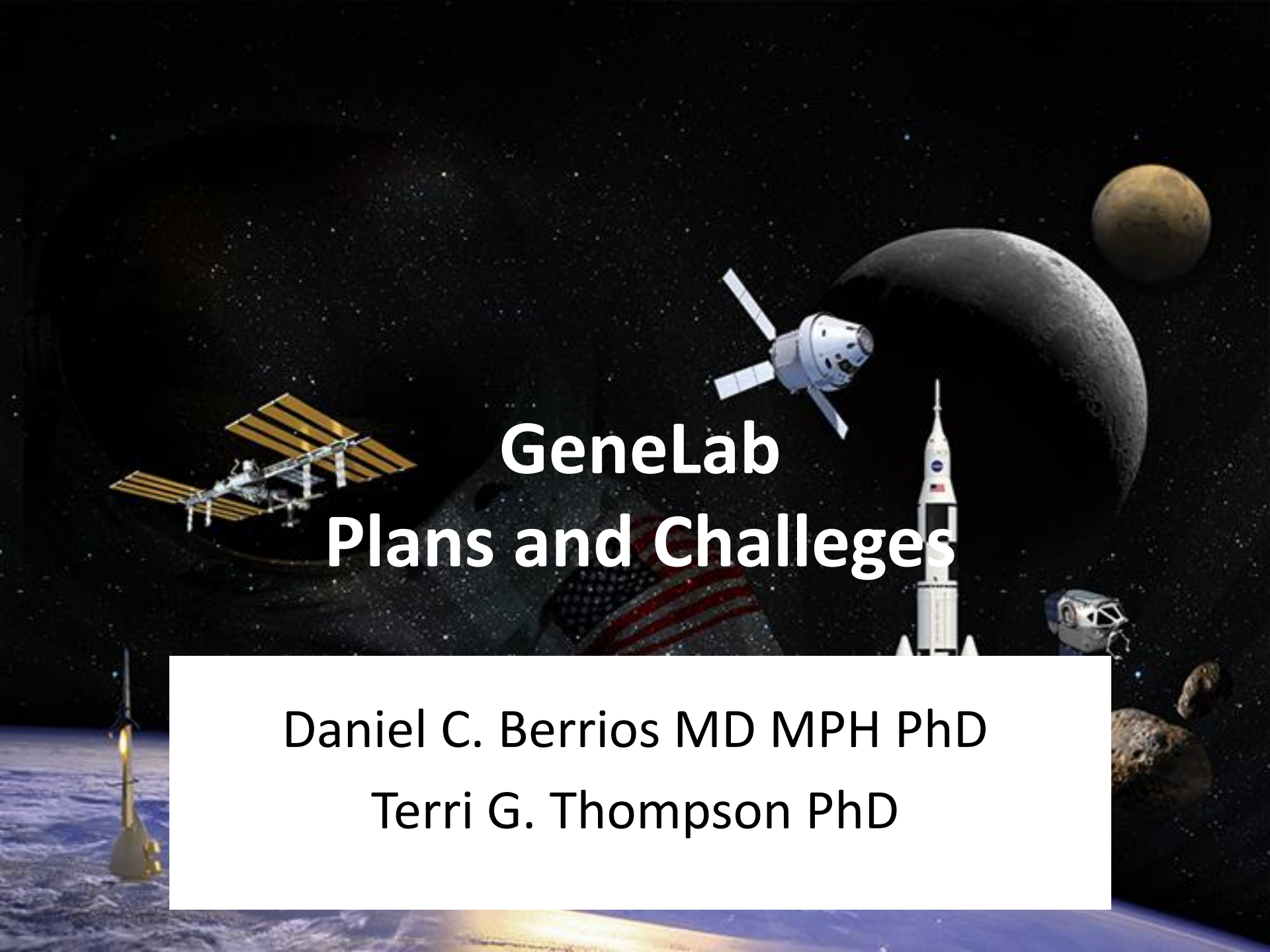


GeneLab
Plans and Challeges

Daniel C. Berrios MD MPH PhD

Terri G. Thompson PhD

Flight Research: International Space Station

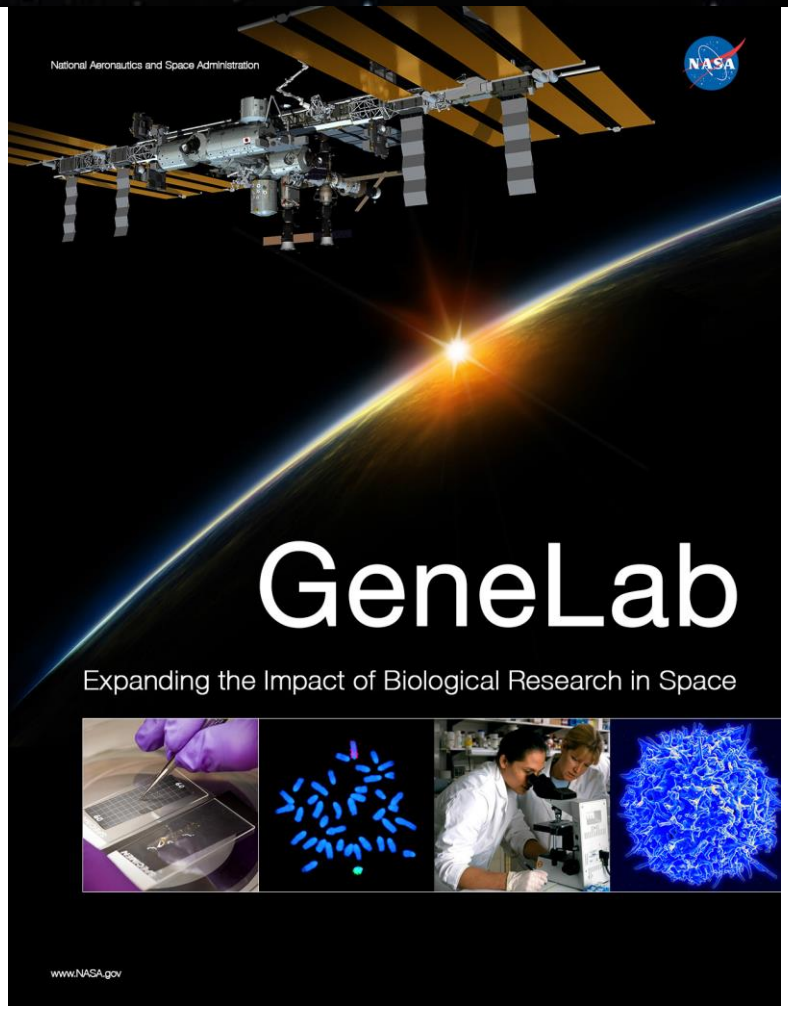
- Almost as soon as the International Space Station was habitable, researchers began using it to study the impact of microgravity and other space effects on several aspects of our daily lives.

- This unique scientific platform continues to enable researchers from all over the world to put their talents to work on innovative experiments that could not be done anywhere else.

GeneLab Strategic Plan

"… there is a significant demand from the scientific community for NASA-funded research opportunities on the ISS that cannot be met using traditional management tactics based on single Principal Investigator (PI)-led investigations. Given the factors that limit the scientific output required to address the significant biological problems required for human exploration beyond low-earth orbit, *NASA will develop and implement a new multi-investigator approach based on high content bioinformatics analytics and with open science and data*."

-GeneLab Strategic Plan, 2014



GeneLab Goals

The GeneLab Goals

1. “Develop an integrated repository and bioinformatics data system for analysis and modeling “

2. “Enable the discovery and validation of molecular networks that are influenced by space conditions through ground-based and flight research using next-generation omics technologies”

3. “Engage the broadest possible community of researchers, industry, and the general public to foster innovation”

4. “Strengthen international partnerships by leveraging existing capabilities and data sharing “

-GeneLab Strategic Plan, 2014

Motivations

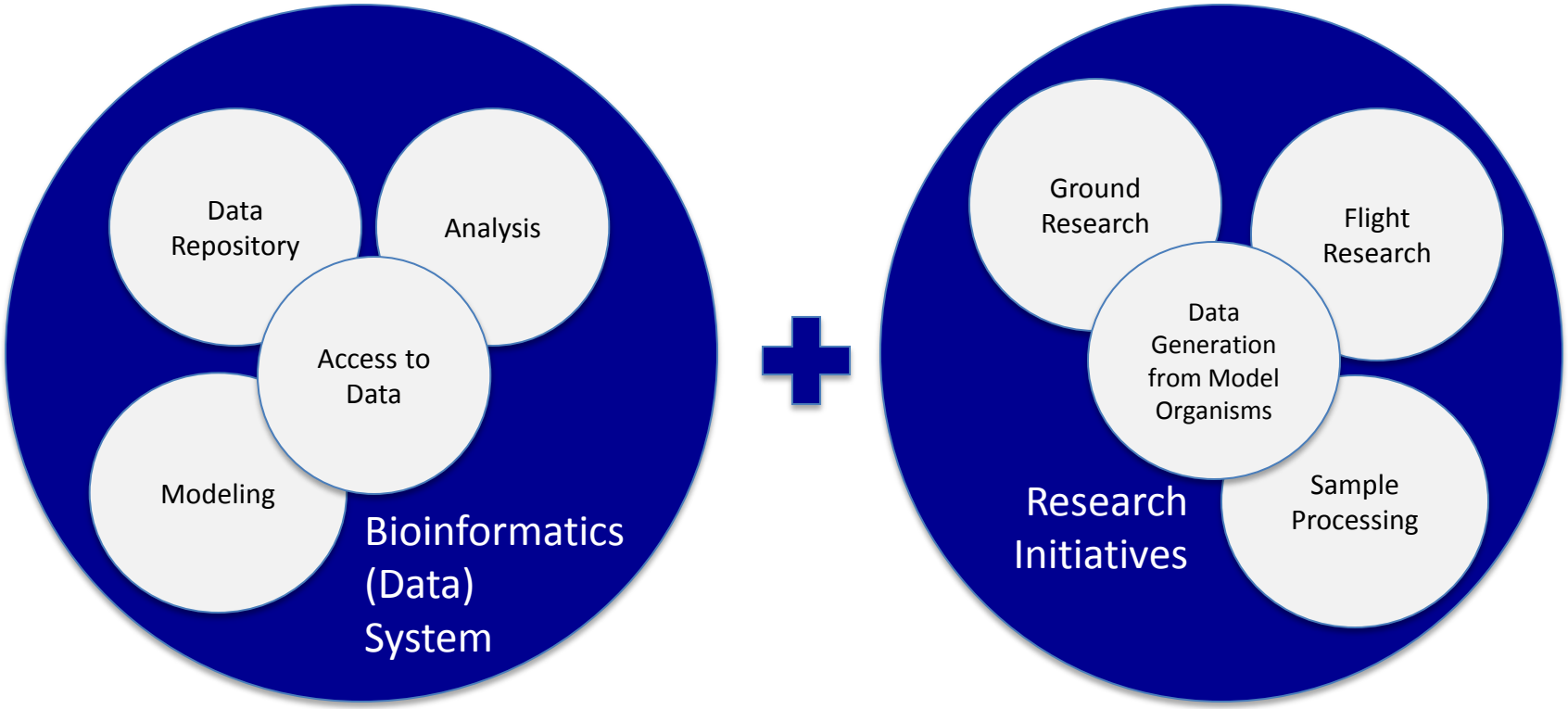
- **Maximize ROI for ISS Utilization**: Open-access, systems-biology spaceflight experiments will provide foundational science that maximizes return-on-investment for rare and costly spaceflight opportunities and remove research "bottleneck"
- **Create a PI Multiplier Effect**: Open access greatly expands the community of researchers using ISS derived data for investigations – ISS research investments will yield numerous follow-up investigations and next generation hypothesis-driven research
- **Leverage NASA and External Partner Strengths**: Brings together NASA's strengths in Space Biology and "big data" analysis with commercial, government and international partners through a scaled and iterative approach that capitalizes on existing databases, analytical tools and biotech capabilities
- **Maximize Utilization of Cutting Edge Bioanalytical Tools and Techniques**: Multiple omics datasets and integrated data system allows scientists to interrogate ISS derived samples using state-of-the-art high throughput genomics, proteomics, metabolomics and bioinformatics tools
- **Speed the Pathway to Translation**: Allows researchers to discover emergent properties in data to identify and understand pathways/macromolecules influenced by space stressors
- Directly responsive to 2011 Decadal Survey Recommendations and OSTP Open Data Initiative.
- Similar goals and efforts NASA ARMD and SMD to increase ROI and PI Multiplier effect and engage broad communities. ARMD/AvS/DashLink, SMD/SS/PDS, SMD/ES/EOS-EOSDIS & NEExchange.

ISS Research

- The unique conditions and workflows aboard the ISS are captured for a flight experiment and associated with the results.
- This information is ingested, stored, indexed and distributed from NASA's Life Science Data Archive.
- NASA GeneLab is expected to capture and distribute 'omics' data and experimental and process conditions most relevant to research community in their statistical and theoretical analysis of NASA's omics data.

NASA astronaut Barry "Butch" Wilmore setting up the Rodent Reseach-1 Hardware in the Microgravity Science Glovebox aboard the International Space Station

Background: GeneLab Vision



Vision: A centralized collaboration space for data deposition, retrieval, analysis and modeling to develop next generation science related to spaceflight

GeneLab Organization

Key:
- HQ and SLPS
- ARC
- JSC

Space Life and Physical Sciences
NASA HQ
Nicki Rayl

ISS Program
Sharon Conover/JSC

Space Biology Project
PM – Jeff Smith

GeneLab
Project Office
PM – Joe Coughlan
DPM – Yung Nguyen

Data Architect
D. Berrios

Data Systems

Design, planning, implementation, testing, verification, and deployment of data systems.

Science & Outreach
Terri Thompson

Managing, directing, and controlling of GeneLab science investigation aspects.

Payload Planning
L Timucin

Development and execution of GeneLab payload plans, timelines, schedules and missions

Sample Processing & Analysis
K. Chakravarty

Designing QC on spaceflight ground samples, cataloging payload samples

Data Processing /Operations
H. Fogle

Quality assessment, processing pipelines, analysis, modeling, algorithm development

GeneLab Users

- **Traditional Space Biology PI Community** – zoologists, botanists, microbiologists, cellular and molecular biologists

- **Non-Traditional Genomics and Systems Biology Communities**

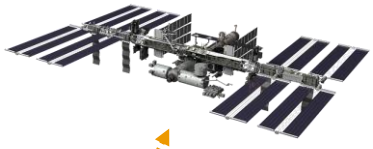
- **NASA Human Research Program Omics Data** – GeneLab ***potential*** host for One Year Twins Study Data – Pilot Study for Human Omics Data. ***Potential*** repository for ground radiation studies using model organisms

- **CASIS** – Opportunity for data mining to identify targets for drug development, personalized medicine and systems biology of model organisms and humans yielding knowledge for terrestrial applications and acceleration of translation pathway
  - CASIS funded investigators encouraged to submit raw data to GeneLab
    - CASIS participated in 2013 GeneLab RFI
    - CASIS representation at September 2013 Omics Workshop
    - CASIS representation on continuing GeneLab Steering Committee and planning efforts

Concept of Operations

Experiment on ISS

Crew performs experimental protocol and harvests tissues.

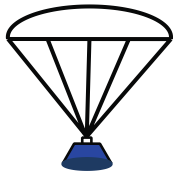


Launch

Experiment is prepared and launched according to approved NRA.

Return to Earth

Material sent back to earth for processing in investigators lab. Controls (ground and/or flight) processed at the same time.

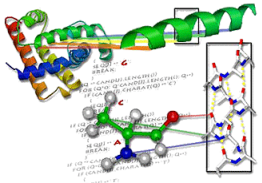


Process Samples

Extracted DNA, RNA and/or protein sent to validated omics center to generate sequence, transcript or protein expression data.

Data Collection & Hosting

Data returned to investigator or GeneLab for analysis. Raw data uploaded into GeneLab database for pubic viewing.



Modeling and Validation

Wet lab validation and computational modeling.

Next Generation Research

Iterative research solicitations for experiments utilizing GeneLab data for ground validation and next generation flight research.

Data Sharing

Data shared with larger scientific community. Results feedback to GeneLab and other databases accelerating scientific discovery by leveraging a bigger community.

Mission Types

Mission	Type	Definition	Example
Dedicated	Reference Data	Mission is entirely dedicated to GeneLab objectives; the Science Definition Team (SDT) defines the experiment and requirements; SDT is selected through the NASA Research Announcement Process	Micro-16 (tissue TBD)
Collaborative	Sample Sharing	GeneLab obtains specimens/samples from the existing PI space flight and ground control experiment	Rodent Research (Mouse) Bioculture System Validation (Mouse cells)
	Augmentation	GeneLab provides supplemental funding to a PI experiment to increase the quantity and/or type of specimens to obtain dedicated sample; augmentation requires NASA SLPS experiment review approval process	BRIC-19 (plant), BRIC-20 (plant) (BRIC=Biological Research in Canisters)

Launch – Flight Experiment Selection

Flight experiment solicitation and selection follows the NASA SLPS peer review process.

- NASA HQ through SLPS Program
- Space Biology Project Office
- Follow Decadal Survey recommendations
- Science Definition Team Solicitation Plan

http://www.nasa.gov/sites/default/files/files/NACRS_SLPSResearch_022414T.pdf

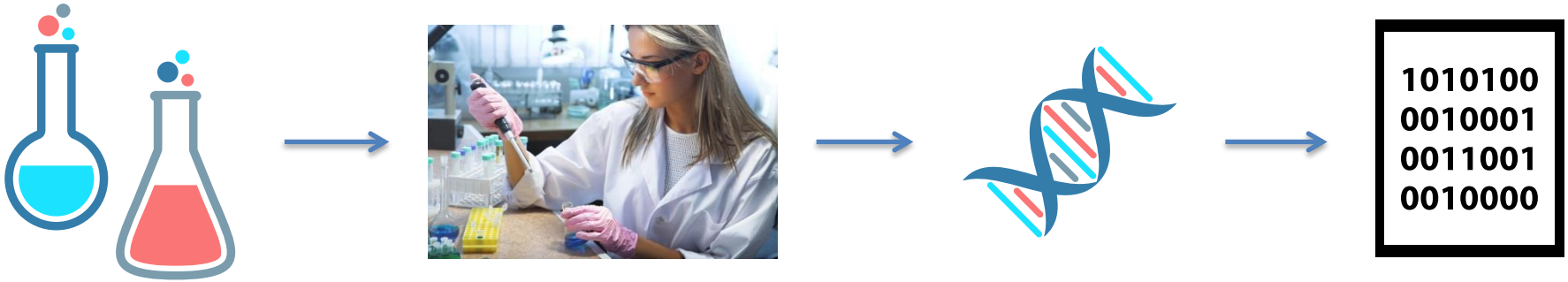
Collaborative Datasets

Year	Payload		Mission Type
2015	BRIC-19	University of Wisconsin	Augmentation, plant
	BRIC-20	Ohio University	Augmentation, Plant
	RR-1*	NASA/CASIS	Sample Sharing, Rodent
	RR-2*	Loma Linda University; Florida State University	Sample Sharing, Rodent
	Bioculture* Validation	NASA	Sample Sharing, mouse cells
	Micro-9*		PI data only

*Additional work proposed for 2015 current out of scope

Process Samples

- Flight and/or ground samples received in Laboratory
- Researcher extracts DNA, RNA and/or Protein
- Extracted material sent to center for data generation
- Data is sent back to researcher from center



Concept of Operations

Overview of the flow of data through the system

- Sample to data
- Data Collection & Hosting
- Information sharing
- Reuse of data

Experiment on ISS

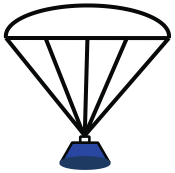
Crew performs experimental protocol and harvests tissues.

Launch

Experiment is prepared and launched according to approved NRA.

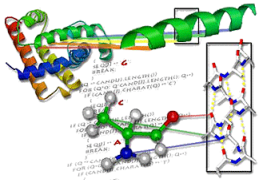
Return to Earth

Material sent back to earth for processing in investigators lab. Controls (ground and/or flight) processed at the same time.



Process Samples

Extracted DNA, RNA and/or protein sent to validated omics center to generate sequence, transcript or protein expression data.



Data Collection & Hosting

Data returned to investigator or GeneLab for analysis. Raw data uploaded into GeneLab database for pubic viewing.

Modeling and Validation

Wet lab validation and computational modeling.

Next Generation Research

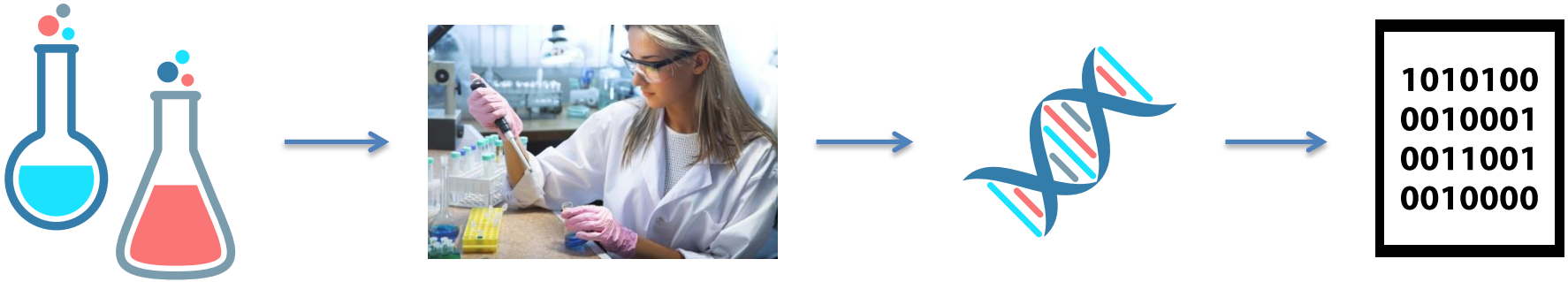
Iterative research solicitations for experiments utilizing GeneLab data for ground validation and next generation flight research.

Data Sharing

Data shared with larger scientific community. Results feedback to GeneLab and other databases accelerating scientific discovery by leveraging a bigger community.

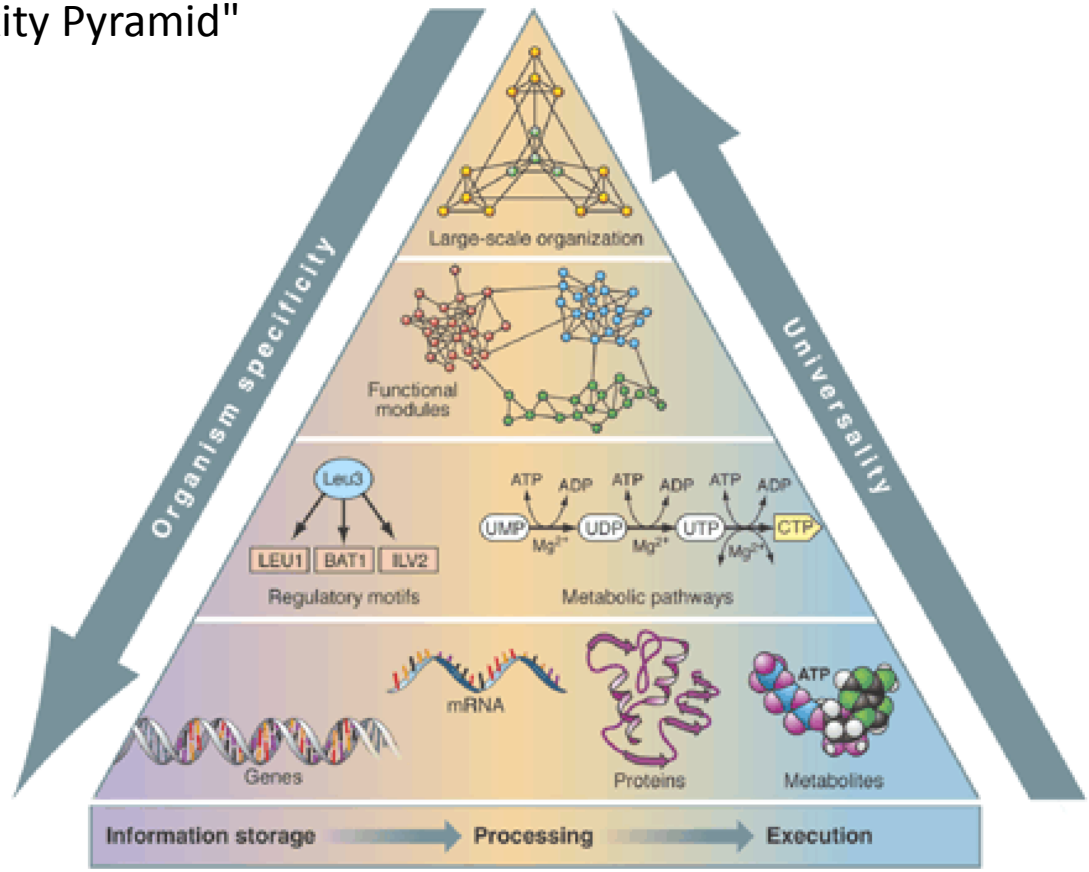
Process Samples

- Flight and/or ground samples received in Laboratory
- Researcher extracts DNA, RNA and/or Protein
- Extracted material sent to center for data generation
- Data is sent back to researcher from center



Process Samples – Omics Data

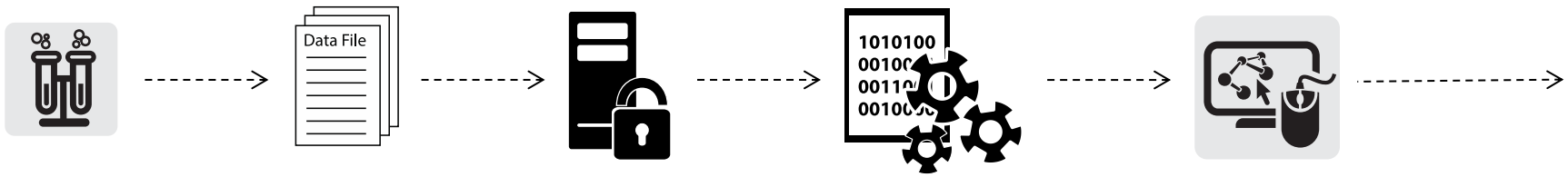
"Life's Complexity Pyramid"



(from Oltvai-Barabasi, Science, Oct 02)

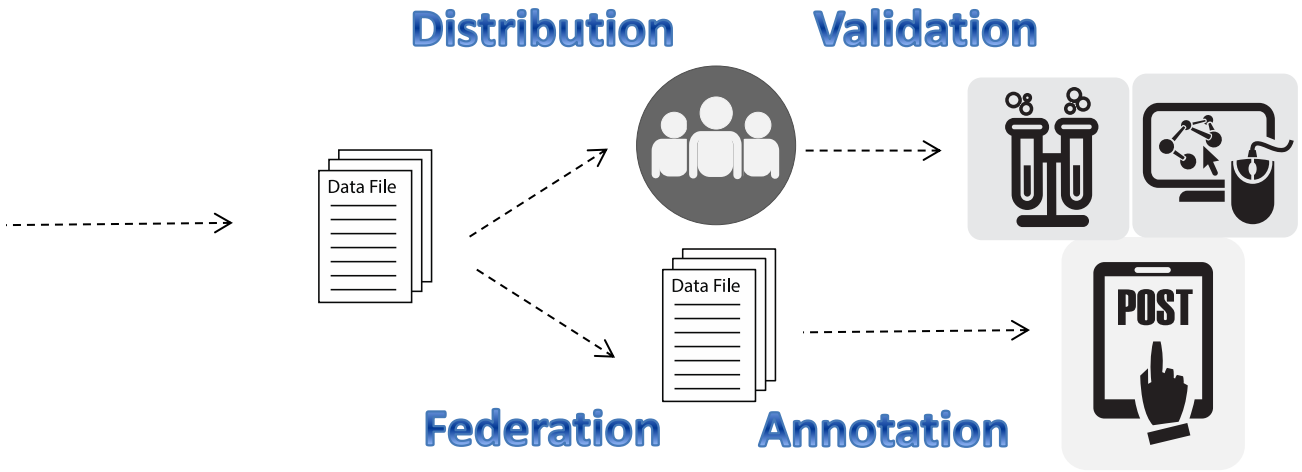
Data Collection and Hosting

- Flight and/or ground samples are processed to generate Omic data
- Raw and metadata are assessed for completeness and quality by GeneLab personnel
- Completed data is uploaded into data system
- Datasets are published in the system and available to the public

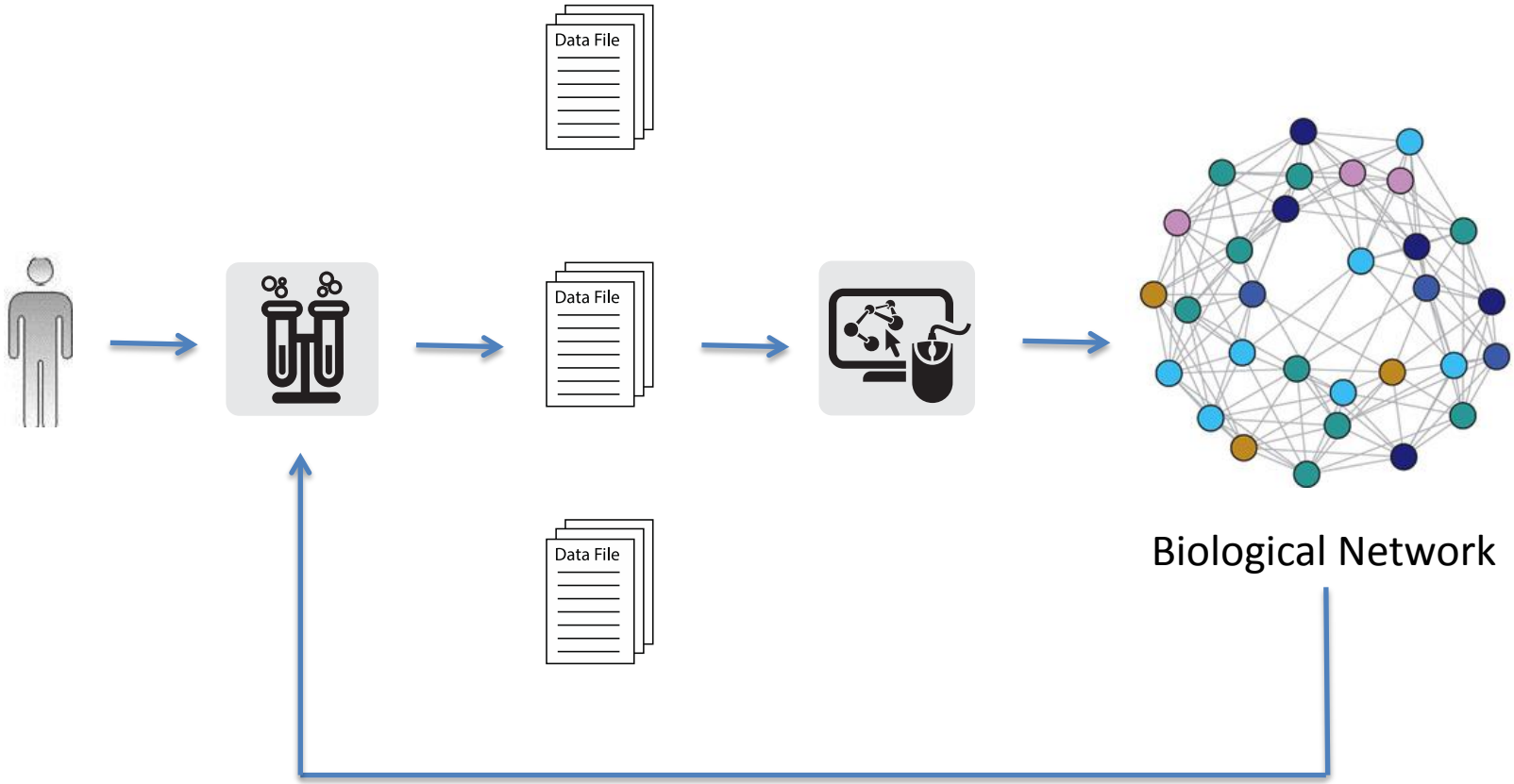


Data Sharing

- Space biology research data hosted by GeneLab is freely available to all
- Researchers can download datasets for analysis, modeling, validation
- Researchers can publish data, methods, results and thoughts

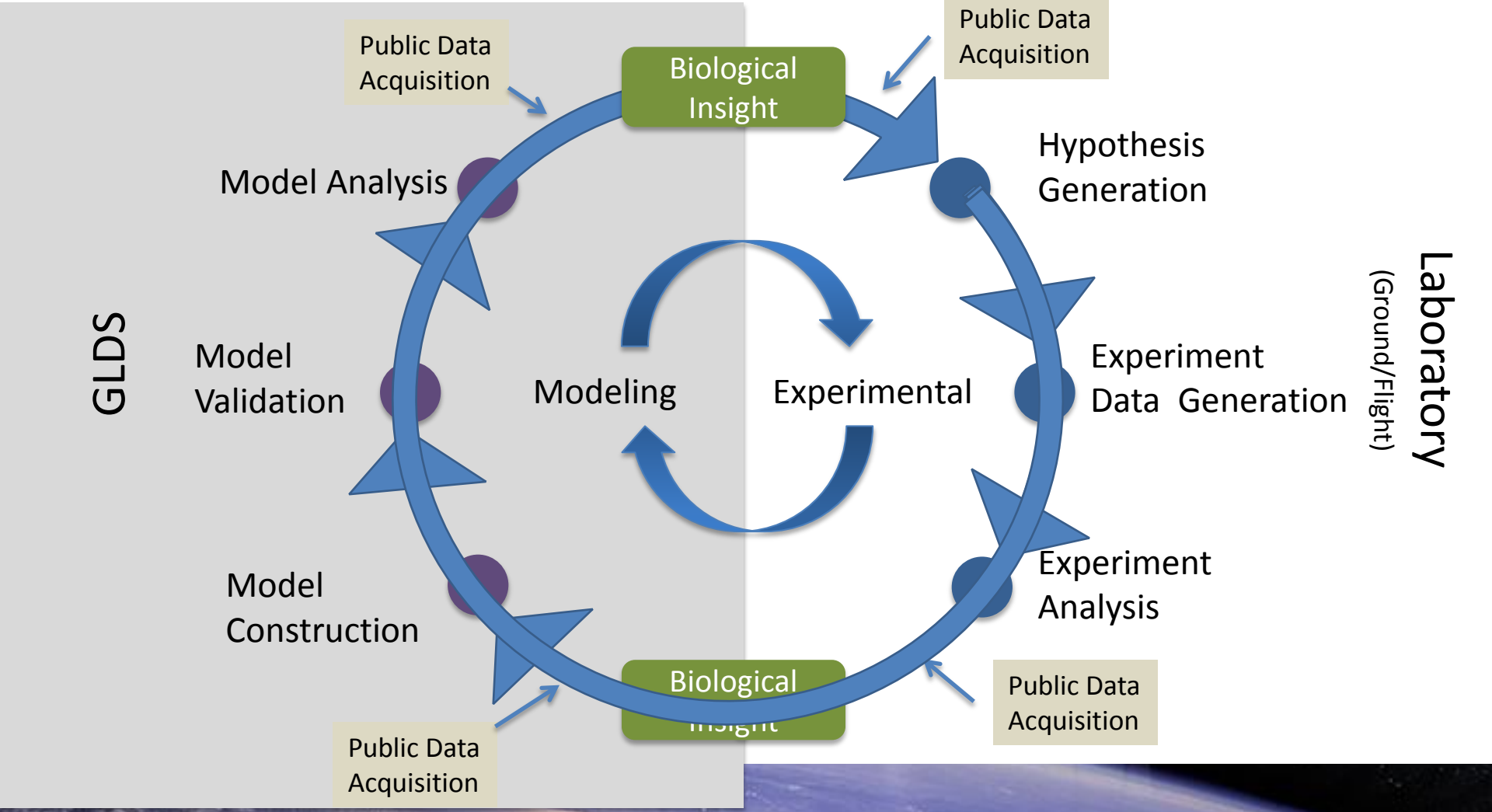


Modeling and Validation



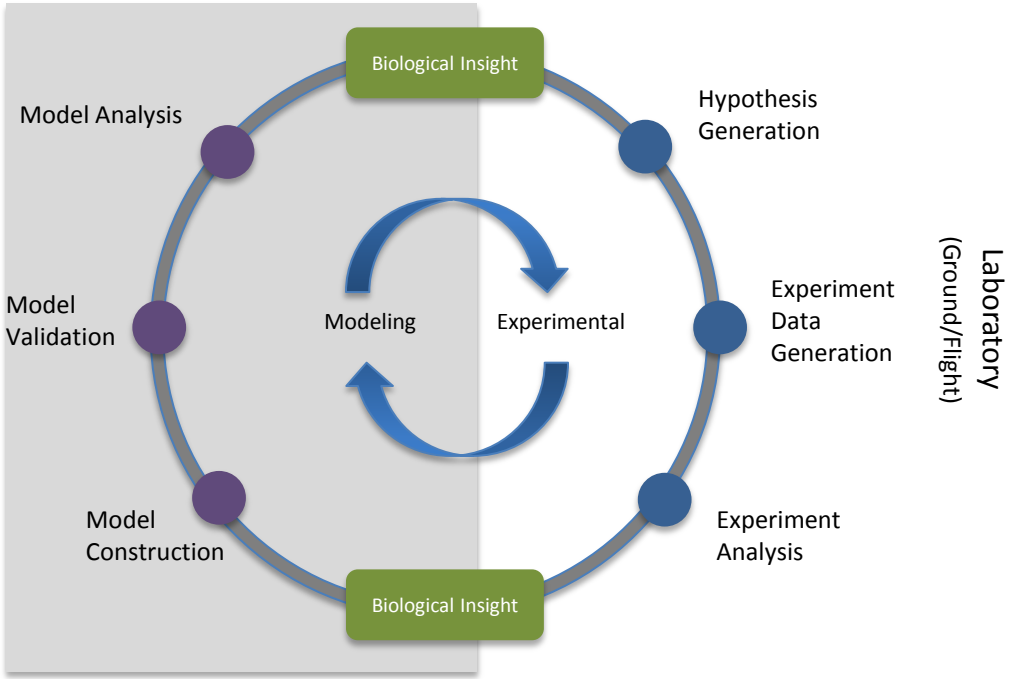
Next Generation Research

Iterative research solicitations for experiments utilizing GeneLab data for ground validation and next generation flight research.



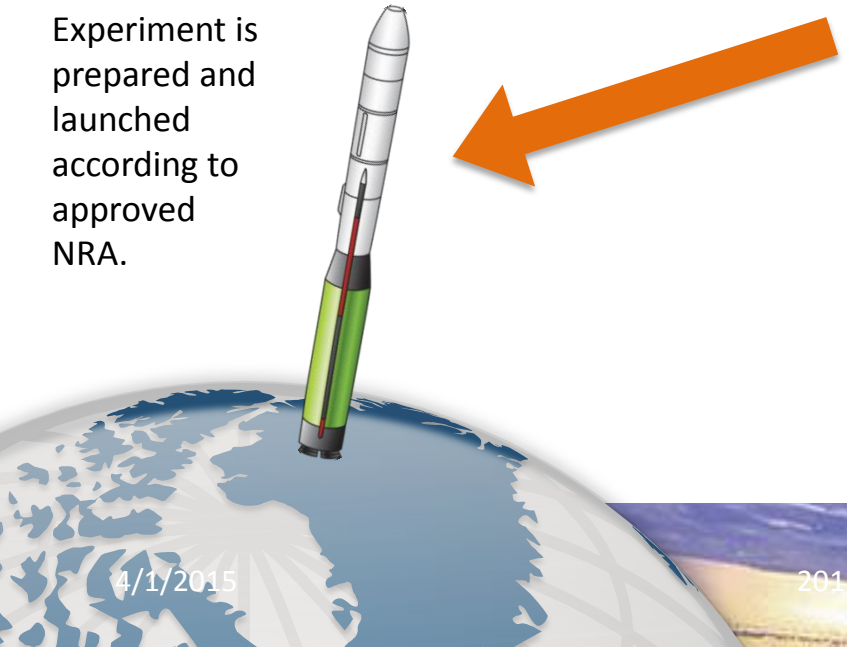
New Grant Opportunities and Missions

- Next generation of grant proposal
- New flight missions



Launch

Experiment is prepared and launched according to approved NRA.



- Computational modeling
- New hypothesis

GeneLab Phased Implementation

Phased Implementation 2014-2020

Phase 1
Searchable Data
FY2014 –2015

IT Systems
- **System Requirements & Architecture**
- **Public Website**
- **Searchable Data Repository**
- Requirements level 1

Science
- Omics Center Solicitation
- Protocol Development
- Data analysis validation
- **Initiate ground controls Collaborate with two manifested flight experiments**
- **SDT Solicitation for Dedicated Flight**

Phase 2
Data Acquisition
FY2015-2016

IT Systems
- Link to Public Databases
- Beta Space Bioinformatics system

Science
- Omics Center Selection
- Data analysis from initial ground studies
- Science Definition Teams Identified
- Outreach Program Plan

Phase 3
System Integration
FY2017 – 2018

IT Systems
- Integrated Platform across model organisms
- Build Community via collaborative science

Science
- Continue ground controls and process enhancement
- Engage with Scientists external to NASA as part of Outreach Program
- Dedicated flight experiments

Phase 4
Implementation
FY2019 – 2020

- Full science community engagement
- Development of analytical and Modeling tools
- Ongoing dedicated flight experiments
- Website and platform sustaining activities
- Continuous improvement

Omics Data Systems

- We have evaluated and rated:
  - KBase (DOE)
  - Globus
  - Intermine
  - GMOD
  - C3/Dashlink
  - dbgap
  - EGA
  - BioWarehouse

- We are aware of:
  - SysBioCube
  - BioCyc
  - Illumina BaseSpace
  - CLC Bioinformatics

Questions / Science

- Regarding GeneLab Dedicated Missions, what is the best way to utilize flight resources to generate maximally relevant omics data wrt
  - Number of different species
  - Multi- vs single omics data type
  - Longitudinal sample collections

Questions / Science

- How can GeneLab best publicize the relevance of flight experiment data to the broader community of ground-based experiments?
- How can GeneLab expand the user-base of the flight experiment data to include more researchers performing related ground-based experimentation?

Questions / Data Systems

- What is the value of providing biocomputing capabilities in terms of increasing scientific throughput?
  - Access to high performance computing resources
  - Biocomputing tools repository
  - Is there analogy to collaborating teams of scientists in other fields (e.g. NASA Earth Science)

Questions / Data Systems

- How can GeneLab best help researchers gain insights from experimental data to create systems biology views?
  - Data Discovery/Federation
  - Data annotation/Knowledge Capture
  - Data visualization
  - In-house developed capabilities vs. integrated externally-developed capabilities vs. direction to external systems and tools

Questions / Data Systems

- How best to represent Omics metadata?
  - Rapidly evolving metadata standards; newer Omics data types have no metadata standards yet
    - Data formats, metadata, protocols and analyses
  - How to capture the unique aspects of space biology experiments as metadata and related data?

GeneLab Staff

Project Manager – Joe Coughlan
Deputy Project Manager – Yung Nguyen
Project Scientist – Terri Thompson
Outreach Lead – Jon Rask
Payload Lead – Linda Timucin
Data Systems Lead/Architect – Daniel Berrios
Lead Developer – Chris Middour
Database System/Web Engineer – Jon Welch
Bioinformatics Scientist – Homer Fogle
Ground Lab Science R&D Lead – Kaushik Chakravarty
Lab Manager – Sam San-Huei Lan
Lab Technician – Rick Chen
Project Analyst/Configuration Manager – Nikita Gilkerson
Project Coordinator – Desireemoi Bridges

Backup Slides

GeneLab
Open Science for Exploration

Home Studies Submit Data Help Search:

All Studies

Transcription profiling of rat keratinocytes exposed to a 56Fe ion beam

Organisms	Factors	Assay Types	Release Date	Description
Rattus norvegicus	irradiate dose	transcription profiling	Nov29-2007	The purpose of the present work was to examine gene expression patterns in a rat keratinocyte line exposed to a 56Fe ion beam Experiment Overall Design The cells were exposed to 1.01 gev/nucleon 56Fe ions generated by the NASA Space Radiat...

Candida albicans response to spaceflight (NASA STS-115)

Organisms	Factors	Assay Types	Release Date	Description
Candida albicans	growth condition	transcription profiling	Nov01-2013	This study presents the first global transcriptional profiling and phenotypic characterization of the major human opportunistic fungal pathogen, Candida albicans, grown in spaceflight conditions. Microarray analysis revealed that C. albican...

Microarray Profile of Gene Expression during Osteoclast Differentiation in Modeled Microgravity

Organisms	Factors	Assay Types	Release Date	Description
Mus musculus	Treatment Group	transcription profiling	Apr07-2010	Microgravity leads to a 10-15% loss of bone mass in astronauts during space flight. Osteoclast is the multinucleated bone resorbing cell. In this study, we used NASA developed ground based Rotary Wall Vessel Bioreactor (RWV), Rotary Cell Cu...

An environment with strong gravitational and magnetic field alterations synergizes to promote variations in Arabidopsis thaliana callus global transcriptional state

Organisms	Factors	Assay Types	Release Date	Description
Arabidopsis thaliana	sample type treatment	transcription profiling	Jan01-2012	Using diamagnetic levitation, we have exposed A. thaliana in vitro callus cultures to five environments with different levels of effective gravity (from levitation i.e. simulated mg* to 2g*) and

GeneLab
Open Science for Exploration

Home Studies Submit Data Help Search:

Transcription profiling of rat keratinocytes exposed to a 56Fe ion beam

2 Datasets available:

ISA-TAB Metadata file from study E-GEOD-6299
Data file from study E-GEOD-6299

GeneLab Accession Number	GLDS-775
Source Accession Number	E-GEOD-6299
Contacts	Name: Ronghe Zhang; Role: submitter; Organization: ; Email:
Submission Date	
Public Release Date	Nov-29-2007
Study Description	The purpose of the present work was to examine gene expression patterns in a rat keratinocyte line exposed to a 56Fe ion beam Experiment Overall Design: The cells were exposed to 1.01 geV/nucleon 56Fe ions generated by the NASA Space Radiation Laboratory facility. Data from Affymetrix rat microarrays (RAT 230_2) were processed by BRB ArrayTools 3.3.0 software, and the Gene Ontogeny (GO) database was utilized to categorize significantly responding genes.
Organisms	Rattus norvegicus
Study Design Factor(s)	Factor: Ionization; Ontology: Concept: Ionization Factor: dose; Ontology: Concept: dose
Assay(s)	

A DNA microarray is a microarray that is used as a physical 2D immobilization matrix for DNA sequences. DNA microarray-bound DNA fragments are used as targets for a hybridizing probed sample

Questions / Science

- What are the data boundaries of scientific relevance vis-a-vis the environment of spaceflight experiments?
  - 60K+ ISS parameters
  - Selections by PIs may overlook parameters
  - Data archived approx. every 2 years, after which retrieval is much more difficult