National Aeronautics and Space Administration

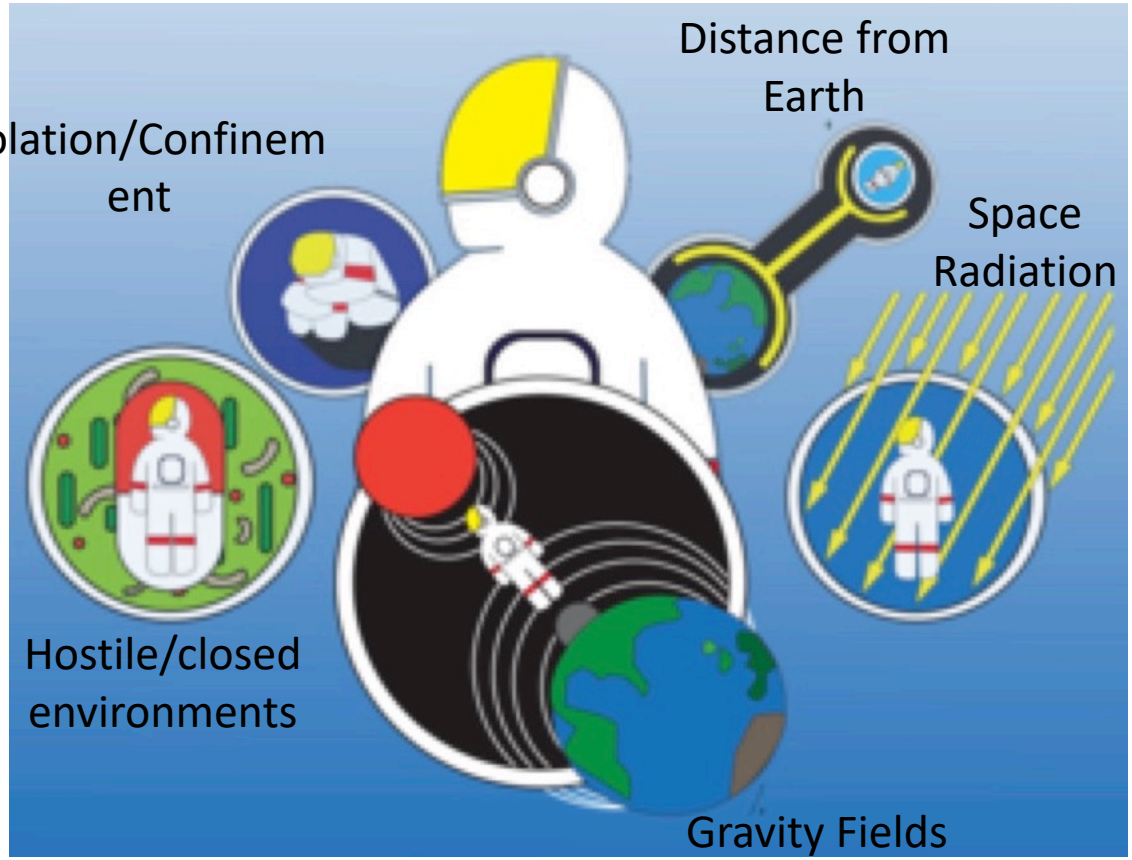
GeneLab: Omics Data System for Space Biology Research

Jonathan Galazka, PhD
GeneLab Project Scientist

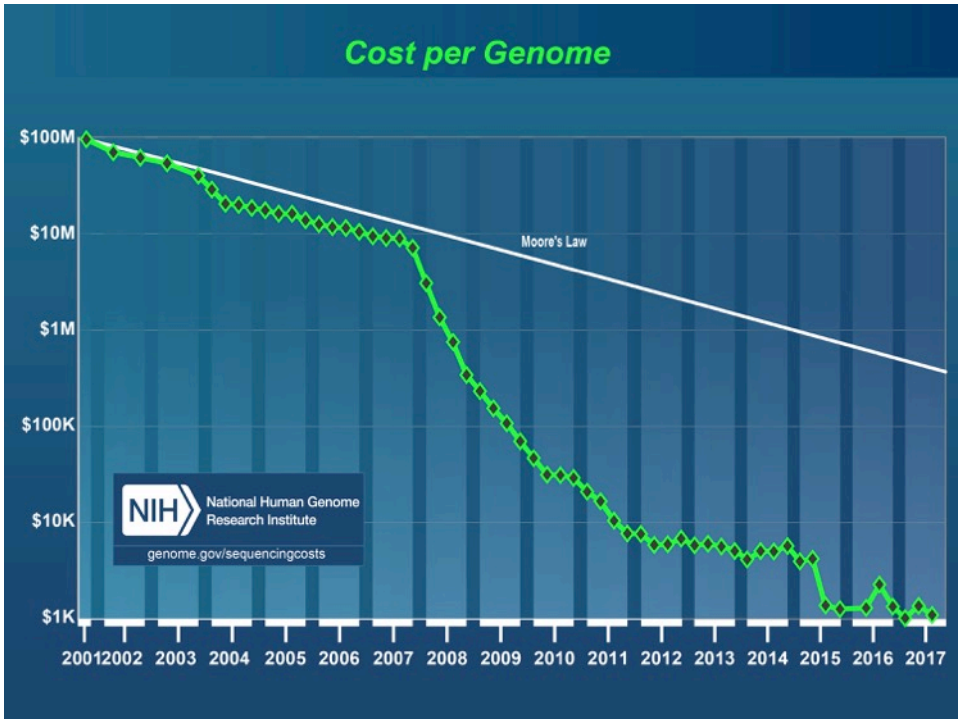
The GeneLab Team (Biologist, Computer Scientists, Data scientists, Science Communications, Project Support)

Challenges of Spaceflight

Complex set of factors ⟶ Acting on a complex system

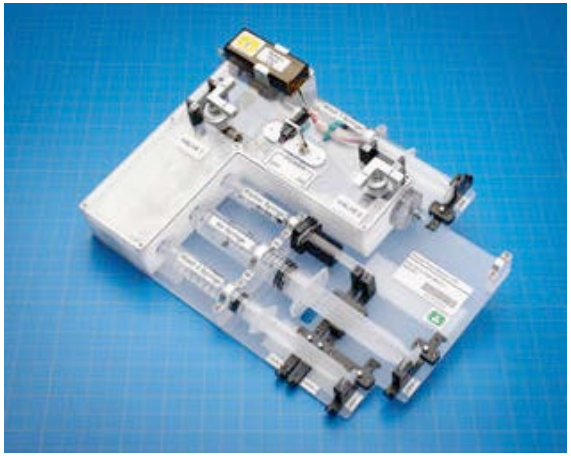


Promise of data driven investigation



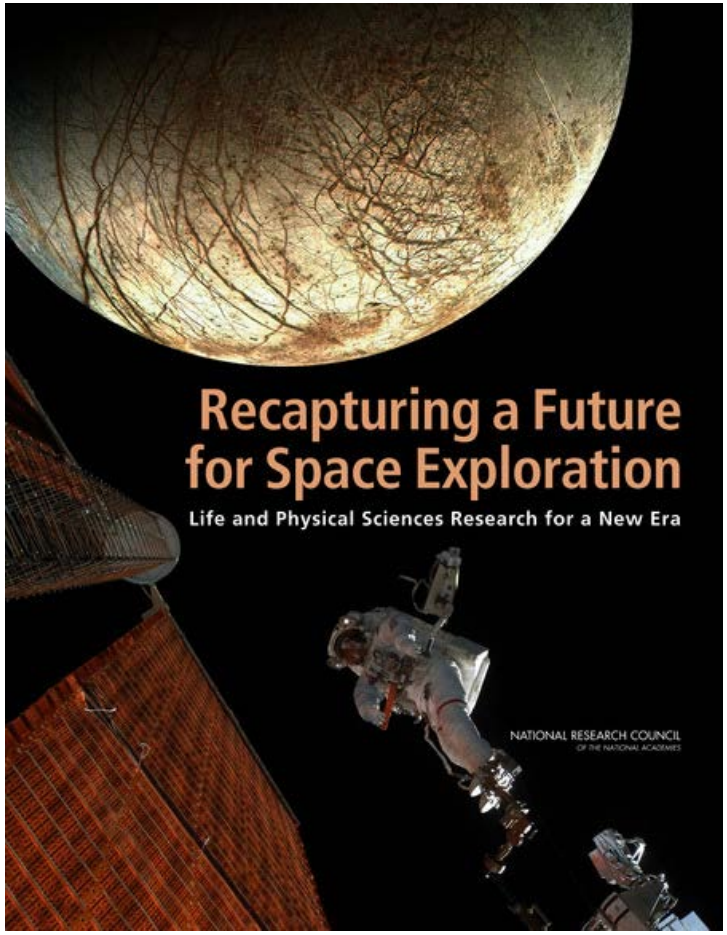
https://www.genome.gov/27565109/the-cost-of-sequencing-a-human-genome/

Oxford Nanopore MinION Gene Sequencer



Sample Preparation Module

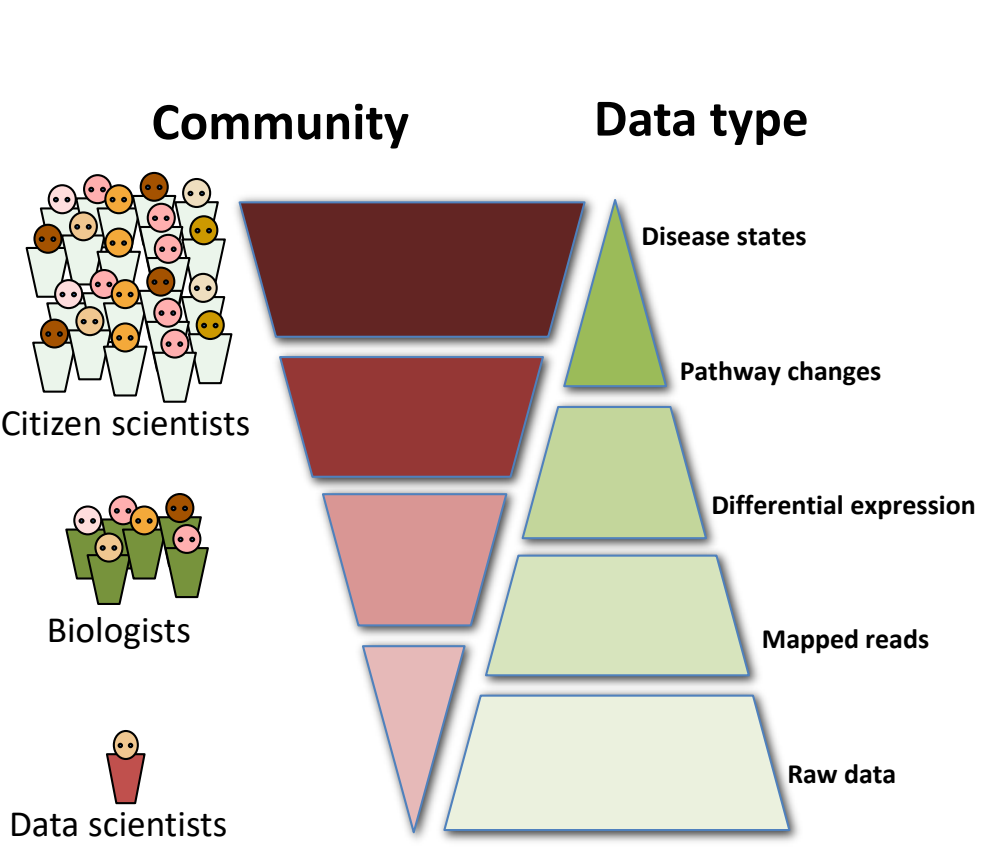
2011 NRC Decadal Survey



*"…**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…"*

GeneLab architecture



New Knowledge

Analysis Working Groups

Galaxy (toolshed)

GenomeSpace (collaborative platform)

GeneLab (core database)

Open Data

EBI PRIDE (proteomics)

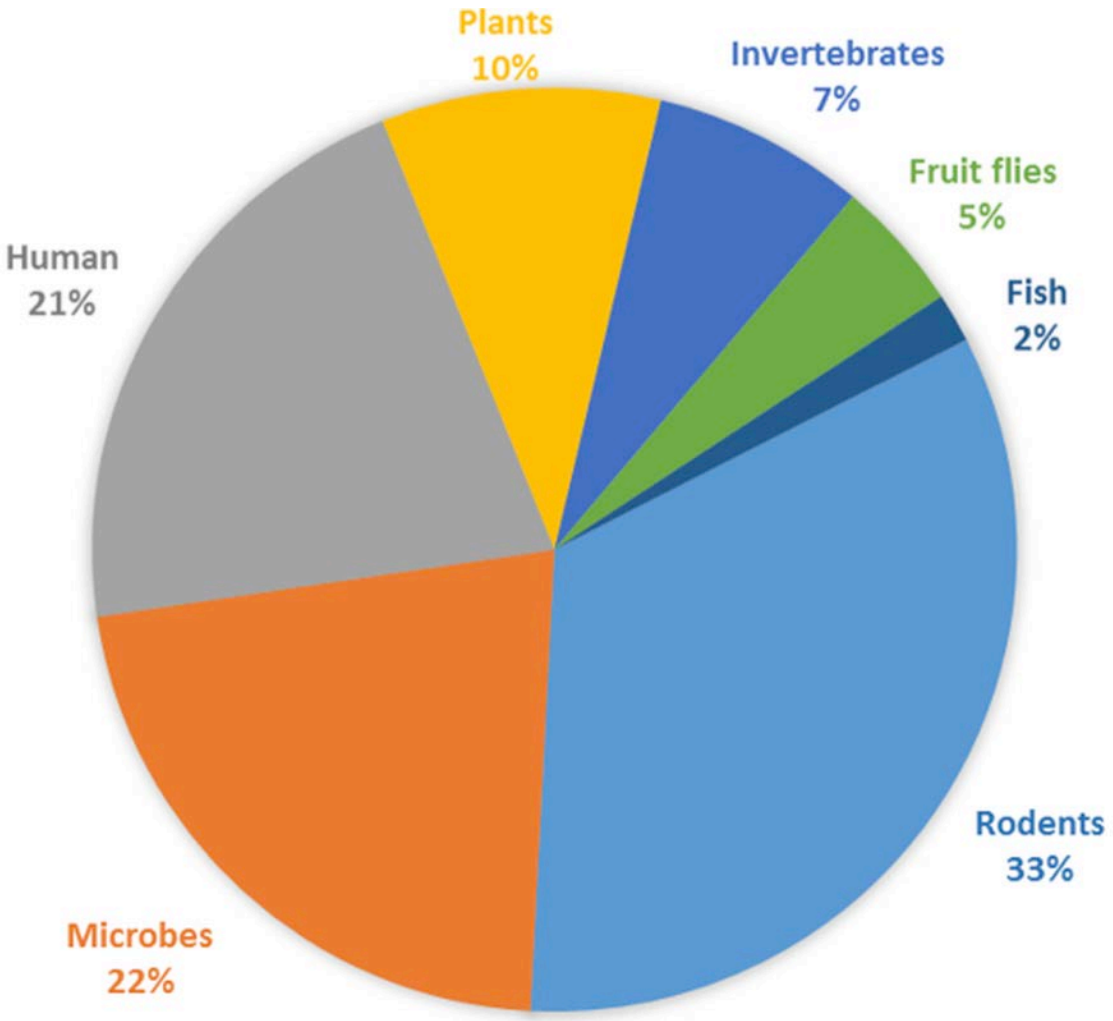
ANL MG-RAST (metagenomics)

NCBI GEO (microarray)

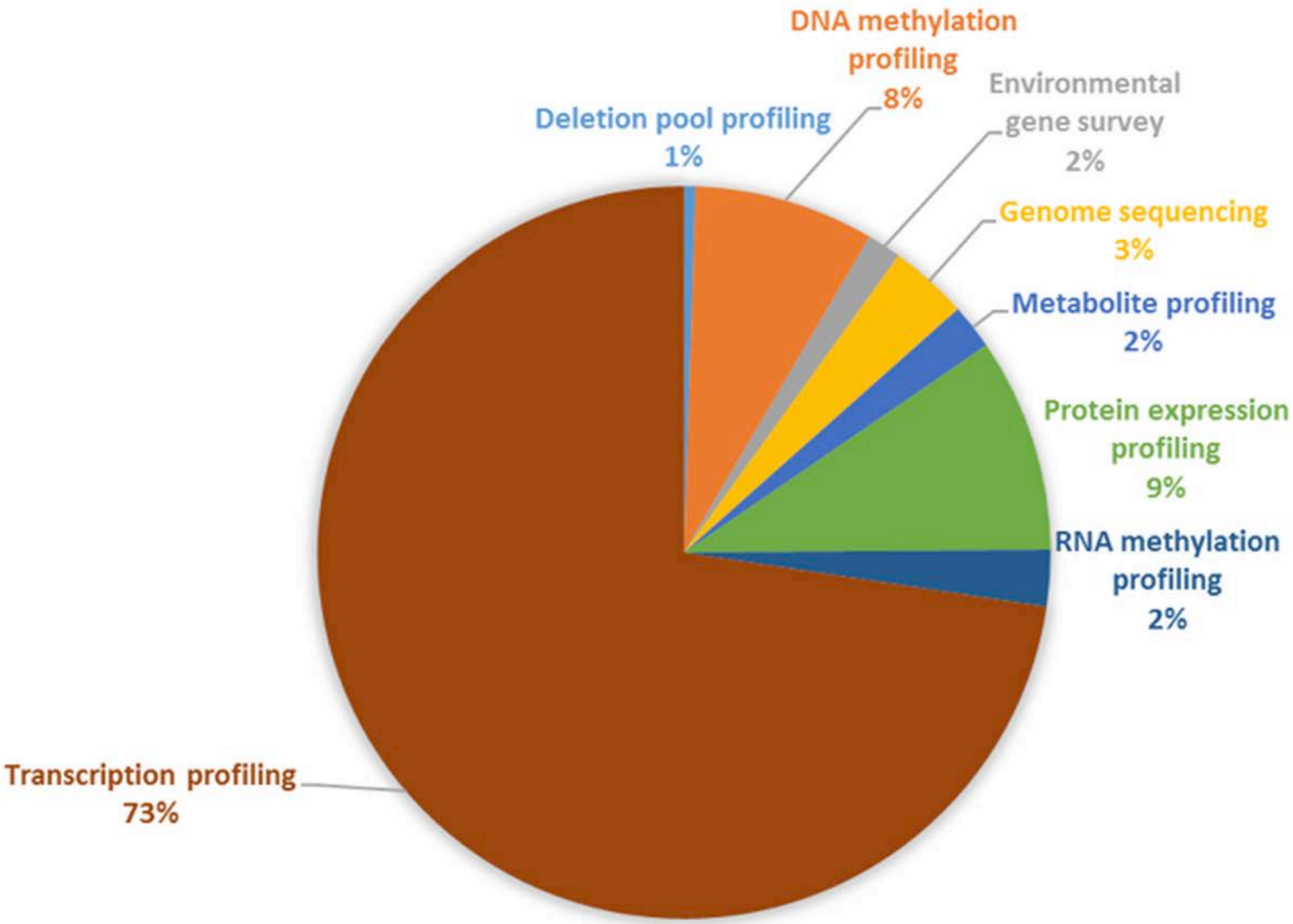
Missions/Experiments

GeneLab Sample Processing

GeneLab Database: 172 data sets

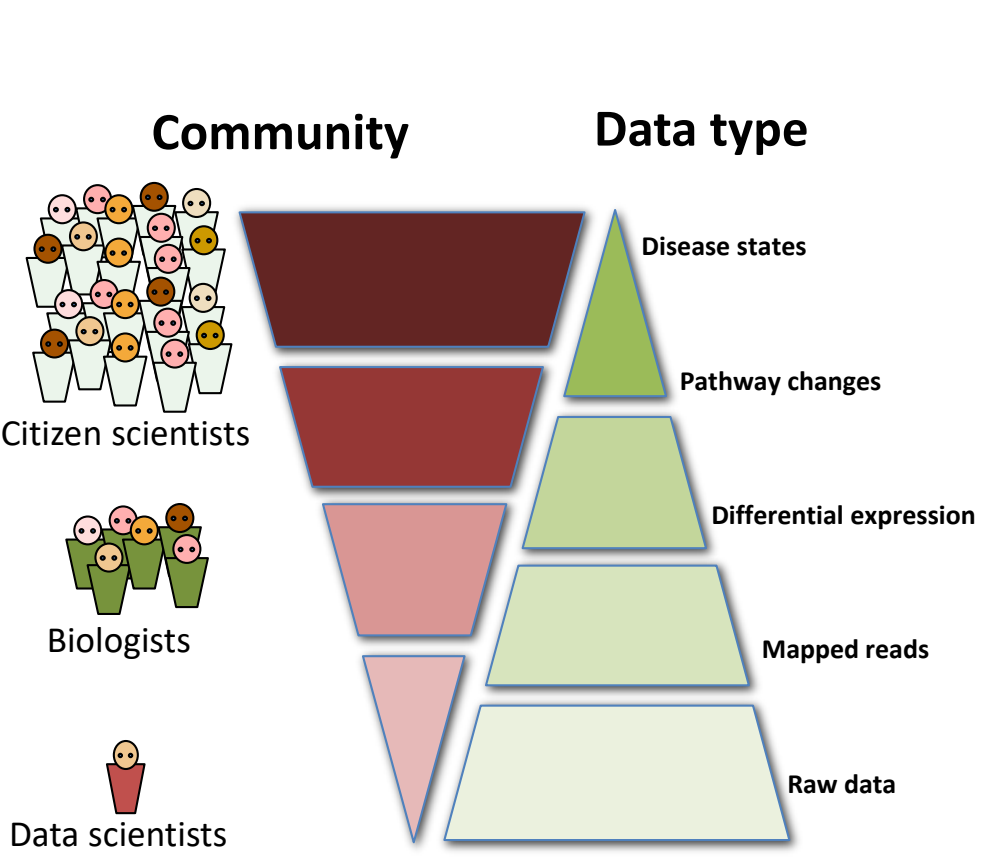


GeneLab Database: 172 data sets



DEMO

GeneLab architecture



New Knowledge

Analysis Working Groups

Galaxy (toolshed)

GenomeSpace (collaborative platform)

GeneLab (core database)

Open Data

EBI PRIDE (proteomics)

ANL MG-RAST (metagenomics)

NCBI GEO (microarray)

Missions/Experiments

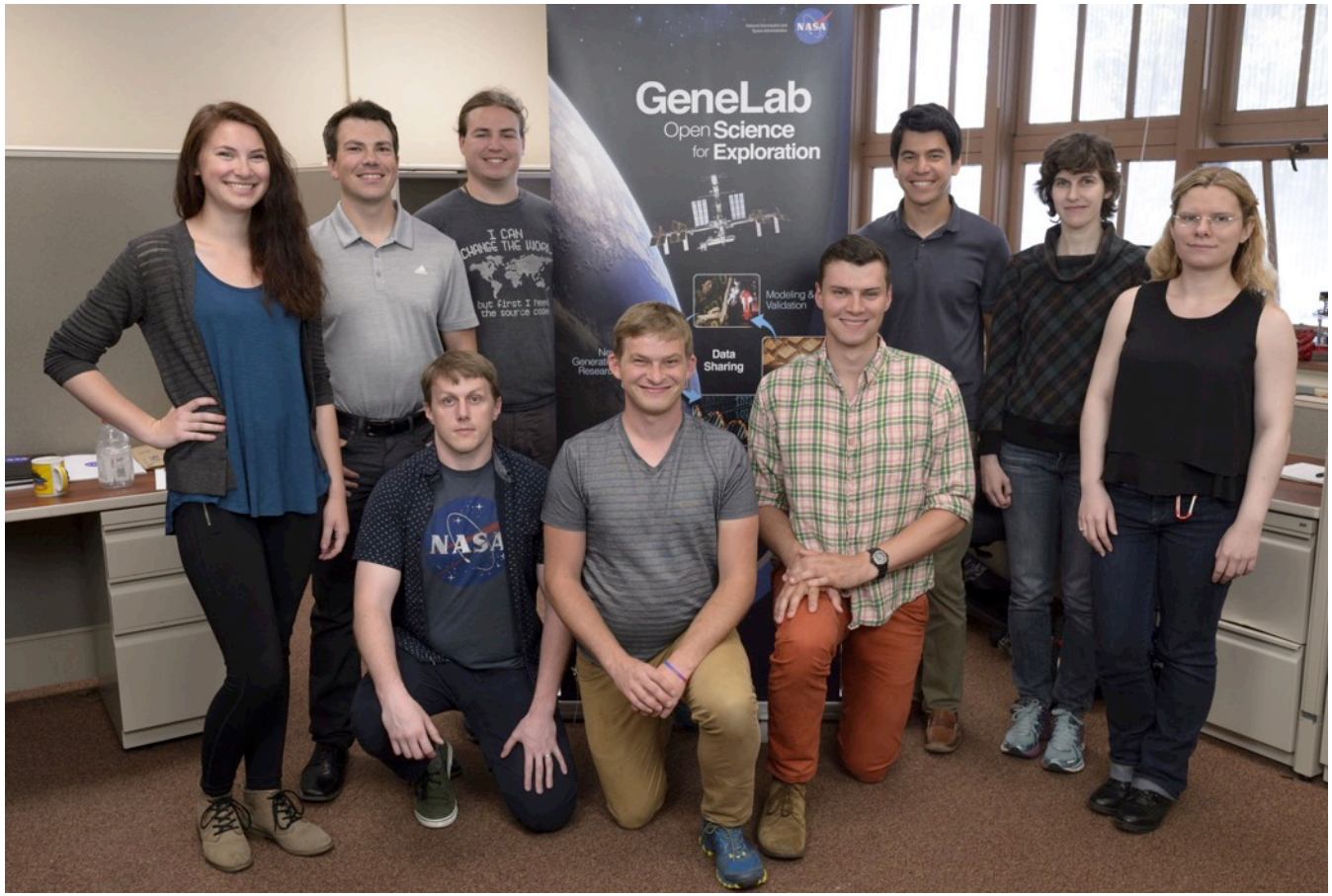
GeneLab Sample Processing

GeneLab Analysis Working Groups

- **~60 individuals**
- **4 Groups: Plants, Microbes, Animals, Multi-omics**
- **Monthly meetings**
- **Deliverables:**
  - Consensus pipelines for primary analysis of data (Microarray, RNASeq, Bisulfite sequencing, Proteomics, 16S metagenomics, Whole genome metagenomics)
  - Recommendations for visualization of data

GeneLab Interns

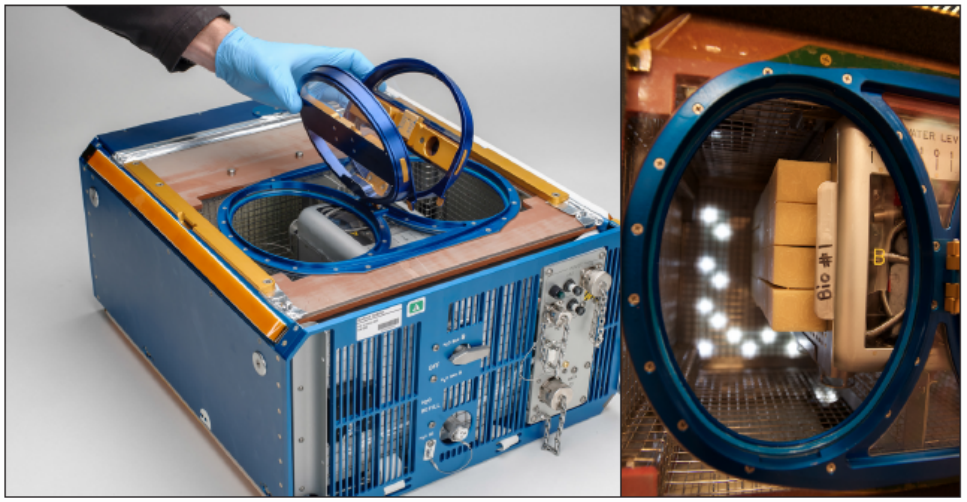
- **10 interns (mostly graduate)**
- **4 Groups: RNA-seq, Microarray, Metagenomics, BS-seq**
- **Developing pipelines and processing data**



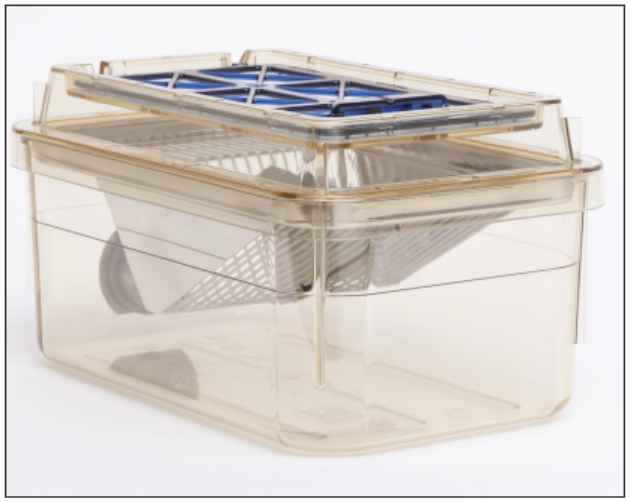
GeneLab Applications

Vivarium vs Rodent Habitat control (AEM) across 5 different rat/mice studies, (no flight samples – CO_2 level matches flight info)

Cage Types

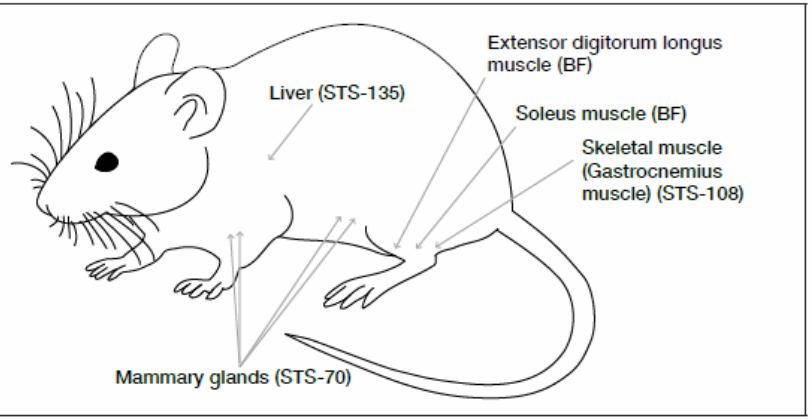


Animal Enclosure Module (AEM)

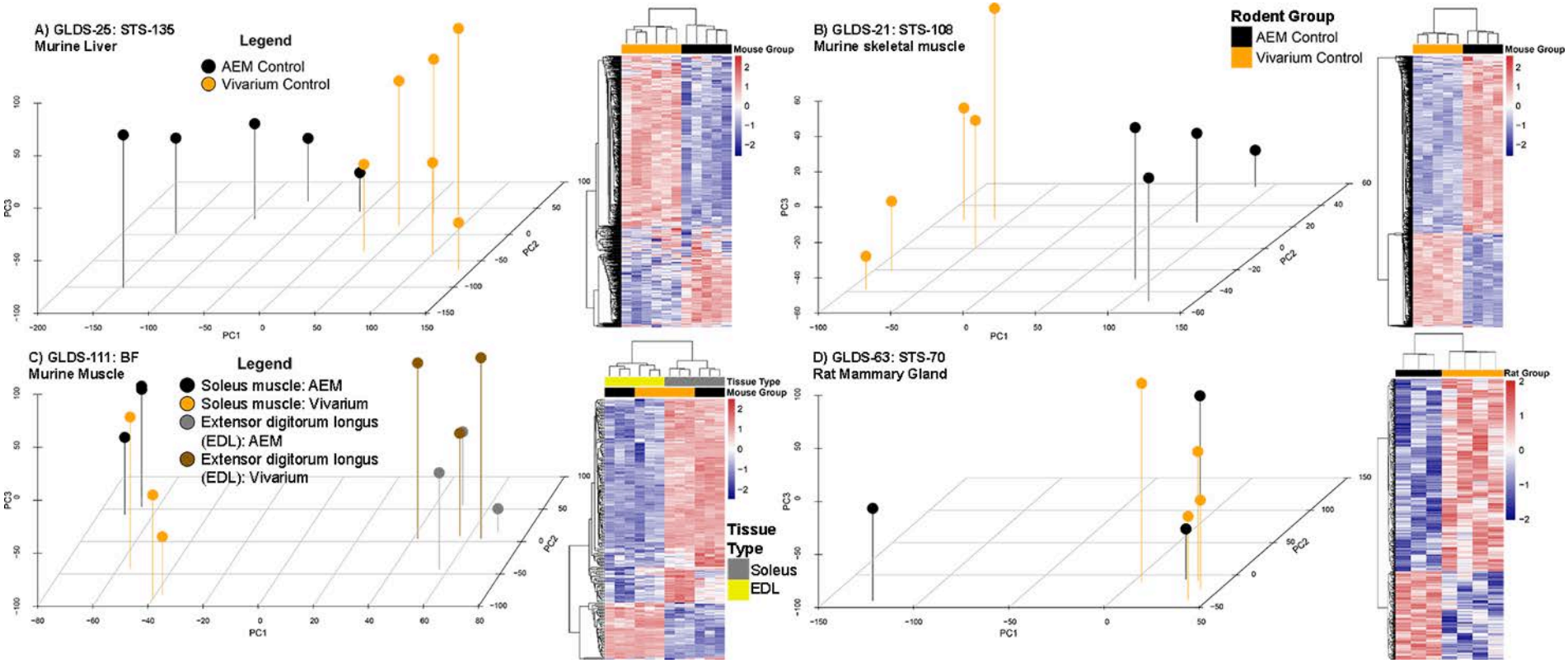


Sample vivarium cage

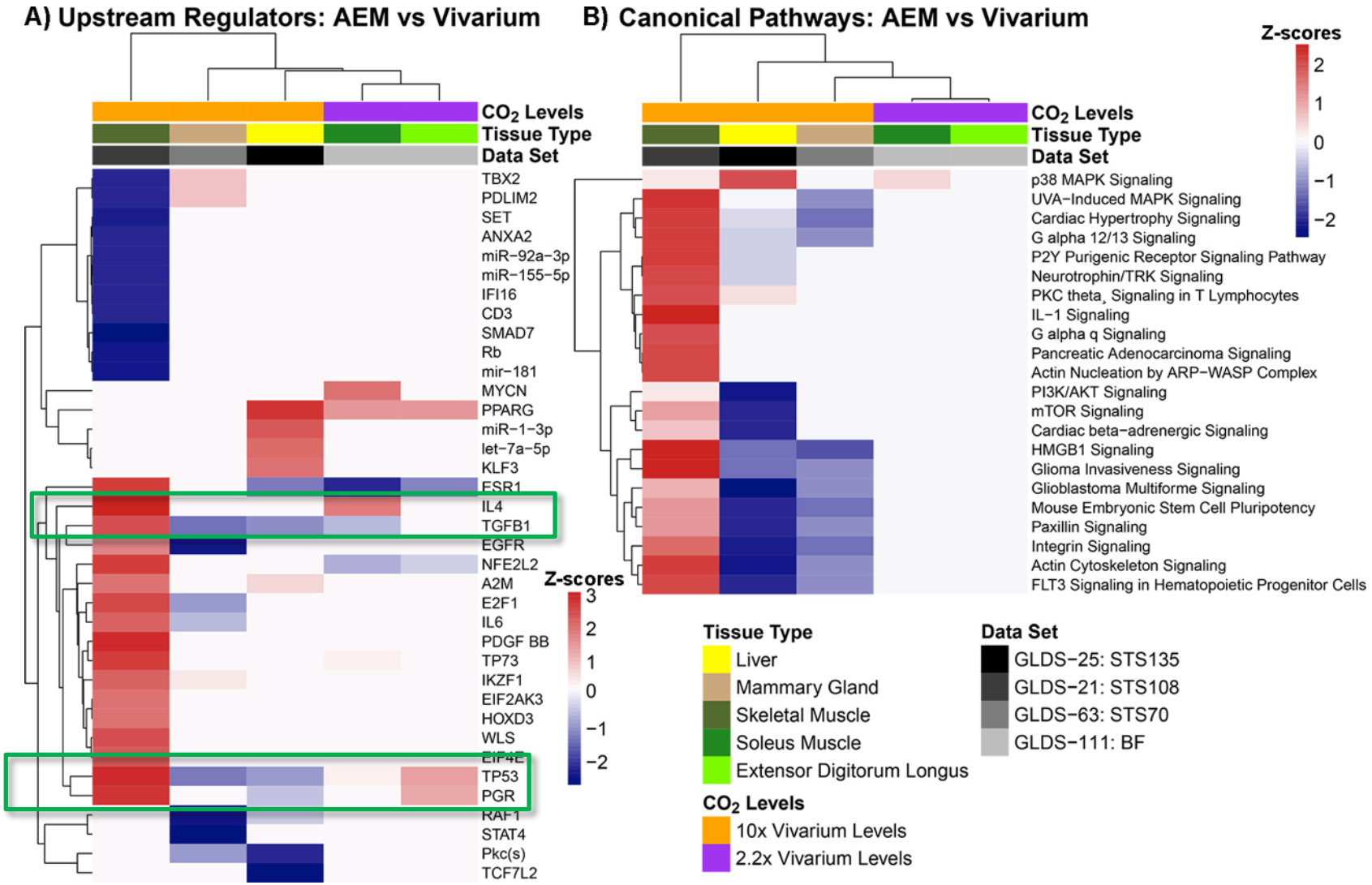
GeneLab study	Mission	Species	CO_2 (ppm)	Tissue type
GLDS-21	STS-108	mouse	~3000	skeletal muscle (gastrocnemius)
GLDS-111	BF	mouse	~600	soleus muscle
GLDS-111	BF	mouse	~600	extensor digitorum longus muscle
GLDS-25	STS-135	mouse	~3000	liver
GLDS-63	STS-70	rat	~3000 (est.)	mammary gland



PCA Plots Suggest Strong Cage Effect



Upstream regulators and canonical pathways show response is tissue specific and highest for high CO_2



GeneLab Sample Processing Laboratory (SPL)

- Expertise:
  - DNA/RNA/protein extraction
  - Animal work

- Develop standards for sample processing (species dependent)
- Responsible for ~50% of GeneLab data by volume