National Aeronautics and Space Administration

A Researcher's Guide to:

INTERNATIONAL SPACE STATION

GeneLab

This International Space Station (ISS) Researcher's Guide is published by the NASA ISS Program Science Office.

Authors:
Olga Stotzky
Peter B. Tran
Sigrid Reinsch, PhD
Sylvain Costes, PhD
Dan Berrios, MD, MPH, PhD
Sandy Dueck

Executive Editor: Joseph S. Neigut
Technical Editor: Susan Breeden
Designer: Cory Duke

Published: May 1, 2018
Revision: (ebook only) December 7, 2018

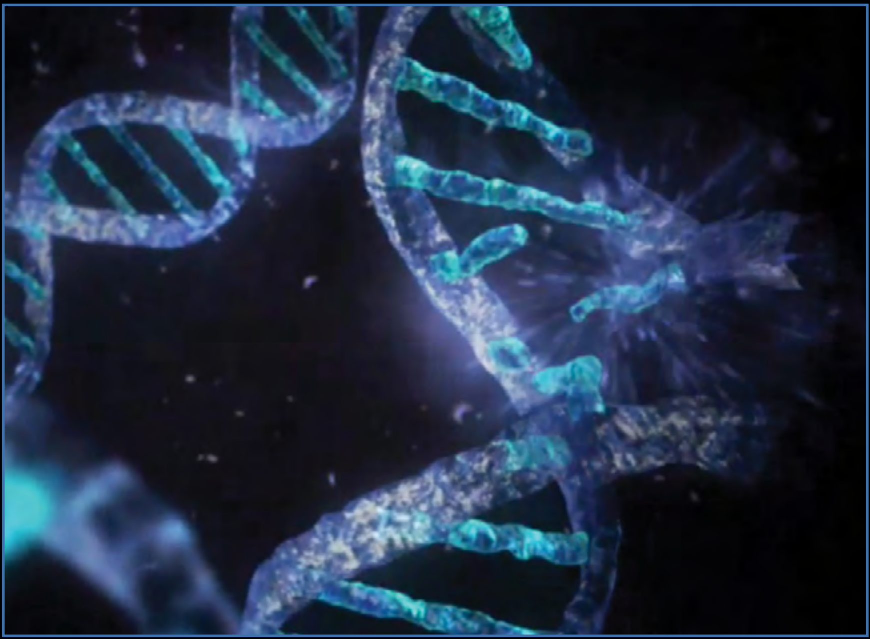
Cover and back cover:

a. GeneLab: Open Science for Life in Space (Image credit: NASA).
b. GeneLab collects genomics data from a variety of model organisms. These studies enable queries of how RNA, DNA, and proteins adapt and respond to the space environment (top). ISS final Configuration (bottom). (Image credit: NASA)

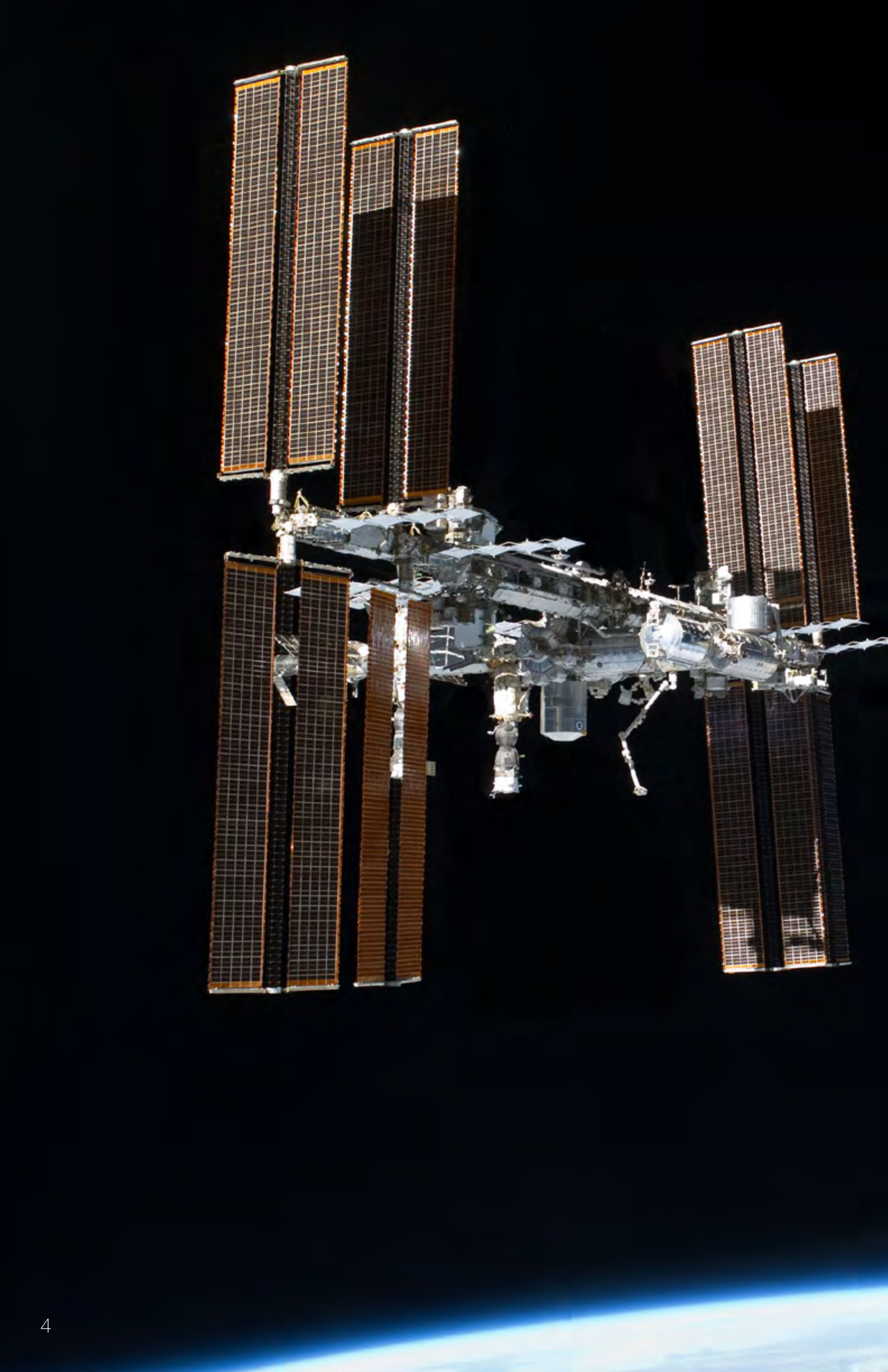
The Lab is Open

Orbiting the Earth at almost 5 miles per second, a structure exists that is nearly the size of a football field and weighs almost a million pounds. The International Space Station (ISS) is a testament to international cooperation and significant achievements in engineering. Beyond all of this, the ISS is a truly unique research platform. The possibilities of what can be discovered by conducting research on the ISS are endless and have the potential to contribute to the benefits of life on Earth and inspire generations of researchers to come.

As we increase utilization of the ISS as a National Laboratory, now is the time for investigators to propose new research and to make discoveries unveiling new knowledge about nature that could not be defined using traditional approaches on Earth.



Artist's rendition of space radiation hitting cell deoxyribonucleic acid (DNA). Human epidemiology studies of exposure to various doses of X-rays or gamma-rays provide strong evidence that cancer and degenerative diseases are to be expected from exposures to galactic cosmic rays or solar particle events. The GeneLab Data System has numerous ground investigations evaluating radiation effects on multiple model systems (NASA).



Unique Features of the ISS Research Environment

1. **Microgravity**, or weightlessness, alters many observable phenomena within the physical and life sciences. Systems and processes affected by microgravity include surface wetting and surface tension, convection, multiphase flow and heat transfer, multiphase system dynamics, solidification, and fire phenomena and combustion. Microgravity induces a vast array of changes in organisms ranging from bacteria to humans, including global alterations in gene expression and three-dimensional aggregation of cells into tissue-like architecture. Experiments and hardware designs must accommodate these different physical dynamics.

2. **Extreme conditions** of the ISS external environment include exposure to extreme heat and cold cycling, ultra-vacuum, atomic oxygen, and high-energy radiation. Few organisms are able to survive exposure to these environmental extremes outside the ISS. Inside the ISS, the Environmental Control and Life Support System (ECLSS) performs numerous functions including controlling atmospheric gas partial pressures, temperature, total pressure and humidity; and scrubbing carbon dioxide, volatile organics, and water vapor from the air.

3. **Space Radiation** is variable with both predictable and unpredictable components, has multiple sources, and affects astronauts, living cells, materials, and devices. The components of space radiation are Trapped Particles, Solar Particles, and Galactic (Cosmic) Rays. The ISS orbital inclination of 51.6 degrees affects the amount of radiation that reaches the ISS. The ISS structure provides some significant shielding for proton fluxes but much less for ionizing radiation. Dosimetry for individual biological experiments is extremely rare and dependent on the design hardware for the experiment.

Table of Contents

The Lab is Open	3
Unique Features of the ISS Research Environment	5
What is Open Data?	7
Why is NASA embracing Open Data?	8
What opportunities does this present for researchers?	9
Open Access, Open Science, Open Inquiry	10
Big Data from Biological Experiments: Systems Biology and Omics	11
GeneLab	12
How GeneLab Works	13
Partnerships	14
On Orbit Research Implementation	15
Follow-on Research	16
GeneLab Team	17
GeneLab Data System	18
Data Submission	20
Funding Opportunities	22
Acronyms	23

What is Open Data?

The National Aeronautics and Space Administration (NASA) has developed an Open Data Plan and associated policy outlining a framework for activities to increase public access to scientific publications and digital scientific data resulting from NASA-funded research.

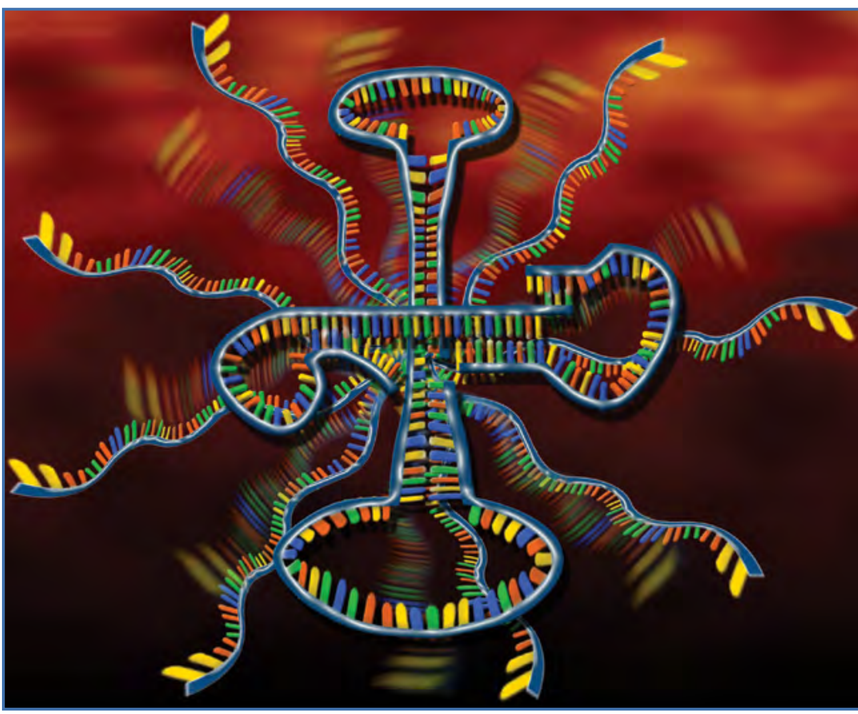
Promoting the full and open sharing of data with research communities, private industry, academia, and the general public is one of NASA's long-standing core values. For example, NASA's GeneLab is an open-access resource for space biology. The GeneLab team routinely processes, archives, and provides omics data to researchers around the globe. NASA's plan expands the breadth of its open-access culture to include data and publications for all of the scientific research that NASA sponsors.

Ten-day-old Arabidopsis thaliana plants growing vertically on the surface of a 10-cm square nutrient gel Petri plate. This experiment platform is routinely used in the Principal Investigators' ground laboratories, and was also used in the Advanced Photovoltaic and Electronic Experiment (APEX) (NASA).

Why is NASA embracing Open Data?

NASA's Open Data Plan was issued in response to the Executive Office of the President's February 22, 2013, Office of Science and Technology Policy (OSTP) Memorandum for the Heads of Executive Departments and Agencies titled "Increasing Access to the Results of Federally Funded Scientific Research." In this memorandum, OSTP directed all agencies with more than $100 million in annual research and development expenditures to prepare a plan for improving the public's access to the results of federally funded research.

NASA invests on the order of $3 billion annually in fundamental and applied research and technology development across a broad range of topics, including space and Earth sciences, life and physical sciences, human health, aeronautics, and technology.

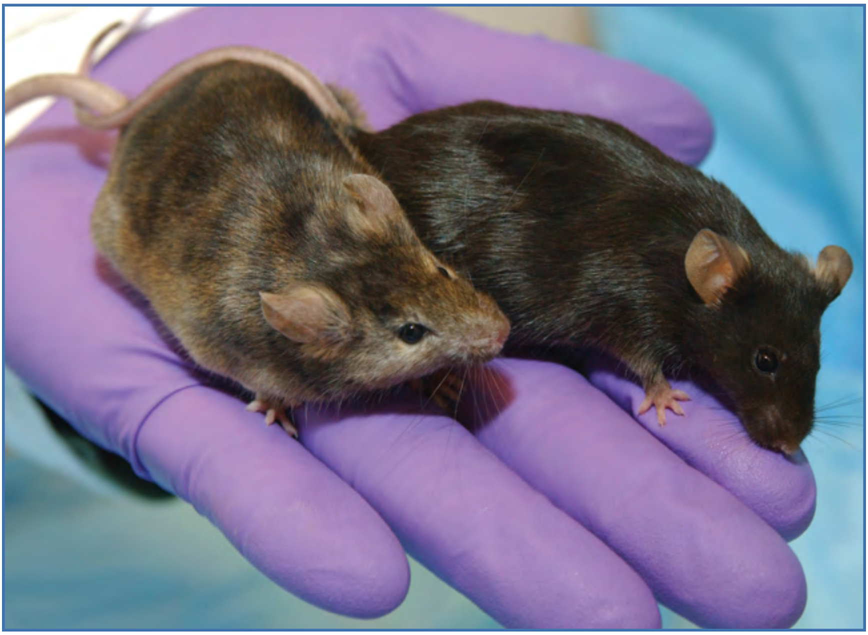


Artist's rendering of ribonucleic acid, molecule. (Image courtesy of Nicolle Rager Fuller, National Science Foundation.)

What opportunities does this present for researchers?

To facilitate access to spaceflight data acquired by investigations on board the ISS or other spaceflight platforms, or resulting from ground-based flight simulation experiments, NASA's Space Life and Physical Sciences Research and Applications (SLPSRA) Division has implemented this open data plan. It is NASA's intention that storing these data in open-access databases will promote their use by researchers to develop new hypotheses to be tested on the ground or in spaceflight, to amplify the results of other studies, or to develop commercial products or other translational tools. NASA issues competitive, ground-based research solicitation of proposals to use these data to identify patterns and trends, and to devise new testable hypotheses and experiments.

NASA also plans to issue a series of solicitations for the analysis of spaceflight-derived datasets. These would provide value-added data interpretations ensuring that all data are maximally "mined" for information relevant to the scientific community.



Laboratory mouse in which a gene affecting hair growth has been knocked out, left, next to a normal lab mouse (National Institutes of Health).

Open Access, Open Science, Open Inquiry

Long-duration human exploration of space faces major hurdles including risks to astronaut health and challenges in environmental control and life support. NASA initiated GeneLab – a multiyear, multiphase project – on the premise that mining of omics data from spaceflight experiments offers an immense opportunity to understand the effects of spaceflight on biological systems, and that this can best be accomplished by ensuring access to these data for as many researchers as possible.

GeneLab captures vast amounts of data from spaceflight and ground samples, and implements open (unrestricted) access to these data. By "democratizing" access, GeneLab aims to increase the return on investment for biological research conducted in microgravity onboard the ISS, and to thereby maximize the scientific impact of each experiment. The ability to conduct such secondary and meta-analyses will help space life sciences take a major leap forward in understanding the effects of microgravity, radiation, and other aspects of the space environment. This approach is intended to increase information sharing rapidly. As the rate of experimentation leads to a faster pace of scientific discovery, GeneLab will be instrumental in producing new knowledge. The discoveries enabled by the GeneLab project may one day lead to cures for diseases on Earth, and help space explorers withstand the rigors of long-duration spaceflight.

Studies on a variety of organisms enable queries of how ribonucleic acid (RNA), deoxyribonucleic acid (DNA) and proteins – the building blocks of life – adapt and respond to the space environment. Because no single analysis can fully unravel the complexities of fundamental systems biology, investigators are designing experiments to provide multiple layers of omics information that can be studied in an integrated fashion. This systems biology approach will help generate a more complete understanding of how biological systems adapt to spaceflight (translational research).

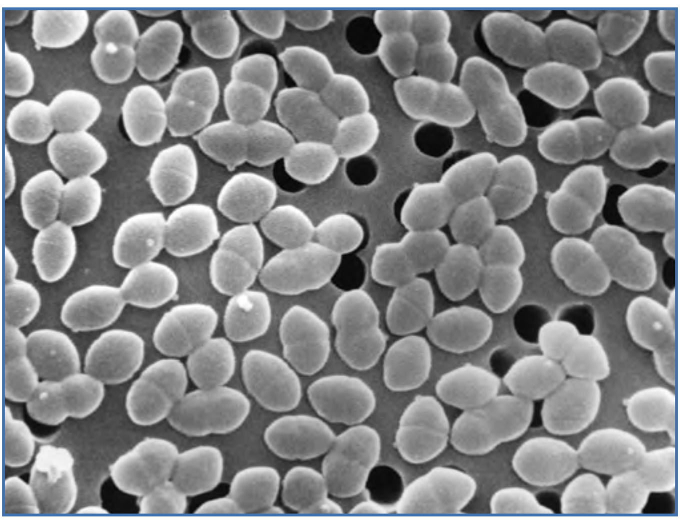
Big Data from Biological Experiments: Systems Biology and Omics

Advances in biological technologies have led to our ability to generate intricate snapshots of the many types of molecules that make up individual cells. It is possible to assay what molecules are present at a given time in single cells, tissues, whole organisms, or even populations of organisms. Understanding the interplay between all of these molecules and how they contribute to the function (or dysfunction) of a cell, tissue, or organism is called systems biology. High-throughput molecular data generated from biological experiments are called omics data sets – i.e., genomics, transcriptomic, proteomic, epigenomic, metagenomics, and metabolomics. Analyses of DNA sequences, or the genome, are genomics; which genes are expressed as RNA is called transcriptomics. Similarly one can study the proteins (proteome), metabolites (metabolome), and even which microbes colonize an organism or environmental niche (microbiome). The data sets that are generated from these analyses are huge and require advanced computational skills and tools to analyze and decipher their biological significance. By analyzing omics data stored in the GeneLab Data System (GLDS), scientists will be able to test hypotheses about how the spaceflight environment affects biological systems and develop better countermeasures for long-duration spaceflight.

Artist's rendition of utilization of omics data collected from spaceflight. (Image credit: NASA)

GeneLab

GeneLab is a project designed to drive the generation of spaceflight and spaceflight-related omics data and a program to curate, annotate, and foster the broadest possible analysis of that data to further the manned exploration of space. How is this done? GeneLab partners with scientists performing spaceflight-relevant biological experiments to gather data from those experiments, and is developing the GLDS as a unique space life sciences data repository to house omics data and associated metadata (information about the conditions of the spaceflight experiment as well as sample processing, etc.). In the coming years, in addition to its data repository function, the GLDS will provide a single integrated collaborative platform where scientists can capture, curate, store, search, share, transfer, analyze, and visualize spaceflight datasets using state-of-the-art bioinformatics tools.



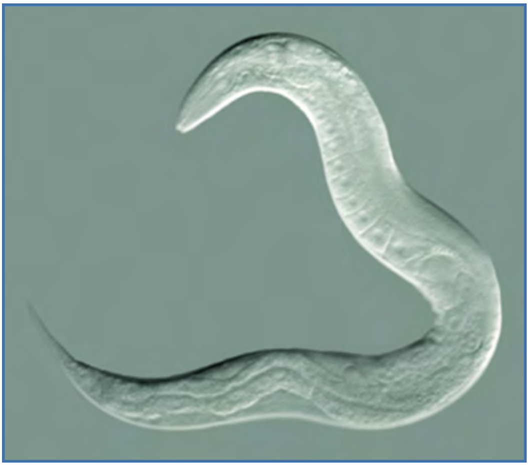
This scanning electron micrograph depicts numbers of bacteria, which are identified as being Gram-positive Enterococcus sp. bacteria. Image courtesy of the Centers for Disease Control and Prevention, Image ID - 209.

How GeneLab Works

When fully implemented, GeneLab will gather and store omics data from spaceflight and spaceflight-relevant experiments, annotate each study with additional space life sciences specific information, interface with other scientific databases for expanded research opportunities, and offer tools to conduct data analysis. Through the GLDS, scientists, researchers, teachers and students will be able to create a place online to connect with their peers and share their results with the world.

As part of GeneLab's objectives, NASA will collaborate with scientists that send organism systems to live aboard the space station. These include various cell lines, bacteria and fungi, plants, fruit flies, worms, fish and rodents—all systems routinely used in genetics research on Earth. DNA, RNA, protein, and metabolite data will be extracted from samples of these organisms collected during space missions.

Samples may be preserved on orbit, and then stored aboard the space station and returned to Earth for further sample processing to extract the various molecules either by GeneLab or by collaborating scientists. Samples can also come from ground experiments that have spaceflight relevance, such as those that examine the biological effects of ionizing radiation, or experiments that simulate microgravity. GeneLab's Sample Processing Laboratory has state-of-the-art equipment and follows stringent processes to properly receive and track samples, perform multi-omics sample processing, conduct analyses, and implement quality control (QC) for given organisms (e.g., rodents, bacteria, and yeast). GeneLab is collaborating with scientists from the National Institute of Standards and Technology (NIST) to implement standards of all kinds. All data derived from GeneLab collaborations are required to be uploaded into the GLDS.

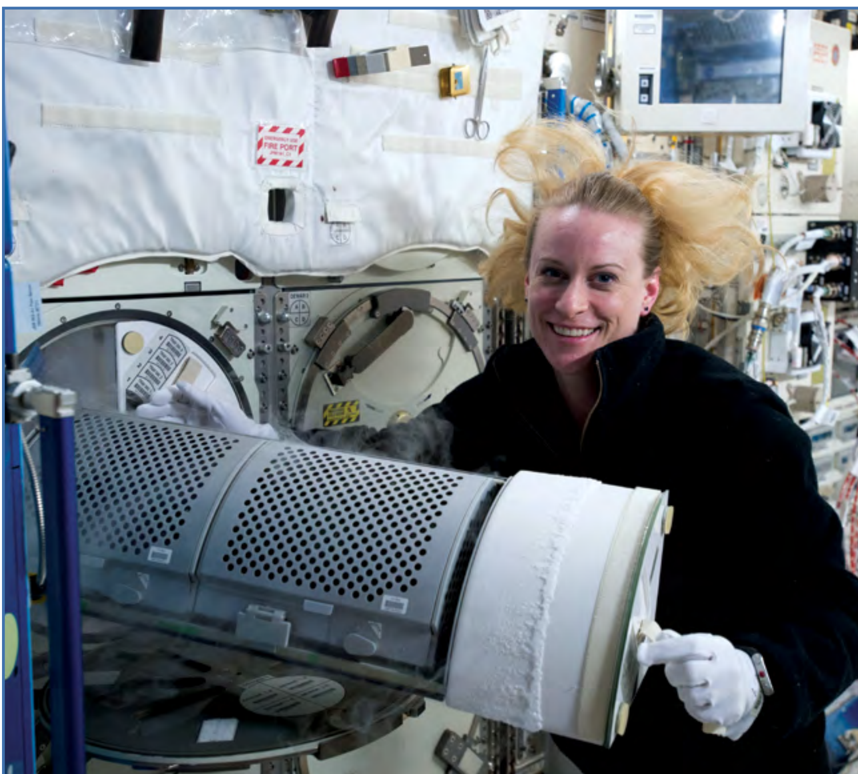


The roundworm Caenorhabditis elegans (C. elegans). Image Credit: Cell Image Library.

Partnerships

One of GeneLab's goals is to maximize the amount of omics data from every spaceflight with biological research payloads through sample sharing or augmentation of experiment samples, to expand omics analyses on spaceflight biological samples. Currently, GeneLab primarily partners with experiments funded through NASA's SLPSRA division, and with the Center for Advancement of Science In Space (CASIS) on experiments.

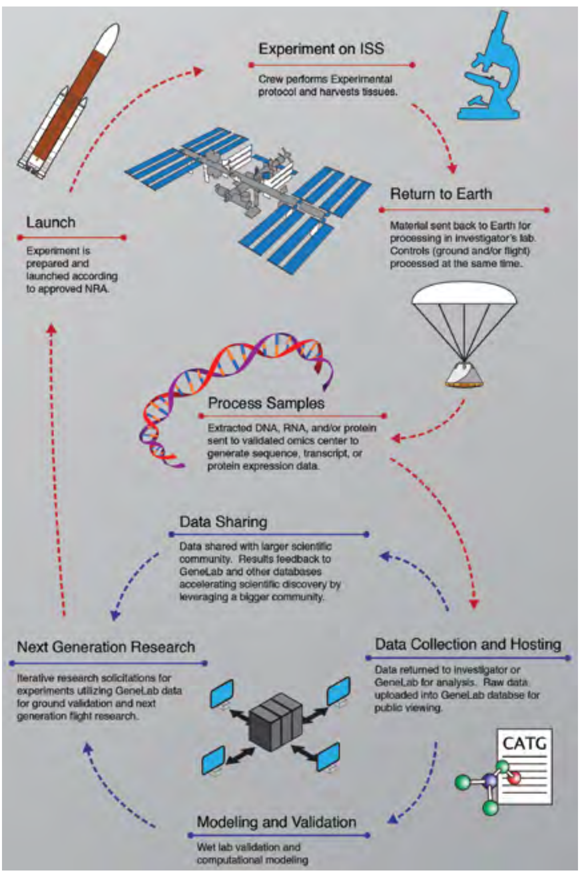
The goals of these partnerships can be sample sharing, additional omics analyses in the originally designed experiment, data sharing, or collaborative data analyses. To date, GeneLab has partnered on multiple flight experiments involving rodents, plants, fruit flies, and microbes.



Astronaut Kate Rubins is photographed removing samples from the Minus Eighty-Degree Laboratory Freezer for ISS (MELFI) 2.

On Orbit Research Implementation

Successful space researchers develop on orbit procedures for research that take into consideration the unique aspects of space station operations. The simplest experiment to perform on Earth will have extremely complicated logistics in microgravity. Every detail of an experiment must be carefully worked through to ensure crew and vehicle safety and that the experiment runs without issues. Launches can change so that critically timed experiments are delayed and fresh samples or reagents might need to be generated in response to launch delays. Space for time-sensitive or environmentally sensitive biologicals is very limited, as is cold stowage both pre- and post-flight. Launches and returns add confounding acceleration variables to experiments with organisms that are actively metabolizing. Once on board station, hands-on crewtime to perform experiments can be limited and sample preservation for downstream analyses is not always ideal.



To maximize the value of data from life science experiments performed in space, and to make the most advantageous use of the remaining ISS research window, GeneLab will employ an iterative approach performing spaceflight experiments, generating, saving, and sharing the datasets derived from biological research in space.

Follow-on Research

Analyses of GLDS data, which explore the network of molecular responses of terrestrial biology to the space environment, will contribute fundamental knowledge of how the space environment affects biological systems. These analyses will yield benefits both on Earth and for space exploration resulting in mitigation strategies to prevent negative effects observed during exposure to the spaceflight environment.

The GeneLab Platform will acquire space-flown samples, process the samples for data generation, and employ an open-access model to expand the scientific audience. Full data integration, community engagement, and availability of open source software in a single web interface will augment scientific research. This model will enable formation of novel hypotheses and follow-on space grants.

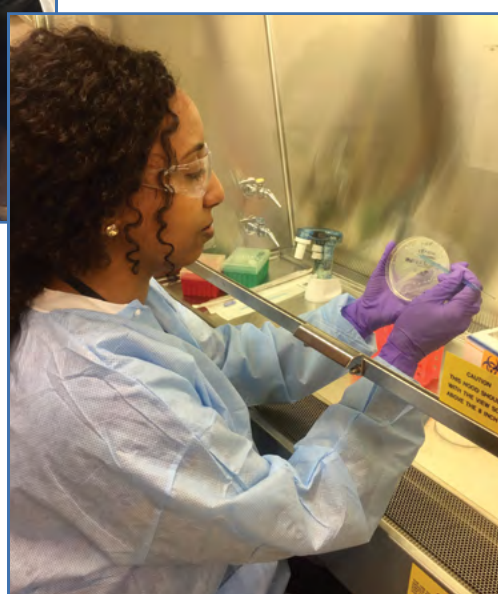
Using the GLDS, researchers will be able to conduct secondary *in silico* data analyses, such as data mining to identify likely targets for drug development.

GeneLab will include open-source bioinformatics tools. In the spirit of open science, GeneLab will encourage the results of such analyses to be shared with the global community of scientists.

GeneLab Team

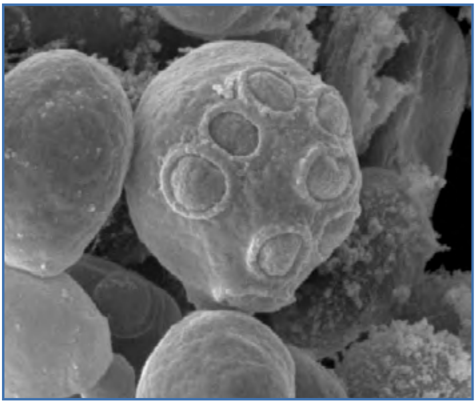
The GeneLab project team is comprised of biologists, computer scientists, software engineers, data architects, data curators, biologists, payload specialists, bioinformaticians, and science communicators at NASA Ames Research Center, Moffett Field, CA. The project also has a state-of-the-art laboratory facility for sample processing. Samples that are processed by GeneLab are sent to various commercial or academic core facilities for omics analysis. The science direction is provided by the SLPSRA program at NASA Headquarters. The project funding is contributed jointly by SLPSRA and the International Space Station Research Integration Office (ISSRIO) at NASA Johnson Space Center, Houston, TX.

GeneLab hosts several exhibit booths at science conferences throughout the year. Team member San-huei Lai Polo listens to college students explain how they use omics data from the GLDS. (NASA)



GeneLab Wetlab team member Sam Gebre practices microbiology methods, including RNA extraction from bacterial samples.

GeneLab Data System



Scanning electrograph image of Saccharomyces cerevisiae cells grown on the International Space Station.

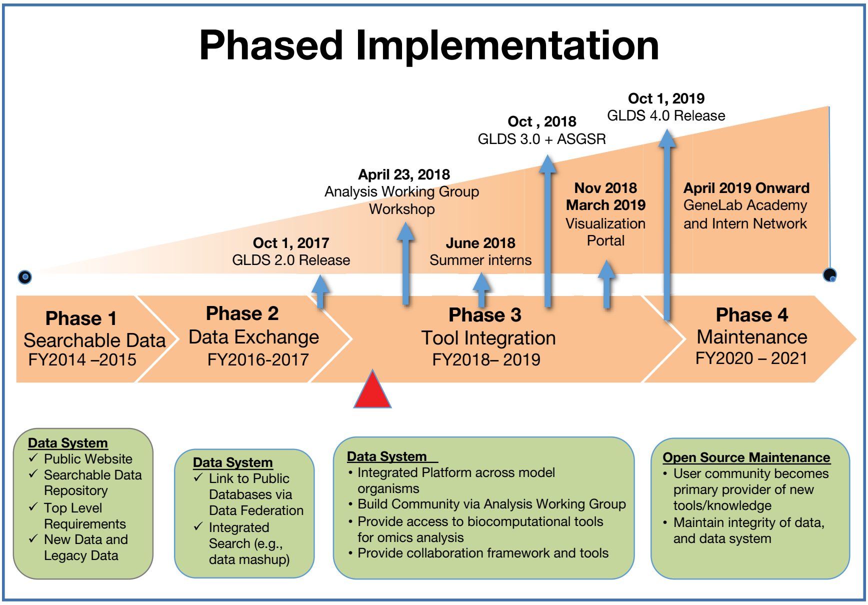
The GLDS is NASA's open-access omics data platform for biological experiments. GLDS houses high-throughput sequencing and other omics data from spaceflight-relevant experiments. GLDS Phase 1 (Version 1.0) went online on April 2015. As of June 2017 (Version 1.0.18), the GLDS contains a total of 133 publicly available datasets and includes search and data download capabilities.

GLDS Phase 2 (Version 2.0) is slated for release at the end of September 2017 and will have integrated search and data federation/integration capabilities leveraging data from other public omics databases, such as the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO), European Bioinformatics Institute (EBI) Proteomics Identifications (PRIDE), and the Argonne National Laboratory (ANL) Metagenomics Analysis-Rapid Annotations using Subsytems Technology (MG-RAST). Future versions of the GLDS within this multiphase project will provide an integrated collaborative platform and biocomputational tools for omics data analysis and sharing.

GeneLab envisions housing other omics data types as new assays are developed and vetted by the larger scientific community.

The GeneLab project is using a multiphased approach to full implementation of the GLDS. During the first two phases, the focus was primarily on developing the database, and on working with the GeneLab stakeholders to develop sound policies for GeneLab collaborative missions. The Genelab Sample Processing Laboratory was set up and protocols for handling and processing spaceflight samples were developed. Omics core facilities were selected and vetted for initial DNA, RNA, protein, and metabolomics analyses. The timeline below shows the accomplishments from the first two phases and the goals for the scientific and data systems teams in the coming years.

Data from experiments worldwide, which explore the biological effects of the spaceflight environment on a wide variety of model organisms, are housed in GLDS including data from rodents, invertebrates, plants, and microbes. Human datasets are currently limited to those with anonymized data (e.g., from cultured cell lines). The datasets include NASA-funded experiments as well as those funded by other international space agencies. GLDS ensures prompt release and open access to high-throughput genomics, transcriptomics, proteomics, and metabolomics data from spaceflight and ground-based simulations of microgravity, radiation, or other space environment factors. The data are meticulously curated to assure that accurate experimental and sample processing metadata are included with each dataset.

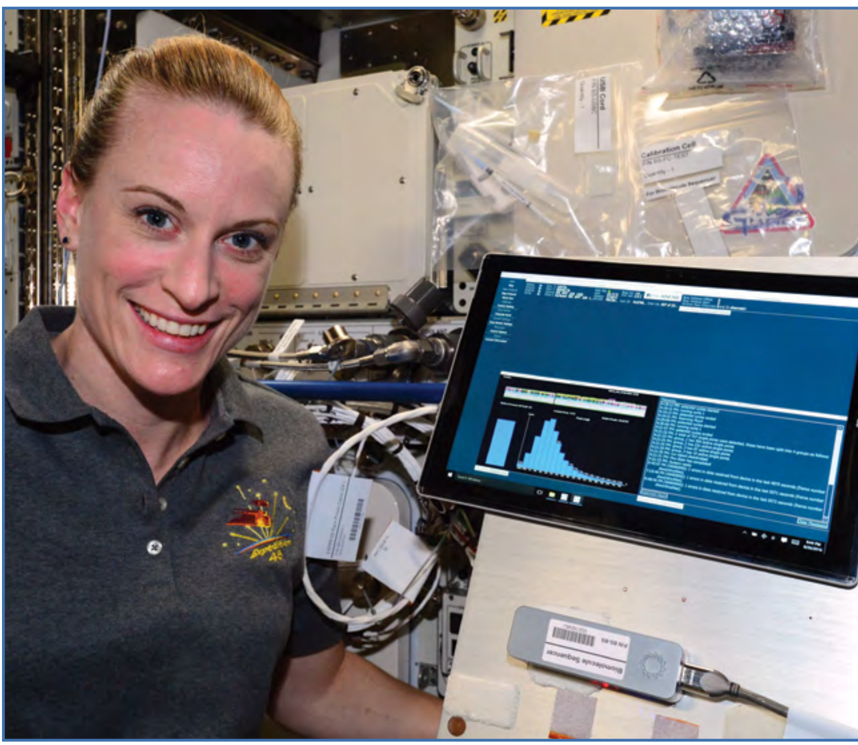


The GeneLab project will follow four distinct phases of implementation. Each phase of the project will be marked by increasing complexity and increasing value and utilization of the GeneLab data system. The final phase of the project will enable full implementation of GeneLab.

Data Submission

For GeneLab's current Phase 1 operational system, any researcher can submit and contribute their spaceflight-related omics data to GeneLab by following a few simple steps and guidelines outlined on the GeneLab website here: https://genelab-data.ndc.nasa.gov/genelab/submissions/.

The data submission process starts with a questionnaire, which includes: Your Full Name; Organizational Affiliation: Your Role in Study; A concise explanation of the relevance of your data to NASA, for example: NASA-funded, ISS-based experiment, space biology experiment, ground analog study, radiation study, etc.; Brief description of all data and metadata that you intend to submit, including; Data Types(s); Approximate size of data file(s).



NASA astronaut Kate Rubins poses for a picture with the minION device during the first sample initialization run of the Biomolecule Sequencer project. Rubins not only became the first person to sequence DNA in space, but she sequenced more than a billion bases during her time aboard the space station.

NASA GeneLab will then review the submission request and send a notification back to the science team. If the submission is approved, science teams will receive more detailed instructions on how to send metadata and assay datasets to GeneLab, including a customized configuration template file. GeneLab is currently accepting metadata in ISA-Tab format (e.g., tabbed delimited text files). The ISACreator tool is an open-sourced desktop General User Interface (GUI) application designed to guide you in filling out the fields correctly. An ISACreator tool version that is already populated with GeneLab-specific metadata fields is available on the GeneLab website: https://genelab-data.ndc.nasa.gov/genelab/. For GeneLab to curate data through NASA's web privacy standards and policies, the project requires certain limited information on citizenship of science team members.

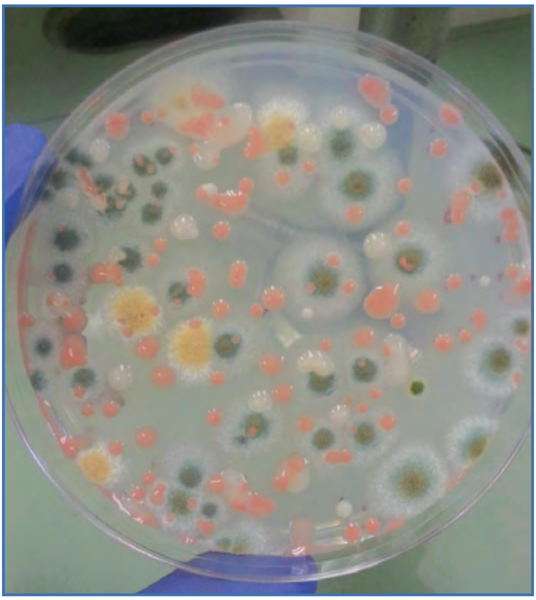
A step-by-step ISACreator tutorial user guide in PDF format can be view and download from here: https://genelab-data.ndc.nasa.gov/genelab/help/GeneLab_Submission_Guide_2.0.pdf.

Drosophila melanogaster (fruit fly). Image courtesy of Ames Research Center.

For later project phases, GeneLab will have key capabilities for researchers to establish their online collaborative workspaces and upload their data for analyses in order to build, foster, and improve scientific collaborations. Contact the GeneLab Data Operations team at arc-dl-genelab-it@mail.nasa.gov for any data submission questions.

Funding Opportunities

Research always reveals new questions as well as answers. Using the NASA Research Announcement process, SLPSRA will support follow-on research for deeper analysis examining fundamental questions. SLPSRA will consider data analyzed from GeneLab as part of the background justification for the importance of the proposed research. Greater detail concerning current funding opportunities for ISS research can be found through the NASA ISS research website: https://www.nasa.gov/mission_pages/station/research/ops/research_information.html.



A petri dish contains colonies of fungi grown from a sample collected aboard the ISS during Microbial Tracking-1, a research investigation that looks at the types of microbes present on the surfaces and in the air of the space station.

The NASA Solicitation and Proposal Integrated Review and Evaluation System (NSPIRES) can be accessed via https://nspires.nasaprs.com/external/.

For more information about GeneLab go to: https://genelab.nasa.gov/.

For more information on NASA-Funded Research Results go to: https://www.nasa.gov/open/researchaccess.

Acronyms

ANL	Argonne National Laboratory
APEX	Advanced Photovoltaic and Electronic Experiment
AWG	Analysis Working Group
CASIS	Center for Advancement of Science In Space
DNA	Deoxyribonucleic Acid
EBI	European Bioinformatics Institute
ECLSS	Environmental Control and Life Support System
GEO	Gene Expression Omnibus
GLDS	GeneLab Data System
GUI	General User interface
ISA	Investigation Study Assay
ISS	International Space Station
ISSRIO	International Space Station Research Integration Office
MELFI	Minus Eighty-Degree Laboratory Freezer for ISS
MG-RAST	Metagenomics Analysis-Rapid Annotations using Subsytems Technology
NASA	National Aeronautics and Space Administration
NCBI	National Center for Biotechnology Infromation
NIH	National Institutes of Health
NIST	National Institute of Standards and Technology
NSPIRES	NASA Solicitation and Proposal Integrated Review and Evaluation System
OSTP	Office of Science and Technology Policy
PRIDE	Proteomics Identifications
QC	Quality Control
RNA	Ribonucleic Acid
SLPSRA	Space Life and Physical Sciences Research and Applications

The Complete ISS Researcher's Guide Series

1. Acceleration Environment
2. Cellular Biology
3. Combustion Science
4. Earth Observations
5. Fluid Physics
6. Fruit Fly Research
7. Fundamental Physics
8. Human Research
9. Macromolecular Crystal Growth
10. Microbial Research
11. Microgravity Materials Research
12. Plant Science
13. Rodent Research
14. Space Environmental Effects
15. Technology Demonstration
16. GeneLab

For more information...

Space Station Science
https://www.nasa.gov/iss-science

Station Facilities
https://www.nasa.gov/stationfacilities

Station Reference Guide: Utilization Edition
https://www.nasa.gov/sites/default/files/atoms/files/np-2015-05-022-jsc-iss-guide-2015-update-111015-508c.pdf

Station Opportunities for Researchers
https://www.nasa.gov/stationopportunities

Station Research Client Helpline
jsc-iss-research-helpline@mail.nasa.gov
(281) 244-6187

National Aeronautics and Space Administration

Johnson Space Center
http://www.nasa.gov/centers/johnson

www.nasa.gov

NP-2017-06-020-JSC