- *Monthly meetings to analyze data in the GeneLab repository*
- *Attend annual meetings and workshops*

Who are the AWGs?



Original AWGs

ANIMAL

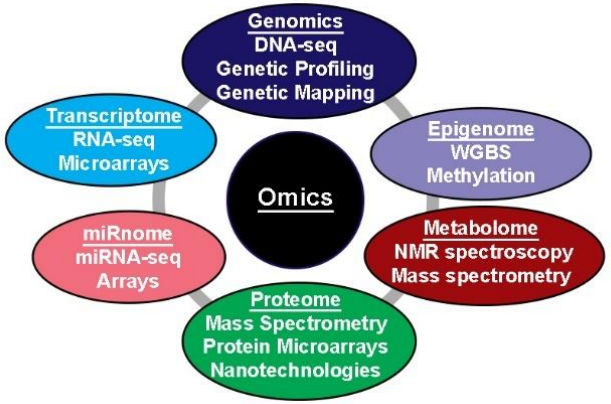
79 members

Facilitates the use of omics in understanding basic mechanisms by which animals and constituent tissues and cells adapt to the spaceflight environment.

MULTI-OMICS

259 members

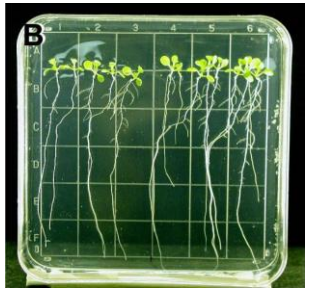
Interactions between the different omics to provide a better understanding of the systemic response.



PLANTS

66 members

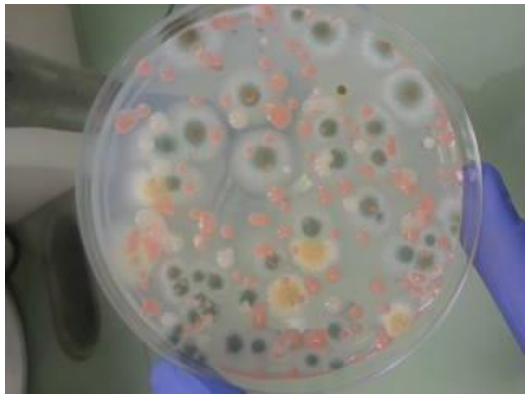
*Share and discuss the latest developments in **AstroBotany** – the discipline of botany concerned with interactions between plant biology and space environment.*



MICROBES

61 members

Focuses on analyzing microbial datasets within GeneLab that includes gene-expression, proteomic, metabolomic and environmental metagenomic datasets.

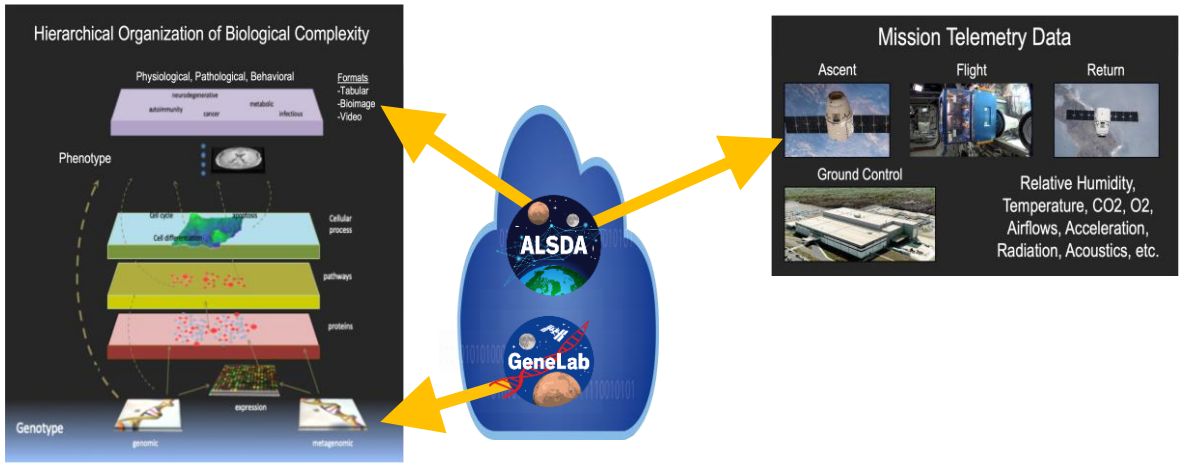


Expanded AWGs

ALSDA

110 members

Feedback on science data and metadata standards for physiological, phenotypic, and behavioral datasets to be reusable. Datasets span from raw to processed-results data, and across tabular, bioimaging, and video formats.



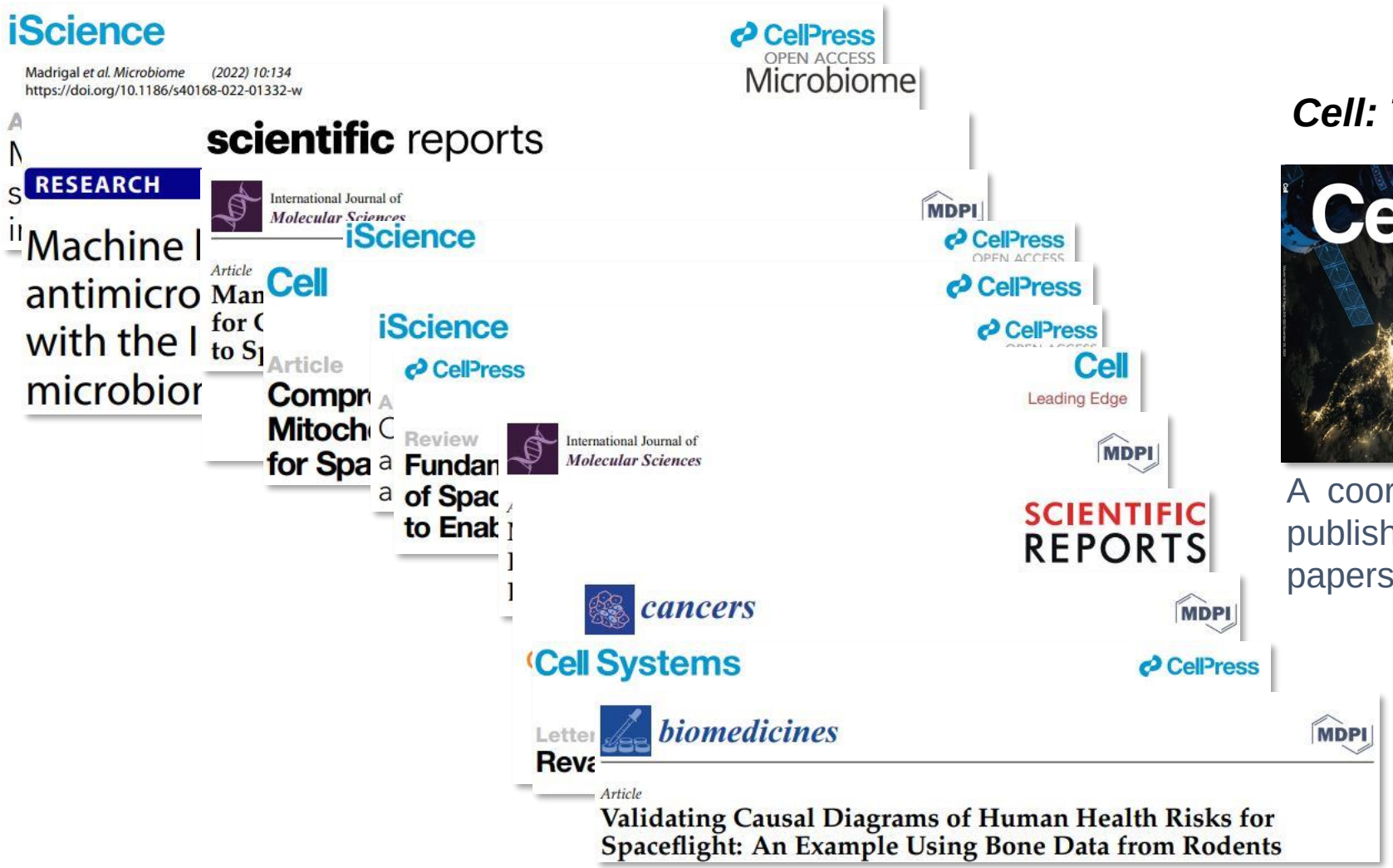
AI/ML

56 members

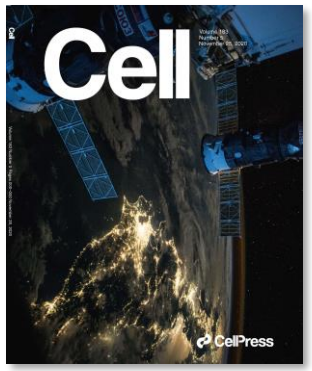
Focuses on developing data AI-readiness guidelines, algorithm and automation development, and developing ethical guidelines to increase trust and explainability surrounding AI in space biology.

Publications

***50** publications (**12 publications produced by the AWGs**) using data available in GeneLab.*



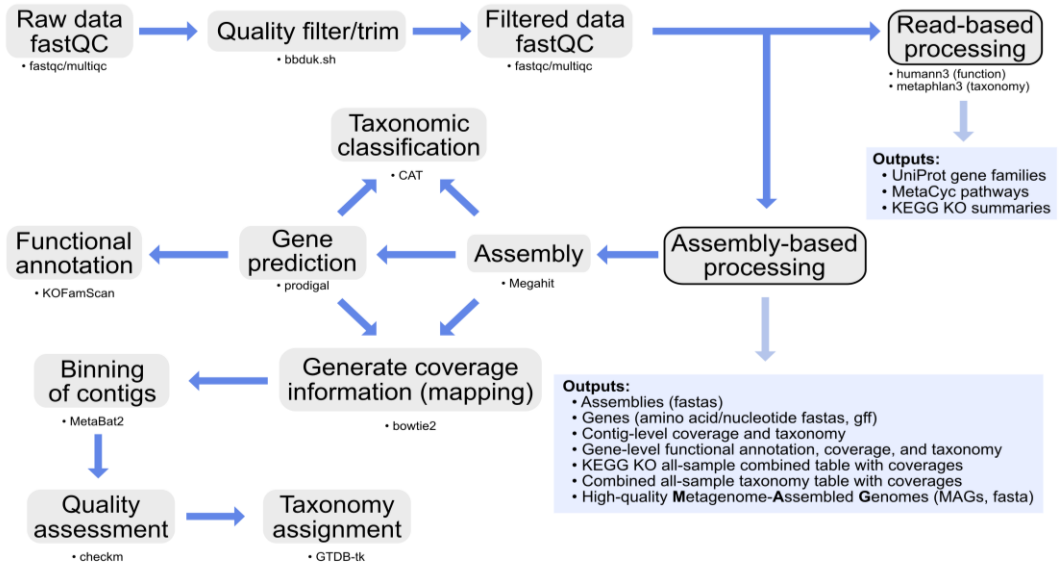
Cell: The biology of spaceflight package



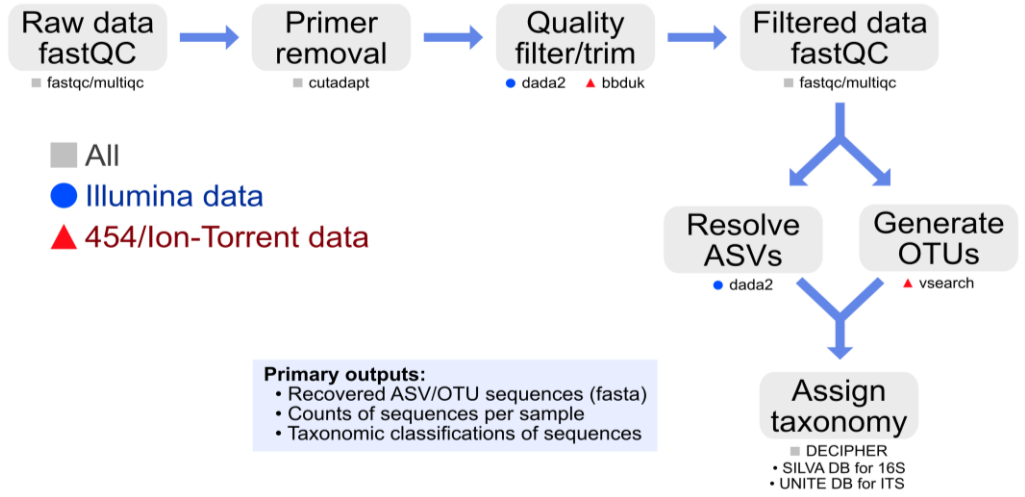
A coordinated package of 29 scientific papers published in five *Cell Press* journals featuring **9** papers utilizing data or resources in GeneLab.

GeneLab Standard Omics Processing Pipelines

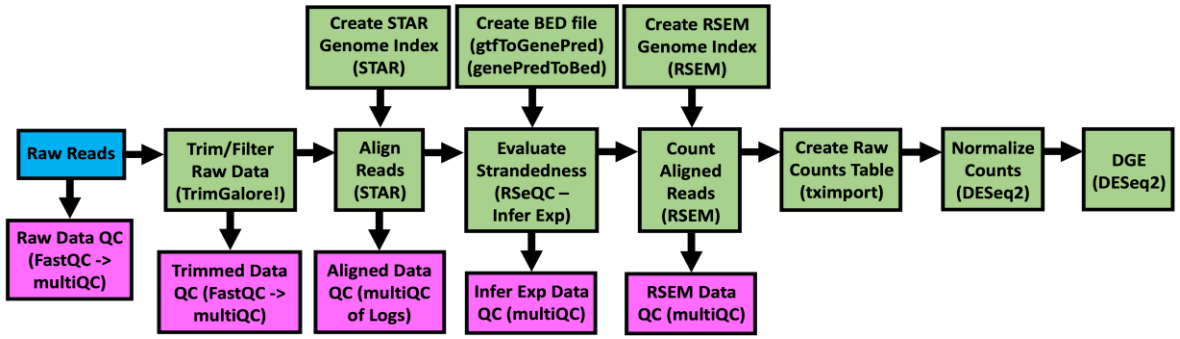
Metagenomics (Illumina) Pipeline



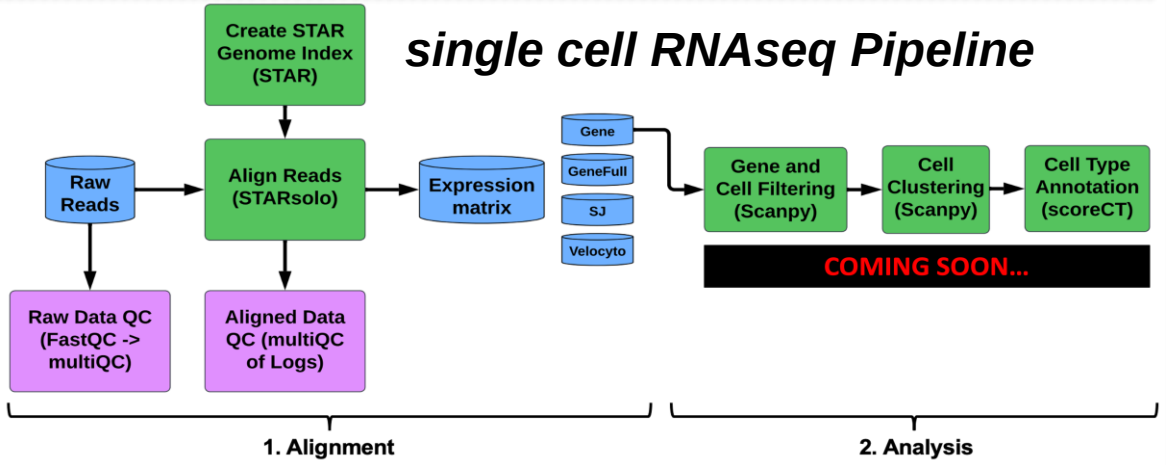
Amplicon Pipeline



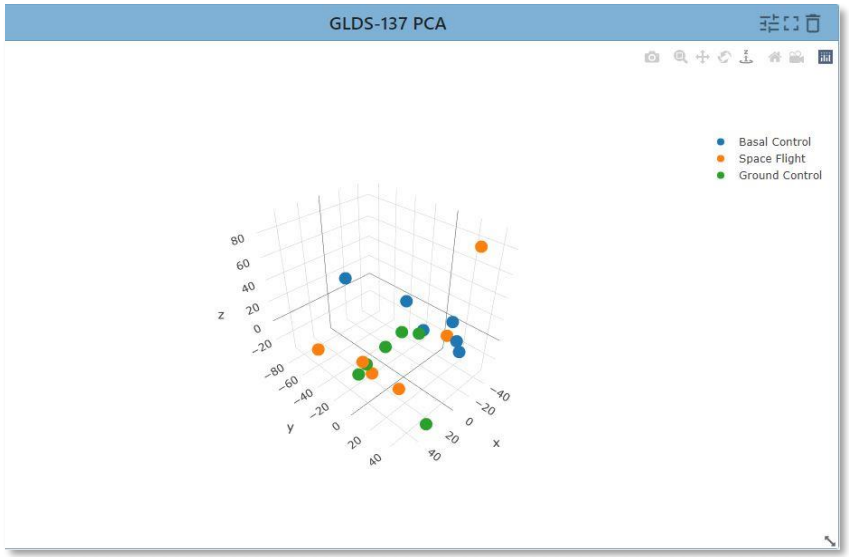
(bulk) RNAseq Pipeline

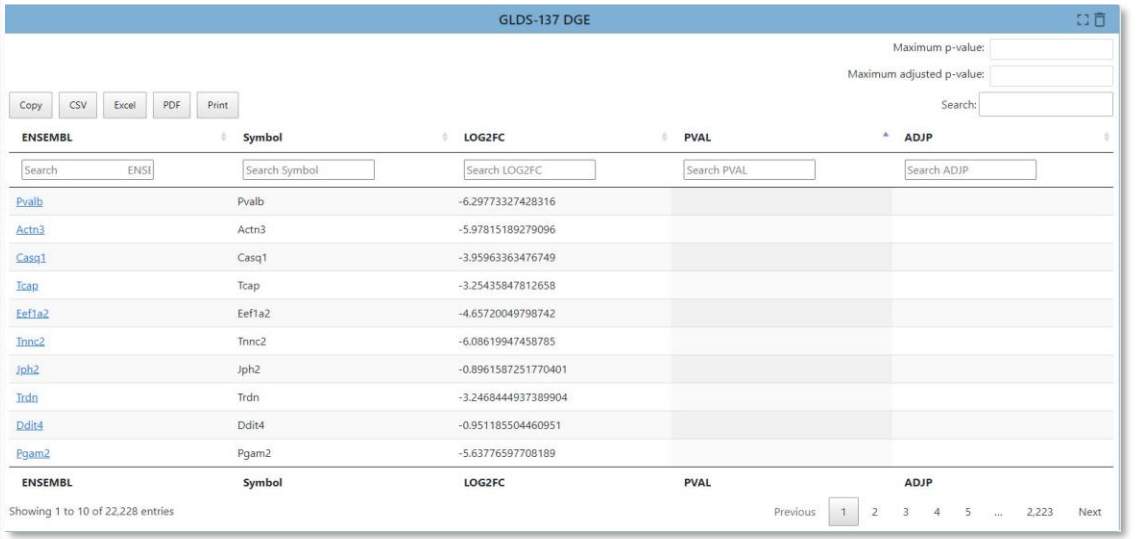


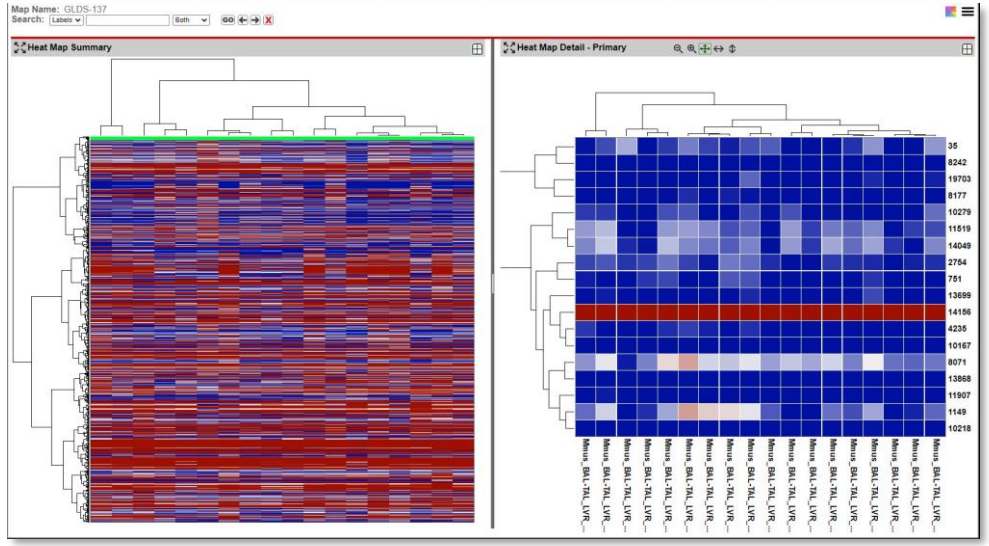
single cell RNAseq Pipeline

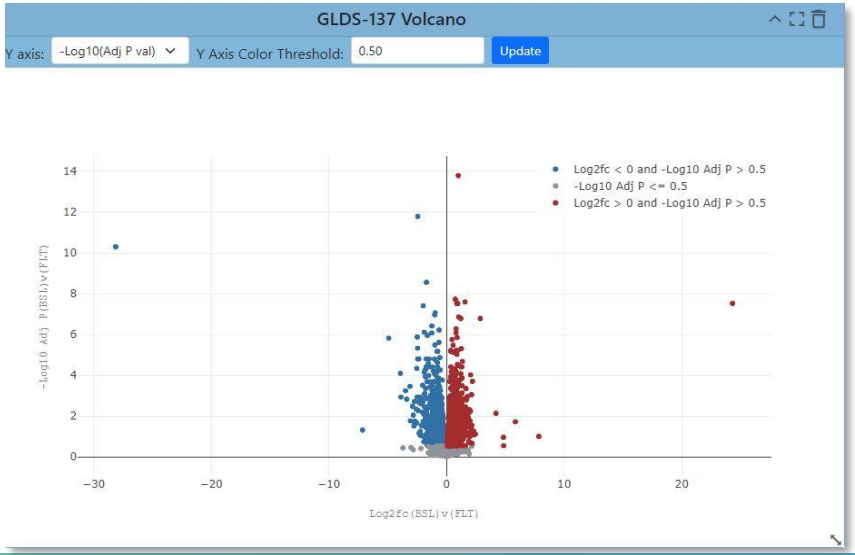


GeneLab Data Visualization Portal (Transcriptomics)









Ames Life Science Data Archive (ALSDA) AWG

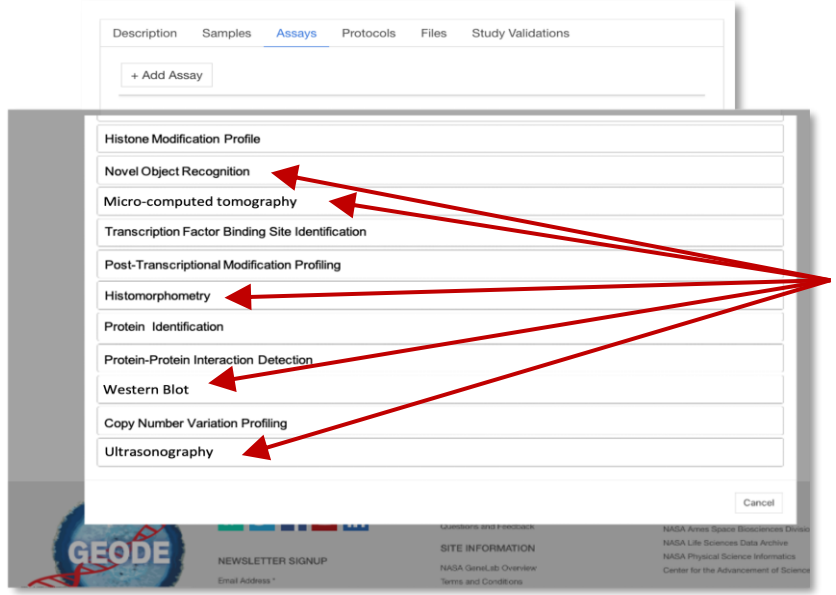
SME Feedback on Assay Metadata "Configurations"

Current Assay Metadata Configurations
Molecular Cellular Imaging (Light/Fluorescence Microscopy)
Behavior (Elevated Plus Maze)
Behavior (Novel Object Recognition)
Calcium Uptake (Spectrofluorimetry)
Protein Quantification (Western Blot)
Flow Cytometry (Flow Cytometry)
Bone Microstructure (Micro-Computed Tomography)
Behavior (Unconstrained Cognitive Flexibility)
Behavior (Attentional Set-Shifting)
Behavior (Object in Place)
Immunoassay (ELISA)
Echocardiogram (Ultrasonography)
Histological Assay (Histomorphometry)
Mechanical Testing (3pt/4pt Bend/Torsion)
Intra-Ocular Pressure (Tonometry)
Behavior (Radial Arm Water Maze)
Behavior (Open Field)
Behavior (3-Chamber Social Interaction)
Behavior (Barnes Maze)
Behavior (Gait)
In Vivo Bone (pQCT)
Mineral Density and Composition (DXA)

Upcoming Configurations
Behavior (Object in Place)
Behavior (Novel Location)
Neurophysiology (Electrophysiological Recordings)
Physiology (Blood Pressue)
Behavior (Prepulse Inhibition)
Metabolites (HPLC with ECD)
Cardiovascular Imaging (MRI)
Neuroimaging (MRI)
Eye Microstructure (microCT)
Grip Force (Force Transducer)
Behavior (Spatial Novelty Y Maze)
Behavior (Tail Suspension Test)
Behavior (Startle Test)
Behavior (Wire Hang)
Behavior (Escher Staircase Track)
Immunophenotyping (Blood Cell Counts)
Infrared Spectroscopy (Infrared Spectroscopy)
Muscle Physiology (Muscle Fiber Typing)
Blood Analysis (Serum Metabolites)
Behavior (Contextual Fear Conditioning)
Behavior (Rotarod)
Physiology (Pulse Oximetry)

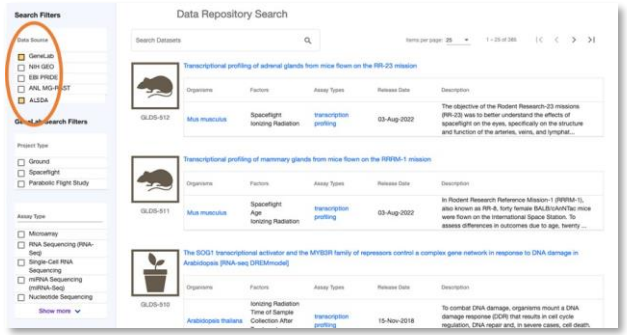
- *22 Configs Created*
- *Feedback on Tabular Data Templates for Reuse Standards per Assay*

Detailed Sample and Assay Metadata Crucial



Configs Used to Add Phenotypic Assays to Submission Portal

This Enables Common Search of Omics & Phenotypic Datasets



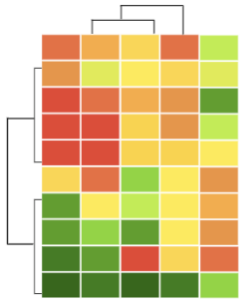
Artificial Intelligence and Machine Learning (AI/ML) AWG Goals

Focus Area 1: AI-Readiness

Pre-processed datasets ready for AI model training

Action items:

- AI-readiness data and metadata guidelines for the Open Science Data Repositories (OSDR)
- Standardized AI-ready dataset structure
- AI-readiness "score"

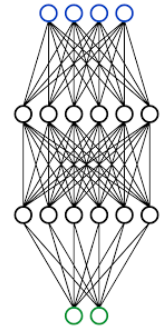


Focus Area 2: Algorithms for Space Biology

Address specific space biology data challenges (low n, high-dim)

Action items:

- Prioritize research questions: Decadal Survey on Life and Physical Sciences Research in Space 2023-2032
- Focus on gaps in current capabilities
- Develop benchmarking tasks for new algorithms

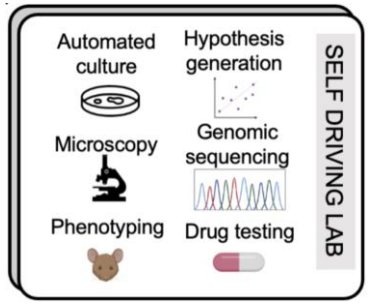


Focus Area 3: Self-Driving Labs

Automate space biological experiments

Action items:

- Prioritize physiological biological questions for automation
- Integrate several automated experimental platforms for comprehensive experimental readout
- *In-situ* analysis capabilities



Focus Area 4: Trust, Ethics & Explainability

Ethical AI guidelines for space biology

Action items:

- Education of next generation of AI scientists
- Consensus document for ethical standards for AI in space biology
- Guidelines for experimental design, metadata, and data

Interested in becoming a member?

Here's how to join: