



# GeneLab: “Omics” Data Systems for Spaceflight and Simulated Spaceflight Environment

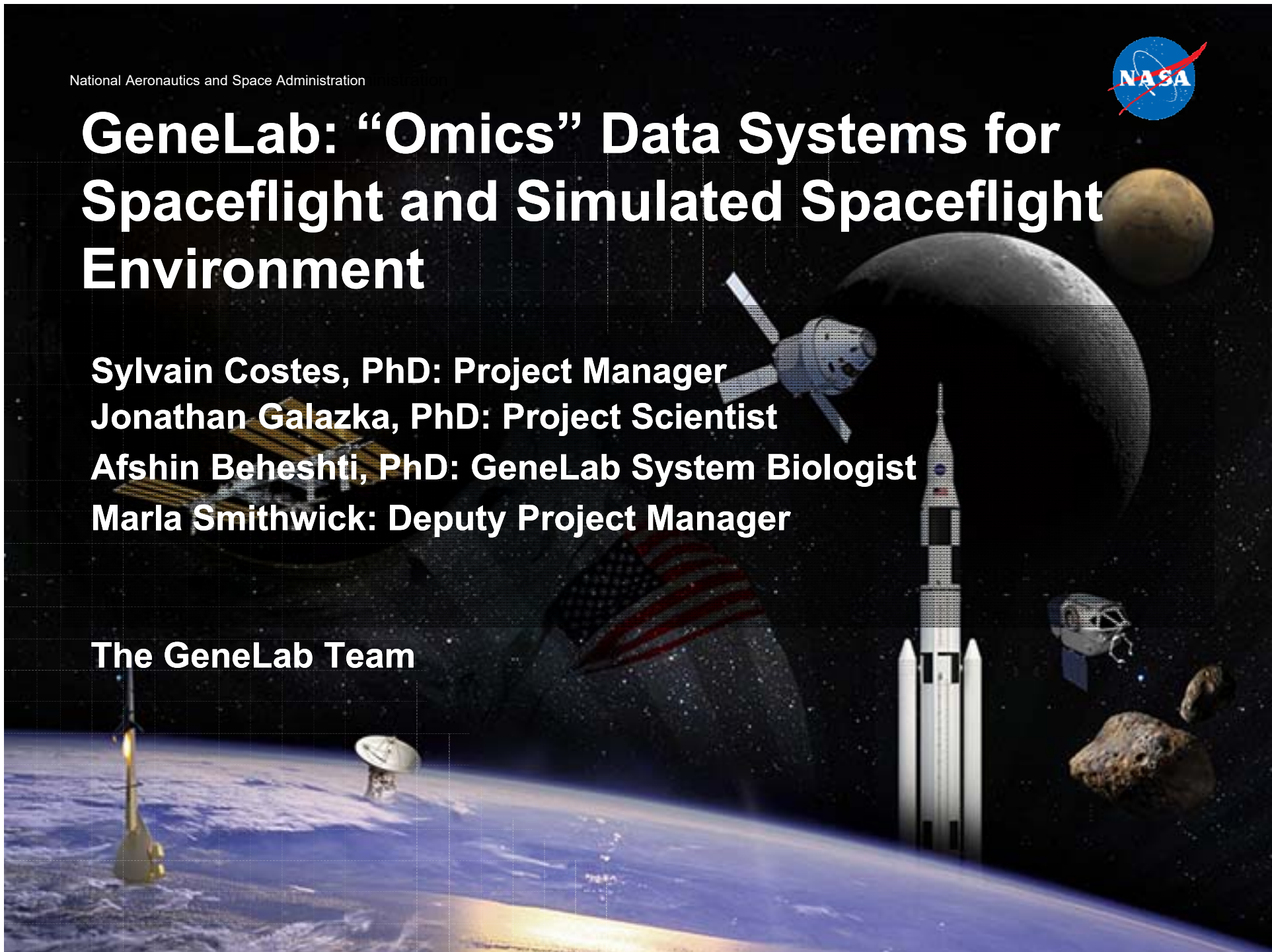
**Sylvain Costes, PhD: Project Manager**

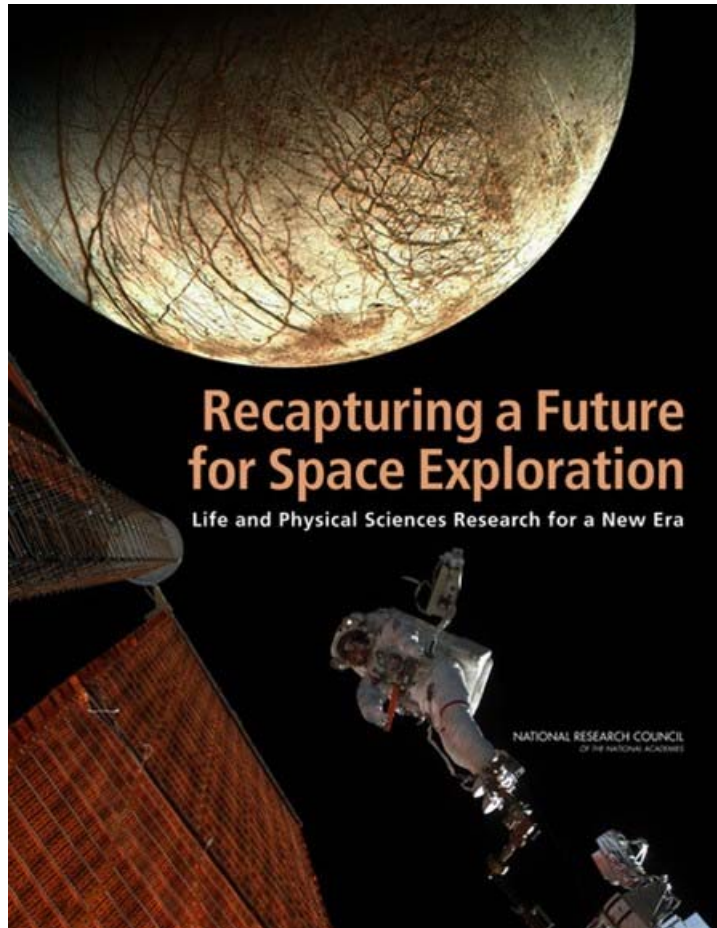
**Jonathan Galazka, PhD: Project Scientist**

**Afshin Beheshti, PhD: GeneLab System Biologist**

**Marla Smithwick: Deputy Project Manager**

**The GeneLab Team**





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

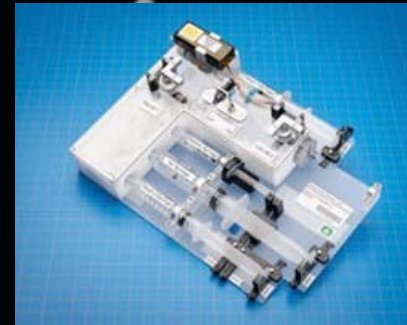
*“...Such techniques generate considerable amounts of **data that can be mined and analyzed** for information by multiple researchers...”*



# Omics Acquisition in Space is Now a Reality



This is truly an exciting time for cellular and molecular biology, omics and biomedicine research on ISS with these amazing additions to the suite of ISS Laboratory capabilities.



**Sample Preparation Module**

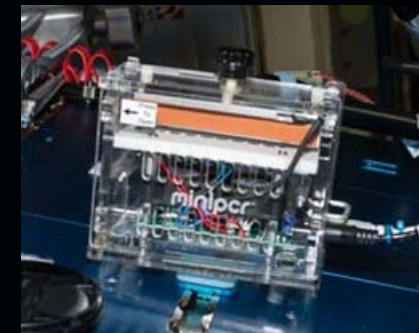


**Oxford Nanopore MinION Gene Sequencer**

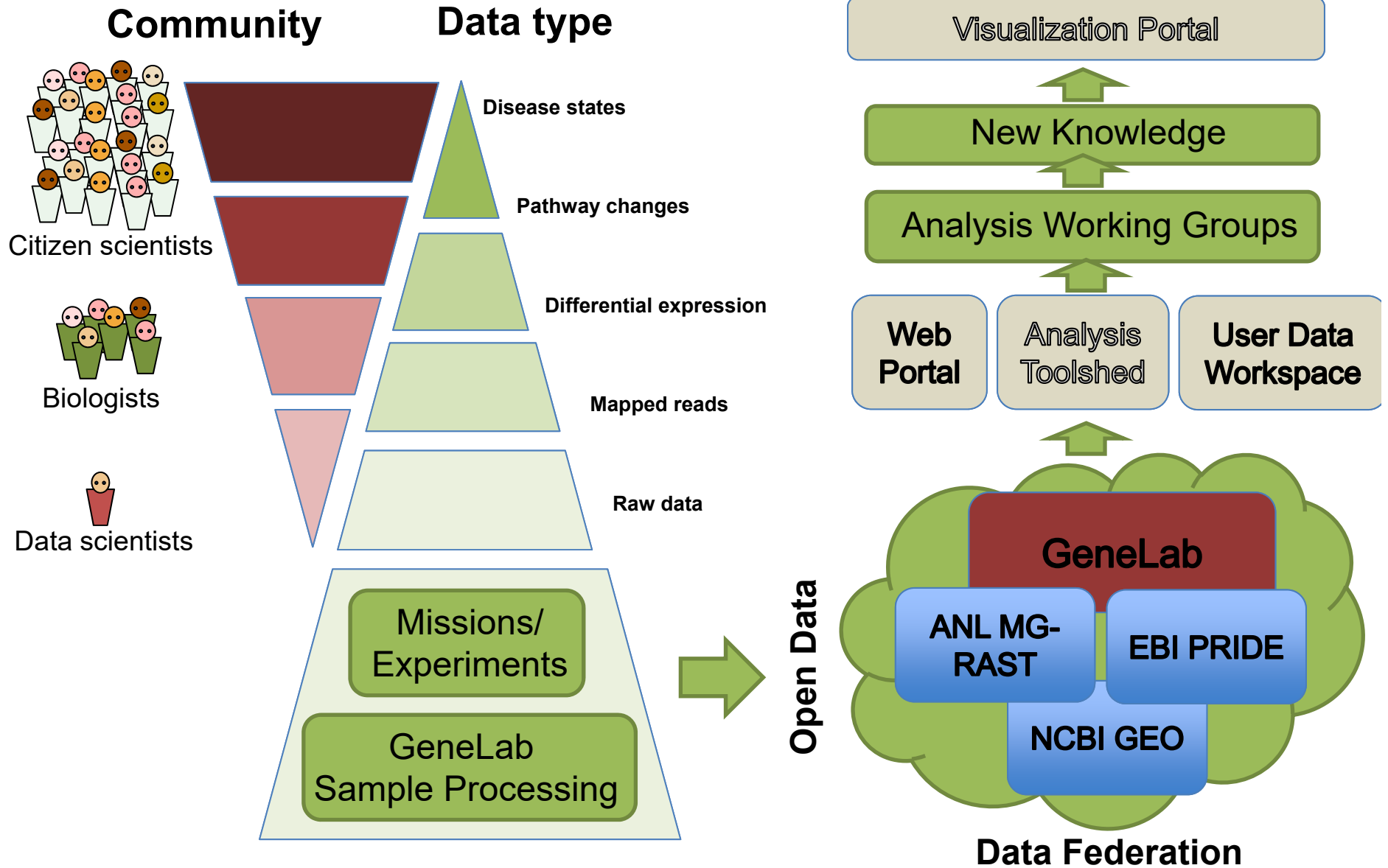
**Cepheid Smart Cycler qRT-PCR**



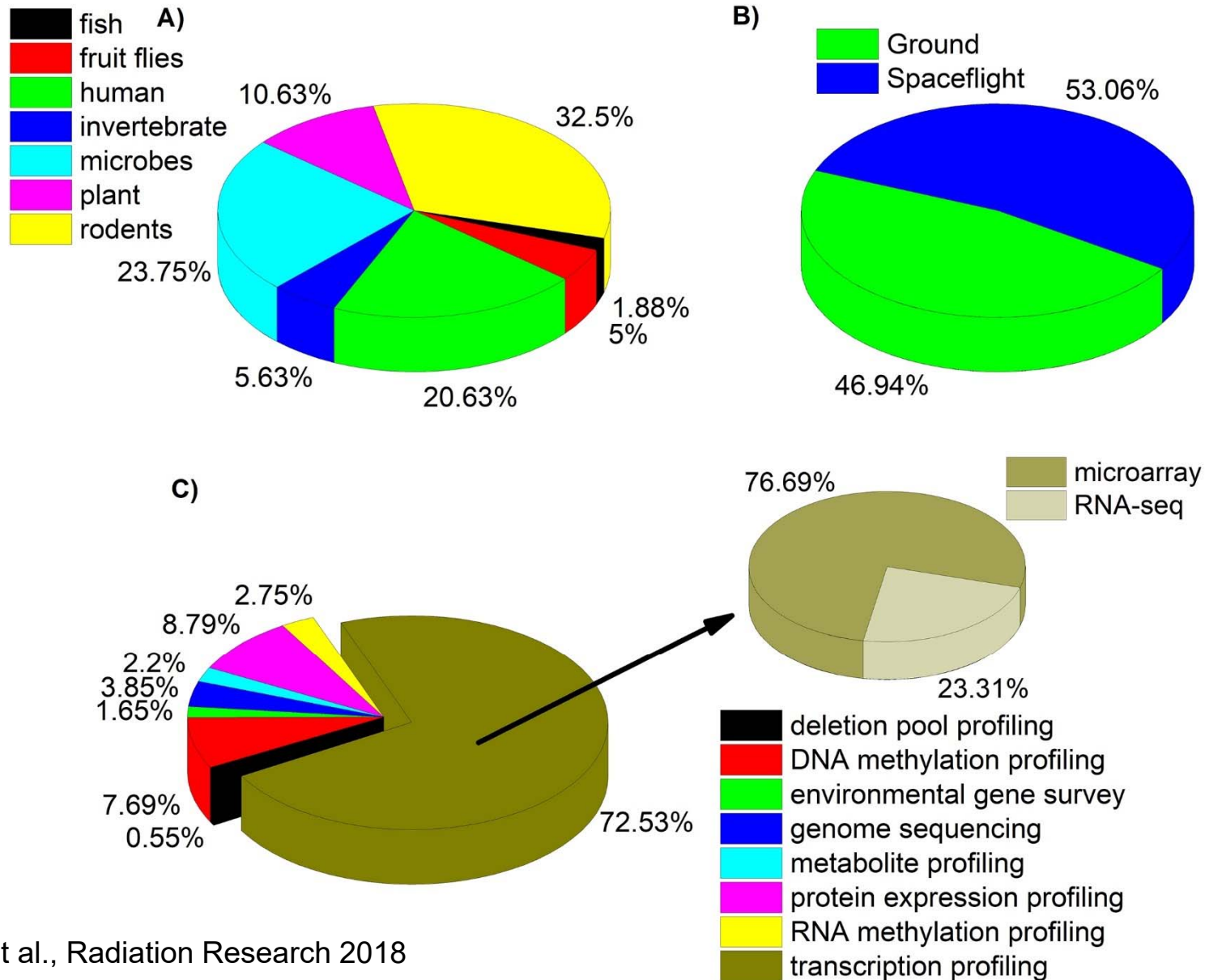
Reaction tube containing lyophilized chemical assay reagent (proprietary)



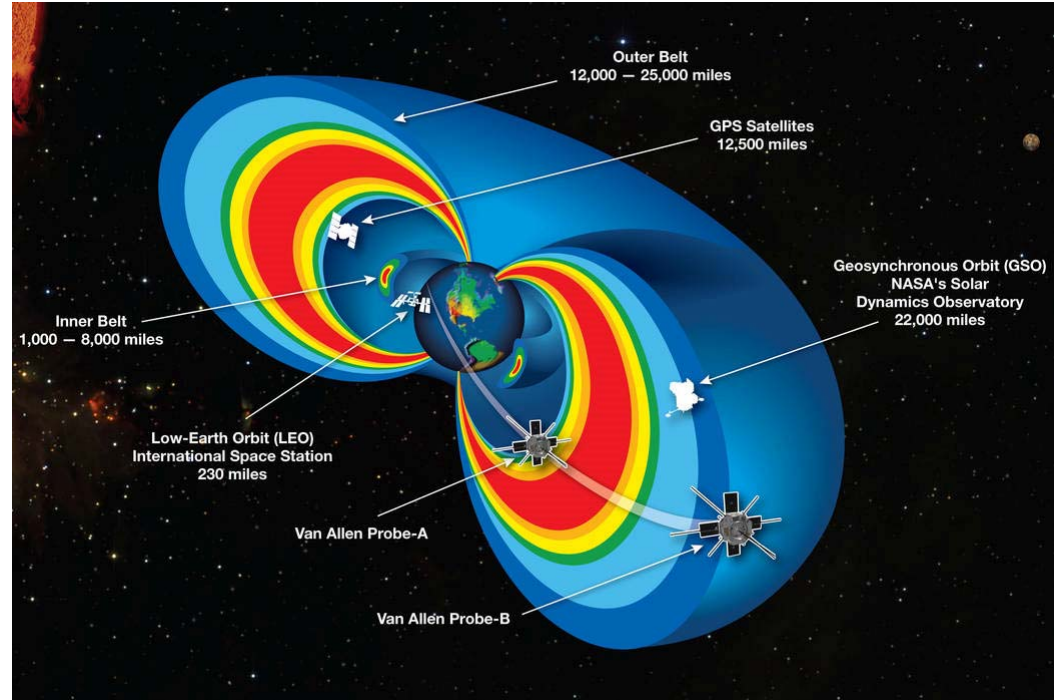
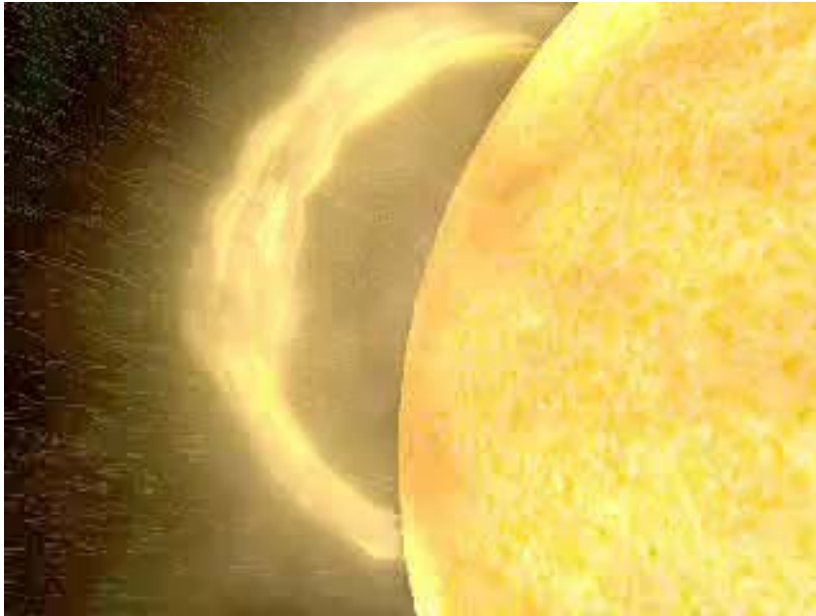
**Mini-PCR**







# Earth's magnetic field protects us from cosmic radiation

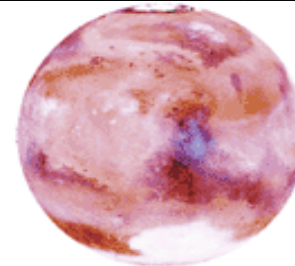


	MILLIREM:
CHEST X-RAY	8 to 50
AVG. YEARLY RADON DOSE	200
U.S. AVG. YEARLY DOSE	350
PET SCAN	1,000
1 YEAR IN KERALA, INDIA	1,300
U.S. NUCLEAR WORKER LIMIT PER YEAR	5,000
APOLLO 14 (9 DAYS)	1,140
SHUTTLE 41-C (18 DAYS)	5,600
SKYLAB 4 (84 DAYS)	17,800
<b>MARS MISSION TOTAL</b>	<b>130,000</b>

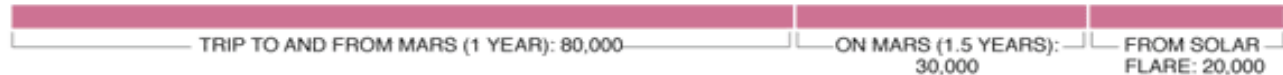
## 2½ Years, 2,600 X-Rays

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

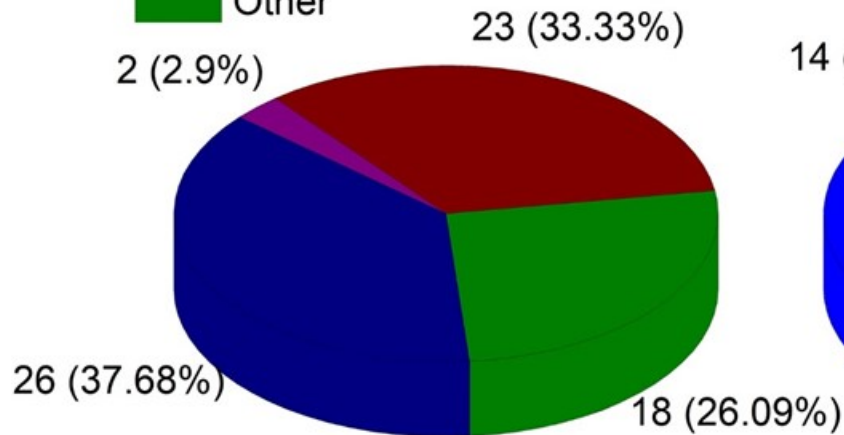
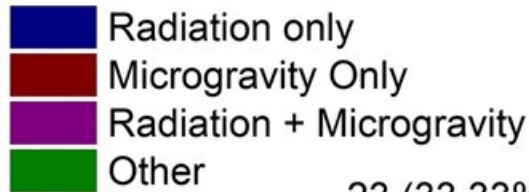


Source: Brookhaven National Laboratory, U.S. Department of Energy

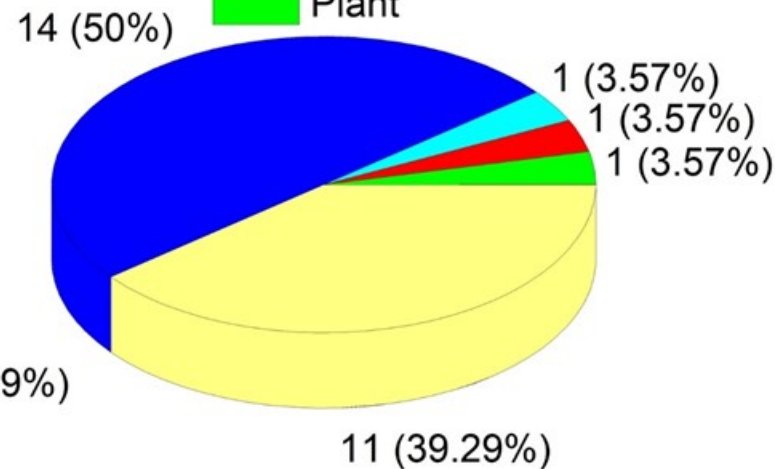
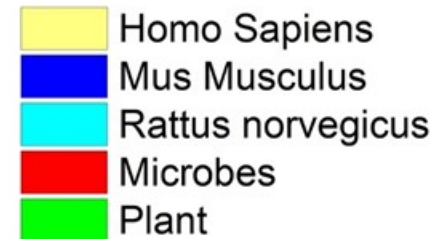
# 69 Ground Data Sets: Radiation and simulated microgravity



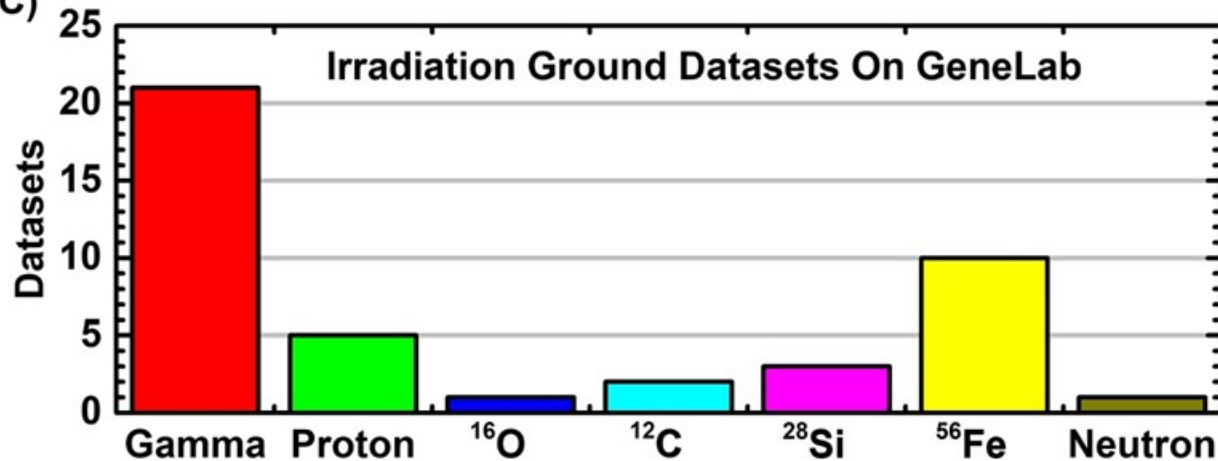
A)



B)

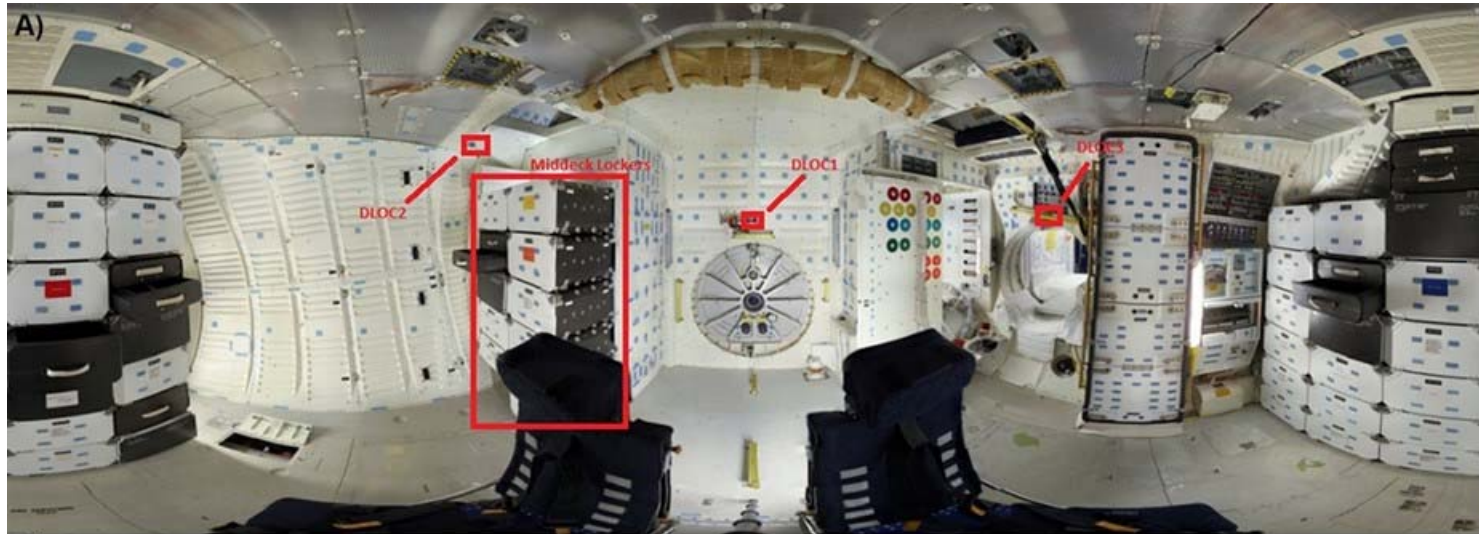


C)

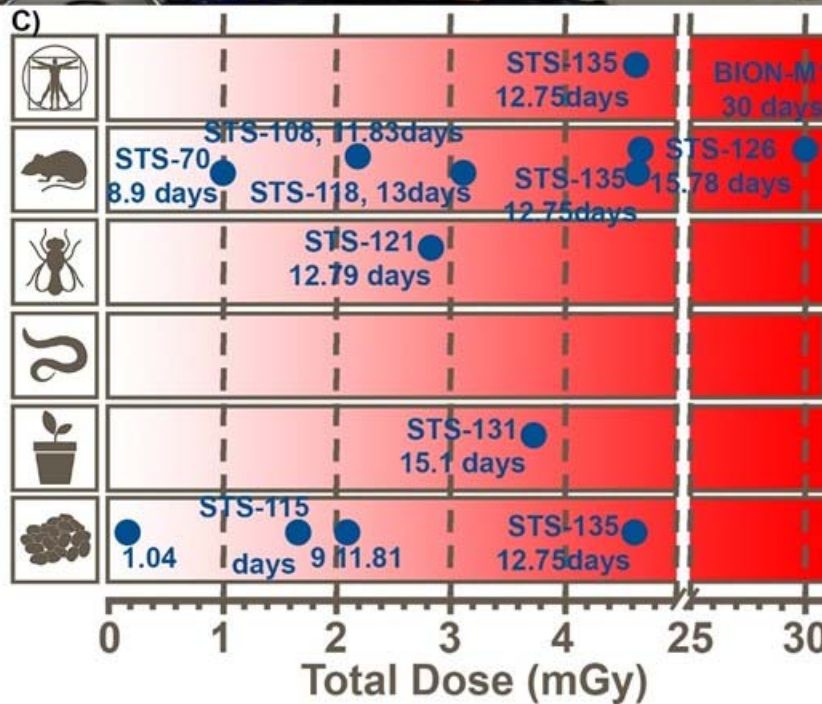
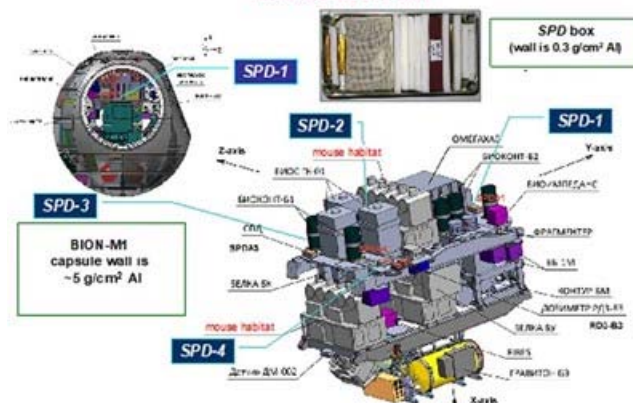




# Radiation Dosimetry for STS samples (ISS to follow)



B) Locations of Radiation Detectors and Animal Holders inside BION-M1





## Data federation/integration with heterogeneous bioinformatics external databases (GEO, PRIDE, MG-RAST)



GLDS-88: Age and Space Irradiation Modulate Tumor Progression: Implications for Carcinogenesis Risk

Source Accession Number [E-GEOD-45606](#)  
Total Data Volume: 31.6 MB

Submitted Date: 28-Mar-2013  
Release Date: 13-Jun-2013

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

### DESCRIPTION

#### Study Description

Age plays a major role in tumor incidence and is an important consideration when modeling the carcinogenesis process or estimating cancer risks. Epidemiological data show that from adolescence through middle age, cancer incidence increases with age. This effect is commonly attributed to a lifetime accumulation of cellular, particularly DNA, damage. However, during middle-age, the incidence begins to decelerate and, for many tumor sites, it actually decreases at sufficiently advanced ages. We investigated if the observed deceleration and potential decrease in incidence could be attributed to a decreased capacity of older hosts to support tumor progression, and whether HZE (high atomic number (Z), high energy (E)) radiation differentially modulates tumor progression in young versus middle-age hosts, issues relevant to estimating carcinogenesis risk for astronauts. Lewis lung carcinoma (LLC) cells were injected into syngeneic mice (143 and 551 days old), which were then subject to whole-body 56Fe irradiation (1GeV/amu). Three findings emerged: 1) among unirradiated animals, substantial inhibition of tumor progression and significantly decreased tumor growth rates were seen for middle-aged mice compared to young mice; 2) whole-body 56Fe irradiation (1GeV/amu) inhibited tumor progression in both young and in middle-aged mice (with greater suppression seen in case of young animals), with little effect on tumor growth rates; and 3) 56Fe irradiation (1GeV/amu) suppressed tumor progression in young mice, to a degree not significantly different than transitioning from young to middle-aged. Thus, 56Fe irradiation (1GeV/amu) acted similar to aging with respect to tumor progression. We further investigated the molecular underpinnings driving the radiation modulation of tumor dynamics in young and middle-aged mice. Through global gene expression analysis, the key players, FASN, AKT1, and the CXCL12/CXCR4 complex, were determined to be contributory. In sum, these findings demonstrate a reduced capacity of middle-aged hosts to support the progression phase of carcinogenesis and identify molecular factors contributory to HZE radiation modulation of tumor progression as a function of age. For genome-wide expression profiling of tumor tissue, Mouse WG-6 BeadArray chips (Illumina, San Diego, CA) were used. Total RNA was amplified with the Ambion Illumina TotalPrep Amplification Kit (Ambion, Austin, TX) and labeled from all replicate biological samples for each condition. The number of tumor sample replicates used from each condition is as follows: 10 samples from young unirradiated mice, 8 samples from young irradiated mice, 7 samples from middle-aged unirradiated mice, 5 samples from middle-aged irradiated mice. Total RNA was isolated and purified using Trizol (Invitrogen) or RNeasy (Qiagen), quantified and qualified using Agilent Bioanalyzer (Agilent) and samples were deemed suitable for amplification and hybridization if they had O.D. 260/280 = 1.7 - 2.1, 26S/18S = 2:1, RIN (RNA integrity number) >7. Total RNA of 500ng per sample was amplified using Ambion TotalPrep (Ambion), and 1.5ug of the product was loaded onto the chips. Following hybridization at 55C, the chips were washed and then scanned using the Illumina iScan (Illumina), and the data were analyzed using GenomeStudio (Illumina). Data were first analyzed for gene expression and then culled for present genes (genes that meet the criteria of detection p-value < 0.05). Expression above background was included in an expressed genes working data set for further analyses. Rank variant normalization was applied to the data before extensive analysis. Differential gene expression analysis was used to compare to the reference group, young unirradiated mice, and genes were then evaluated and validated.

#### Contacts

NAME	ROLE	ORGANIZATION	E-MAIL
Christine E Briggs	submitter	Tufts University School of Medicine	<a href="mailto:cebriggs7135@gmail.com">cebriggs7135@gmail.com</a>
Afshin Beheshti			
Edward Rietman			
Lynn Hlatky	Principal Investigator	Tufts University School of Medicine	
Michael Peluso			
Philip Hahnfeldt			
Rainer K Sachs			

## Federated Search

Home Repository Data Data Mining Tools Submit Data Help Workspa

mouse myostatin x Q

All  GeneLab  NIH GEO  EBI PRIDE  ANL MG-RAST

Search results for: **mouse myostatin** using filter(s):

Sort by Relevance

#### Myostatin inactivation effects on myogenesis in vitro and in vivo

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28986>



Key words: dystrophin, mdx mouse, Duchenne, fibrosis, dystrophy ABSTRACT Stm (MDSC) into myogenic, as opposed to lipofibrogenic, lineages is a promising therapeutic counteracting myostatin, a negative regulator of muscle mass and a pro-lipofibrotic fibrogenic capacity of MDSC from wild...

Organism: *Mus musculus* Accession: [GSE28986](#) PI/Contact: Robert Gelfand Re

#### The transcriptomic signature of myostatin inhibitory influence on the different

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE59674>



GDF8 (myostatin) is a unique cytokine strongly affecting the skeletal muscle phenoty molecular mechanism of myostatin influence on the differentiation of mouse C2C12 m technique. Treatment with exogenous GDF8 strongly affected the growth and devel proliferation and differentiatio...

Organism: *Mus musculus* Accession: [GSE59674](#) PI/Contact: Zofia Wick Relea

#### Development of gene expression signature for defining the cell potency of mu

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE39765>



In order to determine the cell potency, by identification of genes responsible for plur isolated from five week old male wild type(WT), C57Bl6J and another hypertrophied microarray analysis and compared this gene expression to that of a standard mouse and Mstn null mice using an esta...

Organism: *Mus musculus* Accession: [GSE39765](#) PI/Contact: Bipasha Bose Rele

#### Rodent Research-3-CASIS: Mouse liver transcriptomic proteomic and epigen

<https://genelab-data.ndc.nasa.gov/genelab/accession/GLDS-137>



The Rodent Research-3 (RR-3) mission was designed to study the effectiveness of occurs during spaceflight. Myostatin is a protein secreted by myoblasts that inhibits block myostatin cause increases in muscle mass. The RR-3 experiment was spons Advancement of Science in Space and ass...

Organism: *Mus musculus* Factor: Microarray Treatment Assay Type: transcrip





## User Account Mgmt., Access Controls (e.g., Private, Shared, Public Folders)

The screenshot displays the GeneLab web interface. At the top, there is a navigation bar with the GeneLab logo, the text "Open Science for Exploration", and a button labeled "Invite a collaborator". Below this is a menu bar with options: File, Launch, View, Manage, Help, Tools, and GeneLab Data Repository. The main content area shows a file browser view for a folder named "genelab". The browser displays a list of files with columns for "Filename", "Tags", and "Owner". The files listed are GLDS-1 through GLDS-124, all owned by "genelab".

Overlaid on the interface is a "NASA GeneLab OpenID Login" modal. The modal contains the GeneLab logo and the text "Open Science for Exploration". It features a login form with fields for "USERNAME:" and "PASSWORD:", and buttons for "Sign In" and "Cancel". Below the form, there is a link to "Register new NASA GeneLab user" and a link for "Forgot your password?". A disclaimer box at the bottom of the modal states: "This is a US Government system and is for authorized users only. By accessing and using this information system, you acknowledge and consent to the following: You are accessing a U.S. Government information system, which includes: (1) this computer; (2) this computer network; (3) all computers connected to this network; and (4) all devices and storage media attached to this network or to a computer on this network; and (5) cloud and remote information services. This information system is provided for U.S. Government-authorized use only. You have no reasonable expectation of privacy regarding any communication transmitted through or data stored on this information system. At any time, and for any lawful purpose, the U.S. Government may monitor, intercept, search and seize any communication or data transiting, stored on, or traveling to or from this information system. You are NOT authorized to process classified information on this information system. Unauthorized or improper use of this system may result in suspension or loss of access privileges, disciplinary action, and civil and/or criminal penalties."

Filename	Tags	Owner
GLDS-1		genelab
GLDS-10		genelab
GLDS-100		genelab
GLDS-101		genelab
GLDS-102		genelab
GLDS-103		genelab
GLDS-104		genelab
GLDS-105		genelab
GLDS-106		genelab
GLDS-107		genelab
GLDS-108		genelab
GLDS-109		genelab
GLDS-11		genelab
GLDS-110		genelab
GLDS-111		genelab
GLDS-112		genelab
GLDS-113		genelab
GLDS-114		genelab
GLDS-115		genelab
GLDS-116		genelab
GLDS-117		genelab
GLDS-118		genelab
GLDS-119		genelab
GLDS-12		genelab
GLDS-120		genelab
GLDS-121		genelab
GLDS-122		genelab
GLDS-123		genelab
GLDS-124		genelab



# GLD3.0 – Omics Analysis Toolshed



## Barriers to reproducible analysis of omics data:

1. Large files are difficult to move around and process
2. Workflows vary from user to user and details are sometimes poorly documented

## Galaxy platform:

1. Open source, extensible platform for cloud based analysis of omics data
2. Allows any command line tool or script to be run and chained together into workflows
3. Workflows can published, shared and downloaded



Afgan et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Research (2016)



Data Repository Home | Workspace | Analyze Data | **Workflow** | Shared Data | Admin | Help | User

Workflow Canvas | Imported: WF001: Trimmomatic, Bowtie2,picard

**Tools**

search tools

**Inputs**

- Sequencing General
- Epigenomics
- RNA-Seq
- Differential Expression
- Get Data
- Send Data
- Collection Operations
- Lift-Over
- Text Manipulation
- Filter and Sort
- Join, Subtract and Group
- Convert Formats
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Statistics
- Graph/Display Data
- Quantification
- Microbiome Analysis
- Assembly
- Proteomics
- Phylogenetics
- Annotation
- EGSEA easy and efficient ensemble gene set testing
- Pathway / Enrichment Analysis
- Microarray

**Data Manager Tools**

**Workflows**

```

graph LR
    subgraph L002 [Alignment for L002]
        L002_in1[FASTA/Q file #1]
        L002_in2[FASTA/Q file #2]
        L002_in3[Select reference genome]
        L002_out1[output_unaligned_reads_l (fastqsanger)]
        L002_out2[output_aligned_reads_l (fastqsanger)]
        L002_out3[output_aligned_reads_r (fastqsanger)]
        L002_out4[output_unaligned_reads_r (fastqsanger)]
        L002_out5[output (bam)]
        L002_out6[output_sam (sam)]
        L002_out7[mapping_stats (txt)]
    end

    subgraph L001 [Alignment for L001]
        L001_in1[FASTA/Q file #1]
        L001_in2[FASTA/Q file #2]
        L001_in3[Select reference genome]
        L001_out1[output_unaligned_reads_l (fastqsanger)]
        L001_out2[output_aligned_reads_l (fastqsanger)]
        L001_out3[output_aligned_reads_r (fastqsanger)]
        L001_out4[output_unaligned_reads_r (fastqsanger)]
        L001_out5[output (bam)]
        L001_out6[output_sam (sam)]
        L001_out7[mapping_stats (txt)]
    end

    subgraph L003 [Alignment for L003]
        L003_in1[FASTA/Q file #1]
        L003_in2[FASTA/Q file #2]
        L003_in3[Select reference genome]
        L003_out1[output_unaligned_reads_l (fastqsanger)]
        L003_out2[output_aligned_reads_l (fastqsanger)]
        L003_out3[output_aligned_reads_r (fastqsanger)]
        L003_out4[output_unaligned_reads_r (fastqsanger)]
        L003_out5[output (bam)]
        L003_out6[output_sam (sam)]
        L003_out7[mapping_stats (txt)]
    end

    L002_out5 --> MSF
    L001_out5 --> MSF
    L003_out5 --> MSF

    subgraph MSF [MergeSamFiles]
        MSF_in1[Select SAM/BAM dataset or dataset collection]
        MSF_out1[outFile (bam)]
    end
    
```

**Details**

**Edit Workflow Attributes**

**Name:**  
Imported: WF001: Trimmomatic, Bowtie2,picard

**Tags:**

Apply tags to make it easy to search for and find items with the same tag.

**Annotation / Notes:**  
This workflow takes technical replications from Lanes 001, 002, 003 and 007, trims them using Trimmomatic, aligns them using Bowtie, and merges the output alignments using picard mergeSamFiles. Add an annotation or notes to a workflow; annotations are available when a workflow is viewed.



**Total AWG Members: 114**

**AWG Members Per Group:**

Animal	47
Multi-Omics/System Biology	33
Plants	24
Microbes	21

*\*Some members are in multiple groups*



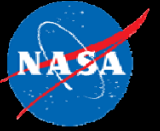
**Annual Workshop (April 2018)**



**2018 Summer Internship:  
Generate all higher order data**

- **Monthly meetings + “Homework”**
- **Deliverables:**
  - Consensus pipelines for primary analysis of data (Microarray, RNASeq, Bisulfite sequencing, Proteomics, 16S metagenomics, Whole genome metagenomics)
  - Recommendations for visualization of data





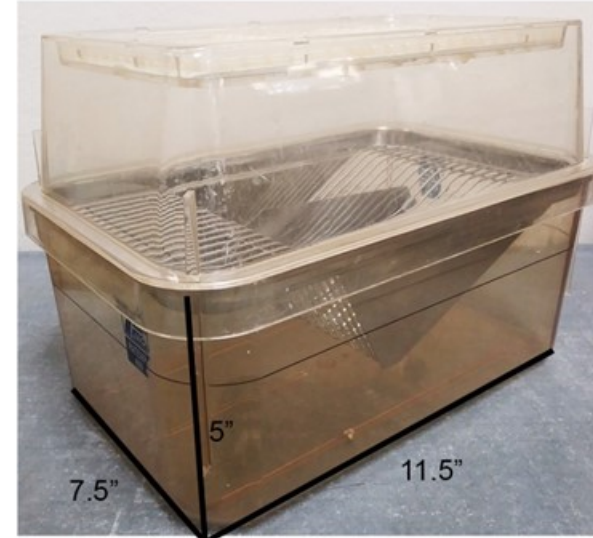
# Cage Effects with rodent experiments: Carbon Dioxide as an Environmental Stressor in Spaceflight

Beheshti A, Cekanaviciute E, Smith DJ, Costes SV. Global transcriptomic analysis suggests carbon dioxide as an environmental stressor in spaceflight: A systems biology GeneLab case study. *Sci Rep.* 2018;8(1):4191. doi: 10.1038/s41598-018-22613-1. PubMed PMID: 29520055; PMCID: PMC5843582.

## A) Cage Types



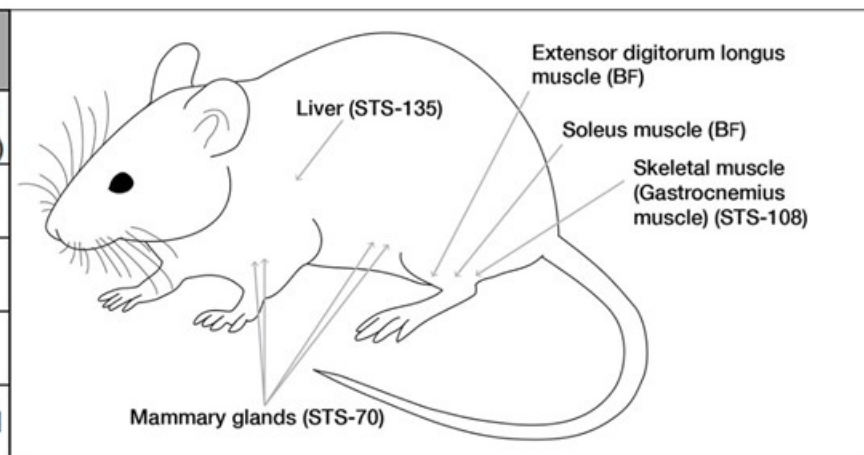
Animal Enclosure Module (AEM)

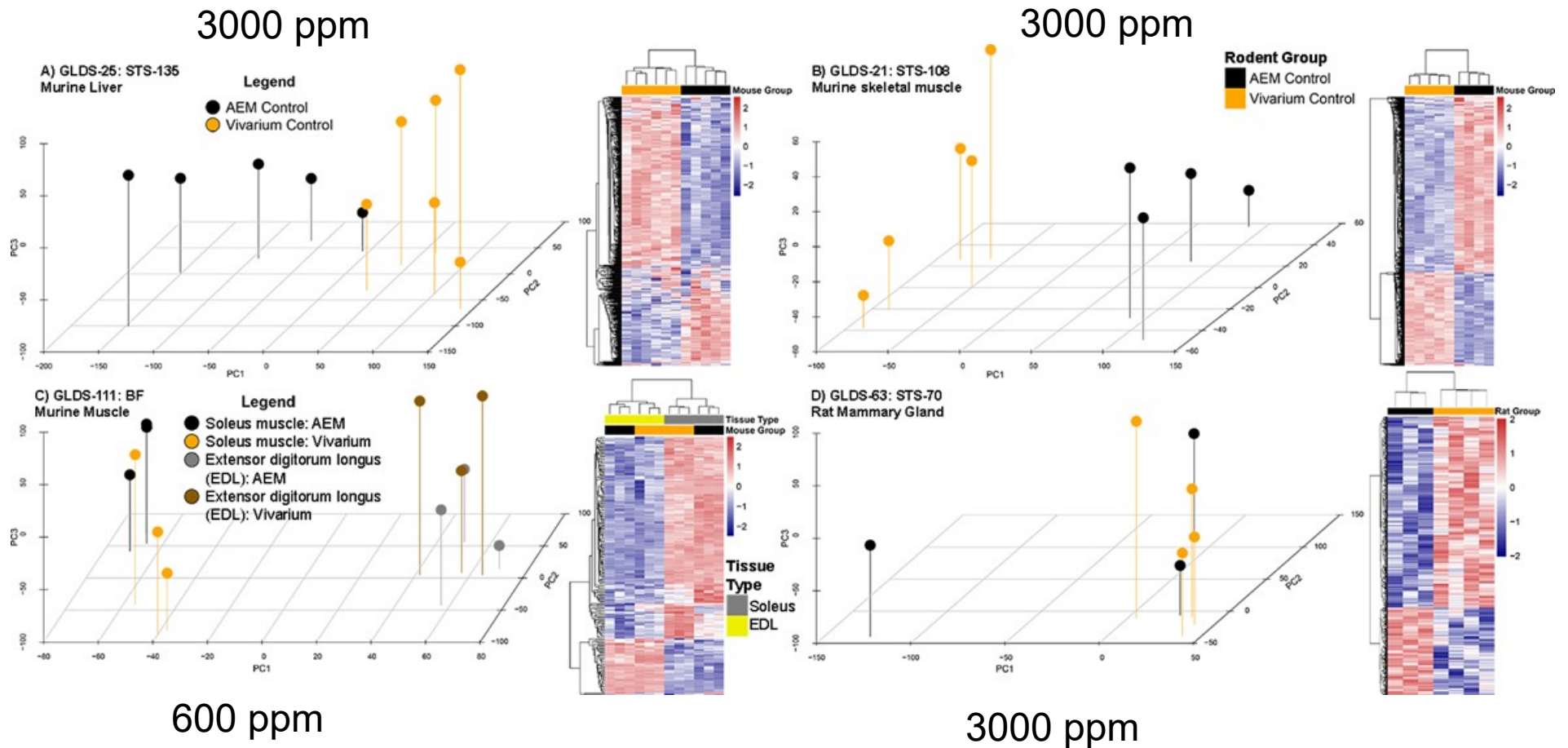


Sample vivarium cage

## B)

GeneLab Study	Mission	Species	CO <sub>2</sub> (ppm)	Duration (days)	Tissue Type
GLDS-21	STS-108	mouse	~3000	11.8	skeletal muscle (gastrocnemius)
GLDS-111	BF	mouse	~600	30	soleus muscle
GLDS-111	BF	mouse	~600	30	extensor digitorum
GLDS-25	STS-135	mouse	~3000	13	liver
GLDS-63	STS-70	rat	~3000 (est)	9	mammary gland





AEM = Animal Enclosure Modules (now referred to as Rodent Habitats)  
 Vivarium = normal ground based rodent cages

Beheshti, et al., Scientific Reports, 2018



## A) Venn Diagram of all significant genes

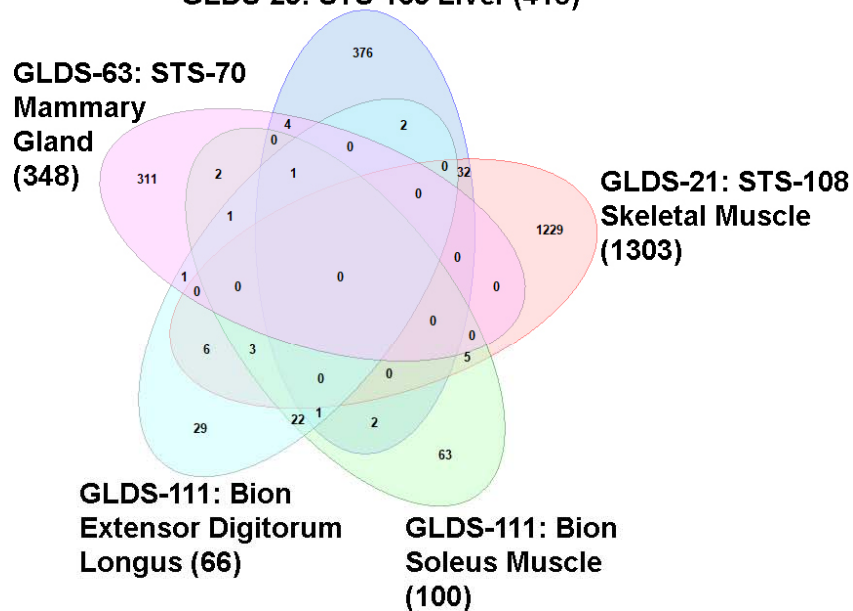
GLDS-25: STS-135 Liver (418)

GLDS-63: STS-70  
Mammary  
Gland  
(348)

GLDS-21: STS-108  
Skeletal Muscle  
(1303)

GLDS-111: Bion  
Extensor Digitorum  
Longus (66)

GLDS-111: Bion  
Soleus Muscle  
(100)



An increase in aldosterone is associated with metabolic syndrome, which is characterized by chronic inflammation; aldosterone secretion can be triggered by hypoxia.

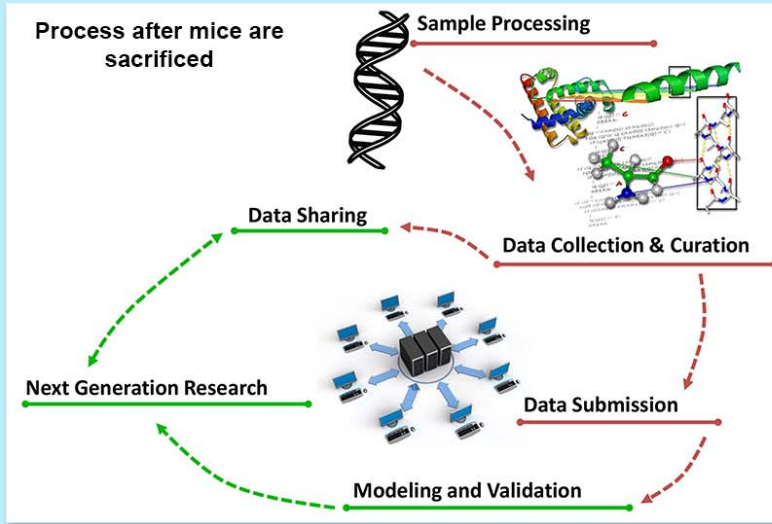
Beheshti, et al., Scientific Reports, 2018



# **Systems Biology analysis reveals biological spaceflight master regulators**

Beheshti, et al., PLOS One, 2018

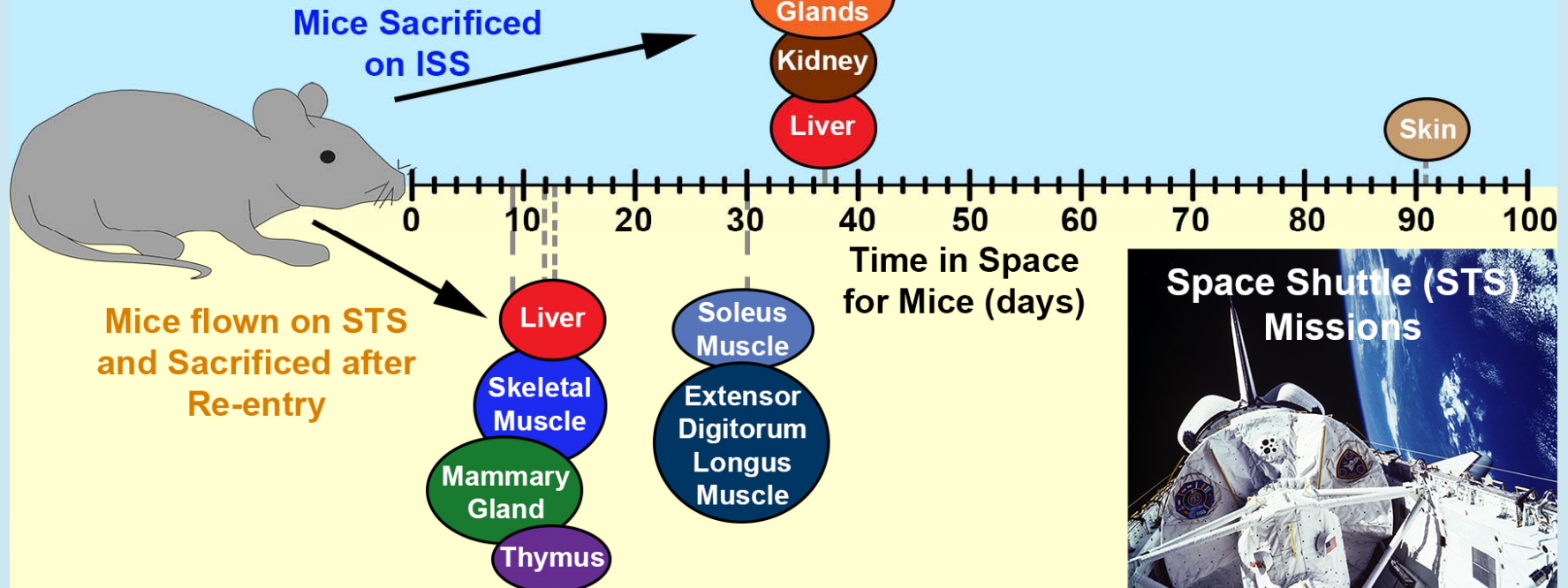
# GeneLab Data Used to Generate Results



- Extensor Digitorum Longus Muscle
- Soleus Muscle
- Gastrocnemius Muscle
- Quadriceps
- Tibialis Anterior Muscle
- Adrenal Glands
- Kidney
- Liver



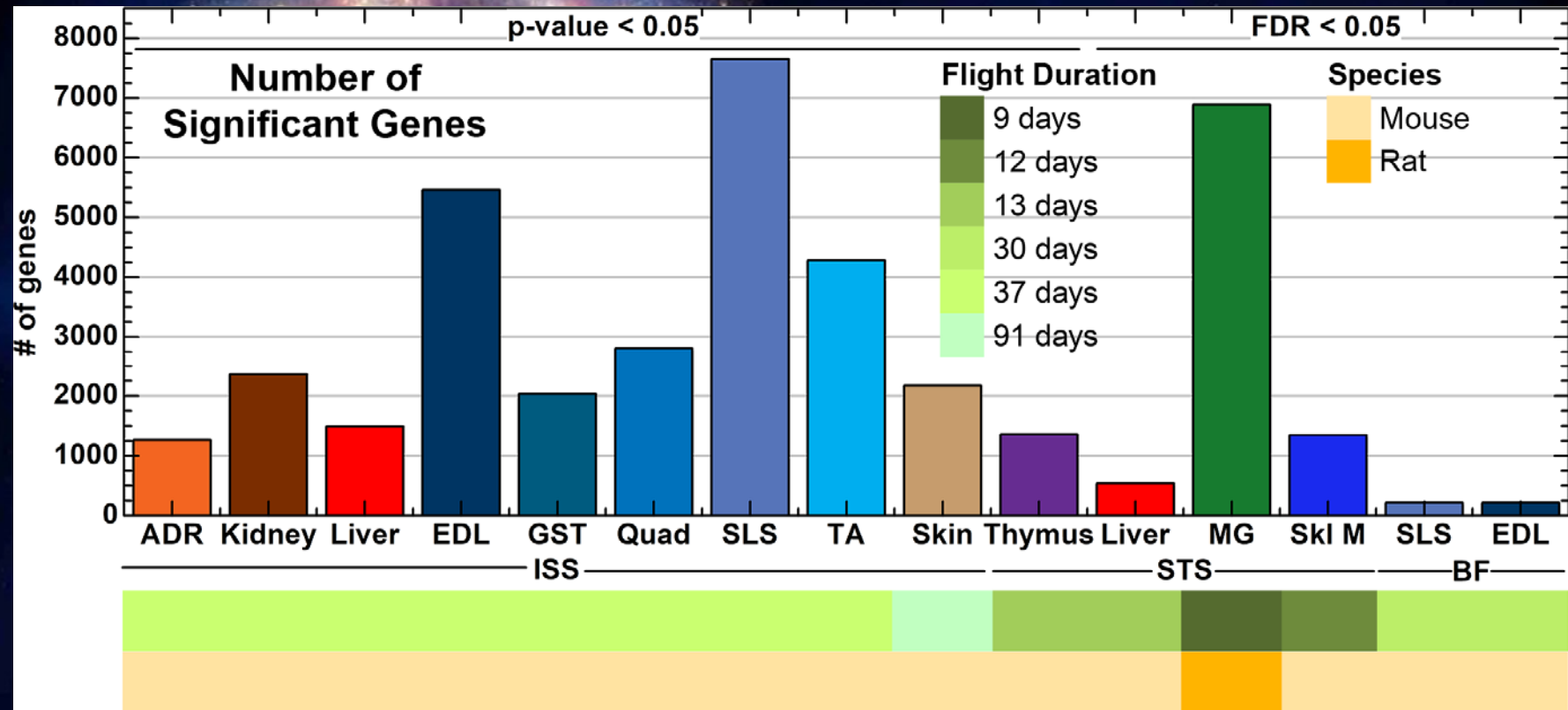
International Space Station (ISS) Missions



Space Shuttle (STS) Missions

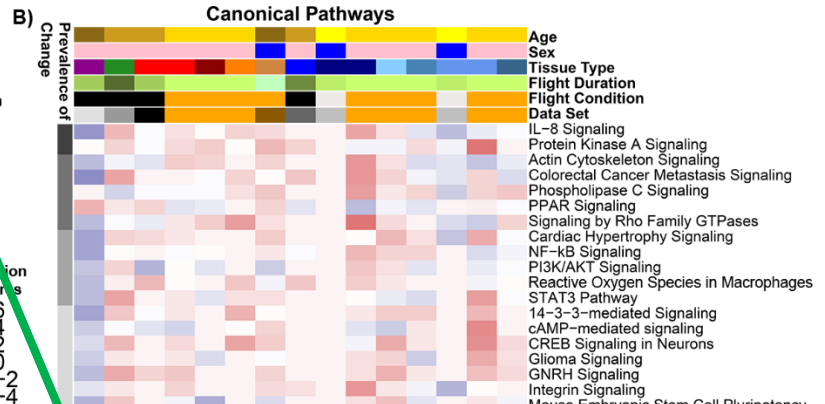
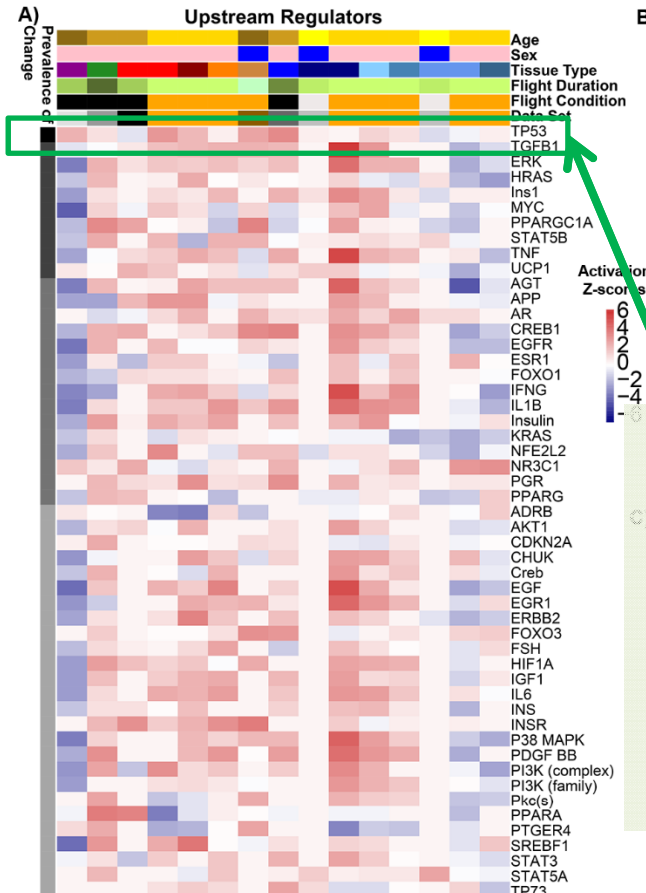


# Number of Significant Genes from Each Dataset

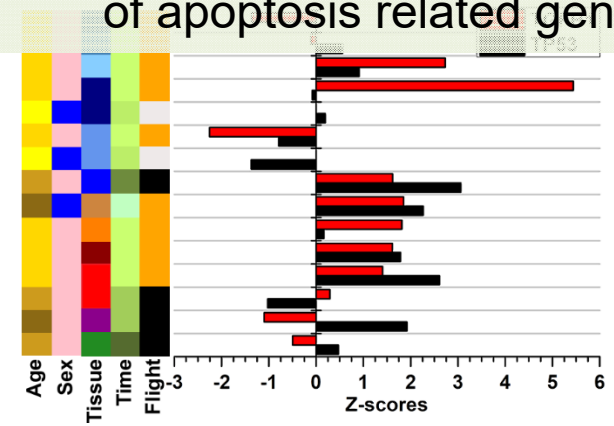
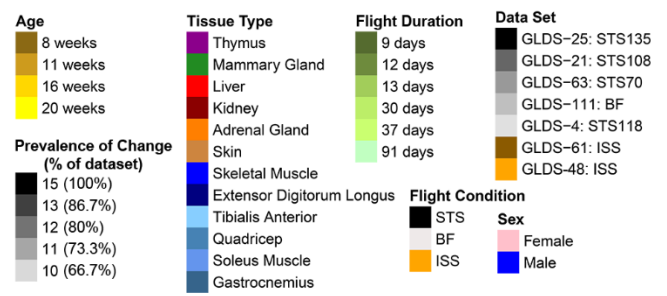


Fold-Change  $\geq$   
| 1.2 |

Pathway/Functional Predictions:  
Ingenuity Pathway Analysis (IPA)  
Gene Set Enrichment Analysis  
(GSEA)



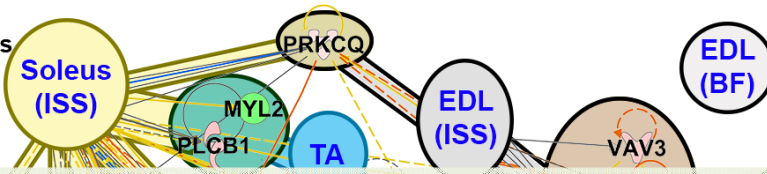
- **p53 found in all tissues**
  - p53 is a transcription factor and in response to genotoxic stress, DNA damage, oncogene activation, and hypoxia, it is recruited to sites in chromatin, thus promoting transcription of apoptosis related genes



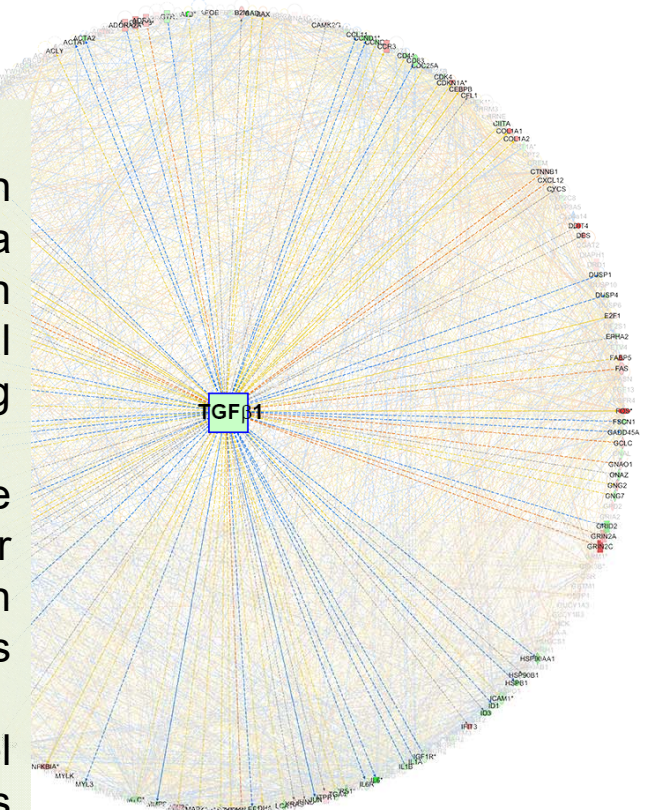




A) Direct Connections for Key Genes for Flight vs AEM



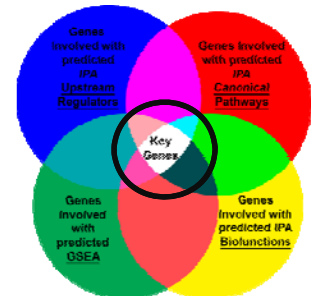
B) Connections Between all Key Genes for all Datasets (Flight vs AEM): Radial Plot with the most Connected Gene in the Middle



• **TGFβ1 found to be central regulator of key genes**

- TGFβ is known to play a context specific role in sustaining tissue homeostasis predominantly via transcriptional regulation of genes involved in differentiation, cell motility, proliferation, cell survival along with regulating immune responses during homeostasis and infection.
- Previous Studies found reduction in gravitational force to diminish TGF-β expression and apoptosis with higher carcinoembryonic antigen expression in 3D human colorectal carcinoma cells, as compared to 3D cultures in unit gravity.
- In another study, differential regulation of blood vessel growth using basic fibroblast growth factor was identified in modeled microgravity with induction early and late apoptosis, extracellular matrix proteins, endothelin-1 and TGFb1 expression

- GLDS-40: ISS, Quad
- GLDS-51: ISS, 9Mn
- GLDS-40: ISS, Tibialis Anterior (TA)
- GLDS-40: ISS, Liver
- GLDS-26: STS 135, Liver
- GLDS-111: BF, Extensor Digitorum Longus (EDL)
- GLDS-4: STS 116, Thymus
- GLDS-33: STS 70, Mammary Gland (MG)

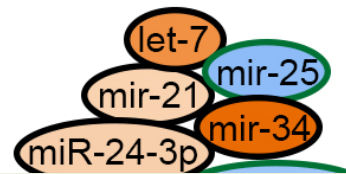




# Predicted miRNAs Involved with Microgravity Effects



## Health Risk Due to miRNAs



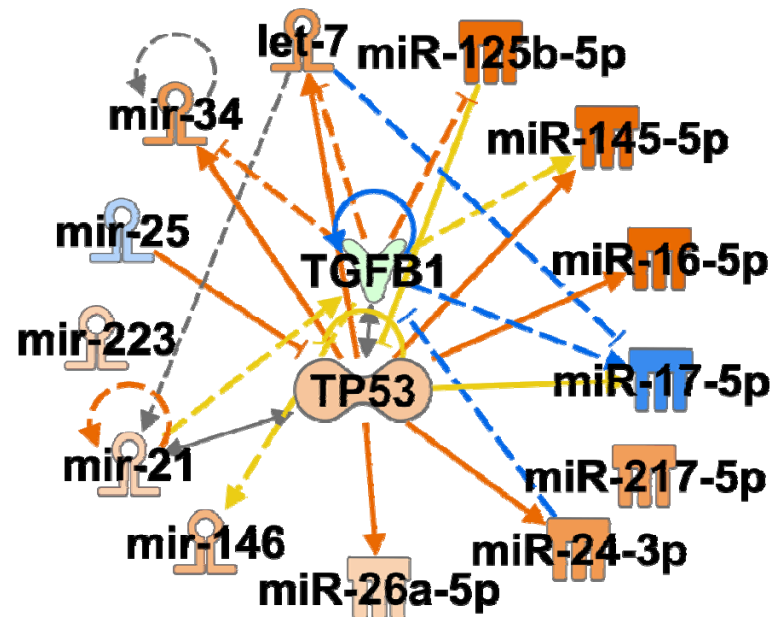
● Predicted Activation

● Predicted Inhibition

○ Negative Impact on Health

- A recent report showed that inactivation of p53 altered TGF- $\beta$  signaling, which ironically displayed both tumor-suppressive and pro-oncogenic functions. p53 functions to integrate crosstalk between Ras/MAPK and TGF- $\beta$  signaling via binding to Smad3, dislocating the Smad3/Smad4 complex formation and differentially regulating subsets of TGF- $\beta$  target genes

*Biological Health  
Risk Increased*





# **Analysis Working Group (AWG) Member related work determines novel systemic biological factors causing damage due to spaceflight**

Work in progress

GeneLab

Open Science for Exploration

# AWG Members Involved



Kathleen Fisch



UNIVERSITY of CALIFORNIA, SAN DIEGO  
SCHOOL OF MEDICINE



Brin Rosenthal



Deanne Taylor



Children's Hospital  
of Philadelphia



Hossein Fazelinia



Perelman  
School of Medicine  
UNIVERSITY of PENNSYLVANIA



Komal Rathi



Helio Costa



STANFORD  
UNIVERSITY



Kathryn Grabek



J. Tyson McDonald



HAMPTON  
UNIVERSITY  
THE STANDARD OF EXCELLENCE



Gary Hardiman



MUSC Health  
MEDICAL UNIVERSITY of SOUTH CAROLINA



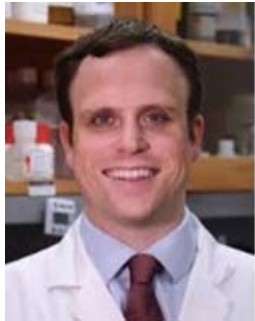
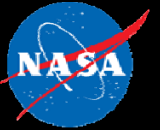
Willian da Silveira



GeneLab

Open Science for Exploration

# AWG Members Involved



Chris Mason



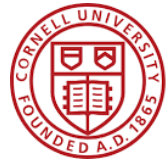
Cem Meydan



Jonathan Foox



Flavia Rius



Cornell University



Yared Kidane



TEXAS  
SCOTTISH RITE HOSPITAL  
FOR CHILDREN



Susana Zanello



Scott Smith



Sara Zwart



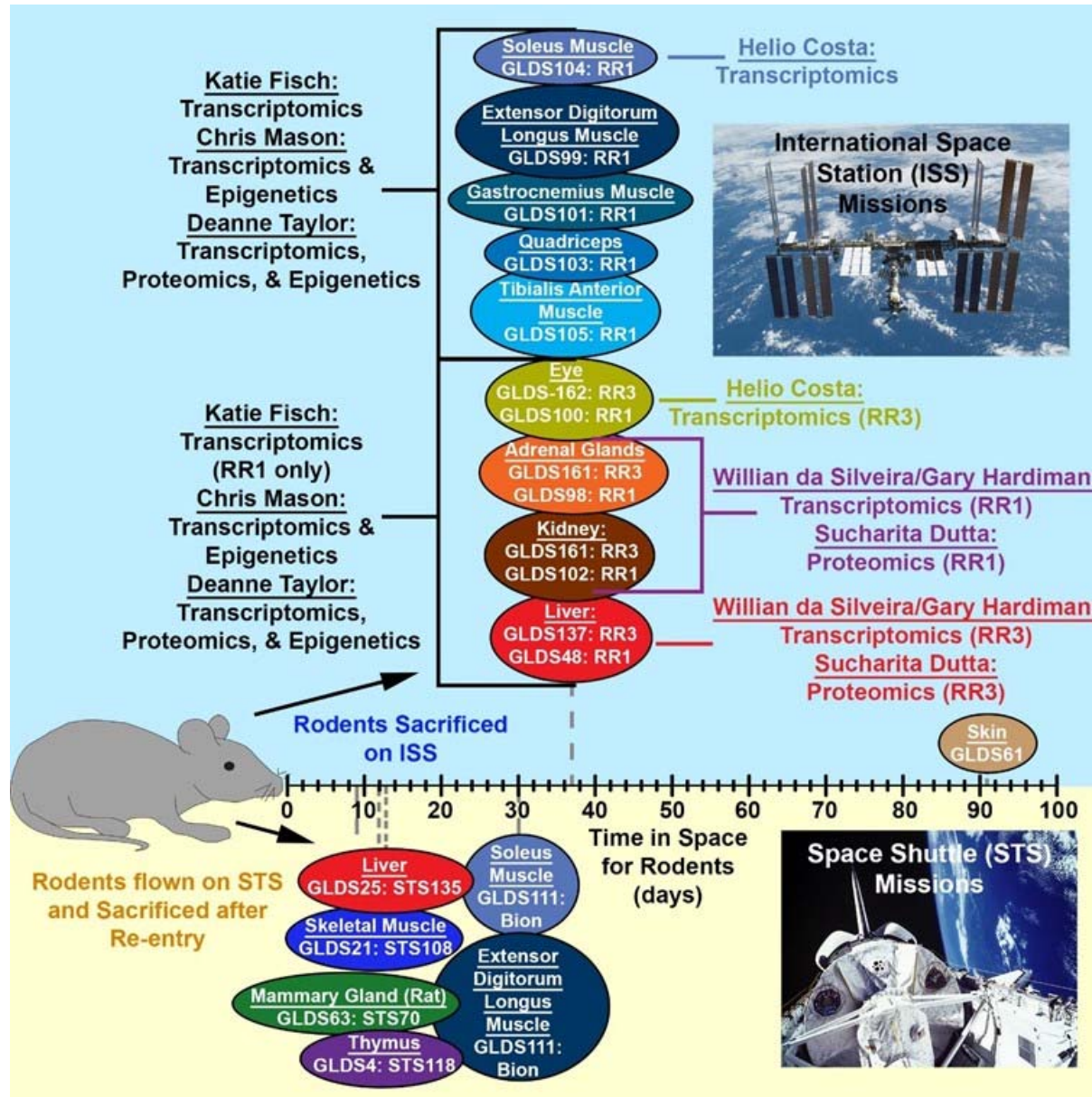
Afshin Beheshti



Sylvain Costes



# Specific Datasets and Tissues AWG Members Analyzed



In addition, human datasets are also included:

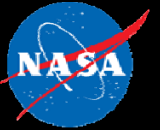
- GLDS-54, GLDS-174, GLDS-86, GLDS-118, GLDS-53, GLDS-54, GLDS-13. GLDS-52, or GLDS-114 (Tyson McDonald and Yared Kidane)



GeneLab

Open Science for Exploration

# Engaging with GeneLab



GeneLab

Open Science  
for Exploration

Home  
About  
Access GeneLab Data  
Analysis Working Groups  
Intern Resources  
Publications  
Newsletters  
FAQ  
Contact Us



© 2015-2018 NASA GeneLab



<https://genelab.nasa.gov>

Participate in GeneLab  
Analysis Working Groups

THANK YOU



Social media :

@NASA Ames Facebook

Twitter #GeneLab

ResearchGate: <https://www.researchgate.net/project/Omics-for-Space-Biology-The-GeneLab-project>