National Aeronautics and Space Administration

GeneLab: "Omics" Data System for Space Biology Research

Afshin Beheshti, PhD: Bioinfomartican and Data Analyst
afshin.beheshti@nasa.gov
Marla Smithwick, MSc, PMP: Deputy Project Manager
marla.m.smithwick@nasa.gov

Space Biosciences Division
NASA Ames Research Center, Moffett Field, CA

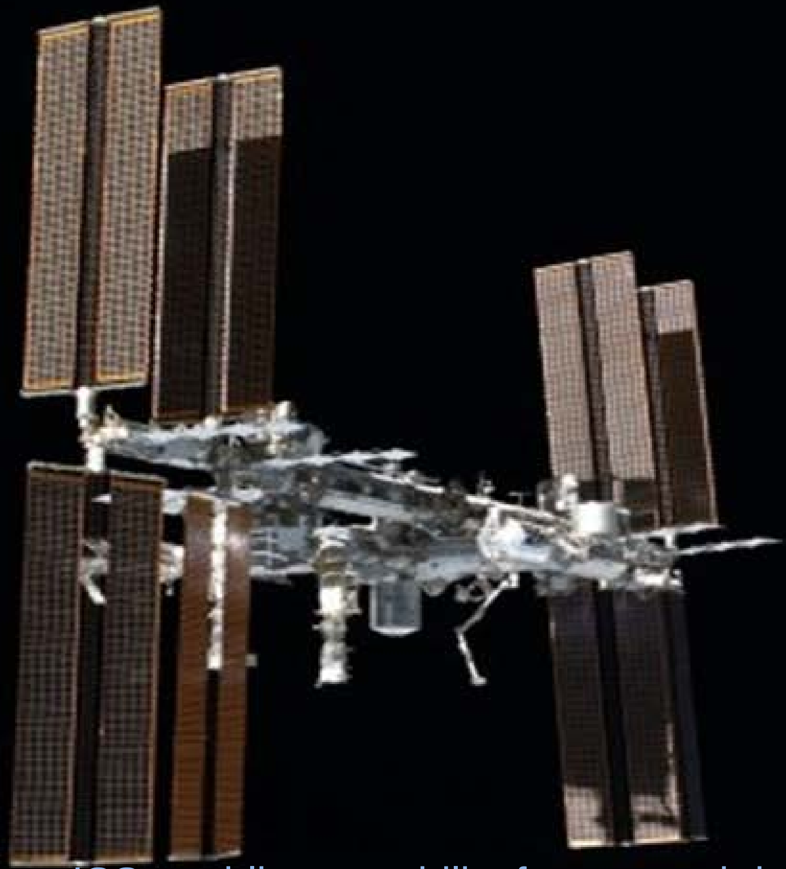
Sylvain Costes, PhD: Project Manager

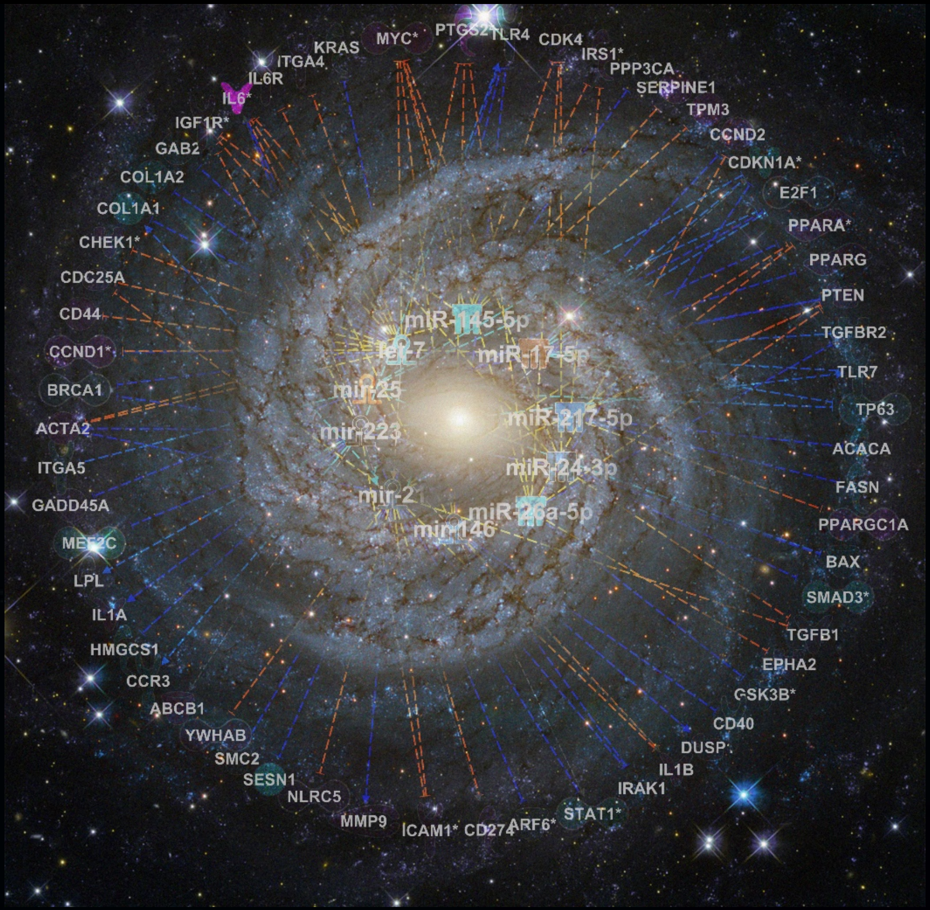
Jonathan Galazka, PhD: Project Scientist

The GeneLab Team

GeneLab
Open Science for Exploration

ISS: Our Orbiting Laboratory



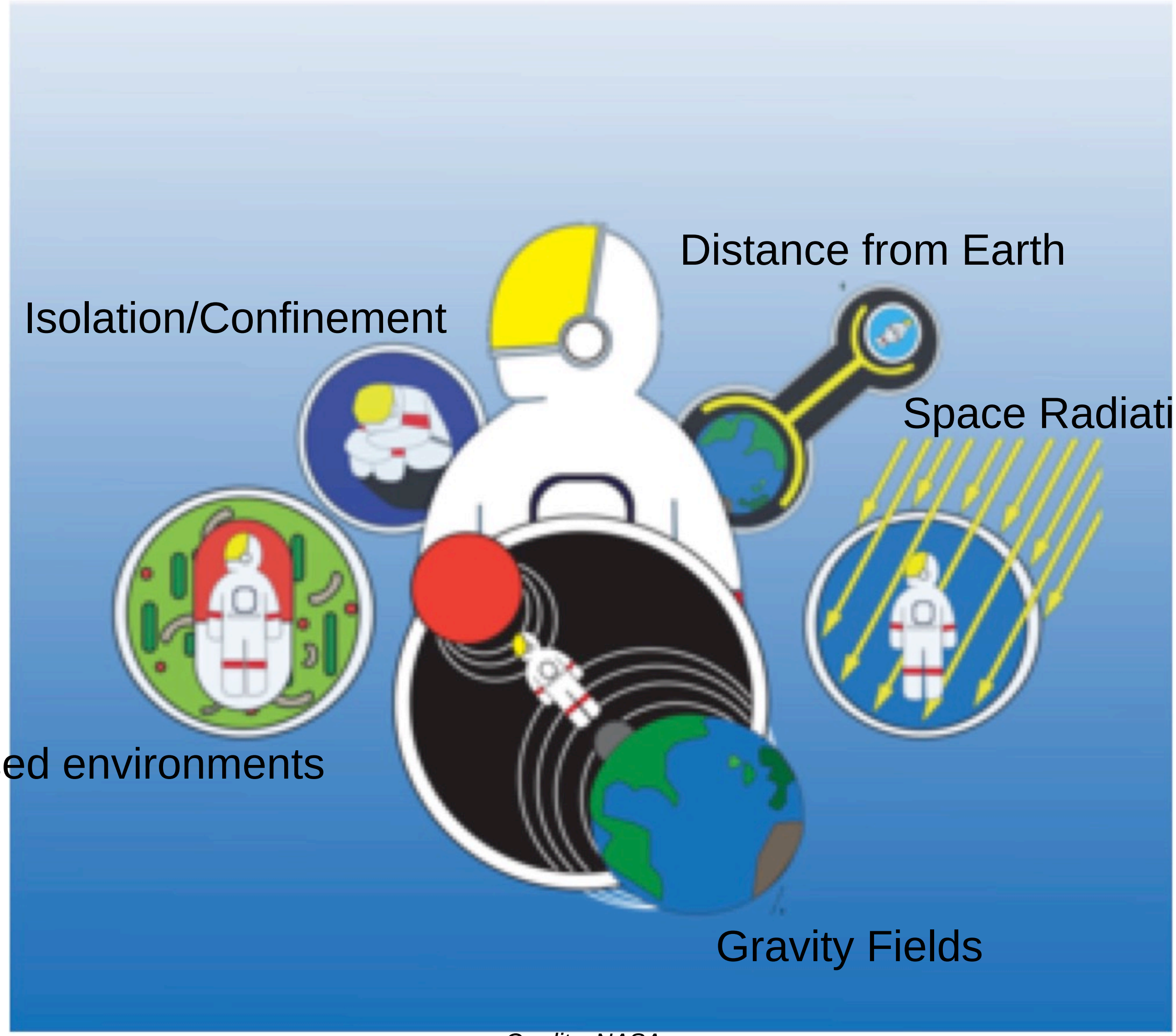


ISS enabling capability for research in cellular and molecular biology includes equipment for *in situ*, on-orbit analysis of biomolecules

Applications of this growing capability range from biomedicine and biotechnology to the growing field of Omics

Challenges of Spaceflight

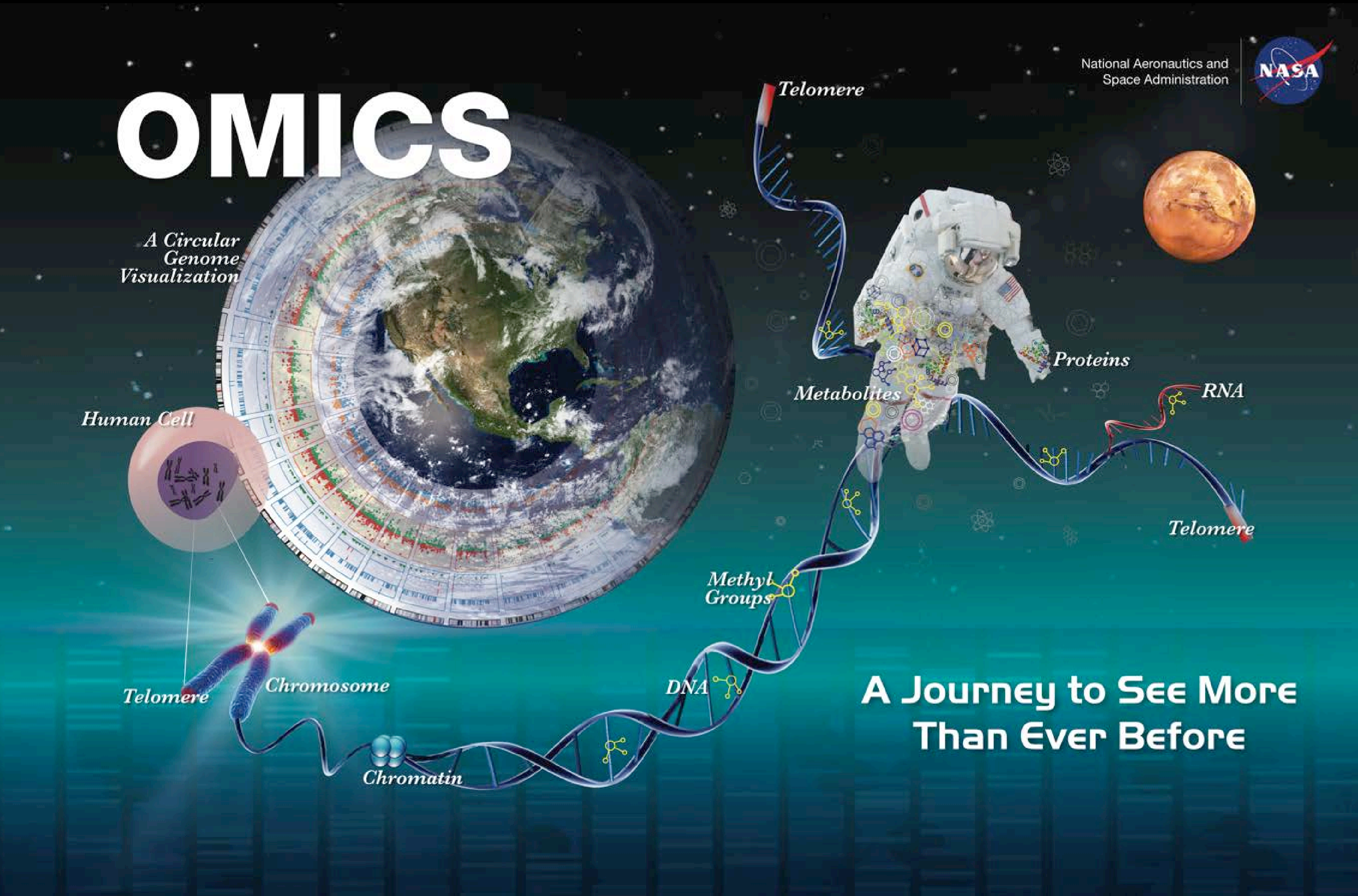
GeneLab
Open Science for Exploration



Credits: NASA

GeneLab
Open Science for Exploration

Opportunities of omics data

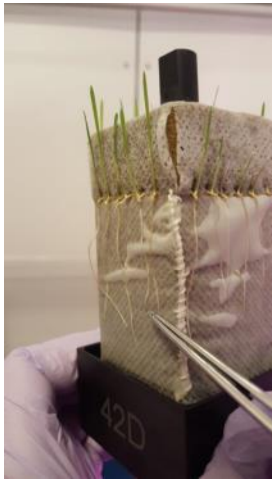


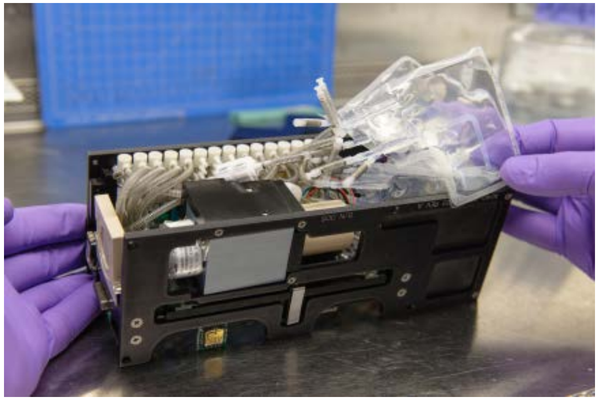
GeneLab
Open Science for Exploration

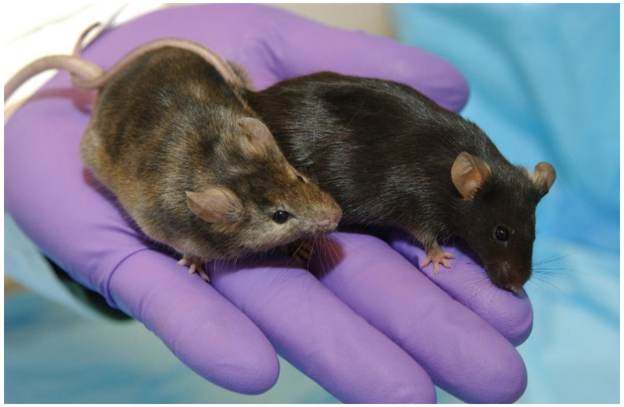
Recent On-Orbit Experiments

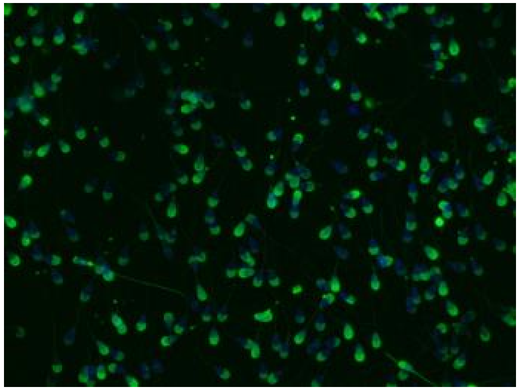
♦ **Space Biology Payloads Launched on SpX-13**

- **Rodent Research-6**
- **Cell Science Validation**
- **Plant Gravity Perception**
- **Microbial Tracking – 2**
- **APEX-06**
- **Fruit Fly Lab**
- **Micro-11**

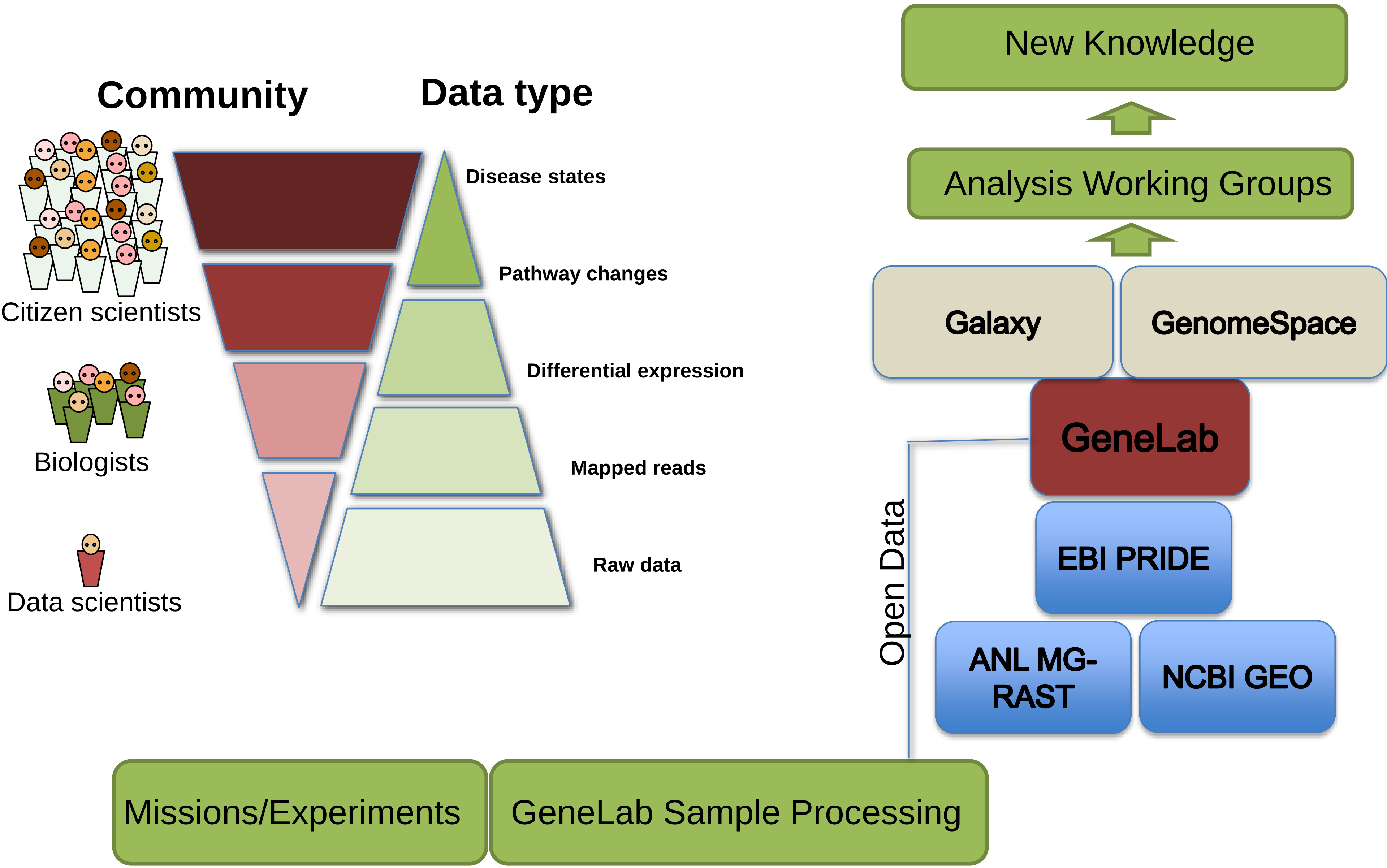








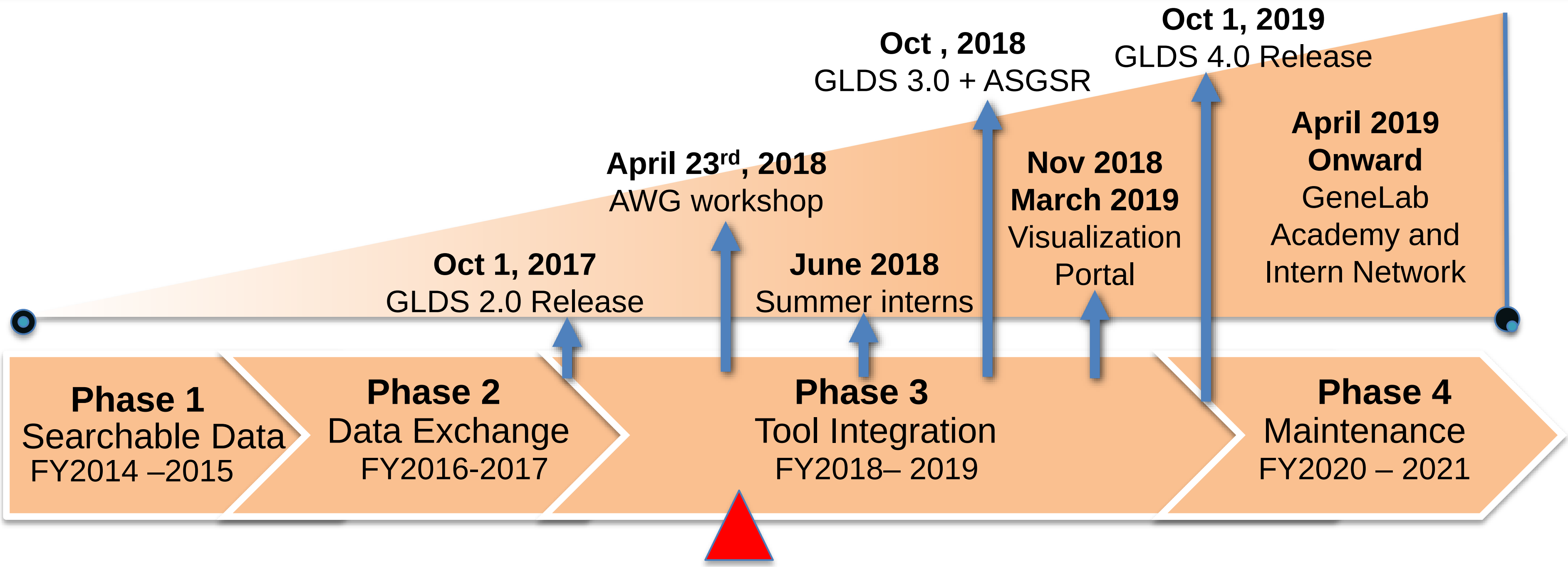
GeneLab



Community
Data type
Citizen scientists
Biologists
Data scientists
Disease states
Pathway changes
Differential expression
Mapped reads
Raw data
New Knowledge
Analysis Working Groups
Galaxy
GenomeSpace
GeneLab
EBI PRIDE
ANL MG-RAST
NCBI GEO
Open Data
Missions/Experiments
GeneLab Sample Processing

Phased Implementation

GeneLab
Open Science for Exploration



Data System
- ✓ Public Website
- ✓ Searchable Data Repository
- ✓ Top Level Requirements
- ✓ New Data and Legacy Data

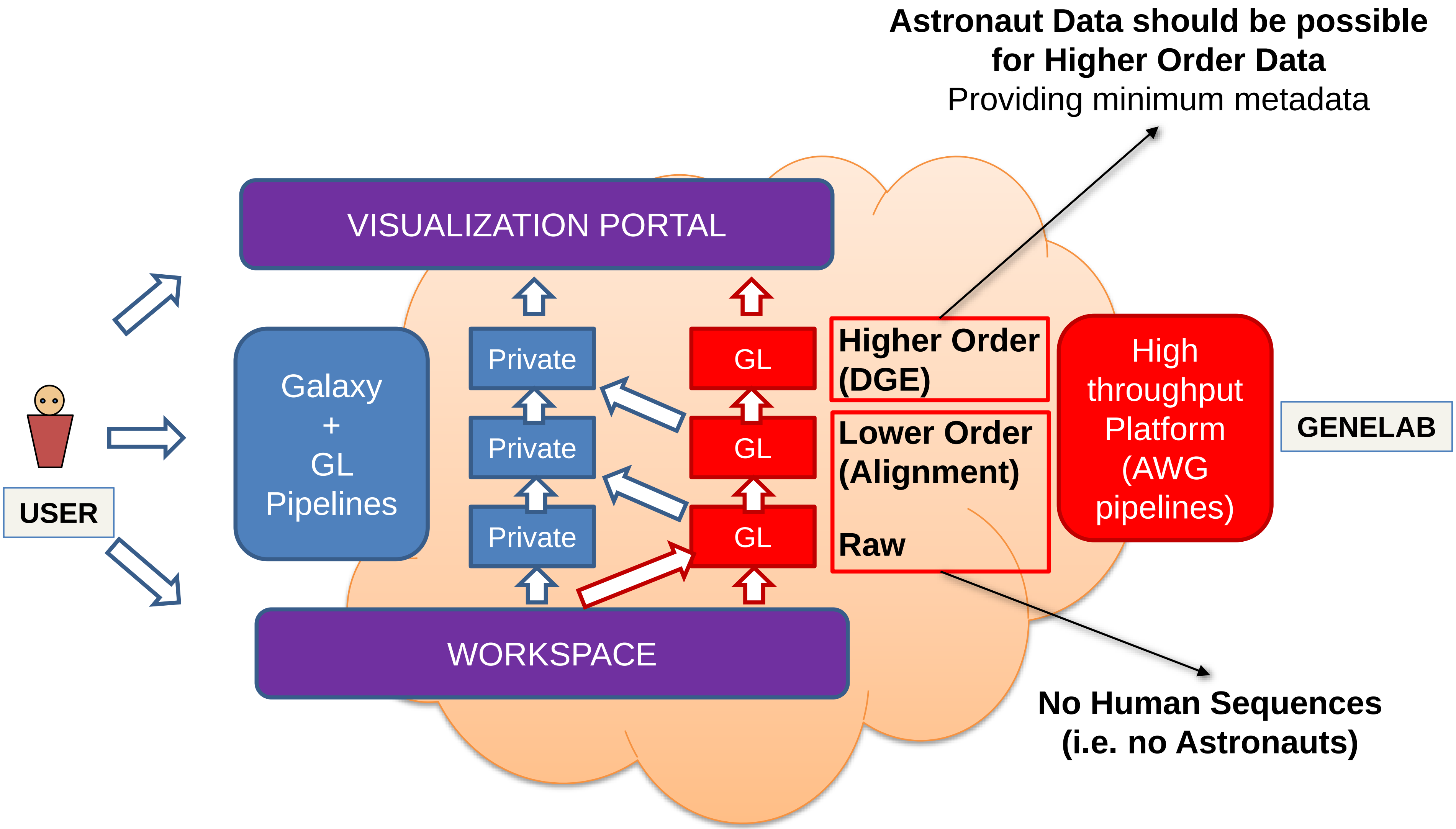
Data System
- ✓ Link to Public Databases via Data Federation
- ✓ Integrated Search (e.g., data mashup)

Data System
- Integrated Platform across model organisms
- Build Community via AWG
- Provide access to biocomputational tools for omics analysis
- Provide collaboration framework and tools

Open Source Maintenance
- User community becomes primary provider of new tools/knowledge
- Maintain integrity of data, and data system

GeneLab
Open Science for Exploration

GLDS 4.0



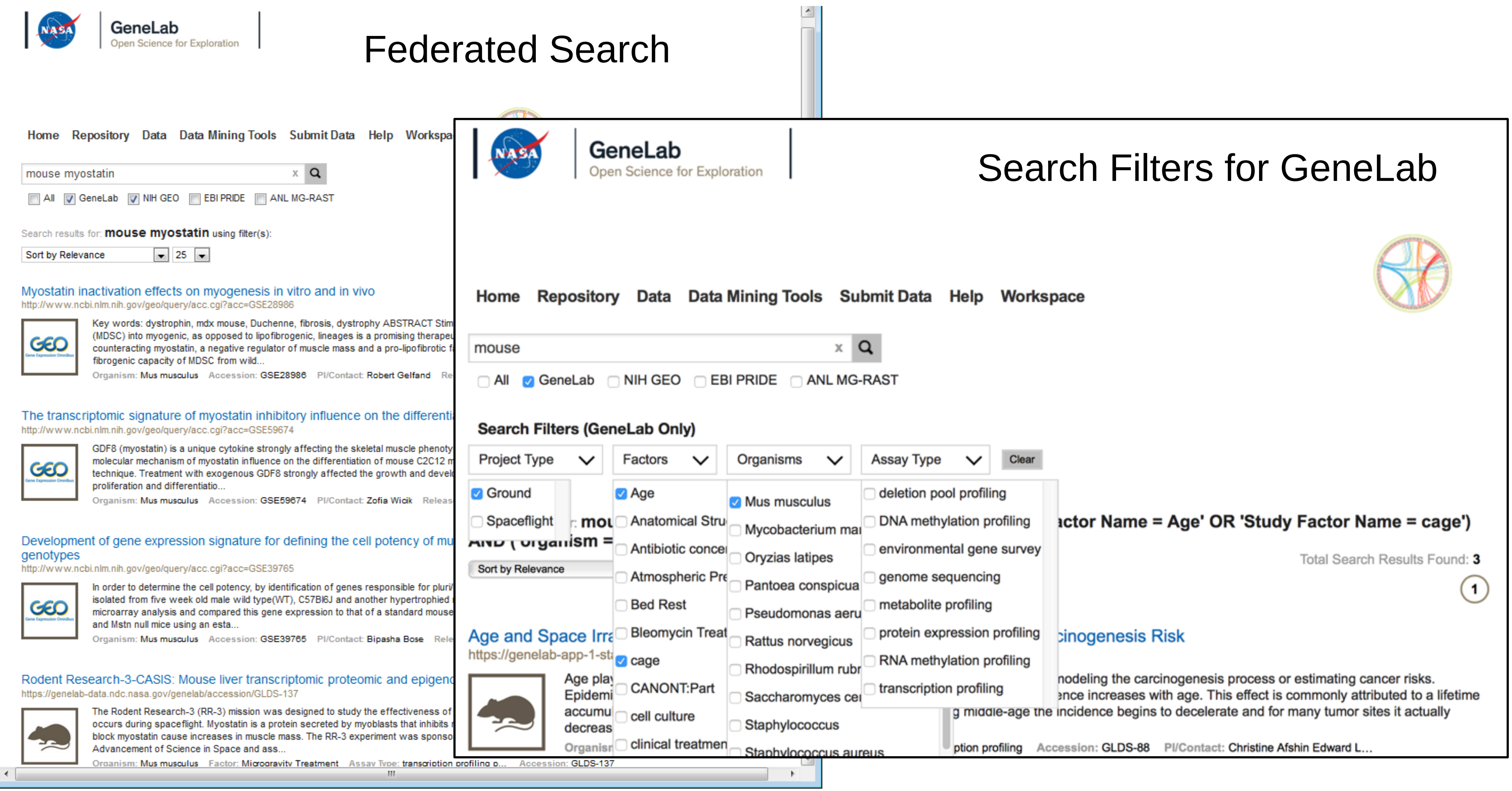
Opportunities

- Analysis Working Group Participation
  - We are actively recruiting individuals and groups with varying degrees of experience to participate, learn, publish!
- Workspace - Upload your own data, use our workspace to compare against our data sets
  - Set up an account, Galaxy tool suite available Oct 2018
- Visualization – create your own visualization tool and pull using our API
- Student interns
  - We bring in students every summer including some spots reserved for international students. We provide funding!
- LSDA allows tissue requests from international labs
  - One source of our tissues is the Life Sciences Data Archive. They have archived tissue samples that can be obtained via an application request available here:
  - https://lsda.jsc.nasa.gov/Request/dataRequest

GLDS Phase 2 (Release 2.0) Google-like Search, Federated Search

GeneLab Open Science for Exploration

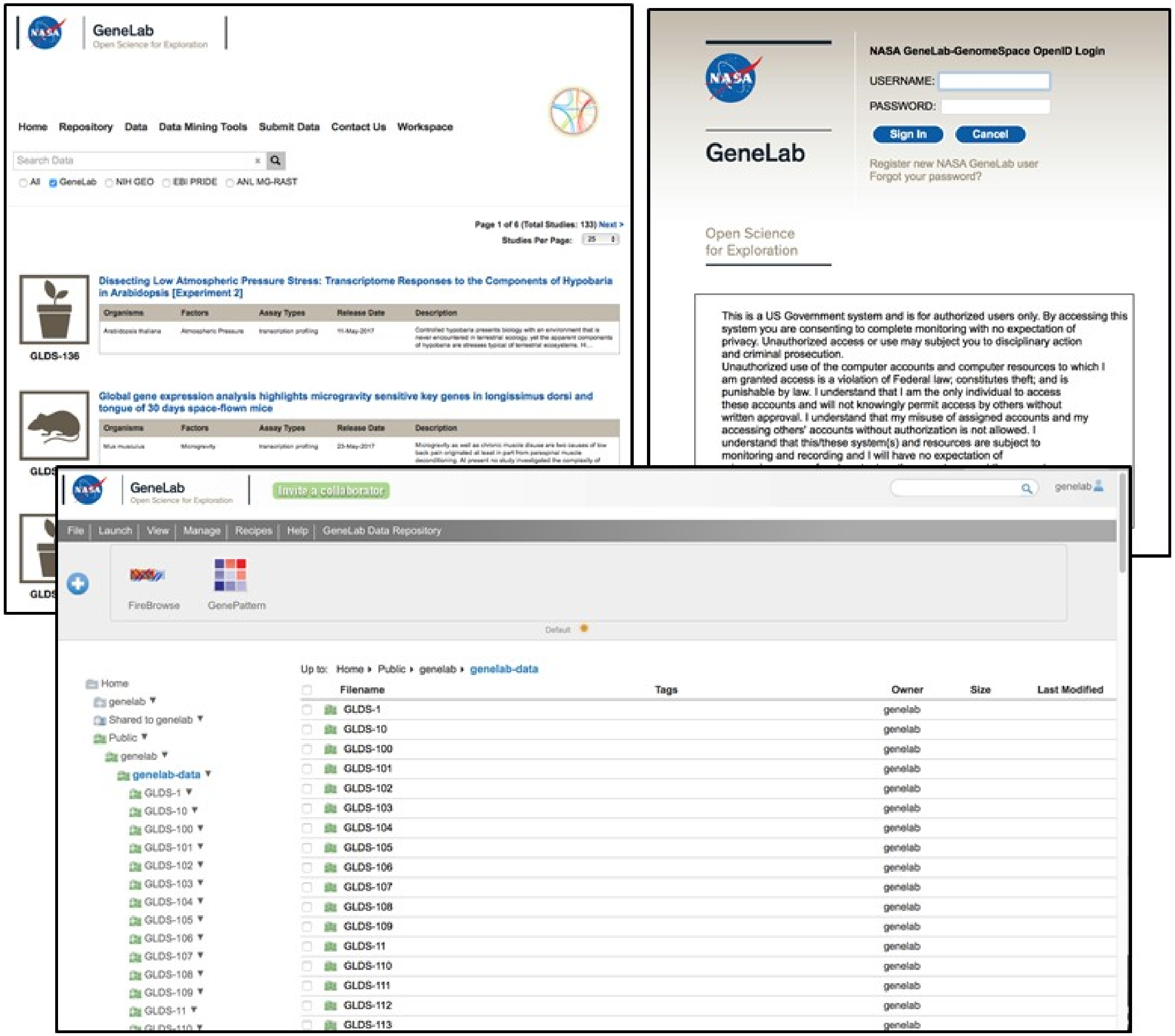
Data federation/integration with heterogeneous bioinformatics external databases (GEO, PRIDE, MG-RAST)



GeneLab
Open Science for Exploration

GLDS Phase 2 (Release 2.0) Customized NASA Collaborative Workspace

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



Galaxy Platform (GLDS 3.0)

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 scripts to be run and chained together into workflows
3. Workflows can published, shared and downloaded

galaxyproject.org

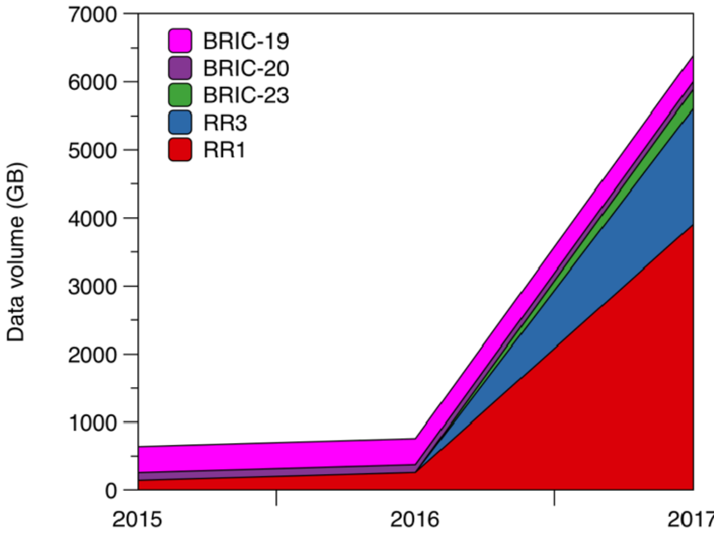
Afgan et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Research (2016)

GeneLab Sample Processing Laboratory (SPL)

GeneLab
Open Science for Exploration

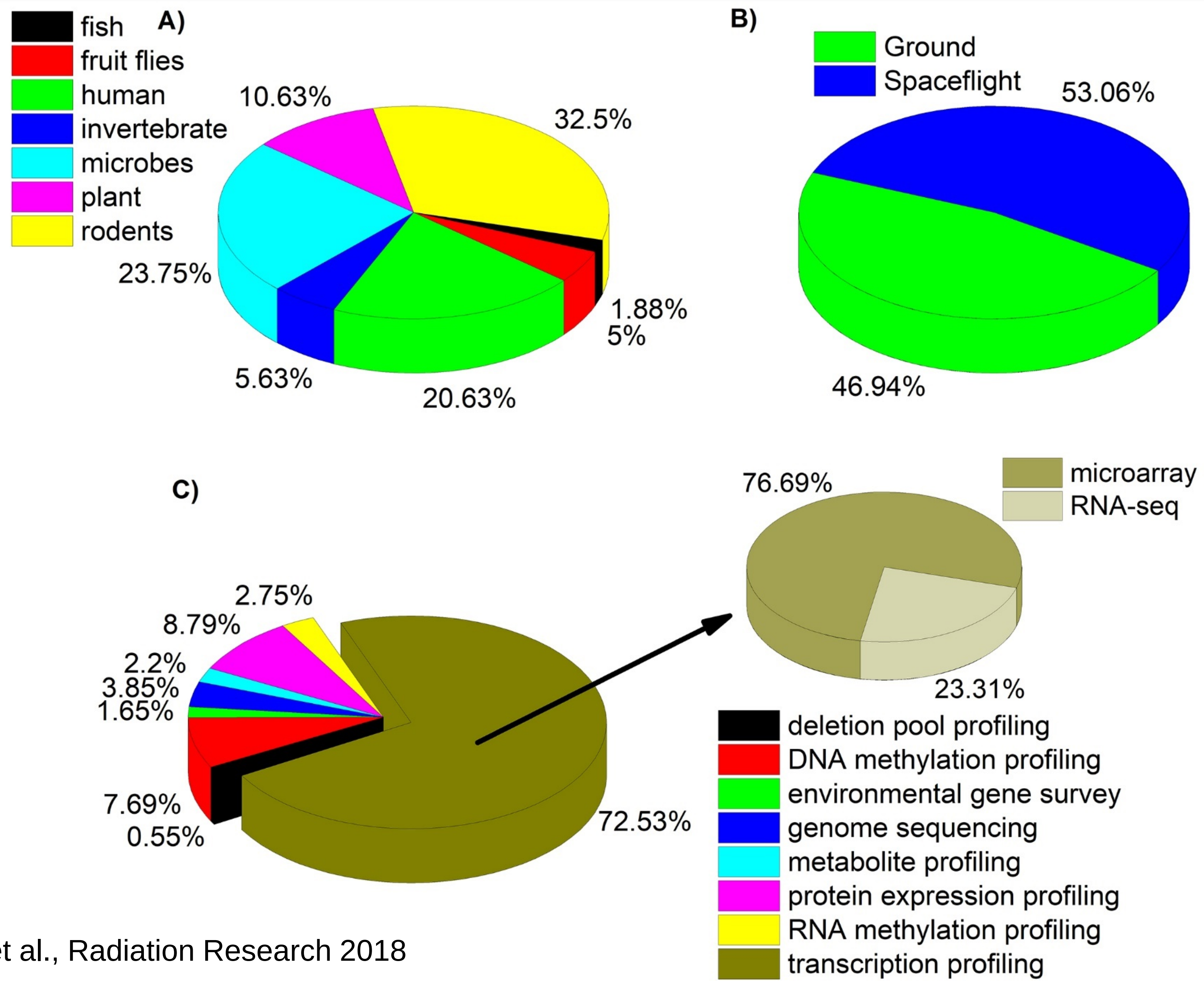
- Expertise:
  - DNA/RNA/protein extraction
  - Animal work

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



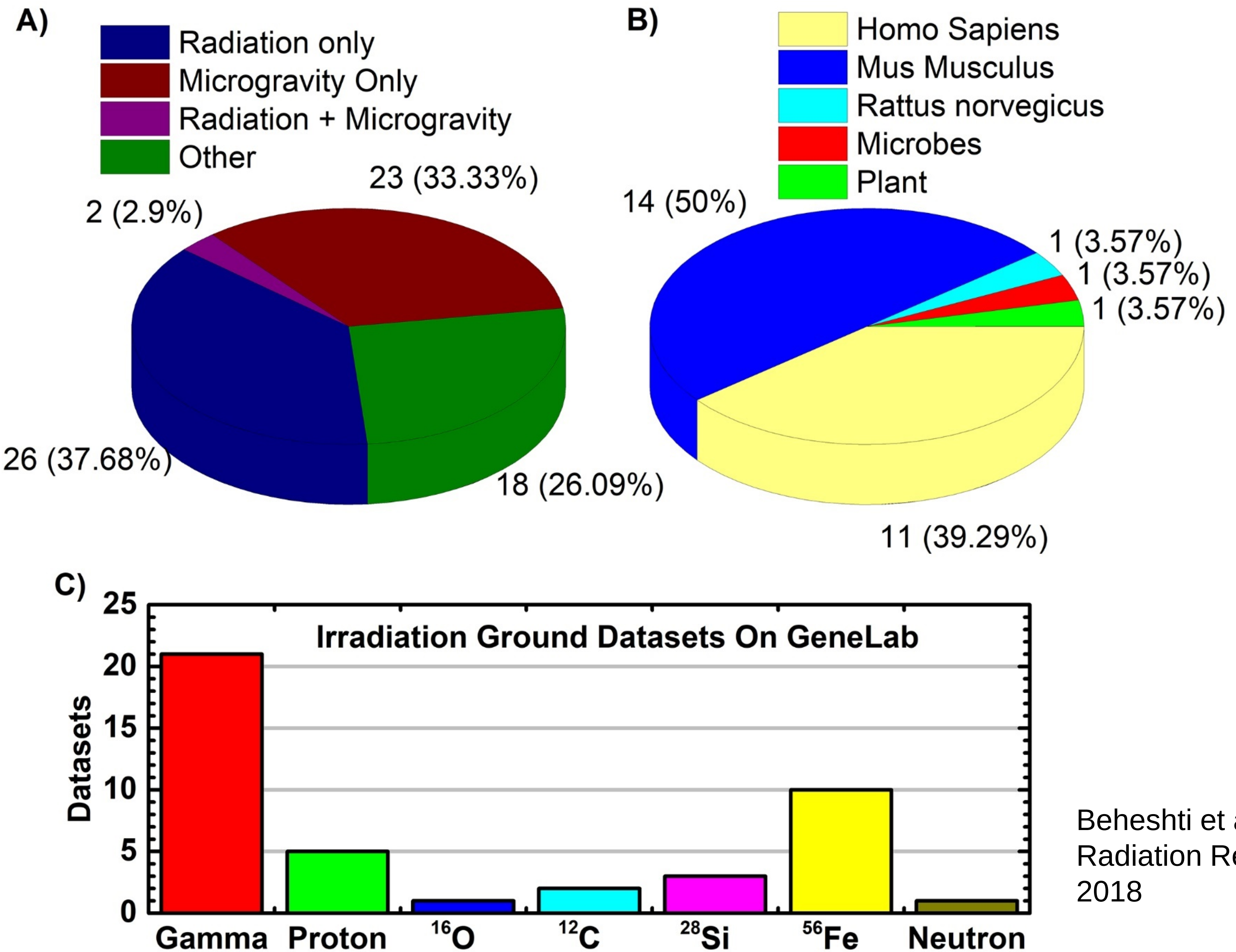
GeneLab
Open Science for Exploration

GeneLab Database: >190 data sets



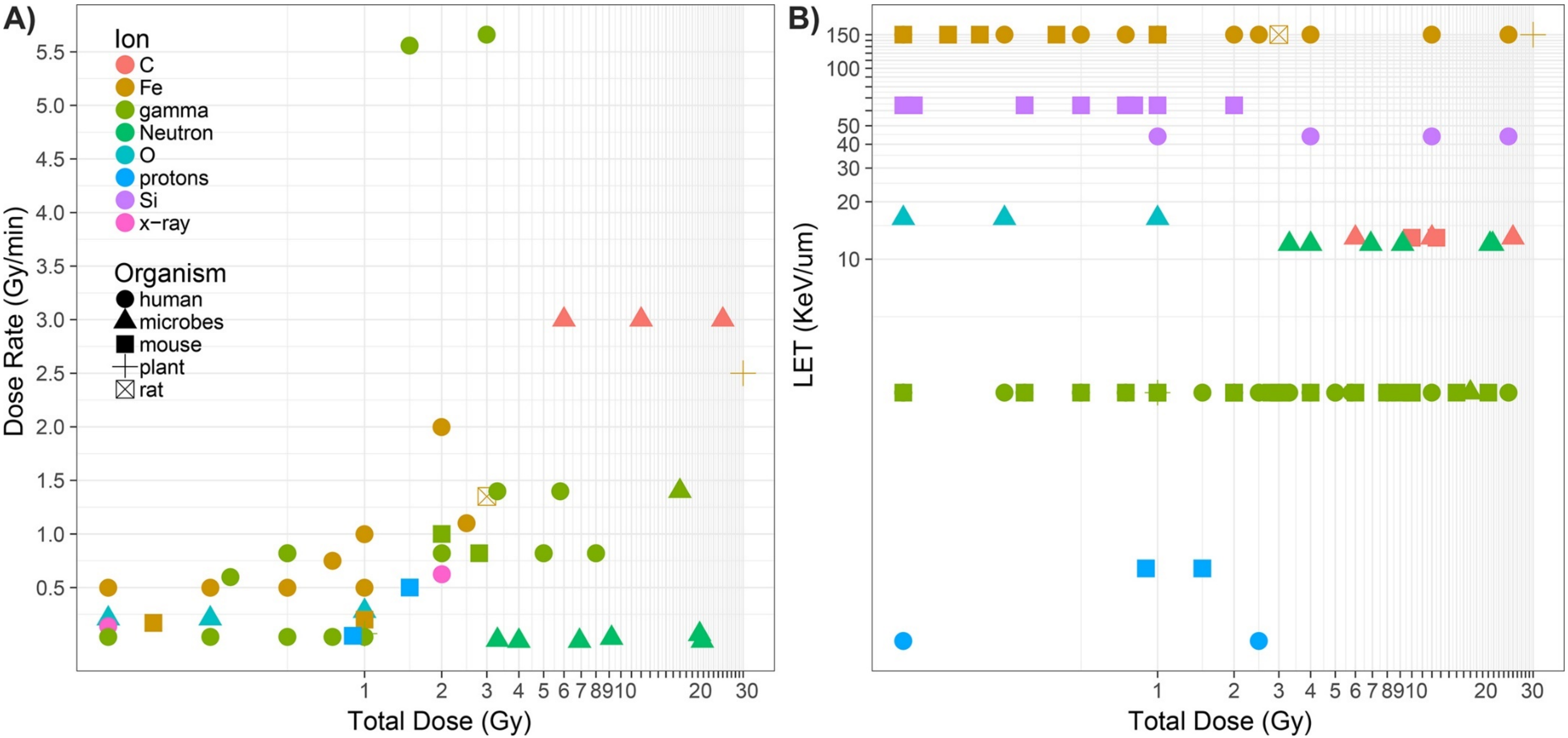
Beheshti et al., Radiation Research 2018

69 Ground Data Sets: Radiation and simulated microgravity



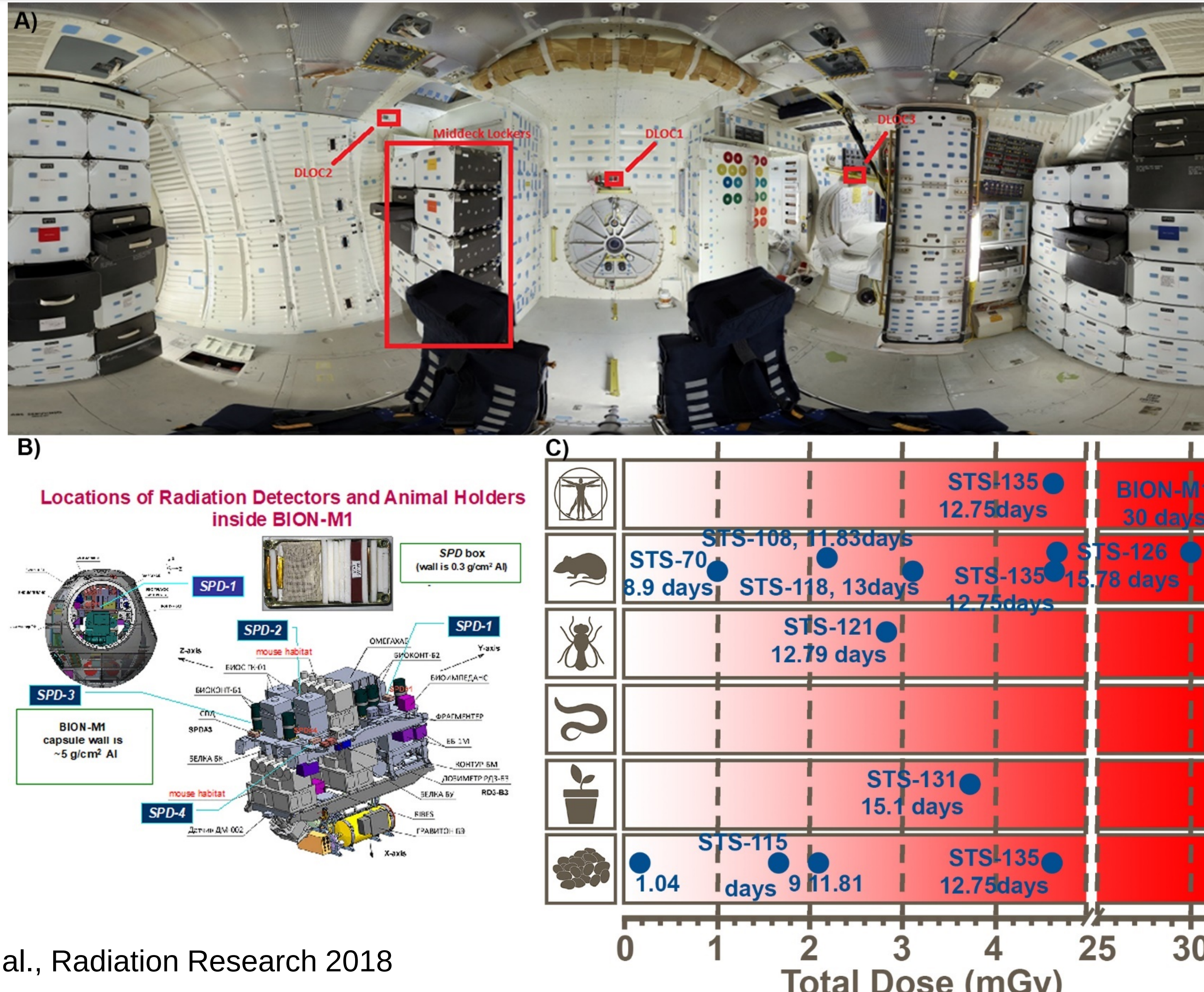
Beheshti et al., Radiation Research 2018

Radiation Ground Studies



Beheshti et al., Radiation Research 2018

Radiation Dosimetry for STS samples (ISS to follow)

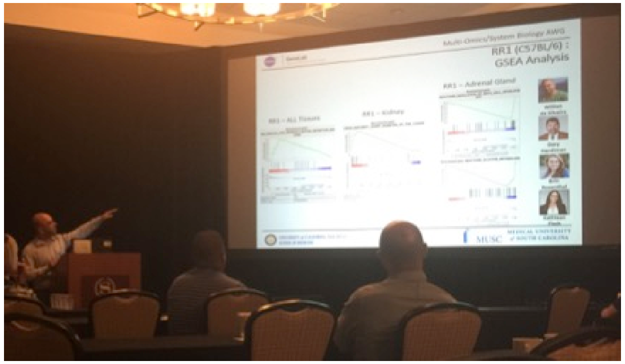


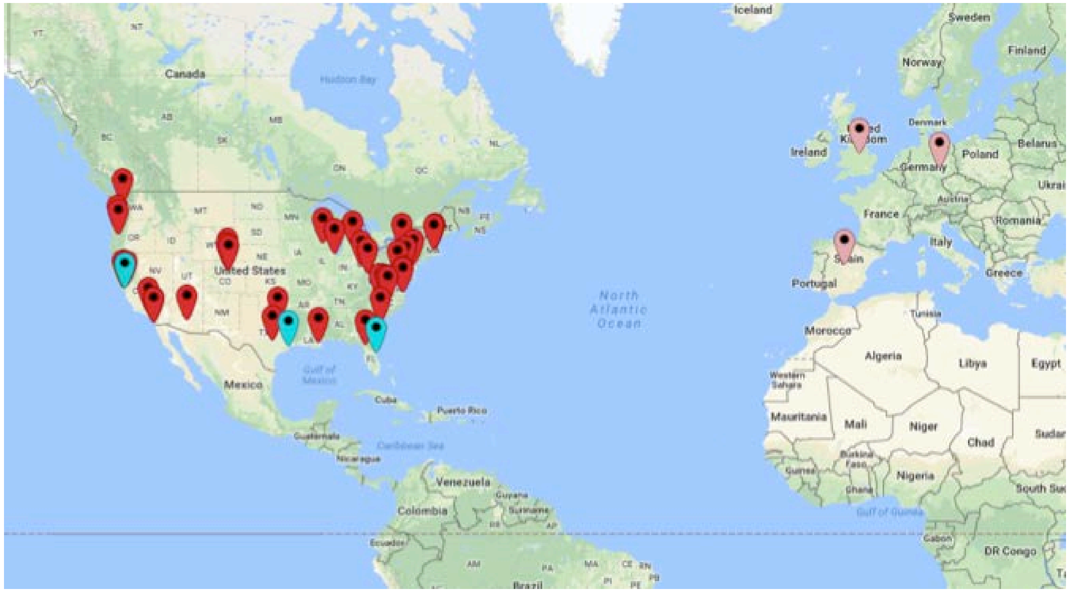
Beheshti et al., Radiation Research 2018

GeneLab
Open Science for Exploration

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
Open Science for Exploration

GeneLab Analysis Working Groups

Questions
What questions can we answer with these datasets?

AWG (now)

Analysis Pipelines
Goal is to identify the best pipeline

Data to be added to GLDS or AWG scope
What additional data do you require to answer you scientific questions?

Metadata to be added to GLDS
What environmental or other metadata to you need to answer your questions?

Student interns (June - Aug)

Implementation (ideally in Galaxy)

Processed data for posting on GeneLab

Post processed data to GeneLab

Public portal for visualization of GeneLab data

AWG (July - ?)

Visualization requirements and systems

Peer reviewed publication (s)

GeneLab
Open Science for Exploration

Hypothesis and Science from GeneLab

- Cage Effects with rodent experiments: Carbon Dioxide as an Environmental Stressor in Spaceflight
- Systems Biology analysis reveals biological spaceflight master regulators
- Space Radiation induces long term impact on the cardiovascular system by the activation of FYN through Reactive Oxygen Species
- AWG related work determines novel systemic biological factors causing damage due to spaceflight

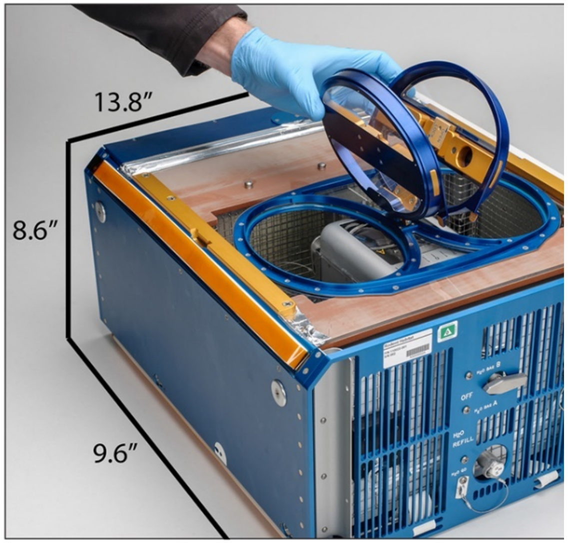
Cage Effects with rodent experiments: Carbon Dioxide as an Environmental Stressor in Spaceflight

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.

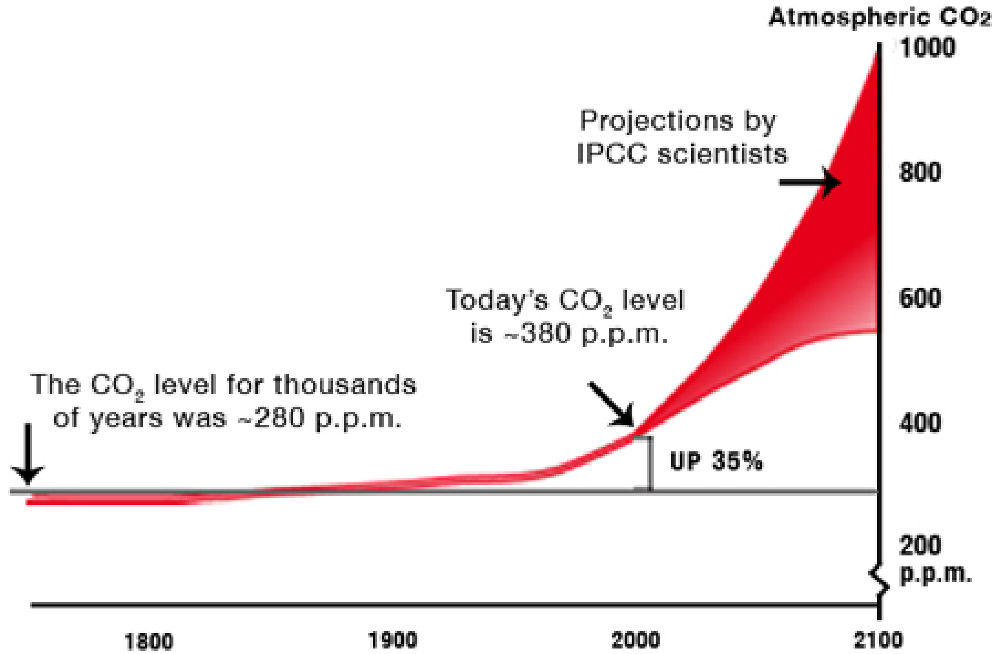
Carbon Dioxide as an Environmental Stressor in Spaceflight

GeneLab
Open Science for Exploration

A) Cage Types



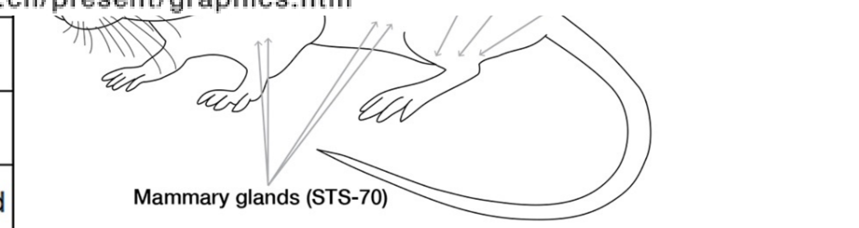
Animal Enclosure Module (AEM)



Source: IPCC
http://www.ipcc.ch/present/graphics.htm

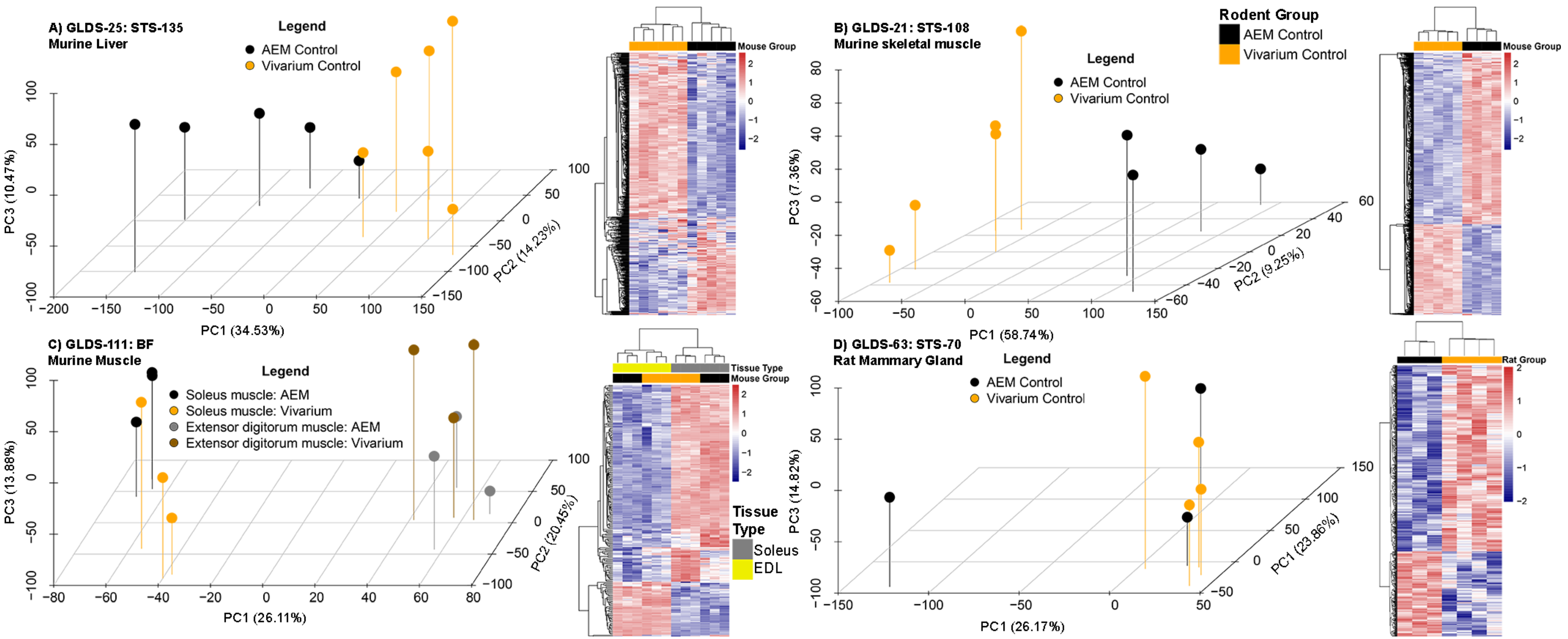
B)

GeneLab Study	Mission	Species	CO_2 (ppm)	Du (d	
GLDS-21	STS-108	mouse	~3000	1	
GLDS-111	BF	mouse	~600		
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



Beheshti, et al., Scientific Reports, 2018

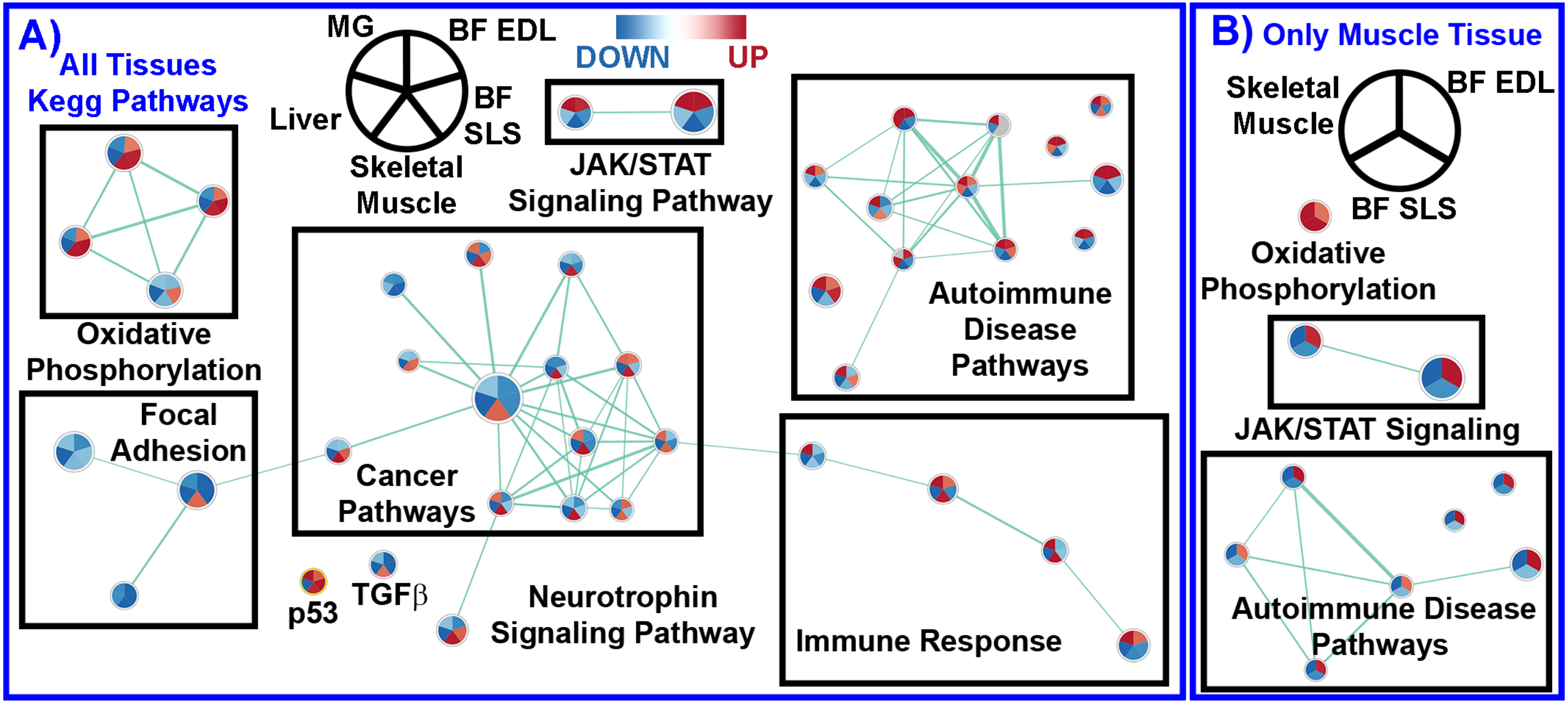
Global Cage Differences



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

Beheshti, et al., Scientific Reports, 2018

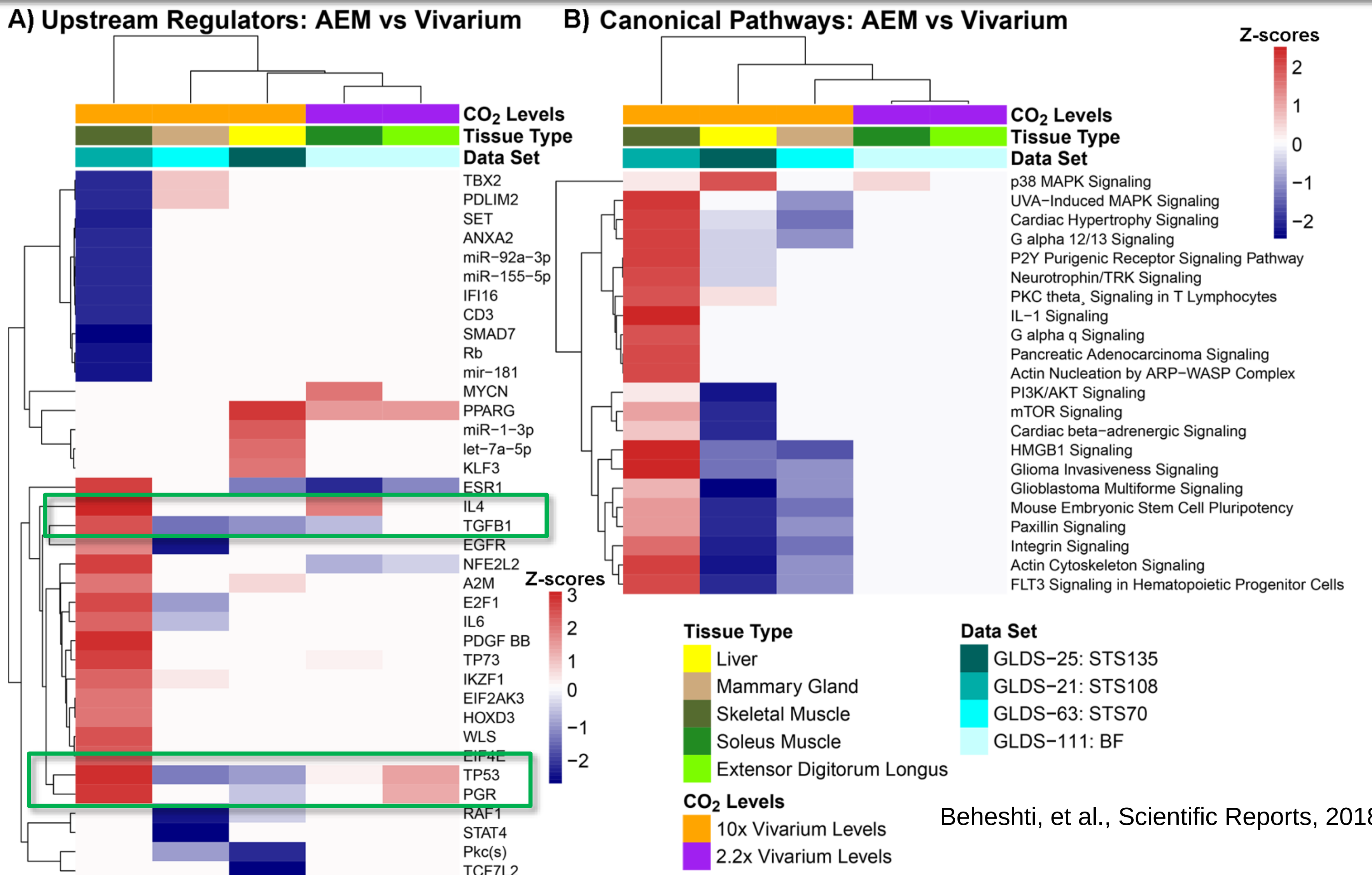
Major Pathways Regulated between Cages



GSEA: Kegg Pathways (network displayed using EnrichmentMap plugin for Cytoscape)

Beheshti, et al., Scientific Reports, 2018

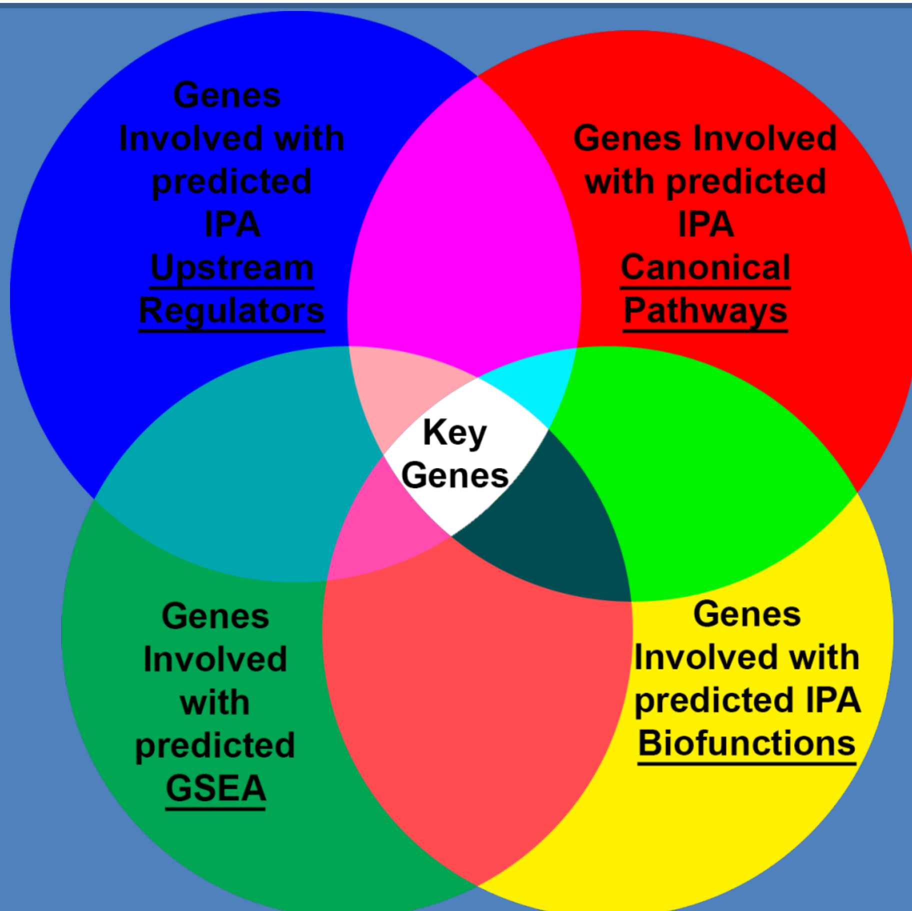
Upstream regulators and canonical pathways show response is tissue specific and highest for high CO_2



Beheshti, et al., Scientific Reports, 2018

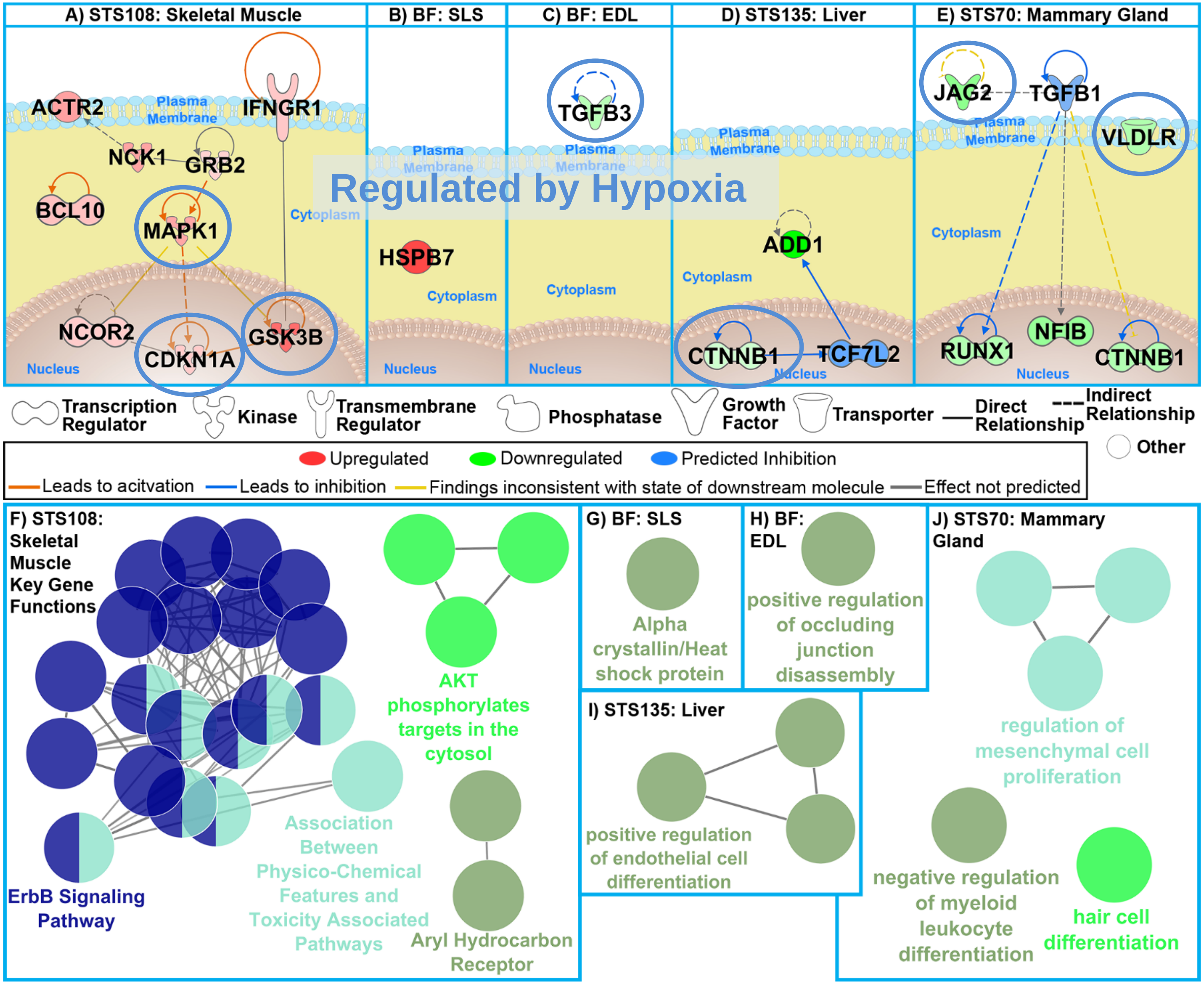
Mild chronic hypoxia due to increased CO_2 levels could explain both the increase in immune responses and a reduction in metabolism – **Need to confirm with AEM experiments at ambient CO_2 levels.**

Determination of Key Driving Genes



Beheshti, et al. Cancer Research 2015
Beheshti, et al. Oncotarget 2015
Beheshti, et al. Cancer Informatics 2015
Beheshti, et al, Radiat Res. 2014 & J. Radiat Res. 2015.
Ravi, Beheshti, et al. Cancer Research 2016
Beheshti, et al. Scientific Reports 2018

Identifying Key Cage-Dependent Drivers



Beheshti, et al., Scientific Reports, 2018

Conclusions for this Study

- Through a systems biology approach we observed global transcriptomic changes in rodents induced by spaceflight-matched environment in AEM cages.
- Identify spaceflight CO_2 levels as a potential environmental stressor that merits experimental investigation
- Systematically changing one environmental aspect at a time (gas concentration, radiation, microgravity, etc.) and analyzing and comparing transcriptional responses could be used to create a network that could predict the most relevant causes and countermeasures for spaceflight-associated conditions, as well as confounding factors for spaceflight experiments.

GeneLab
Open Science for Exploration

NASA

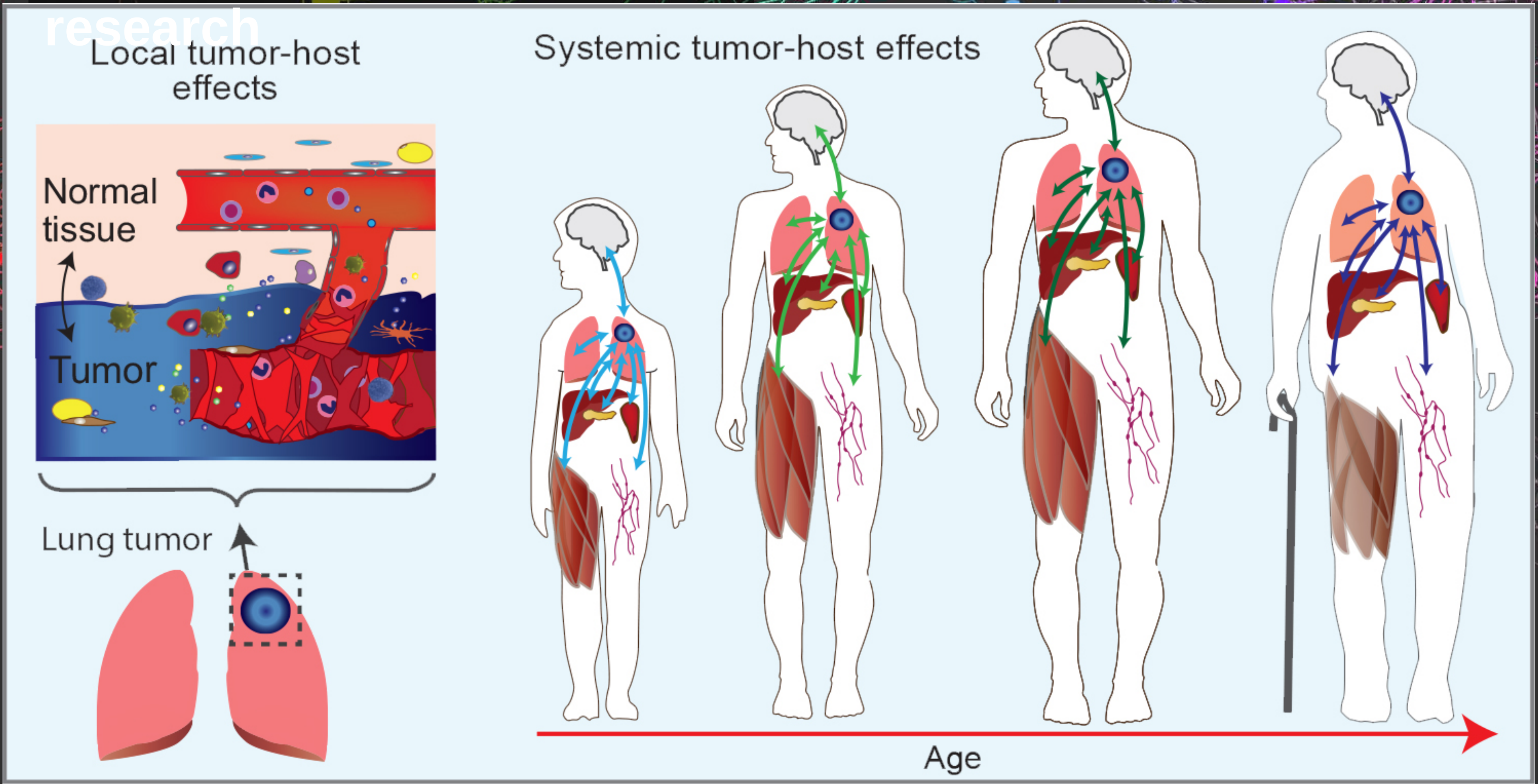
Systems Biology analysis reveals biological spaceflight master regulators

Beheshti, et al., PLOS One, in press

Systems Biology analysis reveals biological spaceflight master regulators

General Approach to Studying a Systematic Response in the Host

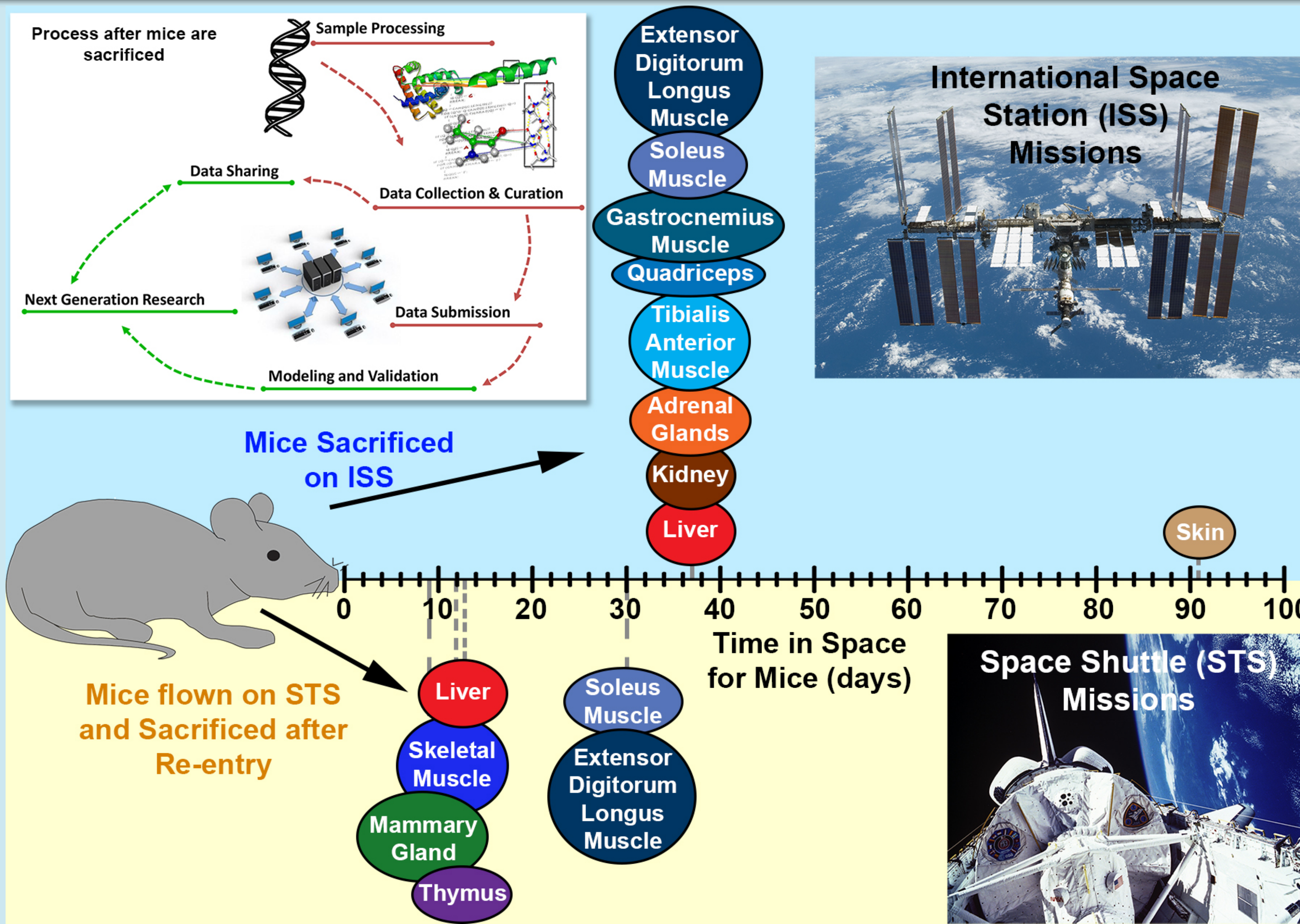
An example for cancer research



Beheshti, et al. Oncotarget, 2015.

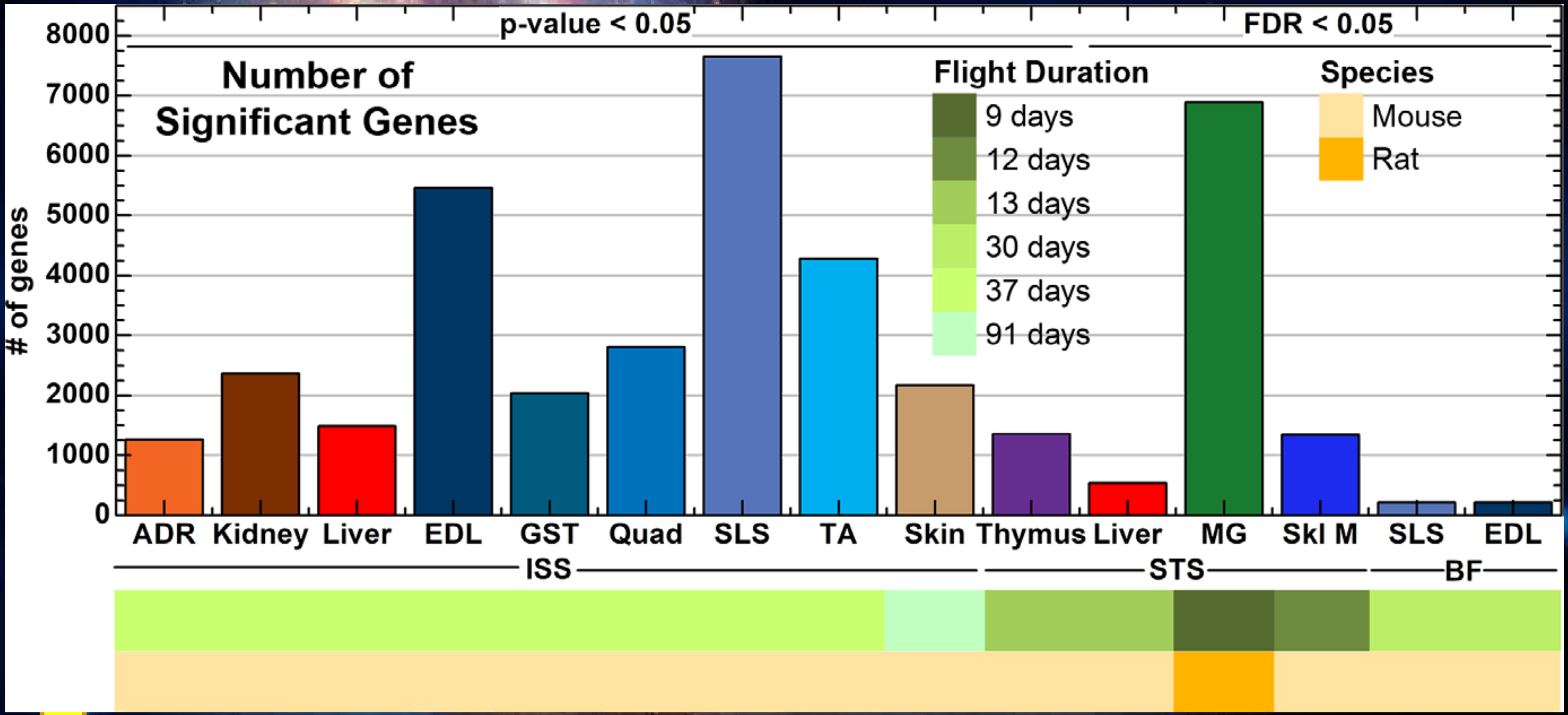
GeneLab Data Used to Generate Results

GeneLab
Open Science for Exploration



Number of Significant Genes from Each Dataset

GeneLab
Open Science for Exploration

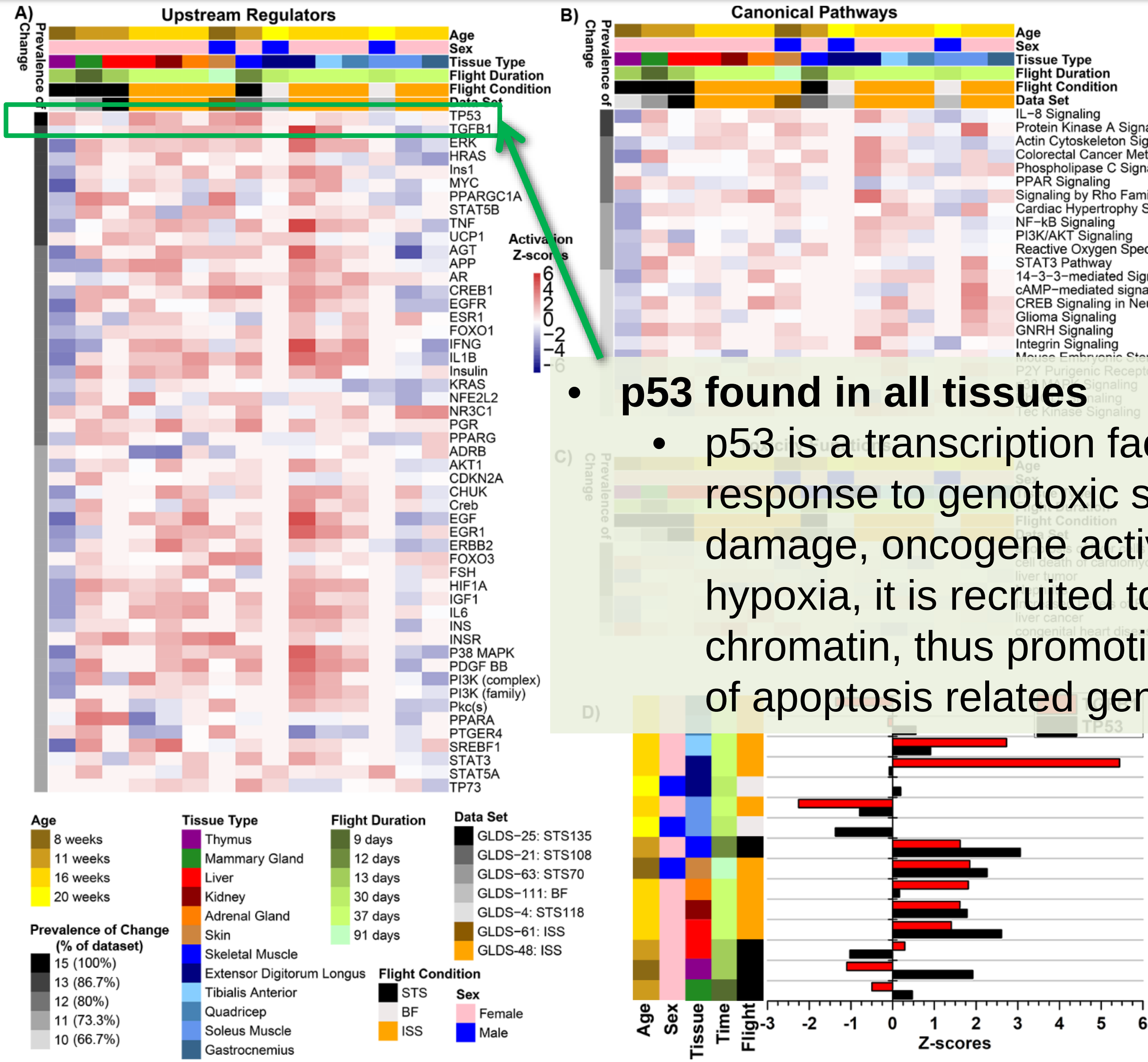


Fold-Change ≥ |1.2|

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

Beheshti, et al., PLOS One, in press

Predicted Master Regulators

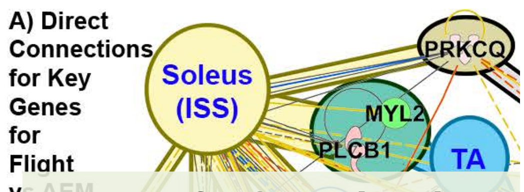


- **p53 found in all tissues**
  - p53 is a transcription factor and in response to genotoxic stress, DNA damage, oncogene activation, and hypoxia, it is recruited to sites in chromatin, thus promoting transcription of apoptosis related genes

Beheshti, et al., PLOS One, in press

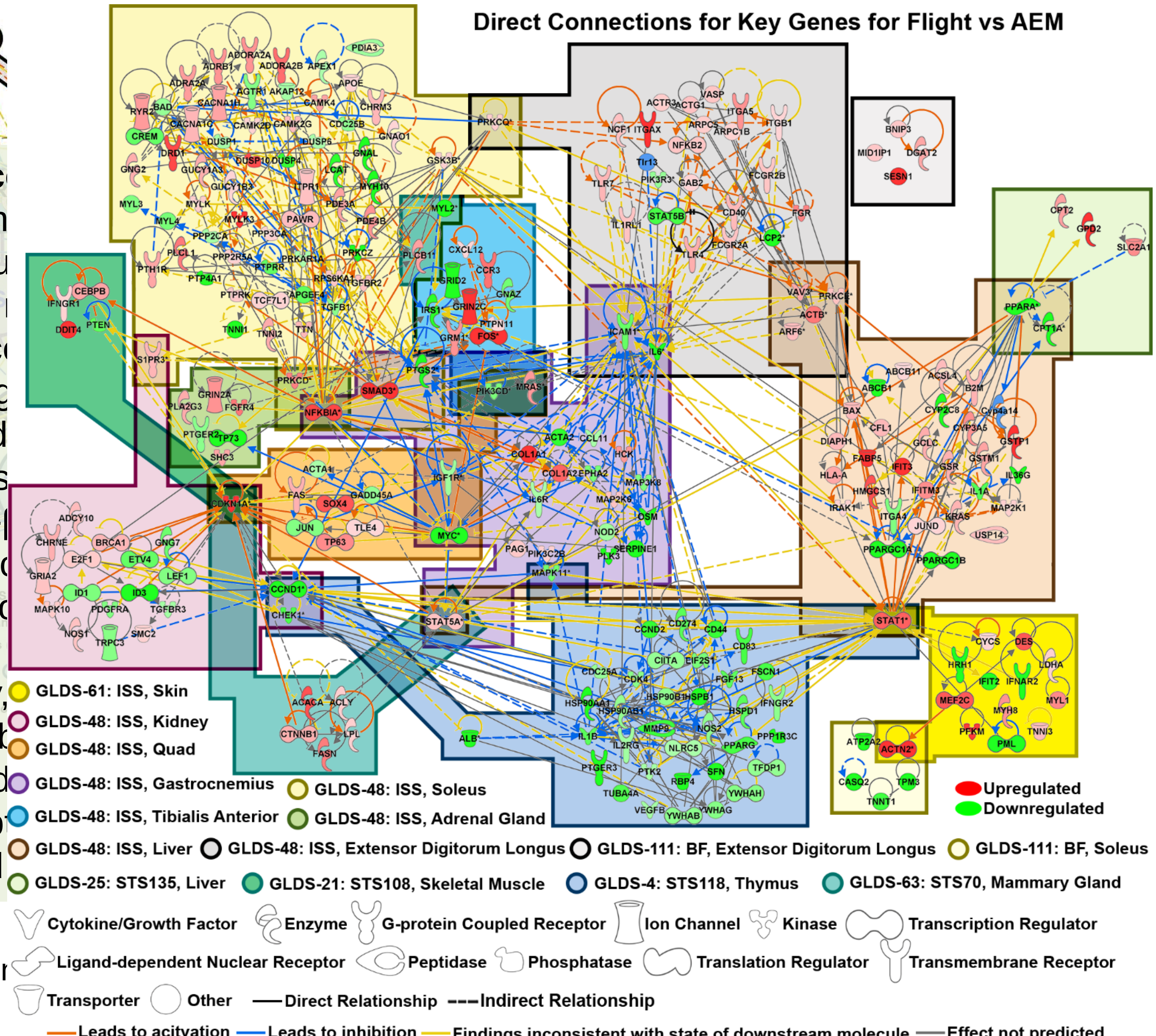
GeneLab
Open Science for Exploration

Key Genes and the Connections



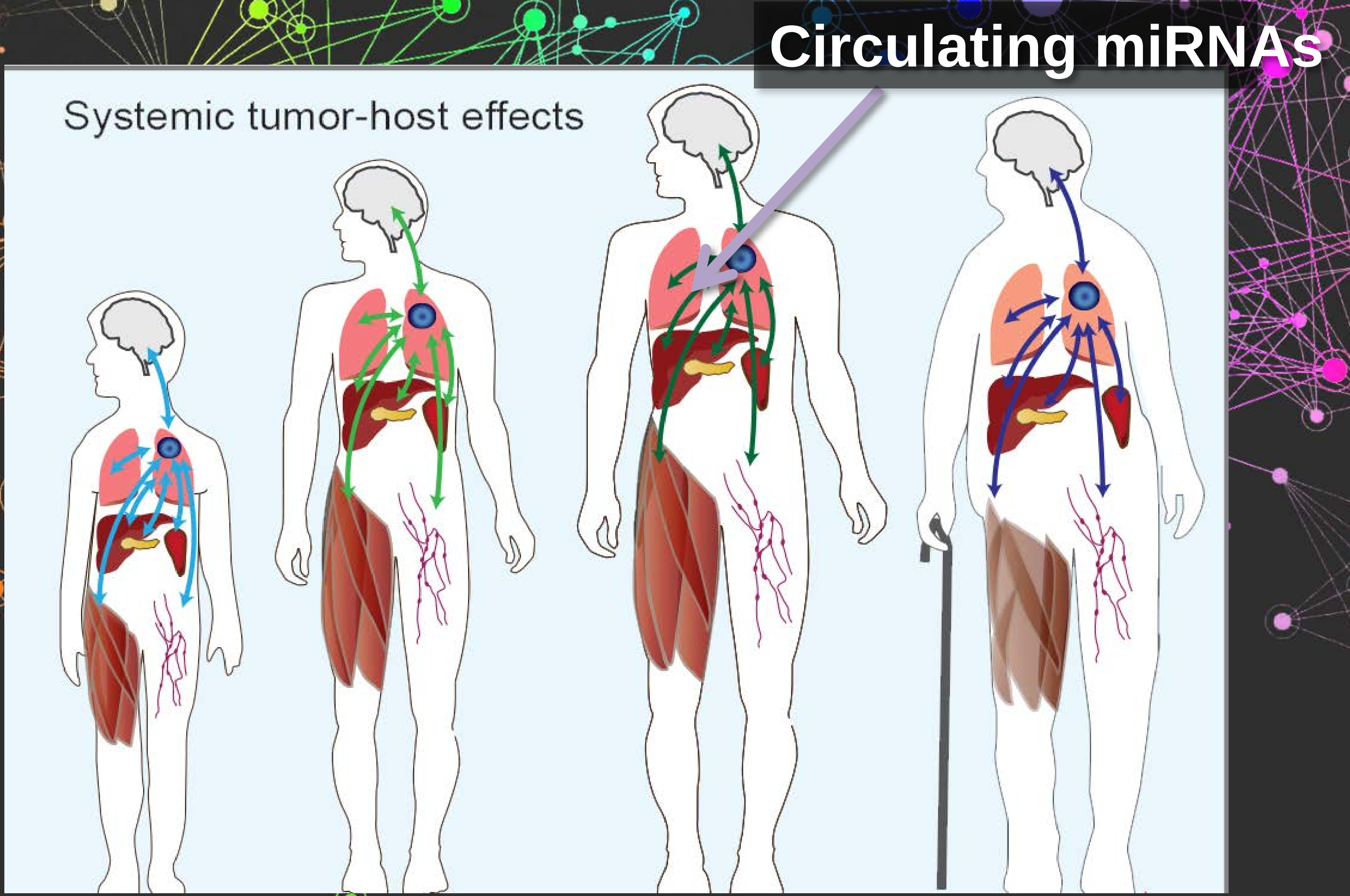
- **TGFβ1 found to be c**
  - TGFβ is known sustaining tissu transcriptional differentiation, c along with reg homeostasis and
  - Previous Studies to diminish TGF-| carcinoembryonic colorectal carcin in unit gravity.
  - In another study, growth using identified in mod and late apop endothelin-1 and

Beheshti, et al., PLOS On

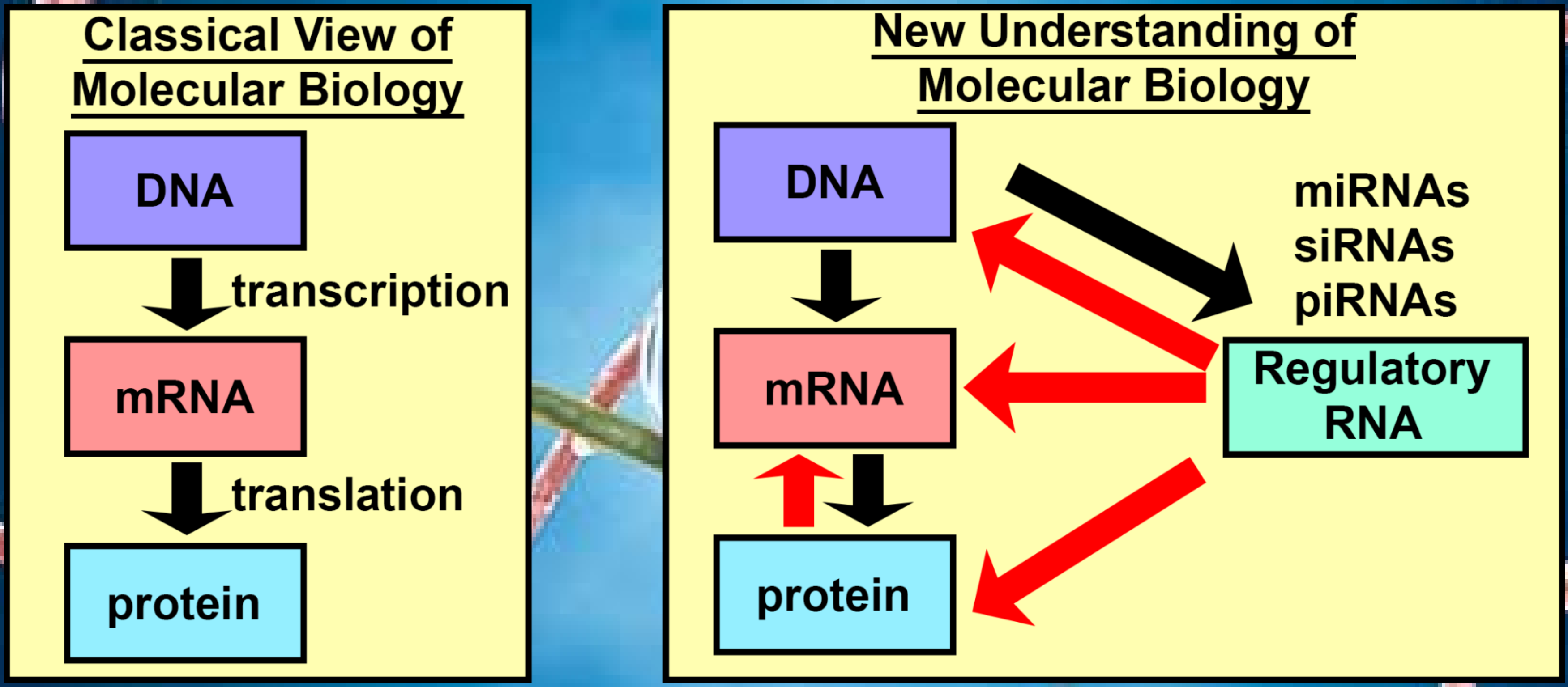


GeneLab
Open Science for Exploration

General Approach to Studying a Systematic Response in the Host



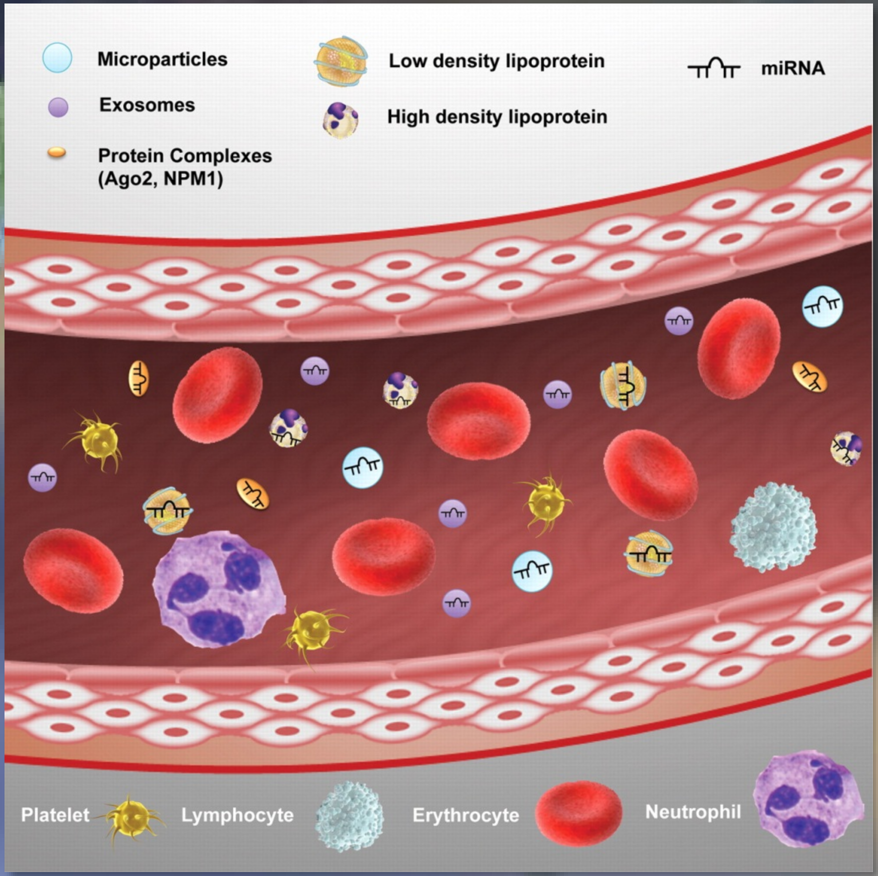
Revised View of Molecular Biology



- A single miRNA has been estimated to regulate up to 500 mRNAs
- miRNAs are single-stranded RNA sequences, of about 22 nucleotides in length, processed from longer transcripts.
- miRNAs are important regulators that repress the translation of mRNA transcripts

Impact of Circulating microRNAs

- Circulating miRNAs can carry signals from organs to other various parts of the body through the blood stream.
- The miRNAs can be transported in Exosomes, microparticles, lipoproteins, and outside any type of packaging.
- Our preliminary data shows that a miRNA signature is carried over from the spleen to the tumor with age.
  - Beheshti, et al. *PLoS ONE 2017*



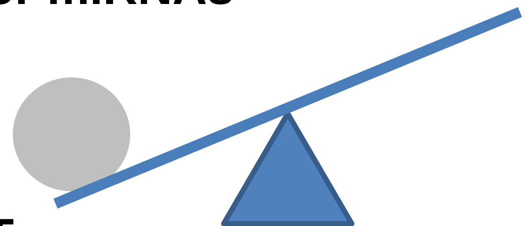
Profiling of circulating microRNAs: from single biomarkers to re-wired networks Anna Zampetaki, Peter Willeit, Ignat Drozdov, Stefan Kiechl, Manuel Mayr. Cardiovascular Research , 2011.

Systems Biology View of miRNAs

GeneLab
Open Science for Exploration

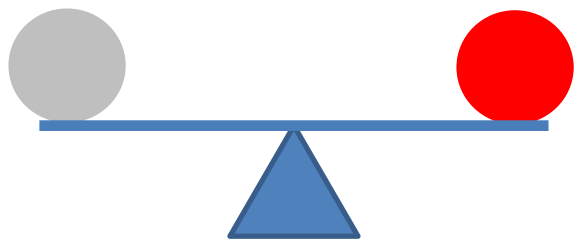
Tumor Suppressor miRNAs

OncomiRNAs

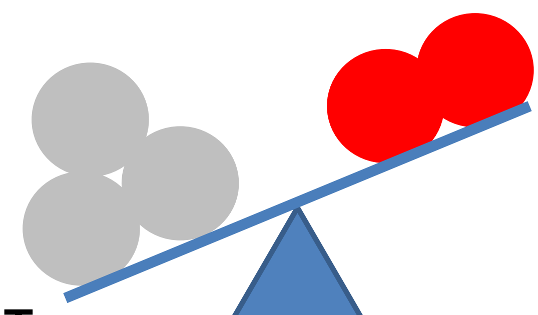


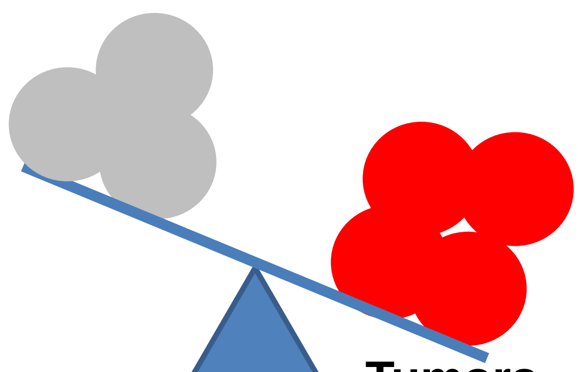
Only looking at a single miRNA

In pursuit of the elusive miRNA...



looking at a pair of miRNAs





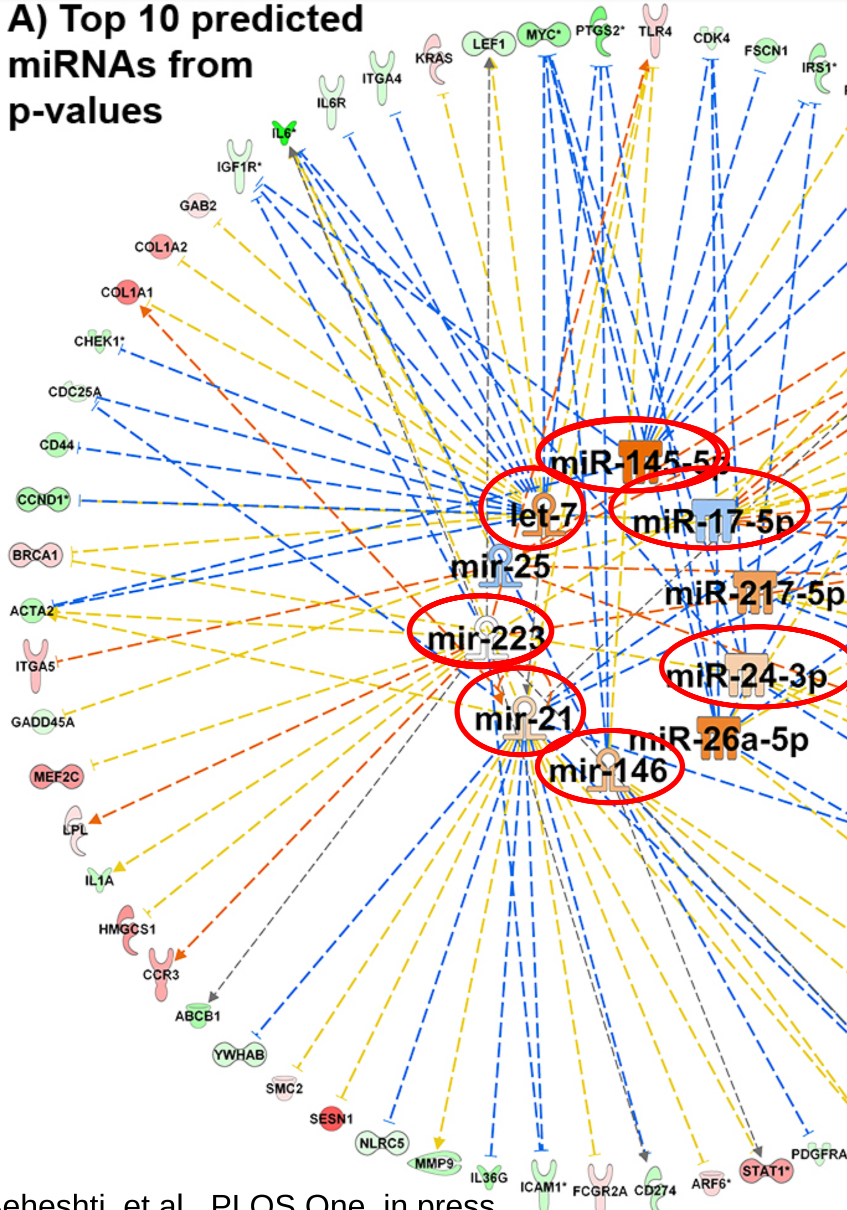
<u>Systems Biology Approach:</u> Looking at how the entire system impacts the most Important miRNAs

Predicted miRNAs Involved with Microgravity Effects

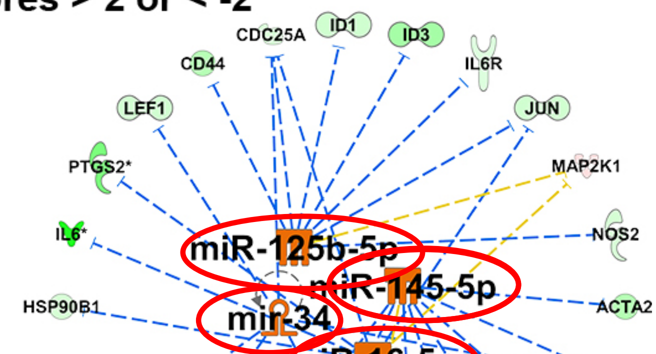
GeneLab
Open Science for Exploration

NASA

A) Top 10 predicted miRNAs from p-values



B) All miRNAs with Z-scores > 2 or < -2



Experimental & Molecular Medicine (2017) 49, e348; doi:10.1038/emm.2017.80

PLOS one

OPEN

Hindawi Publishing Corporation
BioMed Research International
Volume 2014, Article ID 296747, 16 pages
http://dx.doi.org/10.1155/2014/296747

Hindawi

Research Article

Integration Analysis of MicroRNA and mRNA Expression Profiles in Human Peripheral Blood Lymphocytes Cultured in Modeled Microgravity

C. Girardi,[1] C. De Pittà,[1] S. Casara,[1] E. Calura,[1] C. Romualdi,[1] L. Celotti,[1,2] and M. Mognato[1]

[1] *Dipartimento di Biologia, Università degli Studi di Padova, Via U. Bassi 58/B, 35131 Padova, Italy*
[2] *Laboratori Nazionali di Legnaro, INFN, Viale dell'Università 2, Legnaro, 35020 Padova, Italy*

Correspondence should be addressed to L. Celotti; lucia.celotti@unipd.it and M. Mognato; maddalena.mognato@unipd.it

Received 16 April 2014; Revised 22 May 2014; Accepted 22 May 2014; Published 23 June 2014

Academic Editor: Mariano Bizzarri

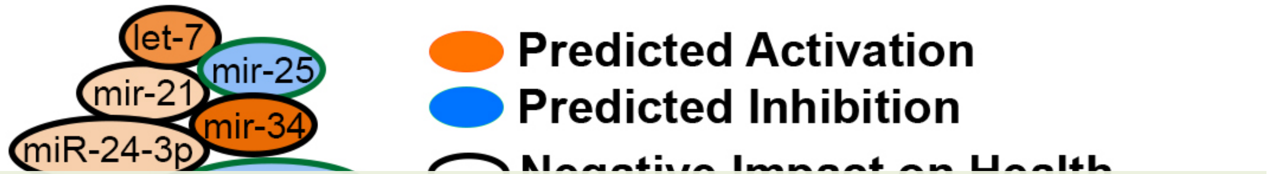
We analyzed miRNA and mRNA expression profiles in human peripheral blood lymphocytes (PBLs) incubated in microgravity condition, simulated by a ground-based rotating wall vessel (RWV) bioreactor. Our results show that 42 miRNAs were differentially expressed in MMG-incubated PBLs compared with 1 g incubated ones. Among these, miR-9-5p, miR-9-3p, miR-155-5p, miR-150-3p, and miR-378-3p were the most dysregulated. To improve the detection of functional miRNA-mRNA pairs, we performed gene expression profiles on the same samples assayed for miRNA profiling and we integrated miRNA and mRNA expression data. The functional classification of miRNA-correlated genes evidenced significant enrichment in the biological processes of immune/inflammatory response, signal transduction, regulation of response to stress, regulation of programmed cell death, and regulation of cell proliferation. We identified the correlation of miR-9-3p, miR-155-5p, miR-150-3p, and miR-378-3p with that of genes involved in immune/inflammatory response (e.g., IFNG and IL17F), apoptosis (e.g., PDCD4 and PTEN), and cell proliferation (e.g., NKX3-1 and GADD45A). Experimental assays of cell viability and apoptosis induction validated the results obtained by bioinformatics analyses demonstrating that in human PBLs the exposure to reduced gravitational force increases the frequency of apoptosis and decreases cell proliferation.

Beheshti, et al., PLOS One, in press

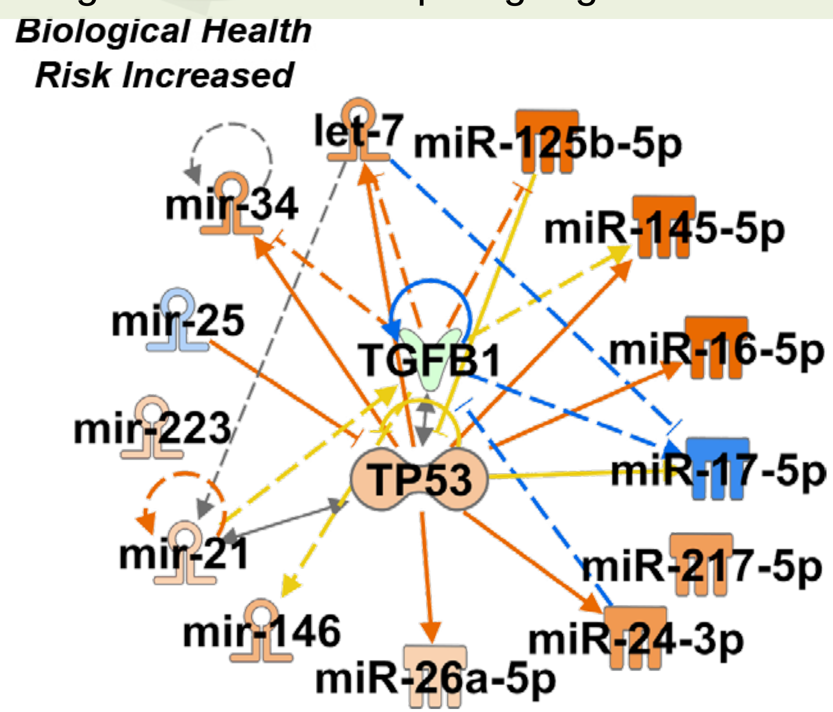
GeneLab
Open Science for Exploration

Predicted miRNAs Involved with Microgravity Effects

Health Risk Due to miRNAs

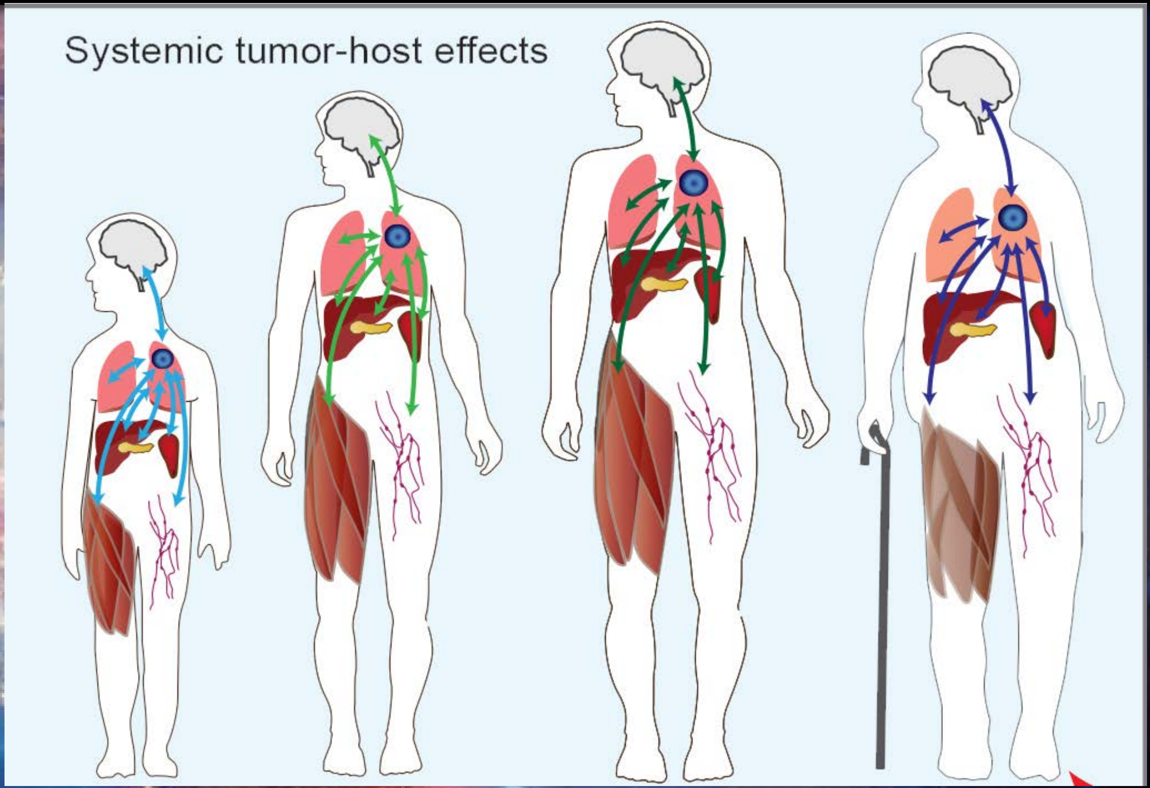


- A recent report showed that inactivation of p53 altered TGF-β signaling, which ironically displayed both tumor-suppressive and pro-oncogenic functions. p53 functions to integrate crosstalk between Ras/MAPK and TGF-β signaling via binding to Smad3, dislocating the Smad3/Smad4 complex formation and differentially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



Beheshti, et al., PLOS One, in press

Conclusions for this Study

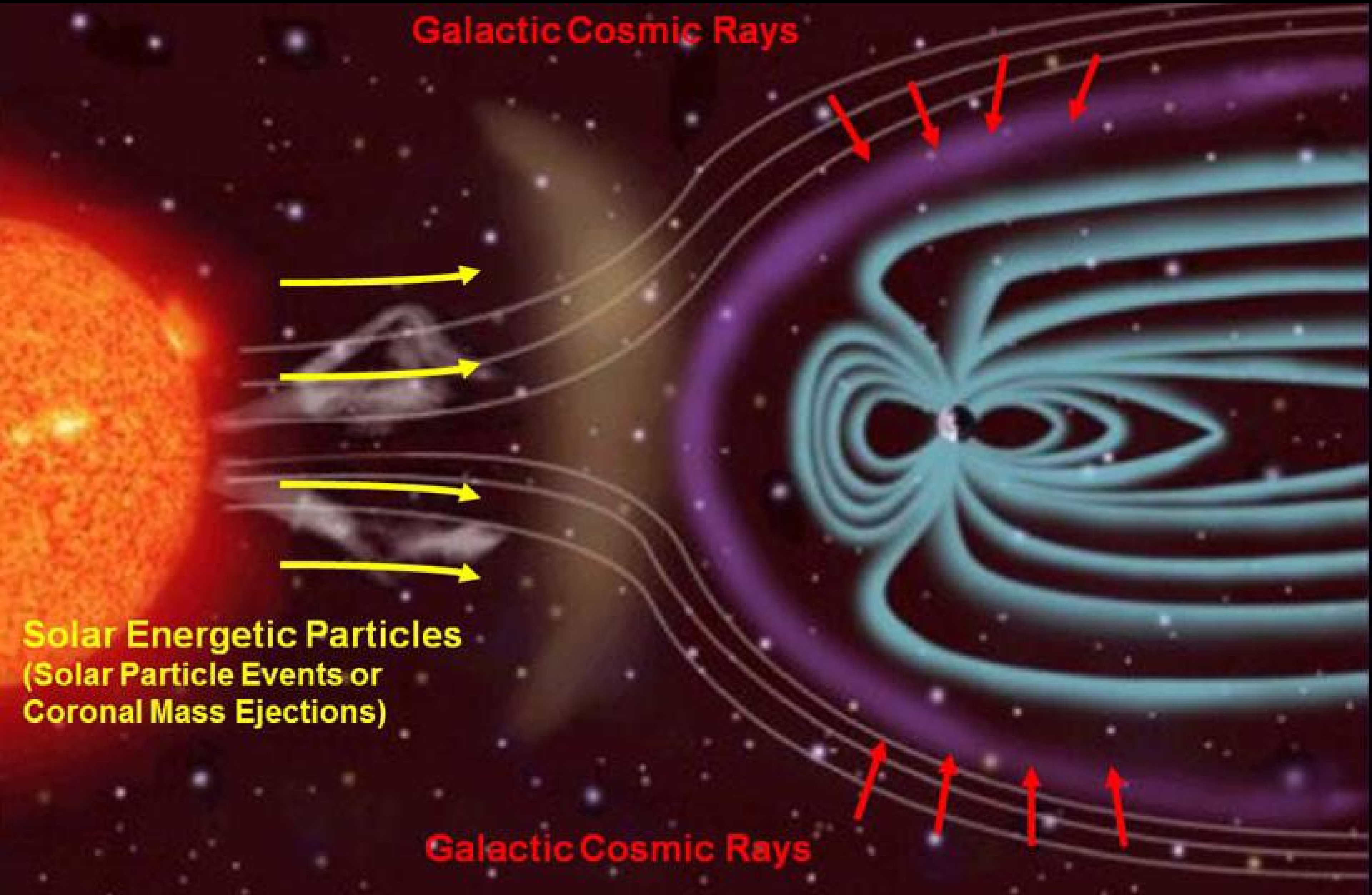


- Systems biology approach allows for systemic understanding of the impact of Microgravity.
- Circulating miRNAs can influence overall progression of health risk to the host.
- miRNAs can potentially be used for novel minimally invasive therapeutics and countermeasures
- GeneLab (genelab.nasa.gov) is a powerful tool to generate hypotheses and direct future space research

Space Radiation induces long term impact on the cardiovascular system by the activation of FYN through Reactive Oxygen Species

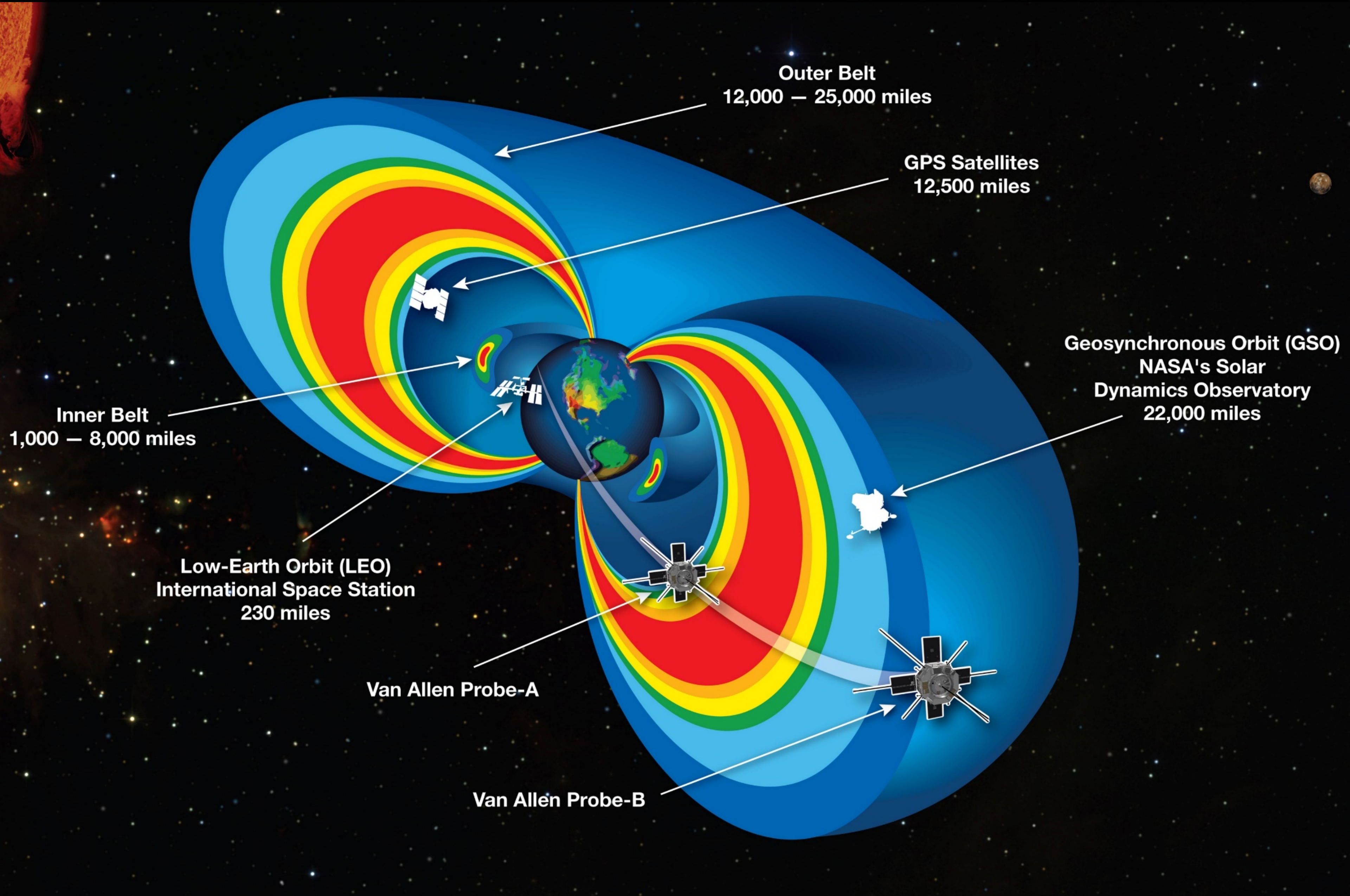
Work in progress

Space Radiation



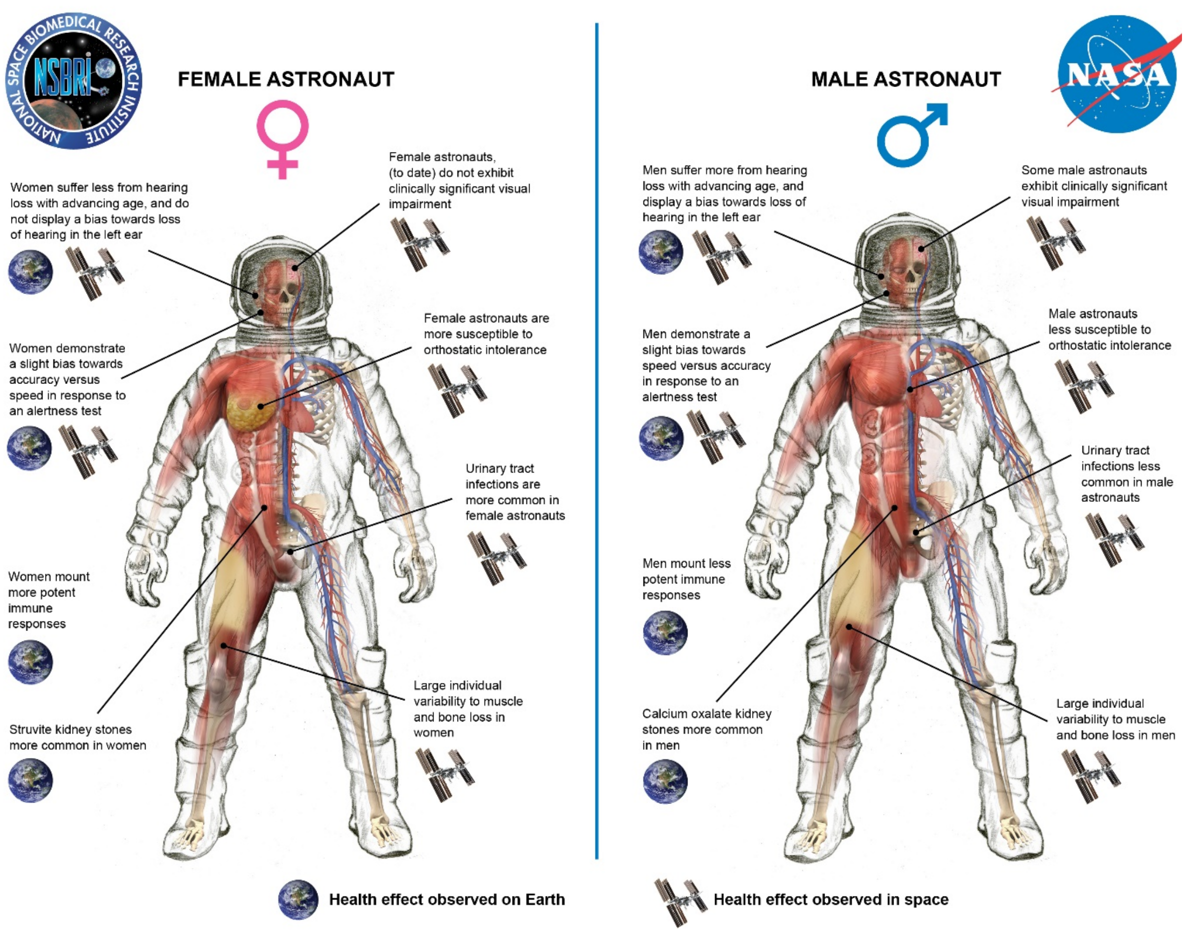
GeneLab
Open Science for Exploration

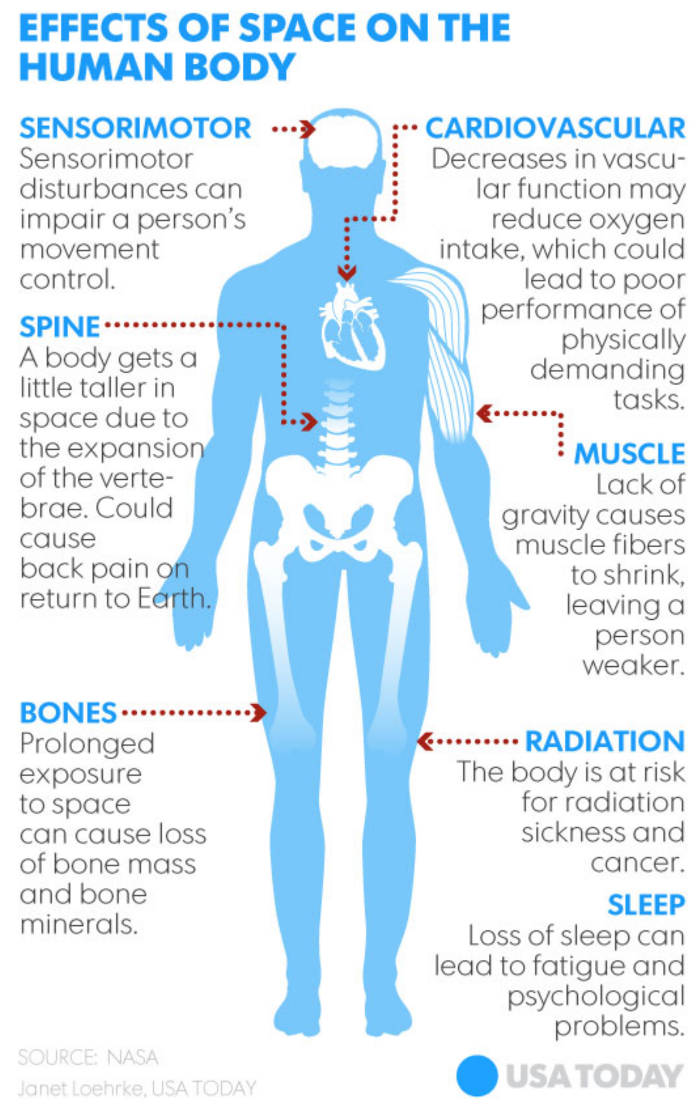
Space Radiation: Van Allen Belt



GeneLab
Open Science for Exploration

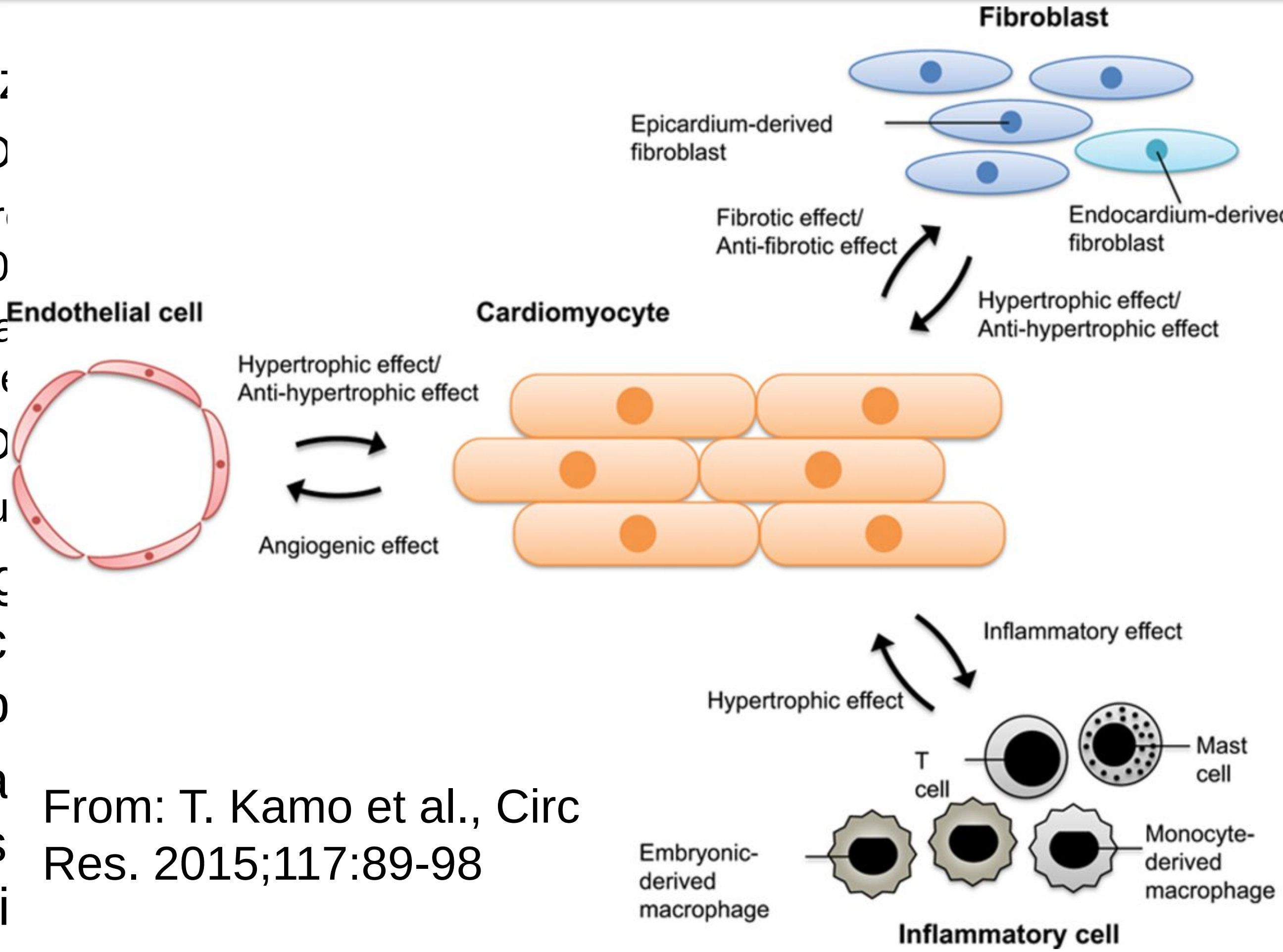
Space Radiation Risk On Astronauts





GeneLab Data Used and Hypothesis

- Analy
  - GLD
    - Ar ... exposure to 90
    - Ca ... rdiac muscle (he
  - GLD
    - hu ... e ISS
- The g ... long-term impac ... 5 different time p
- Our a ... common persis ... n due to radiati

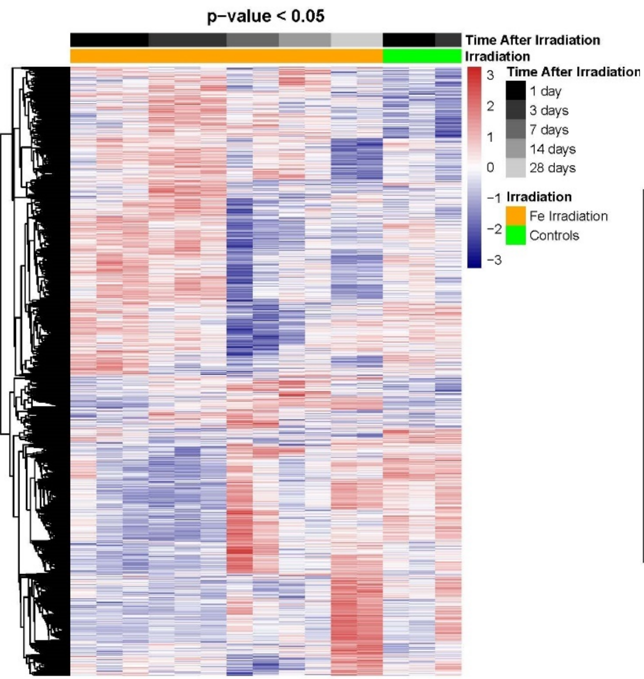


From: T. Kamo et al., Circ Res. 2015;117:89-98

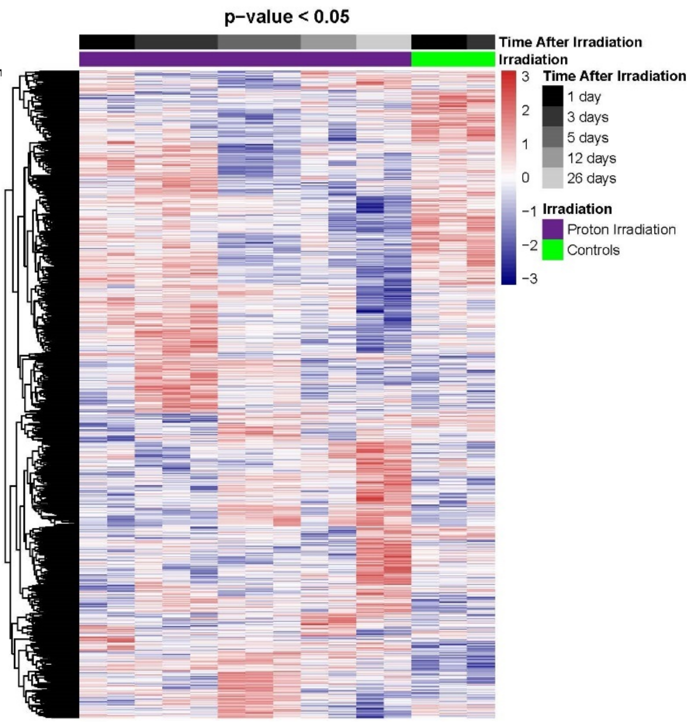
- **Endothelial cells are known to directly regulate the development and activity of cardiomyocytes, and thus their response to spaceflight should be highly correlated with cardiomyocytes**.

Basic Analysis of Data

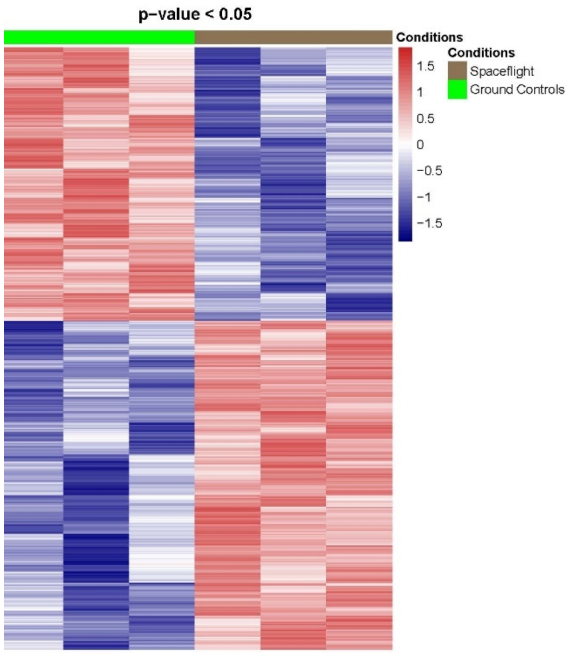
GLDS-109: ^{56}Fe



GLDS-117: Proton

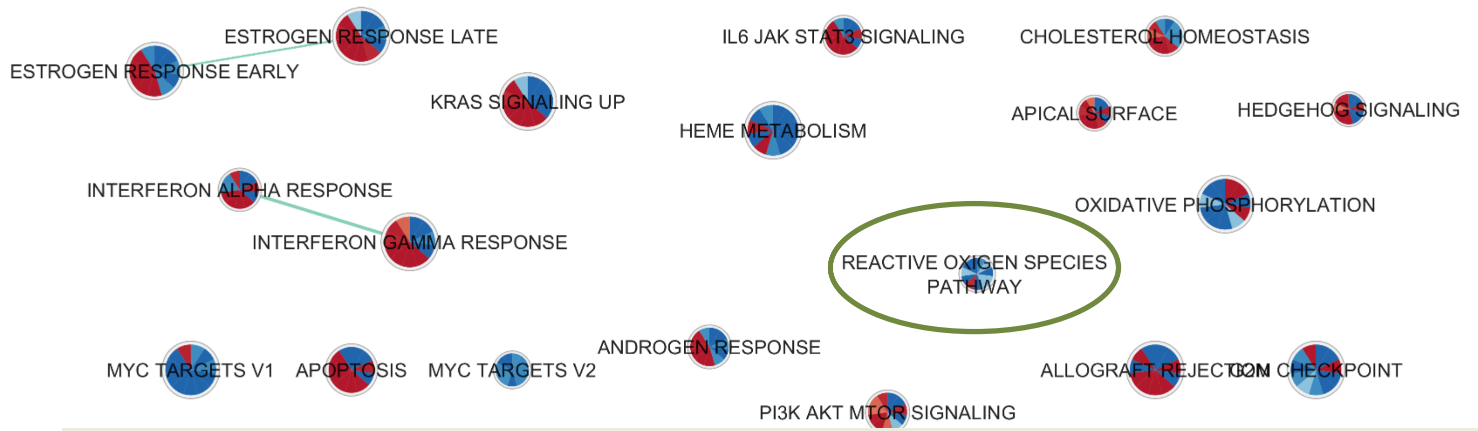


GLDS-52: HUVEC

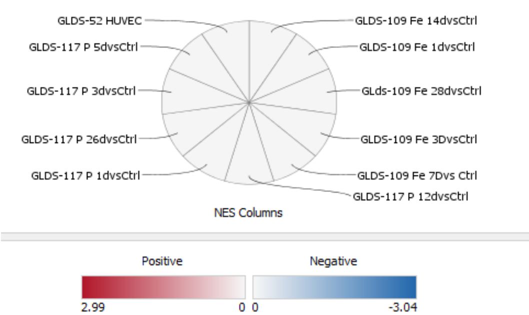


Focused on biological factors that will continue to have impact for all time points after irradiation.

Functional Impact on Health: GSEA

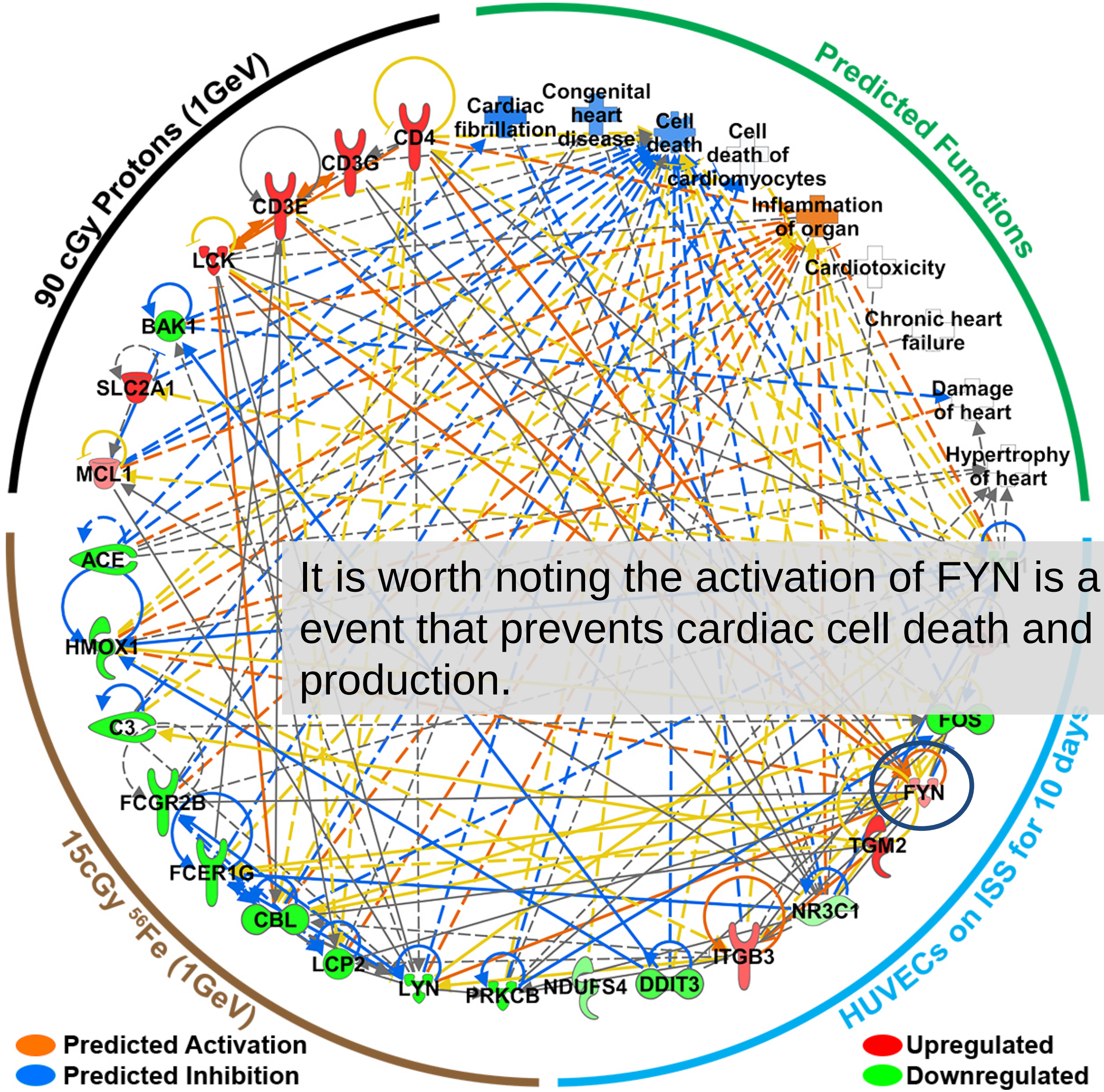


- ROS are formed as a natural byproduct of the normal metabolism of oxygen and have important roles in cell signaling and homeostasis.
- During times of environmental stress (e.g., radiation exposure), ROS levels can increase dramatically.
- This may result in significant damage to cell structures.
- Cumulatively, this is known as oxidative stress.
- ROS are also generated by exogenous sources such as ionizing radiation

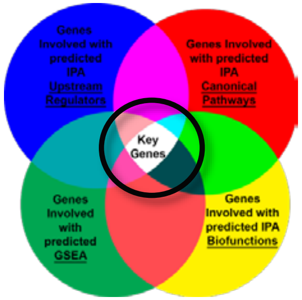


Key Drivers Involved with Space Radiation Induced Cardiovascular Risk

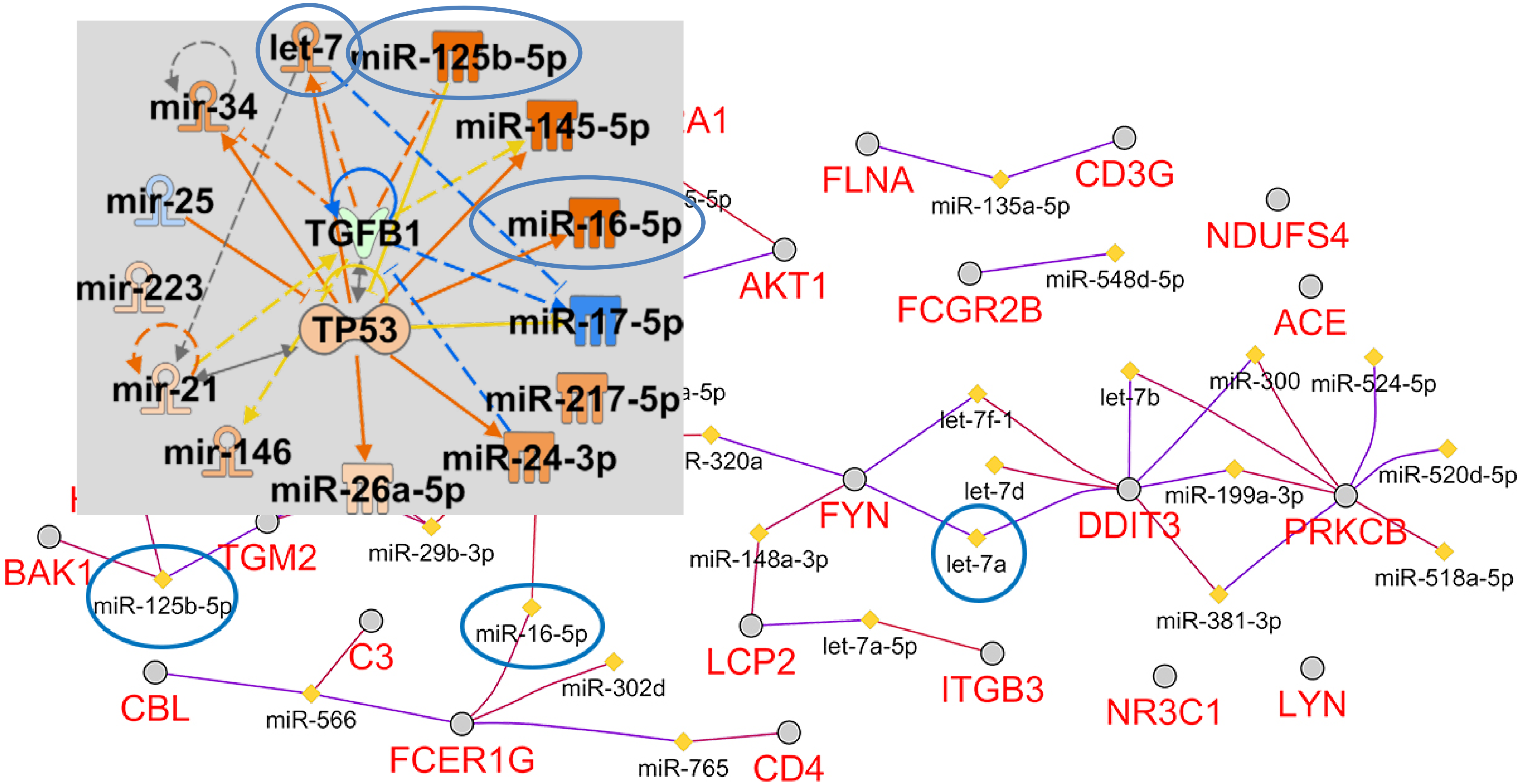
GeneLab
Open Science for Exploration



It is worth noting the activation of FYN is a key event that prevents cardiac cell death and ROS production.



miRNA impact?



Conclusions form this Study

- Space radiation downregulate ROS functions
- Key/driving genes: FYN, LCK, AKT1 are upregulated and LYN and FOS are downregulated with FYN being the central driver/hub for the cardiovascular response to space radiation.
  - FYN is a key event that prevents cardiac cell death and ROS production.
- From our study we thus hypothesize that a feedback loop occurs from the oxidative stress caused by space radiation that upregulates FYN which in turn reduces ROS levels and thus ROS pathways, preventing cardiomyocyte and endothelial cell death and thus protecting the cardiovascular systems.
- We believe that this is a novel mechanism for space radiation induced cardiovascular risk directly linking radiation ground studies to spaceflight.

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

NASA


Kathleen Fisch


Brin Rosenthal

UNIVERSITY *of* CALIFORNIA, SAN DIEGO
SCHOOL OF MEDICINE


Deanne Taylor


Hossein Fazelinia


Komal Rathi

Children's Hospital of Philadelphia

Perelman School of Medicine
UNIVERSITY *of* PENNSYLVANIA


Helio Costa


Kathryn Grabek

STANFORD UNIVERSITY


J. Tyson McDonald

HAMPTON UNIVERSITY
THE STANDARD OF EXCELLENCE


Gary Hardiman


Willian da Silveira

MUSC Health
MEDICAL UNIVERSITY *of* SOUTH CAROLINA

GeneLab
Open Science for Exploration

AWG Members Involved

NASA


Chris Mason


Cem Meydan


Jonathan Foox

Flavia Rius

Cornell University


Yared Kidane

NASA
TEXAS SCOTTISH RITE HOSPITAL FOR CHILDREN


Susana Zanello


Scott Smith


Sara Zwart

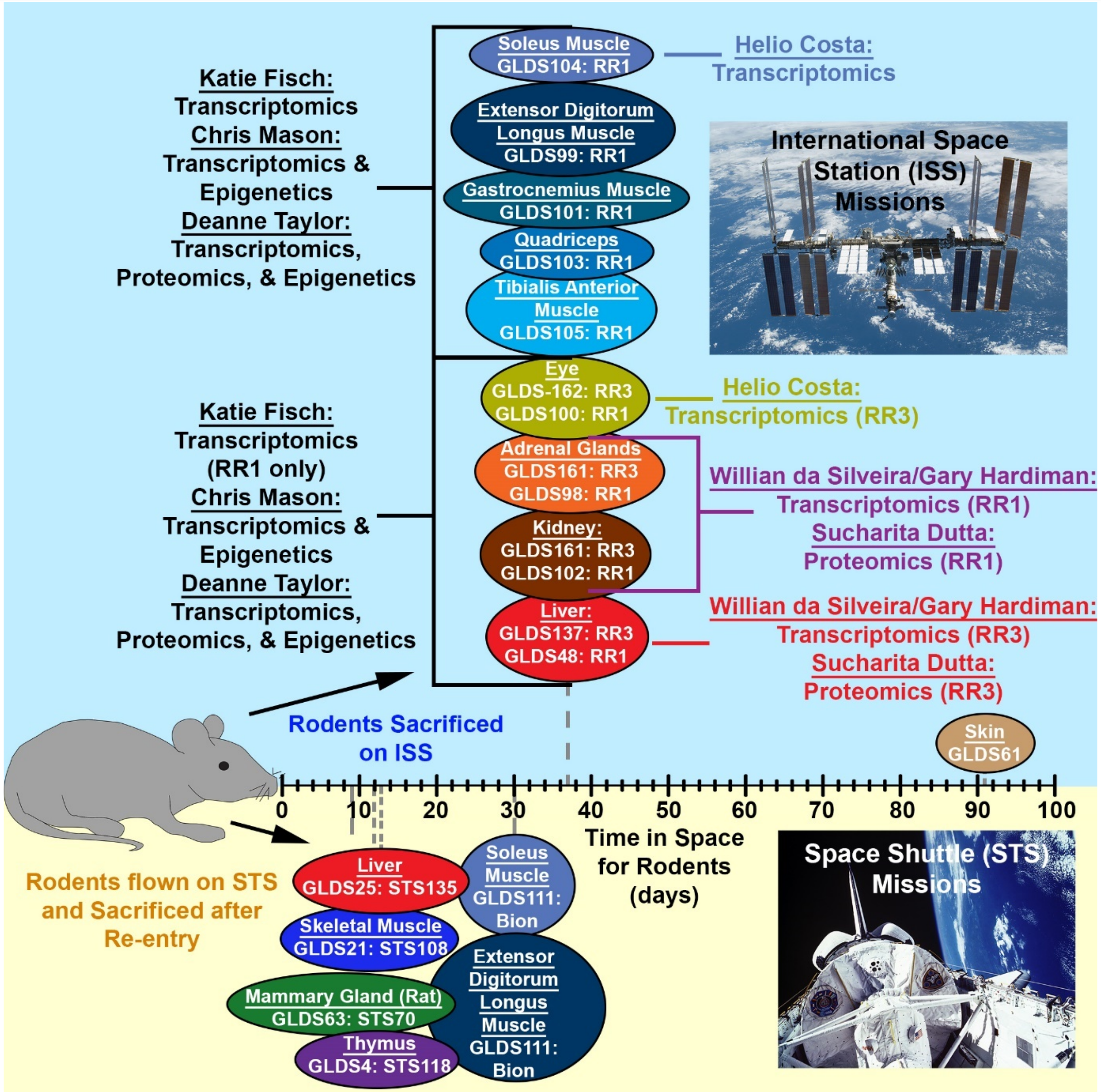
NASA
SPACE CENTER HOUSTON
Manned Space Flight Education Foundation

Afshin Beheshti


Sylvain Costes

NASA
Ames Research Center

Specific Datasets and Tissues AWG Members Analyzed

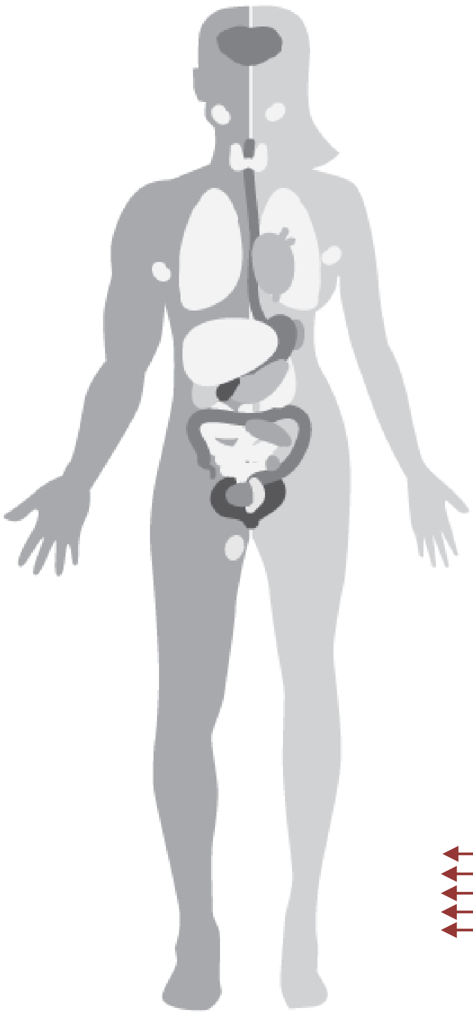


- Additional Datasets that are being analyzed:
  - Human datasets
    - GLDS-54, GLDS-174, GLDS-86, GLDS-118, GLDS-53, GLDS-54, GLDS-13. GLDS-52, or GLDS-114 (Tyson McDonald and Yared Kidane)

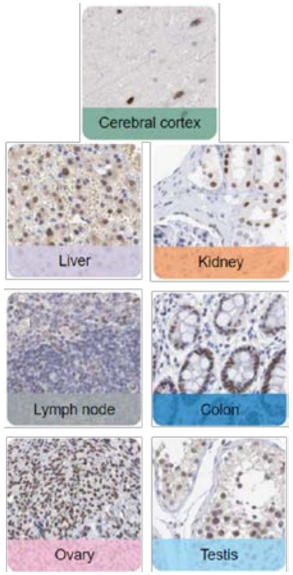
GeneLab
Open Science for Exploration

Structural Hierarchy of the Human Body

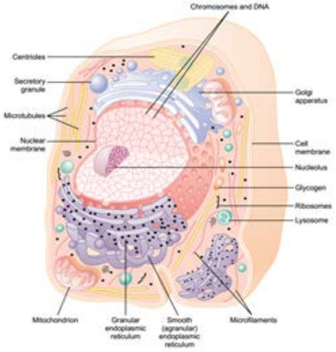
Level 7: The Body
Level 6: Organs



Level 5: Tissues



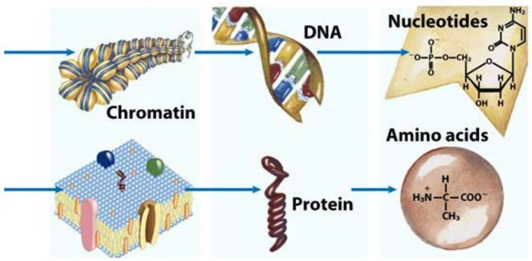
Level 4: The Cell and its Organelles



Level 3: Supra molecular Complexes

Level 2: Macro molecules

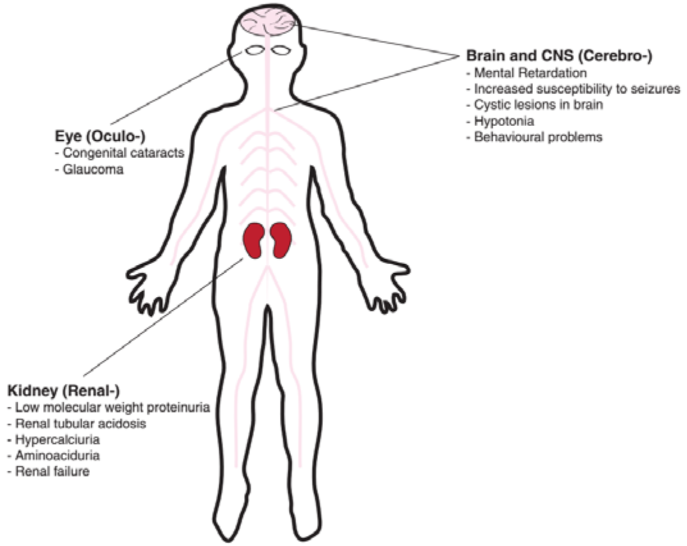
Level 1: Monomeric Units



Hypothesis Developed and Being Worked On

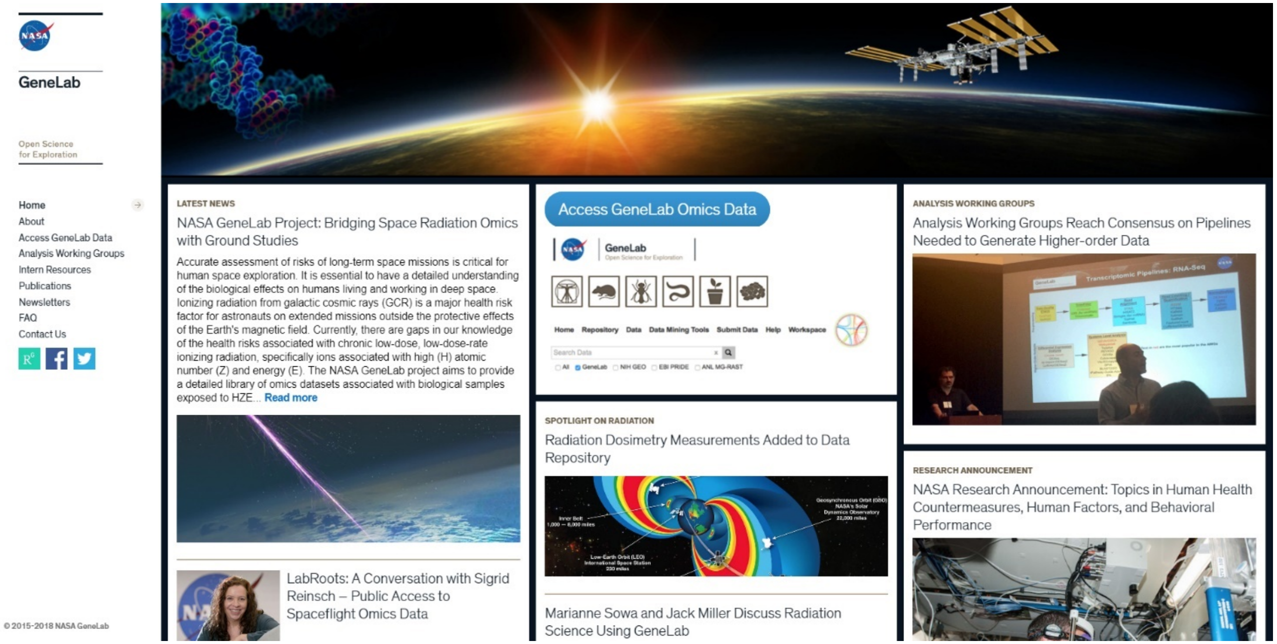
- Spaceflight changes the physical properties of the cell components impacting from the molecular to the whole body level.
- The Mitochondria are the principal cellular component affect.
- The Liver is the principal organ affected in issues related to the metabolism.
- Possible disease that can be associated with liver damage and pathways is: Oculocerebrorenal Syndrome of Lowe
  - "Extensive research has demonstrated that OCRL-1 is involved in multiple intracellular processes involving endocytic trafficking and actin skeleton dynamics. This explains the multi-organ manifestations of the disease."
  - "The classic form of the oculocerebrorenal syndrome of Lowe (OMIM #309000), first described by Lowe et al. in 1952 [1], is characterized by the triad of congenital cataracts, severe intellectual impairment, and renal tubular dysfunction with slowly progressive renal failure"
  - Patients with this disease manifest Cataract, Glaucoma and Muscle hypotonia.

Schematic diagram showing the organs affected in Lowe syndrome



Mehta, Zenobia B et al. "The Cellular and Physiological Functions of the Lowe Syndrome Protein OCRL1." *Traffic* (2014).

Engaging with GeneLab



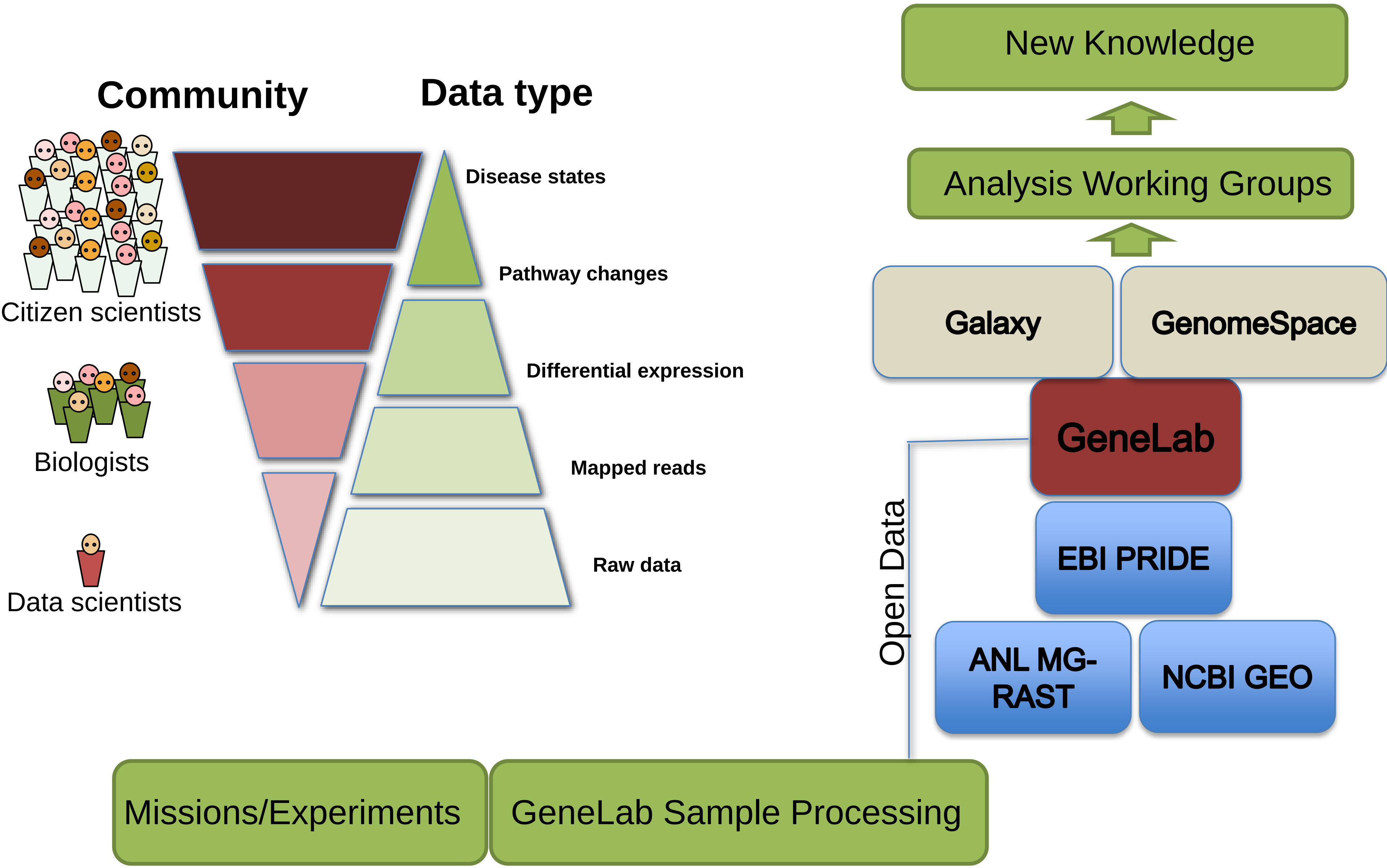
https://genelab.nasa.gov

Participate in GeneLab Analysis Working Groups

- Social media :
  - @NASAAmes **Facebook**
  - **Twitter** #GeneLab
  - **ResearchGate**: https://www.researchgate.net/project/Omics-for-Space-Biology-The-GeneLab-project

GeneLab
Open Science for Exploration

GeneLab



GeneLab
Open Science for Exploration

GeneLab Acknowledgements

Chris Barreras
Afshin Beheshti
Dan Berrios
Valery Boyko
Sonja Caldwell
Jairon Camarillo
Egle Cekanaviciute
John Costa
Sylvain Costes (PM)
Marie Dinh
Sandy Dueck
Homer Fogel
Jon Galazaka (PS)
Samrawit Gebre
Dennis Heher
Lynn Hutchison
Yared Kidane
San-Huei Lai Polo
Tristan Le
Qiang Li
Shu-Chun Lin
Sneha Raghunandan
Shayoni Ray
Sigrid Reinsch
David Smith
Marla Smithwick
Hao Thai
Khai Peter Tran
Andrew Williamson