



Published: March 31, 2024

Citation: Arnold C, Casaletto J, et al., 2024. Spaceflight Disrupts Gene Expression of Estrogen Signaling in Rodent Mammary Tissue, Medical Research Archives, [online] 12(3).

<https://doi.org/10.18103/mra.v12i3.5220>

Copyright: © 2024 European Society of Medicine. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI

<https://doi.org/10.18103/mra.v12i3.5220>

ISSN: 2375-1924

Spaceflight Disrupts Gene Expression of Estrogen Signaling in Rodent Mammary Tissue

Cleoma Arnold, MS^{*1}; James Casaletto, PhD²; Philip Heller, PhD¹

¹ Department of Computer Science, San Jose State University, 1 Washington Sq, San Jose, CA 95192

² Artificial Intelligence for Life in Space, NASA Ames, Mountain View, CA 94043

***Corresponding author:** cleoma.arnold@gmail.com

ABSTRACT

Life aboard spacecraft poses various dangers to astronaut health, with hazards including microgravity, radiation, and enclosed spaces. Research into mitigating these health issues includes analysis of the transcriptome of rodents sent to the International Space Station. This project investigates the effects of spaceflight on the gene expression of mammary tissue of female mice of two age groups, 10-12 and 32 weeks, in order to assess the impact of age and spaceflight on gene expression. Spaceflight-induced changes to gene expression in rodent mammary tissue could contribute to the characterization of the impact of spaceflight on female astronaut health, which has been historically underserved. Analysis of the OSD-511 dataset from NASA's Open Science Data Repository utilized a containerized implementation of their RNA-Seq pipeline on the San José State University High Performance Computing Cluster. Seven genes were found to be differentially expressed across all comparison groups; one gene, *Greb1*, is implicated in hormone mediated disease. Age appears to influence biological pathways affected by spaceflight in mammary tissue, with young mice experiencing metabolic changes while older mice undergo changes to inflammatory pathways. Further research is needed to determine the mechanism of spaceflight-induced gene expression changes.

Keywords: Spaceflight, Differential Gene Expression, RNA-Seq, Mouse, Mammary, *Greb1*

Introduction

Space is inhospitable to life. Our species, as well as all life on Earth, evolved to adapt to a narrow range of conditions found on the Earth's surface, from the air we breathe to the ground we walk on. Excursions from those conditions can potentially be damaging or deadly. These limitations have not prevented humanity from exploring the frontier above us, but this exploration must be done without excessive risk to human health. In 1967, at the start of the Apollo program, NASA operated without the concept of "permissible doses" of hazards such as radiation in order to prioritize success of the mission, but this is no longer the case; now we have the tools and technology to make space travel safer without putting human lives in danger¹.

Other progress is being made as well: in 1967, all astronauts were male; from 1998 to 2013, 85% of the people who traveled to the International Space Station, or ISS, were male; within the next decade a woman and a person of color will travel to the moon for the first time as part of the Artemis mission¹⁻³. Historically, there has been little characterization of the health risks for female astronauts; because inclusion of female workers has been a recent trend in this industry, there is little data available on long term health effects⁴. Extrapolation from the occupational outcomes of similar professions has been problematic: male astronauts are well modeled by airplane pilots, as there is significant overlap between those two groups in terms of job duties and exposures, but prior research has compared female astronauts to flight attendants, who do not perform similar tasks^{2,5}. As travel to space becomes more equitable, research needs to be done in order to make sure that a wider variety of health needs are met. Ideally that research would be done without risking human lives as test subjects and instead would use cultured cells or model organisms such as rodents so that regardless of sex, astronauts are able to accomplish both their mission and a safe return to their families on Earth.

One such series of experiments, the Rodent Research Reference Missions, has flown rodents to the ISS in order to best study the effects of spaceflight on mammalian biology without putting astronauts in unnecessary danger⁶. This allows study of longer spaceflights; rodents have shorter lifespans than humans, and are therefore able to spend a greater percentage of their lives in space. As the needs of the organism change during spaceflight, different genes are turned on or off in a given cell in order to supply gene products, either RNA or protein, to the cell, tissue, or body as a whole. Messenger RNA can be sequenced and

compared to controls on Earth to find out what genes are being transcribed during the rodent's time in space.

This sequencing data, as well as many other datasets, is available to the public through the Open Science Data Repository (OSDR, formerly known as GeneLab). OSDR is NASA's open-access database operated for the purpose of sharing data from spaceflight experiments with the broader scientific community⁷. OSDR hosts experimental data from academic and government sources, as well as establishing standardized analysis pipelines for analysis of that data so that experimental results are more easily compared regardless of source⁷. Previous analyses of differential gene expression have used a series of individual bioinformatics tools, but NASA has recently created a containerized pipeline incorporating those tools into one package and made it available to the public.

This project investigates the output of the containerized analysis pipeline, focusing on changes to the transcriptome of female mouse mammary tissue in space. First, transcribed genes are compared between spaceflight and Earth controls to see what genes are differentially expressed during spaceflight. This particular mission, known as RR-8, sent two age groups of mice in order to assess the impact of spaceflight on mice that are young or middle aged⁶. The transcribed genes and differentially expressed gene pathways of these two age groups are compared to see whether old mice react to space differently than young ones. Special emphasis is placed on gene expression associated with mammary tissue disease, as breast cancer is a significant health concern in female populations and most studies of spaceflight focus on muscle or liver tissue⁸. While prior studies of cultured breast cancer cells show a less aggressive phenotype when grown in microgravity, occupations with high radiation exposure such as astronauts may anticipate higher rates of cancer relative to the general population^{1,9}. Characterization of that risk in rodents could potentially improve the safety of future female astronauts.

The incorporation of multiple bioinformatics tools into a single analysis pipeline has the potential to simplify analysis of transcriptomics data by naive users, opening up analysis of publicly available spaceflight RNA-Seq data to more members of the public. With more equitable, democratic access to analysis tools and information on the effects of spaceflight on more demographics, we will be one step closer to making space travel safer for a

wider range of astronauts from different backgrounds.

THE HAZARDS OF SPACEFLIGHT

Spaceflight is associated with a variety of risks to health in both the short and long term. Microgravity is associated with negative long-term changes to the musculoskeletal system, as well as short-term fluid shifts that affect the cardiovascular system and can induce systemic inflammation¹⁰. Long distance flights outside Earth's protective magnetic field can expose astronauts to high levels of radiation that can damage DNA, leading to epigenetic changes and potential long-term development of cancer¹⁰. Missions on the ISS remain in low Earth orbit but still expose astronauts to the equivalent of two x-rays for every year spent on the station, or approximately 8-10 times the occupational limit for radiation workers in the United States^{10,11}. The carbon dioxide levels in enclosed spaces such as spacecraft trend considerably higher than the atmosphere of Earth - as much as ten times higher - and can result in health issues from chronic hypoxia¹².

The combination of these different hazards can also result in health issues unique to astronauts; for example, the combination of fluid shift, oxidative stress and chronic low-level radiation exposure can result in long term damage to the retina¹⁰. Mitochondrial dysfunction in response to reactive oxygen species is one of the most common physiological changes associated with spaceflight, resulting in alteration of lipid metabolism which in turn can lead to a variety of negative long-term health outcomes^{8,10,13}. As missions are expected to last longer and range further from Earth's protective magnetic field, these negative health outcomes need to be well understood and managed in order to ensure the safety of the crew.

THE RR-8 MISSION AND THE OSD-511 DATASET

In Rodent Research Reference Mission 1 (RRRM-1 or RR-8), female BALB/cAnNTac mice of two different age groups (10-12 and 32 weeks, or "Young" and "Old") were split into three categories

according to experimental and housing conditions: Spaceflight, Ground Control, and Vivarium Control (see Table 1, below)¹⁴. The control mice were divided according to their housing type - ground control mice were housed in the same type of housing as the mice sent to space, while vivarium mice were housed in standard vivariums¹⁴. Ground control mice were housed in densities replicating the living conditions experienced by the spaceflight LAR (live animal return) cohort¹⁴. In the experimental Spaceflight cohort, mice were flown aboard the ISS for 40 days in low orbit, then returned to Earth and allowed to recover for several days prior to sacrifice and processing of their tissues for analysis¹⁴. Prior work by Beheshti et al concluded that cage type used for transcriptome analysis on the ISS does influence gene expression based on the tissue type being studied, so both sets of comparisons between Spaceflight and control groups (Ground Control and Vivarium Control) were retained for this project¹².

The stated goal of the RR-8 mission was to understand the physiology of the previously observed phenomenon of accelerated aging in space by analyzing the behavior, tissues and gene expression profiles of young and old mice⁶. The number of mice involved in the mission was greater than those included in the OSD-511 dataset; in the RR-8 mission a subset of the mice were sacrificed onboard the ISS, but those mice are not included in the dataset despite being described in the experimental details of the OSD-511 page¹⁴. Musculoskeletal tissues and livers of the RR-8 mice were used for other studies, but OSD-511 contains only data obtained via dissection and RNA sequencing of the mammary glands of a subset of RR-8 mice (and not livers as indicated in the study description)¹⁴. The dose of radiation received by the mice throughout the entire trip was approximately 10 mGys, as much as an abdominal computed tomography (CT) scan or three years of background radiation, so it is unlikely that this dose would significantly impact the health of the mice in the study^{11,14}.

Table 1: Identification numbers of mice assigned to each group from OSD-511 dataset

	Old	Young	Total
Spaceflight	FL 3, 4, 8, 11 , 15	FL 2, 5, 6, 14, 17	10
Ground Control	GL 3, 4, 7, 8, 11, 12, 16	GL 1, 2, 5, 6, 13, 14, 17, 18	15
Vivarium Control	VL 3 , 4, 7, 8, 11, 12, 15, 16, 19, 20	VL 1, 2, 5, 6, 9, 10, 13 , 17	18
Total	22	21	43

Numbers in **bold/underline** indicate the three mice which are missing EERC spike-in used to normalize counts to sequencing run control.

Methods

THE CONTAINERIZED RNA-SEQ ANALYSIS PIPELINE
Prior pipelines used to analyze RNA-Seq data from OSDR have incorporated a variety of bioinformatics tools for raw data quality control (QC), trimming, alignment and annotation, which NASA has consolidated into a containerized pipeline using a Nextflow implementation of Singularity^{15,16}. This project used version NF RCP-F 1.0.3 of the OSDR RNA-Seq analysis workflow (Figure 1)¹⁵. This workflow is not only streamlined,

but also accommodates inexperienced users who may not be familiar with the intricacies of different individual tools. The containerized analysis pipeline can be executed using a single command using a high-performance computing cluster (HPC) such as the San José State University College of Science High Performance Computing Cluster, or COS-HPC. Initiation of the workflow on the COS-HPC by a simple bash script enables analysis of OSDR data without the need to download individual sequencing files. Parallelization of analysis pipeline steps is managed via the SLURM job scheduler.

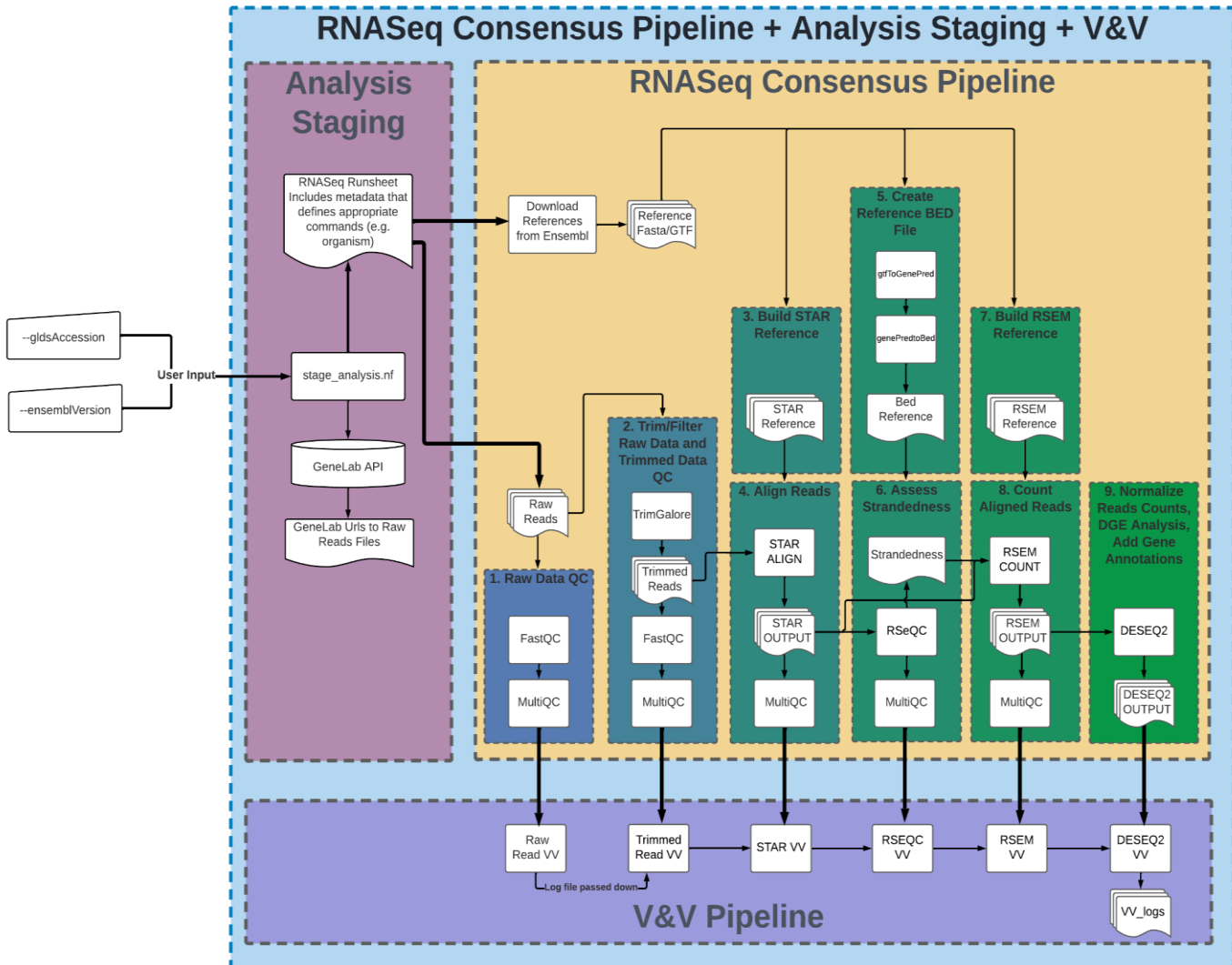


Figure 1: Overview of containerized OSDR RNA-Seq analysis pipeline (figure reprinted with author's permission)¹⁵. This pipeline was implemented via Nextflow on the San José State University College of Science HPC.

The containerized analysis pipeline automates the retrieval and download of sequencing data from OSDR¹⁵. It then trims, aligns and analyzes the quality of the RNA-Seq data retrieved¹⁵. In parallel, a reference genome for the organism specified is downloaded from ENSEMBL and used to create a STAR/RSEM reference for downstream

analysis of alignment, strandedness and counts of the processed reads¹⁷. These reference files are cached and may be reused on later runs in order to expedite analysis. Sequences are aligned to the reference and counted, with normalization occurring both with and without ERCC. Finally, differential gene expression, or DGE, analysis takes place,

comparing counts between all combinations of groups. As part of these comparisons, the DGE step of analysis fits each sample count to a negative binomial distribution and performs multiple hypothesis testing using Wald's test to generate a p-value for each sample.

Analyzed RNA-Seq count datasets are very large: OSD-511 consists of counts of over 29,000 annotated genes from 43 different mice¹⁴. As the number of genes increases, so does the likelihood of encountering false positive results, also known as the false discovery rate (FDR). To account for this, the p-values generated in the DGE step are adjusted to reduce the FDR. The DGE pipeline adjusts p-values using the Benjamini-Hochberg method, which ranks p-values of each gene and retains those smaller than (more significant than) a calculated rejection threshold, depending on the dataset¹⁸.

Finally, the genes are annotated to add ENSEMBL IDs, gene names, and some descriptions of gene function and associated pathways¹⁷. Most genes in the analysis pipeline seem to have associated descriptions, but approximately 8.5k of the 29k genes tested lacked annotation. Manual analysis and annotation of these genes using their ENSEMBL IDs to parse annotations from the latest release of the mouse genome revealed that most are pseudogenes (non-coding RNA that is produced by the cell but does not produce a protein, some of which are thought to play a role in the regulation of gene expression), but several were legitimate genes that code for a product but were not annotated. Further work on the annotation process of this pipeline may yield more useful information for later users.

Output is generated at each major step of the pipeline such as QC, trimming, alignment, normalized counts and DGE. Most of these output files are text or csv files, and several analysis tools such as FastQC and MultiQC generate html files that can be opened in a web browser. This project used JMP version 17 to analyze the majority of the data generated from the pipeline, but the tool used to filter and analyze counted transcripts should not influence the genes found to be differentially expressed.

METHODS FOR ANALYZING THE OSD-511 DATASET

The OSD-511 dataset consists of 43 samples prepared from 300bp ribodepleted RNA libraries with dual index UMI adapters from IDT (IDT, 384-well xGen Dual Index UMI Adapters)¹⁴. They were sequenced using an Illumina NovaSeq 6000 with a

target depth of 60M clusters, producing 149 bp paired-end reads¹⁴. These paired-end forward and reverse reads are denoted by the suffixes “_1” and “_2” respectively¹⁴. Since the total dataset sums to 1.66 TB, automation of sample download and processing by the COS-HPC is a significant improvement over a manual workflow¹⁴.

The normalized count data was retrieved and filtered to remove rows with less than five counts per sample (215) counts across all 43 mice, which removed approximately 8k rows from the 29k row dataset. The rationale of removing genes with low counts is to reduce the number of false positive differentially expressed genes found as compared to the standard DESeq analysis package threshold of ten counts¹⁹. If a gene has an extremely low count, it is difficult to say whether that is a result of an error in the sample preparation, failure to sequence the region adequately, failure to identify the region sequenced, or whether it is truly biologically relevant. The significance level was set at an adjusted p-value of 0.05 based on the FDR adjustment performed in the DGE step, and the threshold for differential expression was set at an absolute value of log base 2 fold change greater than or equal to 1.

DOWNSTREAM GENE ONTOLOGY ANALYSIS

Three tools were selected for downstream analysis of the differentially expressed genes found as a result of filtering: the Molecular Signatures Database (MSigDB), Gene Set Enrichment Analysis (GSEA), and the Database for Annotation, Visualization and Integrated Discovery (DAVID)²⁰⁻²². These tools are in widespread use by the bioinformatics community due to their ease of use and ability to highlight patterns of gene function from otherwise large and complicated datasets. All three are able to take an input list of differentially expressed genes by identifiers such as the names or IDs, returning affected pathways and gene clusters of interest based on gene function and families.

The list of differentially expressed genes for each comparison group were used as the input for each tool, as well as log base 2 fold change in expression as needed. The fold change in expression allows GSEA to output which pathways are up- or down-regulated, which can reveal more precise patterns of affected physiological functions. Output pathways from GSEA were included if they met the recommended FDR q-value cutoff of 0.25²¹. The list of gene clusters from DAVID and MSigDB were selected based on p-value and FDR cutoffs of less than 0.05, as no recommendations were given regarding thresholds.

Results

SEQUENCE QUALITY CONTROL METRICS

Sequence quality was good across all primarily 150 bp samples with Phred scores > 30, indicating 99.9% or better accuracy. MultiQC indicated that no samples in the dataset contained adapter contamination above 0.1%. STAR alignment scores of all samples to genomic reference indicate that the majority of reads are uniquely mapped or map to multiple loci.

DIFFERENTIAL GENE EXPRESSION

Venn diagrams (Figure 2) of DGEs describe how

comparisons of counts between spaceflight and control groups overlap with one another. The largest overlap is in the Young (10-12 week) group of mice, with 225 differentially expressed genes found in common across both the Ground Control vs Spaceflight and Vivarium Control vs Spaceflight comparison groups. A large number of differentially expressed genes were found in the Old/Ground Control vs Spaceflight group that are unique to that group, but very few genes were found to be differentially expressed in the Old/Vivarium vs Spaceflight group.

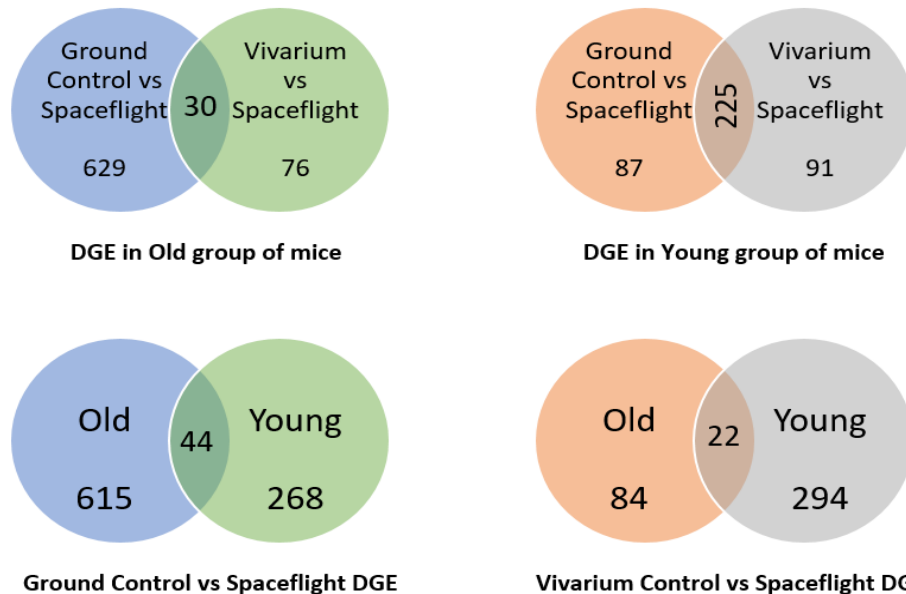


Figure 2: Venn diagrams indicating the number of DGEs found across different comparison groups.

Seven genes were found to be differentially expressed across all comparison groups (Table 2). Three genes (Figure 3) are upregulated in Spaceflight relative to all Earth controls, while four genes are downregulated. Threshold for differential expression was set at absolute log base 2 fold change in expression greater than 1, while significance was set at adjusted p-value less than or equal to 0.05 as per the recommendations from OSDR¹⁵. One differentially expressed pseudogene, *Gm37933*, was removed because it lacked annotation beyond its identity as noncoding RNA.

One gene, *Dipk1b*, does not quite pass the DGE cutoff in one comparison, but it was included because it passes the cutoff in the three other comparisons and was very close to meeting that same criterion in the last group. In Figure 3, the Young Ground Control vs Spaceflight comparison log fold change in expression was multiplied by -1 in order to normalize log fold change in expression between comparison groups (such that the log fold change gives the change in expression of the Spaceflight group relative to control, rather than the control expression relative to Spaceflight).

Table 2: Genes Differentially Expressed Across All Spaceflight/Control Comparisons

ENSEMBL ID	Description	Gene Name
ENSMUSG00000026532	spectrin alpha, erythrocytic 1	Spta1
ENSMUSG00000028785	hippocalcin	Hpca
ENSMUSG00000036186	divergent protein kinase domain 1B	Dipk1b
ENSMUSG00000036523	gene regulated by estrogen in breast cancer protein	Greb1
ENSMUSG00000047298	potassium channel, subfamily V, member 2	Kcnv2
ENSMUSG00000066000	zinc finger protein 979	Zfp979
ENSMUSG00000073830	major urinary protein 14	Mup14

Log Fold Change in Expression of Differentially Expressed Genes in Spaceflight Mice Relative to Control Mice (p-value < 0.05)

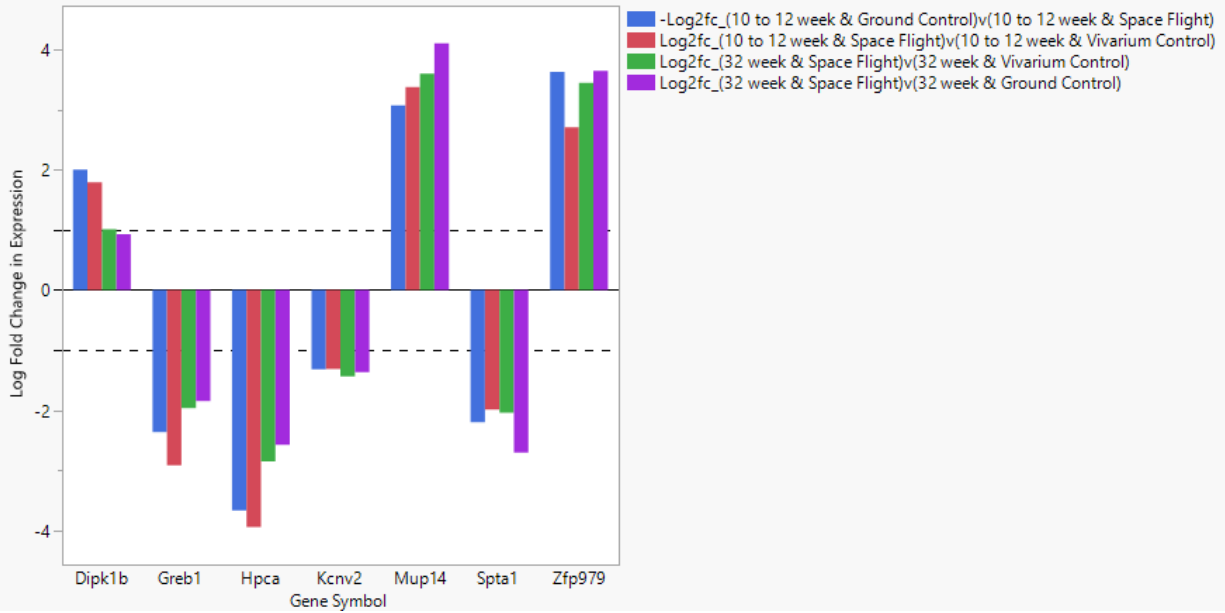


Figure 3: Log base two-fold expression of the seven differentially expressed genes found across all comparison groups.

GENE ONTOLOGY RESULTS

Table 3 shows the top ten pathways identified by MSigDB for both Young control comparison groups, while Tables 4 and 5 show the pathways identified for Old mice. There is significant overlap in the

pathways affected in Young mice, with approximately half of the pathways disrupted in both control comparisons. All calculated FDR q-values for the MSigDB output (Tables 3-5) are below the significance threshold of 0.05²⁰.

Table 3: MSigDB Gene Ontology results for each Young comparison group

Unique to Ground Control vs Spaceflight Group	<ul style="list-style-type: none"> •GOBP CELLULAR LIPID METABOLIC PROCESS •GOBP HOMEOSTATIC PROCESS •GOBP ORGANOPHOSPHATE METABOLIC PROCESS •GOBP ORGANIC HYDROXY COMPOUND BIOSYNTHETIC PROCESS
Both Spaceflight comparison groups (Young mice)	<ul style="list-style-type: none"> •GOBP LIPID METABOLIC PROCESS •GOBP SMALL MOLECULE METABOLIC PROCESS •GOBP LIPID BIOSYNTHETIC PROCESS •GOBP SMALL MOLECULE BIOSYNTHETIC PROCESS •GOBP ORGANIC HYDROXY COMPOUND METABOLIC PROCESS •GOBP STEROL BIOSYNTHETIC PROCESS
Unique to Vivarium Control vs Spaceflight Group	<ul style="list-style-type: none"> •GOBP ALCOHOL METABOLIC PROCESS •GOBP ORGANIC ACID METABOLIC PROCESS •GOBP MONOCARBOXYLIC ACID METABOLIC PROCESS •GOBP STEROL METABOLIC PROCESS

Table 4: MSigDB Results for Old Ground Control vs Spaceflight Comparison

Gene Set Name	Number of Genes in Gene Set	P-Value	FDR q-value
GOCC EXTRACELLULAR SPACE	45	8.98E-10	9.66E-06
GOBP REACTIVE OXYGEN SPECIES METABOLIC PROCESS	13	4.51E-07	2.35E-03
GOMF CYTOKINE BINDING	10	6.55E-07	2.35E-03
GOBP RESPONSE TO INORGANIC SUBSTANCE	18	1.41E-06	3.78E-03

Gene Set Name	Number of Genes in Gene Set	P-Value	FDR q-value
GOBP INFLAMMATORY RESPONSE	22	5.26E-06	1.13E-02
GOBP SECOND MESSENGER MEDIATED SIGNALING	13	6.50E-06	1.16E-02
GOBP MULTICELLULAR ORGANISMAL LEVEL HOMEOSTASIS	23	1.20E-05	1.85E-02
GOBP HOMEOSTATIC PROCESS	35	2.36E-05	3.17E-02
GOBP CELL MOTILITY	34	3.10E-05	3.70E-02
GOBP NERVOUS SYSTEM PROCESS INVOLVED IN REGULATION OF SYSTEMIC ARTERIAL BLOOD PRESSURE	4	3.58E-05	3.75E-02

Table 5: MSigDB Results for Old Vivarium Control vs Spaceflight Comparison

Gene Set Name	Number of Genes in Gene Set	P-Value	FDR q-value
GOBP MUSCLE STRUCTURE DEVELOPMENT	13	2.45E-08	2.64E-04
GOMF CYTOSKELETAL PROTEIN BINDING	14	1.06E-07	5.70E-04
GOCC EXTRACELLULAR SPACE	18	2.96E-07	8.88E-04
GOBP MUSCLE TISSUE DEVELOPMENT	10	3.30E-07	8.88E-04
GOMF ACTIN BINDING	9	1.04E-06	2.11E-03
GOBP REGULATION OF ANATOMICAL STRUCTURE SIZE	10	1.17E-06	2.11E-03
GOMF UDP GLYCOSYLTRANSFERASE ACTIVITY	6	1.45E-06	2.23E-03
GOBP POSITIVE REGULATION OF PROTEIN MODIFICATION PROCESS	12	1.94E-06	2.61E-03
GOBP REGULATION OF PHOSPHORUS METABOLIC PROCESS	14	2.85E-06	3.13E-03
GOBP ANIMAL ORGAN MORPHOGENESIS	13	3.05E-06	3.13E-03

The results from DAVID analysis give greater than 15 clusters for each comparison group; many of these clusters are primarily composed of low significance findings. All comparison groups have significantly differentially expressed genes that belong to clusters related to extracellular organization or cell-cell interactions, but the most clusters from this group were found in Old/Vivarium vs Spaceflight mice. Both Young comparisons found at least four different clusters related to lipid metabolism and the endoplasmic reticulum, while both Old groups had clusters related to G-protein coupled receptors and calcium metabolism. The Old Ground Control/Spaceflight comparison identified differentially expressed genes associated with cardiomyopathy and immune response, similar to the results seen in Table 4, as well as a cluster of genes involved in heme and iron binding. The Young Ground Control/Spaceflight comparison contained several clusters associated with glucose metabolism

and insulin signaling that were not identified in any other comparison group.

GSEA identified upregulated pathways in three of the four comparison groups and a single downregulated pathway in the remaining comparison (Table 6). It did not identify any significantly downregulated gene sets from the differentially expressed genes found in the Young/Ground Control vs Spaceflight or either of the Vivarium vs Spaceflight comparisons, at least not with the recommended FDR (false discovery rate) q-value cutoff of 0.25²¹. No significantly upregulated gene sets were found in the Old/Ground Control vs Spaceflight comparison. Note that normalized enrichment scores (NES) reflect comparisons to control relative to spaceflight, so a negative number represents upregulation.

Table 6: GSEA Differentially Expressed Hallmark Gene Sets from Comparisons of Control Habitats and Spaceflight from All Age Groups

Control Habitat	Age Group	Gene Set Name	Number of Genes found from set	NES	FDR	FWER p-value
Ground	Young	HALLMARK CHOLESTEROL HOMEOSTASIS	15	-1.95	0.055	0.037
Vivarium	Young	HALLMARK EPITHELIAL MESENCHYMAL TRANSITION	11	-2.40	0.002	0.003
Ground	Old	HALLMARK HEME METABOLISM	6	2.05	0.003	0.003
Vivarium	Old	HALLMARK MYOGENESIS	5	-2.01	0.002	0.002

Discussion

Spaceflight appears to induce a variety of gene expression changes. There is significant overlap in differentially expressed genes and gene ontology in the Young mouse groups, while the Old mice are more distinct from each other. This may have been caused by the significant difference in controls for the Old cohort; the Vivarium control for the Old group could be the source of this dissimilarity, as evidenced by the amount of overlap between the different groups (Figure 2). Because the controls in the Young group are more similar, they have more differentially expressed genes in common (Table 3). The gene expression results seen in this study are divided into three sections: metabolic pathway and gene expression changes, age-related changes in gene expression, and changes to gene expression that relate to hormone production.

METABOLIC PATHWAY AND GENE EXPRESSION CHANGES AS A RESULT OF EXPOSURE TO SPACEFLIGHT

Prior studies of spaceflight-induced changes in gene expression have concluded that exposure to space conditions such as microgravity, chronic radiation and enclosed spaces result in changes to lipid metabolism, mitochondrial dysfunction and oxidative stress responses^{8,10,12,13}. Based on MSigDB pathway analysis of the Young cohort, those same pathways appear to be disrupted in the mammary glands of mice. Gene expression associated with lipid metabolism was altered in the Spaceflight mice relative to controls for both young and old mice, especially in younger mice. Beheshti et al observed that mammary gland gene expression in rats showed changes related to inflammatory and oxidative stress following nine days in space¹². Three out of the top five pathways identified by MSigDB as differentially expressed in Old/Spaceflight vs. Ground Control mice pertain to inflammation and oxidative stress, implying that this same expression pattern also occurs in mice.

Several of the genes differentially expressed in all Spaceflight comparison groups are associated with disease states. *Kcvn2*, which is downregulated in all Spaceflight groups, codes for a voltage-gated potassium pump called KV8.2²³. Nonsense mutation of this gene in mice results in degenerative retinopathy, while upregulation in the hippocampus results in increased susceptibility to seizures in model epileptic mice^{23,24}. It is not known what would induce downregulation of *Kcvn2* in mammary tissue, nor what effect that may have. Short and long-term dysfunction of the retina and central nervous system (CNS) are associated with time spent in space, but it is unknown whether *Kcvn2* or a human ortholog plays a role in those health issues⁸.

Spta1, which is also downregulated in all four comparison groups, codes for erythrocytic spectrin alpha, a protein that attaches the cell plasma membrane to external actin and provides structural support for the cell²⁵. *Spta1* is involved in the KEGG apoptosis pathway and is mutated in 10% of all human cancers^{25,26}. However, it is primarily upregulation of *Spta1* that commonly occurs in cancer; downregulation of *Spta1* in clinically relevant conditions such as hereditary spherocytosis causes hemolytic anemia via deformation of red blood cells, making them fragile²⁷. Spaceflight has been shown in prior studies to induce hemolytic anemia, but this phenomenon has largely been attributed to the effects of microgravity, with no known mechanism²⁸. No association between *Spta1* downregulation and spaceflight induced hemolysis is known, but it is possible that *Spta1* could potentially play some role in this process. It is unknown why this gene would experience transcription changes in the mammary gland, but it is also possible that blood was sequenced in addition to sections of the intended tissue.

Other differentially expressed genes support well-characterized effects of spaceflight. Hippocalcin, encoded by the gene *Hpca*, is a calcium-binding

protein that prevents neurological degeneration via mediation of Caspase-12²⁹. Deletion of the hippocalcin gene has been shown to result in increased physiological aging of the brain²⁹. Spaceflight appears to induce downregulation of hippocalcin in all comparison groups in this study, which may support an explanation of the phenomenon of increased aging in space, as well as neurological issues associated with space travel^{13,30}. Expression changes of hippocalcin in mammary tissue, however, have not been well characterized.

In contrast, several other genes found to be differentially expressed in this study are poorly understood. MUP (major urinary protein) proteins originally were discovered in mouse urine and were thought to be involved in solely social behavior such as reproductive signaling and territorial marking, but have since been revealed to play a more complicated role in mouse metabolic processes³¹. *Mup14* is highly upregulated in all Spaceflight mouse comparison groups, but it is unclear whether that is due to social stressors such as double housing and rough landing upon return to Earth, or whether it is due to metabolic changes as a result of spaceflight. The impact of this particular differentially expressed protein is unknown at this time.

Two other genes with altered expression in the Spaceflight mice are *Dipk1b* and *Zfp979*, which code respectively for an uncharacterized protein kinase and a zinc finger binding protein. Neither of these proteins has a known function; the zinc finger protein encoded by *Zfp979* is thought to play a role in transcriptional RNA polymerase binding of regulatory regions but its exact purpose is unknown³². Upregulation of *Zfp979* could play a role in the variety of epigenetic changes that happen as a result of spaceflight, but it is unknown what those changes could be. Further research may

explain the gene expression changes seen in this study as the functions of these proteins are characterized.

GREB1 EXPRESSION CHANGES

Of the genes found to be differentially expressed across all Spaceflight comparisons of mouse mammary tissue, only one encodes a protein nominally associated with mammary tissue. *Greb1*, or growth regulation by estrogen in breast cancer 1, codes for an estrogen regulated cytoplasmic glycosyltransferase that stabilizes ER α (estrogen receptor alpha) via post-translational modification^{33,34}. Both proteins play a vital role in endocrine signaling in reproductive tissues; *Greb1* is highly conserved in eukaryotes, predating evolution of the ER α protein itself³⁴. *Greb1* is upregulated in multiple hormone driven reproductive cancers in the breast, ovaries and prostate, and is required by the process by which hormone-responsive cancers respond to the ER α signaling pathway³³.

Downregulation of *Greb1* in mammary tissue, as is seen in both Spaceflight groups of mice (Figure 4) can occur in a healthcare setting when estrogen-sensitive cancers develop acquired resistance to ER α pathway-targeting therapeutics such as Tamoxifen, disrupting hormone sensitivity³⁵. Knockout of *Greb1* in mice impairs both normal growth and fertility due to destabilization of ER α ³⁴. It is possible that exposure to space-related factors may confer temporary infertility or other endocrine issues, as *Greb1* is required for normal estrogen signaling. Downregulation of *Greb1* could potentially be an issue if an astronaut, male or female, were seeking care for a hormone-responsive cancer, as *Greb1* is a marker required for response to therapeutics prescribed for hormone-sensitive cancers of the breast, ovaries and prostate³⁵.

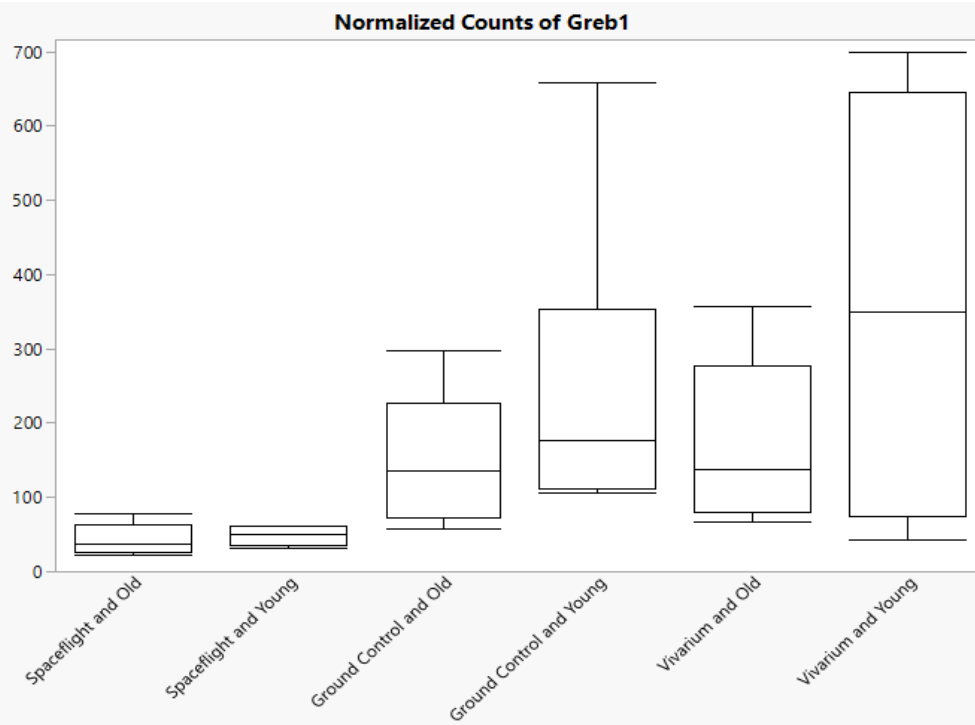


Figure 4: Boxplot of normalized counts of *Greb1*, gene regulated by estrogen in breast cancer 1. Significant (p -value <0.025) downregulation of *Greb1* is seen in both Spaceflight groups regardless of age, while (on average) lower expression of *Greb1* is seen in Old mice relative to Young mice.

Genetic variants of *Greb1* and subsequent disruption of estrogen signaling have been linked to bone mineral loss and osteoporosis, which are also major health concerns for astronauts^{10,36}. Previous studies have attributed space-related bone density changes to microgravity, but it is possible that hormone disruption may play a role as well; prior work has demonstrated that testosterone levels are reduced in space which could potentially lead to bone density loss^{30,36,37}. However, these findings have been disputed, as other studies have found that testosterone levels are unaffected by space travel³⁸.

It is also possible that the downregulation of *Greb1* in mammary tissue is due to some aspect of accelerated physiological aging in space. Old mice tend to express *Greb1* at lower levels than young mice (Figure 4) and *Greb1* expression levels in humans tend to spike during puberty and fertile portions of the menstrual cycle, and decrease during menopause³³. Mice don't go through menopause like humans do; rodents go through a similar process called estropause during which their estrous cycles and ovulation usually cease³⁹. The *Greb1* downregulation seen in Spaceflight mice could potentially indicate an induced state of reproductive senescence, as menopause and estropause are both induced by changes in the CNS and endocrine signaling that result in apoptosis of reproductive tissue³⁹. Spaceflight has been

associated with CNS dysregulation and unexpected apoptosis of some tissues, but it is unknown whether spaceflight could induce menopausal-like changes to the reproductive system⁸. Literature on menopause and space travel is sparse and focuses more on radiation induced destruction of immature follicles in the ovary; maintaining the fertility of female astronauts seems to be a higher priority for the scientific community than retaining hormonal homeostasis⁴⁰.

It is unclear what the mechanism of mammary tissue *Greb1* downregulation in space may be and whether space travel-induced bone density loss is due to mechanical or hormonal causes, or some combination thereof. It is also unknown whether spaceflight-induced *Greb1* downregulation in mammary tissue is indicative of systemic downregulation of *Greb1* gene expression, or whether estrogen levels may be disrupted in spaceflight by this mechanism. Further research may reveal the extent and cause of this disruption, as well as whether *Greb1* expression is able to eventually normalize long term following return to Earth.

AGE-RELATED CHANGES IN GENE EXPRESSION

Responses to spaceflight in mammary tissue appear to vary based on the age of the mice in the study. Young mice seem to experience more significant changes in lipid metabolism associated with

spaceflight (Table 3) while Old Spaceflight mice experienced immune system dysregulation and oxidative stress as compared to the Ground Control group (Table 4-5). The Old cohort of mice seemed to experience spaceflight related gene expression changes more similarly to those documented in prior studies of other mouse tissues and human astronauts, with more similar patterns of disrupted pathways⁸. Old mouse mammary cells experience greater disruption of pathways associated with the extracellular matrix, which replicates results seen in studies of mouse muscle tissue⁸. Inflammation and inflammatory pathway disruption is seen in both human astronauts and the Old cohort; these are well characterized effects of spaceflight physiological stressors^{8,10,13}.

Prior work has postulated that spaceflight accelerates physiological aging - perhaps the effects of spaceflight are more systemic in a group that is already older, extending changes seen in more sensitive tissues such as the liver, into less sensitive tissues such as that of the mammary⁴¹. Like the majority of astronauts, the Old group of mice is not elderly but is middle aged^{2,42}. Modeling the effects of space travel may be more accurate using animal models that are matched with astronauts in species-adjusted age. While the dose of radiation received by the mice of RRM-1 was relatively low, it has been documented that responses to radiation vary with age⁴³. As travel in space ranges further from Earth's protective magnetic field, exposure to radiation will increase dramatically¹⁰. Middle aged animals (and presumably middle aged human astronauts) have less protection from the negative health effects of radiation than young ones, so continued work with this demographic may prove to be helpful in determining the extent and effects of the hazard of space radiation⁴³.

Conclusion

This implementation of the containerized analysis pipeline on the SJSU COS-HPC was able to successfully identify differentially expressed genes from the OSD-511 dataset, as well as changes to expression of sets of genes associated with biological pathways. Many genes were found to be differentially expressed with a significance threshold of 0.05 across each of the different combinations of Spaceflight and two different mouse housing controls, Ground Control and Vivarium. Differential expression of seven genes was found across all combinations of Spaceflight/control comparison groups from this study, some of which may relate to health effects seen in prior spaceflight studies. Many of these

gene transcripts such as *Mup14*, *Dipkb1*, and *Zfp979* are not well characterized in disease states or even normal physiological function, and future work may reveal the role they play in the mammary tissue of mice in space.

One of those genes found in all Spaceflight comparison groups has particular relevance to mammary tissue - *Greb1*. Given *Greb1*'s influence on hormone-sensitive cancer development and response to therapeutics as part of cancer treatment, this change of expression levels in mammary cells could have potential physiological impact. Estrogen is, after all, a hormone not exclusively involved in female or reproductive biological processes (and reproduction is not solely a female process), but also affects growth and bone remodeling over time^{34,36}. Further research into the mechanism of *Greb1* downregulation in mouse mammary tissue, as well as the overall hormone and health impact of this change in expression, could prove beneficial to astronauts regardless of gender.

Analysis of the pathways that are altered in the young and old cohorts of mice in this study has revealed that different pathways are altered across age groups. Younger mice seem to reflect more of the metabolic changes associated with exposure to space, while older mice appear to undergo gene expression changes associated with inflammation, immune system dysregulation and extracellular matrix dysfunction. Older mice may be better models for astronaut health as flights range further from Earth, assuming astronaut demographics continue to trend towards middle aged astronauts². As both missions and astronaut demographics begin to explore new, unexplored territory, further research is needed to identify the problems that will be faced by the next generation of astronauts.

Conflicts of Interest Statement

The authors have no conflicts of interest to declare.

Acknowledgements

Thanks to Jonathan Oribello for his extensive work helping troubleshoot and operate NASA's containerized analysis pipeline, and to Yue Wang and Kevin Ross for their work establishing the Singularity installation on the SJSU COS-HPC. Thanks as well to Dr. Wendy Lee for her help editing the thesis on which this publication was based.

References

1. Cucinotta FA, Durante M. Risk of radiation carcinogenesis. Human health and performance risks of space exploration missions. *NASA SP-2009-3405*. Houston: National Aeronautics and Space Administration. 2009:119-70.
2. Goel N, Bale TL, Epperson CN et al. Effects of sex and gender on adaptation to space: behavioral health. *J of Women's Health*. 2014;23(11):975-86. Doi:10.1089/jwh.2014.4911
3. National Aeronautics and Space Administration. *Artemis - nasa*. Accessed October 20, 2023. <https://www.nasa.gov/humans-in-space/artemis/>
4. Reynolds R, Little MP, Day S et al. Cancer incidence and mortality in the USA Astronaut Corps, 1959–2017. *Occ and env Med*. 2021;78(12):869-75. Doi:10.1136/oemed-2020-107143
5. Barr YR, Bacal K, Jones JA, Hamilton DR. Breast cancer and spaceflight: risk and management. *Aviation, space, and env med*. 2007;78(4):A26-37.
6. National Aeronautics and Space Administration. *Rodent research-8 (rr-8)*. Accessed November 16, 2023. <https://www.nasa.gov/mision/station/research-explorer/investigation/?#id=7713>
7. Berrios DC, Galazka J, Grigorev K, Gebre S, Costes SV. NASA GeneLab: interfaces for the exploration of space omics data. *Nucleic acids research*. 2021;49(D1):D1515-22.
8. da Silveira WA, Fazelinia H, Rosenthal SB et al. Comprehensive multi-omics analysis reveals mitochondrial stress as a central biological hub for spaceflight impact. *Cell*. 2020;183(5):1185-201. Doi:10.1016/j.cell.2020.11.002
9. Nassef MZ, Melnik D, Kopp S et al. Breast cancer cells in microgravity: new aspects for cancer research. *Internat J of Mol Sci*. 2020;21(19):7345. Doi: 10.3390/ijms21197345
10. Afshinnekoo E, Scott RT, MacKay MJ et al.. Fundamental biological features of spaceflight: advancing the field to enable deep-space exploration. *Cell*. 2020;183(5):1162-84. Doi: 10.1016/j.cell.2020.10.050
11. Brenner DJ, Doll R, Goodhead DT et al. Cancer risks attributable to low doses of ionizing radiation: assessing what we really know. *Proceed of Nat Acad of Sci*. 2003;100(24):13761-6. Doi: 10.1073/pnas.2235592100
12. 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. *Scientific reports*. 2018;8(1):4191. Doi: 10.1073/pnas.2235592100
13. Garrett-Bakelman FE, Darshi M, Green SJ et al. The NASA Twins Study: A multidimensional analysis of a year-long human spaceflight. *Science*. 2019;364(6436):eaau8650.
14. Galazka JM, Roberts M, Stodieck L et al. Transcriptional profiling of mammary glands from mice flown on the RRRM-1 mission. *NASA Open Sci Data Repository, Version 2*. Updated August 3, 2022. Accessed September 12, 2023. Doi:10.26030/wdpr-va45 <https://osdr.nasa.gov/bio/repo/data/studies/OSD-511>
15. Oribello J, Saravia-Butler A, Costes S, Gebre S, Galazka, J. GeneLab bioinformatics processing pipeline for Illumina RNA-seqencing data. *GitHub*. Updated August 18, 2022. Accessed September 12, 2023. https://github.com/J-81/GeneLab_Data_Processing/blob/DEV_NF_RCP-F/RNAseq/Pipeline_GL-DPPD-7101_Versions/GL-DPPD-7101-F.md
16. Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C. Nextflow enables reproducible computational workflows. *Nature biotechnology*. 2017;35(4):316-9. Doi: 10.1038/nbt.3820
17. Cunningham F, Allen JE, Allen J et al. Ensembl 2022. *Nucleic acids research*. 2022;50(D1):D988-95. Doi: 10.1093/nar/gkab1049
18. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J of Royal statistical society: series B (Methodological)*. 1995;57(1):289-300. Doi: 10.1016/s0166-4328(01)00297-2
19. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*. 2014;15(12):1-21. Doi: 10.1186/s13059-014-0550-8
20. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database hallmark gene set collection. *Cell sys*. 2015;1(6):417-25. Doi:10.1016/j.cels.2015.12.004
21. Subramanian A, Tamayo P, Mootha VK et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceed of Nat Acad of Sci*. 2005;102(43):15545-50. Doi:10.1073/pnas.0506580102
22. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists

- using DAVID bioinformatics resources. *Nature protocols*. 2009;4(1):44-57. Doi: 10.1038/nprot.2008.211
23. Shahi PK, Srinivasan A, Pattnaik BR. A novel *Kcnv2* nonsense mutation mouse model of Cone Dystrophy with Supernormal Rod Response. *Investigative Ophthalmology & Visual Sci*. 2022;63(7):1784-F0333.
24. Jorge BS, Campbell CM, Miller AR et al. Voltage-gated potassium channel *KCNV2* (Kv8. 2) contributes to epilepsy susceptibility. *Proceed of Nat Acad of Sci*. 2011;108(13):5443-8. Doi:10.1073/pnas.1017539108
25. Swanton C. My Cancer Genome: a unified genomics and clinical trial portal. *The Lancet Oncology*. 2012;13(7):668-9.
26. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*. 2000;28(1):27-30.
27. Chonat S, Risinger M, Sakthivel H et al. The spectrum of *SPTA1*-associated hereditary spherocytosis. *Front in physiology*. 2019;10:815. Doi: 10.3389/fphys.2019.00815
28. Trudel G, Shahin N, Ramsay T, Laneville O, Louati H. Hemolysis contributes to anemia during long-duration space flight. *Nature Med*. 2022;28(1):59-62. Doi:10.1038/s41591-021-01637-7
29. Korhonen L, Hansson I, Kukkonen JP et al. Hippocalcin protects against caspase-12-induced and age-dependent neuronal degeneration. *Mol and Cellular Neurosci*. 2005;28(1):85-95. Doi:10.1016/j.mcn.2004.08.015
30. Strollo F, Rioldino G, Harris B et al. The effect of microgravity on testicular androgen secretion. *Aviation, space, and env med*. 1998;69(2):133-6.
31. Greve S, Kuhn GA, Saenz-de-Juano MD, Ghosh A, von Meyenn F, Giller K. The major urinary protein gene cluster knockout mouse as a novel model for translational metabolism research. *Scientific Reports*. 2022;12(1):13161. Doi:10.1038/s41598-022-17195-y
32. Wheeler DL, Barrett T, Benson DA et al. Database resources of the national center for biotechnology information. *Nucleic acids research*. 2007;36(suppl_1):D13-21. Doi: 10.1093/nar/gkm1000
33. Hodgkinson K, Forrest LA, Vuong N, Garson K, Djordjevic B, Vanderhyden BC. *GREB1* is an estrogen receptor-regulated tumour promoter that is frequently expressed in ovarian cancer. *Oncogene*. 2018;37(44):5873-86. Doi:10.1038/s41388-018-0377-y
34. Shin EM, Huynh VT, Neja SA et al. *GREB1*: An evolutionarily conserved protein with a glycosyltransferase domain links *ERα* glycosylation and stability to cancer. *Sci Advances*. 2021;7(12):eabe2470. Doi:10.1126/sciadv.abe2470
35. Wu Y, Zhang Z, Cenciarini ME et al. Tamoxifen resistance in breast cancer is regulated by the *EZH2-ERα-GREB1* transcriptional axis. *Cancer research*. 2018;78(3):671-84. Doi: 10.1158/0008-5472.CAN-17-1327
36. Hegarty KG, Drummond FJ, Daly M, Shanahan F, Molloy MG. *GREB1* genetic variants are associated with bone mineral density in Caucasians. *J of bone and mineral metabolism*. 2018;36:189-99. Doi:10.1007/s00774-017-0823-x
37. Fitts RH, Trappe SW, Costill DL et al. Prolonged space flight-induced alterations in the structure and function of human skeletal muscle fibres. *J of physiology*. 2010;588(18):3567-92. Doi: 10.1113/jphysiol.2010.188508
38. Smith SM, Heer M, Wang Z, Huntoon CL, Zwart SR. Long-duration space flight and bed rest effects on testosterone and other steroids. *J of Clin Endocrinology & Metabolism*. 2012;97(1):270-8. Doi:10.1210/jc.2011-2233
39. Koebele SV, Bimonte-Nelson HA. Modeling menopause: The utility of rodents in translational behavioral endocrinology research. *Maturitas*. 2016;87:5-17. Doi:10.1016/j.maturitas.2016.01.015
40. Rose BI. Female astronauts: Impact of space radiation on menopause. *Euro J of Obstetrics & Gynecology and Reprod Bio*. 2022;271:210-3. Doi:10.1016/j.ejogrb.2022.02.022
41. Strollo F, Gentile S, Strollo G, Mambro A, Vernikos J. Recent progress in space physiology and aging. *Frontiers in physiology*. 2018;9:1551. Doi:10.3389/fphys.2018.01551
42. Hagan C. When are mice considered old? *JAX Blog*. November 7, 2017. Accessed October 14, 2023. <https://www.jax.org/news-and-insights/jax-blog/2017/November/when-are-mice-considered-old>
43. Shuryak I, Sachs RK, Brenner DJ. Cancer risks after radiation exposure in middle age. *J of Nat Cancer Instit*. 2010;102(21):1628-36. Doi:10.1093/jnci/djq346