Select Pathways Revealed in A. Thaliana During Spaceflight Via Meta-Analysis of National Aeronautics and Space Administration Genelab Data

Zachary Thomas Baham

Follow this and additional works at: https://digitalcommons.latech.edu/theses
SELECT PATHWAYS REVEALED IN A. *THALIANA* DURING SPACEFLIGHT
VIA META-ANALYSIS OF NATIONAL AERONAUTICS AND SPACE
ADMINISTRATION GENELAB DATA

by

Zachary Baham, B.S. in Biology

A Thesis Presented in Partial Fulfillment
of the Requirements of the Degree
Master of Science

COLLEGE OF APPLIED AND NATURAL SCIENCES
LOUISIANA TECH UNIVERSITY

May 2023
LOUISIANA TECH UNIVERSITY

GRADUATE SCHOOL

April 11, 2023

Date of thesis defense

We hereby recommend that the thesis prepared by

Zachary Thomas Baham

entitled

Select Pathways Revealed in A. Thaliana During Spaceflight via Meta-Analysis of National Aeronautics and Space Administration GeneLab Data

be accepted in partial fulfillment of the requirements for the degree of

Master of Science in Biology

Joshua P. Vandenbrink
Supervisor of Thesis Research

William J. Campbell
Head of Biological Sciences

Thesis Committee Members:
Joshua P. Vandenbrink
Jamie Newman
Terri Maness

Approved:
Gary A. Kennedy
Dean of Applied & Natural Sciences

Approved:
Ramu Ramachandran
Dean of the Graduate School
ABSTRACT

In this study, a combination analysis using the pathview and pathfindR packages for the R programming language was performed on a collection of openly available Arabidopsis thaliana sequencing data from the National Aeronautics and Space Administration (NASA) GeneLab to discover broad connections between three primary spaceflight stressor types and altered pathway regulation to provide focused targets for future research. Studies were divided into groups based on their exposure to microgravity, cosmic radiation, or a theorized heat-stress factor caused by reduced convection in certain technologies. Through this study, three major pathways were identified: the phenylpropanoid biosynthesis, proteasome, and starch and sucrose metabolism pathways. In addition, the ribosome and photosynthesis pathways were found to be upregulated in some datasets yet downregulated in others, while the endocytosis and ribosome biogenesis pathways were found to be upregulated and downregulated respectively. This study identifies these pathways and discusses possible reasons for their altered regulation, proposes future directions for study, and addresses open data quality and curation.
APPROVAL FOR SCHOLARLY DISSEMINATION

The author grants to the Prescott Memorial Library of Louisiana Tech University the right to reproduce, by appropriate methods, upon request, any or all portions of this Thesis. It is understood that “proper request” consists of the agreement, on the part of the requesting party, that said reproduction is for his personal use and that subsequent reproduction will not occur without written approval of the author of this Thesis. Further, any portions of the Thesis used in books, papers, and other works must be appropriately referenced to this Thesis.

Finally, the author of this Thesis reserves the right to publish freely, in the literature, at any time, any or all portions of this Thesis.

Author _____________________________

Date _____________________________
DEDICATION

This thesis is dedicated to my mother, my sister, and all my friends who believed in me and supported me on my journey. I couldn’t have done it without every last kind word and well-wish.
TABLE OF CONTENTS

ABSTRACT ........................................................................................................................................ iii
APPROVAL FOR SCHOLARLY DISSEMINATION ........................................................................ iv
DEDICATION .................................................................................................................................. v
LIST OF FIGURES .................................................................................................................... ix
LIST OF TABLES ...................................................................................................................... xii
ACKNOWLEDGMENTS ............................................................................................................. xiii
CHAPTER 1 INTRODUCTION ..................................................................................................... 1
  1.1 Space Exploration Today .............................................................................................. 1
  1.2 Critical Elements and their Lunar Abundance ......................................................... 2
  1.3 Space-Grown Plants for Long-Term Missions ....................................................... 3
CHAPTER 2 BACKGROUND ....................................................................................................... 5
  2.1 Challenges for Space-Grown Plants ......................................................................... 5
  2.2 Space Genomics using Arabidopsis Thaliana ........................................................... 6
    2.2.1 A. thaliana as a Model Plant ........................................................................... 6
    2.2.2 Known Effects of Microgravity as a Stressor .................................................. 7
    2.2.3 Known Effects of Cosmic Radiation as a Stressor ......................................... 7
    2.2.4 Convection and the Potential for Heat Stress in Airtight Containers .......... 9
    2.2.5 The Need for Broad Analysis ........................................................................ 10
CHAPTER 3 METHODS ............................................................................................................... 12
  3.1 Data Collection ........................................................................................................... 12
  3.2 Data Filtering and Selection ...................................................................................... 12
6.2.2 Plant Immune Function in Sterile Environments ........................................ 42
6.2.3 Effects of Proteasome Upregulation on Proteomics ............................... 42
6.2.4 Starch and Sucrose Metabolism and Heat Stress ................................. 42
6.2.5 Data Availability, Formatting, and Metadata .......................................... 43
6.2.6 Closing Remarks ................................................................................... 44

APPENDIX A SUPPLEMENTAL FIGURES ......................................................... 46
A.1 Phenylpropanoid Biosynthesis Pathview Diagrams ................................. 47
A.2 Starch and Sucrose Metabolism Pathview Diagrams ............................... 53
A.3 Proteasome Pathview Diagrams .............................................................. 55
A.4 Plant-pathogen Interaction Pathview Diagrams ....................................... 62
A.5 Photosynthesis Pathview Diagrams .......................................................... 67
A.6 Ribosome Pathview Diagrams ................................................................. 78

APPENDIX B SUPPLEMENTAL DATA .............................................................. 95
B.1 R Programs and Packages ........................................................................ 96
B.2 Additional Analysis Information ............................................................. 98

BIBLIOGRAPHY .............................................................................................. 101
## LIST OF FIGURES

| Figure 3-1: | Data stratification of usable datasets | 14 |
| Figure 3-2: | Data processing pipeline before pathway analysis | 17 |
| Figure 4-1: | Summary figure of all significant pathways of altered regulation by group | 19 |
| Figure 4-2: | PathfindR enrichment diagram for HSRC | 22 |
| Figure 4-3: | PathfindR enrichment diagram for microgravity | 25 |
| Figure 4-4: | PathfindR enrichment diagram for radiation | 28 |
| Figure A-1: | Pathview diagram of GLDS-38 phenylpropanoid biosynthesis pathway | 47 |
| Figure A-2: | Pathview diagram of GLDS-120 (PhyD mutant, dark growth) phenylpropanoid biosynthesis pathway | 48 |
| Figure A-3: | Pathview diagram of GLDS-314 (red light) phenylpropanoid biosynthesis pathway | 49 |
| Figure A-4: | Pathview diagram of GLDS-469 (transgenic plants, 8-hour light period) phenylpropanoid biosynthesis pathway | 50 |
| Figure A-5: | Pathview diagram of GLDS-469 (wild-type, 8-hour light period) phenylpropanoid biosynthesis pathway | 51 |
| Figure A-6: | Pathview diagram of GLDS-469 (wild-type, 16-hour light period) phenylpropanoid biosynthesis pathway | 52 |
| Figure A-7: | Pathview diagram of GLDS-38 starch and sucrose metabolism pathway | 53 |
| Figure A-8: | Pathview diagram of GLDS-120 (Col-0, light growth) starch and sucrose metabolism pathway | 54 |
| Figure A-9: | Pathview diagram of GLDS-7 (hypocotyl) proteasome pathway | 56 |
| Figure A-10: | Pathview diagram of GLDS-7 (shoot) proteasome pathway | 57 |
| Figure A-11: | Pathview diagram of GLDS-213 (0G) proteasome pathway | 58 |
Figure A-12: Pathview diagram of GLDS-469 (transgenic plants, 16-hour light period) proteasome pathway ................................................. 59
Figure A-13: Pathview diagram of GLDS-469 (wild-type, 8-hour light period) proteasome pathway ............................................... 60
Figure A-14: Pathview diagram of GLDS-469 (wild-type, 16-hour light period) proteasome pathway .................................................. 61
Figure A-15: Pathview diagram of GLDS-120 (Col-0, light growth) plant-pathogen interaction pathway ............................................. 63
Figure A-16: Pathview diagram of GLDS-469 (transgenic plants, 16-hour light period) plant-pathogen interaction pathway ......................... 64
Figure A-17: Pathview diagram of GLDS-469 (wild-type, 8-hour light period) plant-pathogen interaction pathway ........................................ 65
Figure A-18: Pathview diagram of GLDS-469 (wild-type, 16-hour light period) plant-pathogen interaction pathway ........................................ 66
Figure A-19: Pathview diagram of GLDS-7 (whole-plant) photosynthesis pathway ..... 67
Figure A-20: Pathview diagram of GLDS-38 photosynthesis pathway ................. 68
Figure A-21: Pathview diagram of GLDS-213 (0G) photosynthesis pathway .......... 69
Figure A-22: Pathview diagram of GLDS-213 (1G) photosynthesis pathway .......... 70
Figure A-23: Pathview diagram of GLDS-7 (hypocotyl) photosynthesis pathway....... 71
Figure A-24: Pathview diagram of GLDS-7 (shoot) photosynthesis pathway .......... 72
Figure A-25: Pathview diagram of GLDS-120 (Col-0, light growth) photosynthesis pathway ........................................................................ 73
Figure A-26: Pathview diagram of GLDS-314 (Red light) photosynthesis pathway ...... 74
Figure A-27: Pathview diagram of GLDS-469 (transgenic plants, 8-hour light period) photosynthesis pathway ............................................. 75
Figure A-28: Pathview diagram of GLDS-469 (transgenic plants, 16-hour light period) photosynthesis pathway ............................................. 76
Figure A-29: Pathview diagram of GLDS-469 (wild-type, 16-hour light period) photosynthesis pathway ............................................. 77
Figure A-30: Pathview diagram of GLDS-7 (whole-plant) ribosome pathway .......... 79
Figure A-31: Pathview diagram of GLDS-38 ribosome pathway ........................................ 80
Figure A-32: Pathview diagram of GLDS-120 (WS, light grown) ribosome pathway ... 81
Figure A-33: Pathview diagram of GLDS-251 ribosome pathway ........................................ 82
Figure A-34: Pathview diagram of GLDS-314 (dark) ribosome pathway ............................. 83
Figure A-35: Pathview diagram of GLDS-314 (dark) ribosome pathway ............................. 84
Figure A-36: Pathview diagram of GLDS-7 (hypocotyl) ribosome pathway ....................... 85
Figure A-37: Pathview diagram of GLDS-7 (shoot) ribosome pathway .............................. 86
Figure A-38: Pathview diagram of GLDS-120 (Col-0, light grown) ribosome pathway .......... 87
Figure A-39: Pathview diagram of GLDS-120 (WS, dark grown) ribosome pathway ... 88
Figure A-40: Pathview diagram of GLDS-205 ribosome pathway ...................................... 89
Figure A-41: Pathview diagram of GLDS-213 (1G) ribosome pathway ............................. 90
Figure A-42: Pathview diagram of GLDS-469 (transgenic plants, 8-hour light period) ribosome pathway ................................................................. 91
Figure A-43: Pathview diagram of GLDS-469 (transgenic plants, 16-hour light period) ribosome pathway ................................................................................................. 92
Figure A-44: Pathview diagram of GLDS-469 (wild-type, 8-hour light period) ribosome pathway ................................................................. 93
Figure A-45: Pathview diagram of GLDS-469 (wild-type, 16-hour light period) ribosome pathway ................................................................................................. 94
LIST OF TABLES

**Table 4-1:** Upregulated Pathways from the HSRC Stressor Group ................................ 21

**Table 4-2:** Downregulated Pathways from the HSRC Stressor Group ............................. 21

**Table 4-3:** Upregulated Pathways from the Microgravity Stressor Group ..................... 23

**Table 4-4:** Downregulated Pathways from the Microgravity Stressor Group ................. 24

**Table 4-5:** Upregulated Pathways from the Radiation Stressor Group ......................... 26

**Table 4-6:** Downregulated Pathways from the Radiation Stressor Group ...................... 27

**Table B-1:** Packages Used in Pathview Analysis .......................................................... 96

**Table B-2:** Packages Used in PathfindR Analysis ......................................................... 97

**Table B-3:** Commonality Thresholds for Filtering Upregulated Pathways .................... 98

**Table B-4:** Commonality Thresholds for Filtering Downregulated Pathways ............... 99

**Table B-5:** Number of Utilized Datasets by Study ....................................................... 99
ACKNOWLEDGMENTS

I’d like to thank Dr. Joshua Vandenbrink for assisting me in getting my project started and supporting me with his experience throughout the process. I would also like to thank Dr. Jamie Newman for being a fantastic committee member and someone I could rely on throughout my graduate experience. I want to also offer my thanks to Dr. Terri Maness for the different perspective she brought to the table, and the important criticisms she provided for refining my thesis. I would also like to thank Louisiana Tech University for giving me this opportunity to nourish my mind with knowledge. My colleagues in the biology graduate school also deserve many thanks, as their kind words and encouragement gave me much motivation to carry on.

Finally, and most importantly, I must thank my mother Kelli Baham and my sister Emily Baham. They have been unrelentingly positive and supportive in my journey, and I could not have done it without them at my side.
CHAPTER 1

INTRODUCTION

1.1 Space Exploration Today

Space is regarded as humanity’s final frontier; an environment that is so hostile to life that elaborate technology must be developed to visit it even briefly. Regardless of the danger, people continue to push boundaries to explore outer space and learn more about our world through it. The micro- and zero-gravity conditions only easily experienced in space offer us an opportunity to explore multiple areas of research, including areas of human health as well as advancements in resource acquisition. For instance, studies under microgravity aboard the International Space Station (ISS) on amyloid formation – a process tied to neurodegenerative diseases such as Alzheimer’s disease – have resulted in more detailed characterization of how these proteins crystallize. Furthermore, protein crystallization studies in microgravity have given researchers more detailed structural data for approximately 40-50% of analyzed proteins, leading to new drugs that treat muscular dystrophy, breast cancer, or periodontal disease. From a more economic point of view, research in space has also expanded the field of biomining, which is the practice of using microorganisms to leach valuable material, such as rare earth elements, from rock to then be extracted for our use. Space exploration pushes science consistently further, and even now, many people talk of establishing a lunar outpost and a “gateway”
system – simply designated as the Gateway program by the National Aeronautics and Space Administration (NASA) - that could serve as a launching platform for trips to and from the Moon or even further into our solar system\(^6,7\). For what purpose would such a gateway be needed, however? One of the strongest reasons today is to acquire resources typically rare or unavailable on our planet. In fact, recent research has revealed ways to utilize microbes in the acquisition of specifically rare earth metals\(^3\).

1.2 Critical Elements and their Lunar Abundance

Many of the modern technologies we use daily rely on elements like indium, cobalt, or tritium. Each of these elements are considered high-criticality, meaning that they have a vulnerable supply yet are commonly used or relied on\(^8\). High criticality can be a result of political control over a given element – such as the Democratic Republic of Congo controlling most of the global cobalt supply – or ease of acquisition such as in the case of indium – found in small quantities only as a byproduct in zinc smelting – or tritium, which naturally occurs in the atmosphere\(^9-12\). To illustrate their importance, it should be noted that cobalt is a key component in lithium-ion batteries found in numerous electronics and electric cars, indium is necessary in the compound indium tin oxide – a transparent semiconductor found in touch-screens – and tritium is used as a nonhazardous self-luminescing compound and candidate fuel for fusion reactors\(^11,13-16\). While these elements are high-criticality on Earth, their presence upon our closest celestial neighbor is of great consideration. Lunar regolith – the thin, soil-like uppermost layer of the moon’s surface – contains concentrations of approximately 1.4 to 15 parts-per-billion (ppb) tritium on its sunlit areas, and may contain as much as 50 ppb in permanently dark
In addition, sections of regolith are designated as KREEP – an acronym that stands for potassium (K), rare earth elements (REE), and phosphorous (P) – which contains potential reserves of other rare earth elements labeled as high-criticality on Earth\textsuperscript{19}. Unmanned or temporary missions to the Moon could recover these materials, but it would also be feasible to establish a lunar colony to mine them. In fact, resources mined from the Moon’s surface could instead be used to construct many of the technologies required on a long-term habitation, reducing the number of supply ships such a base might require for materials. However, there is another unique limiting factor to using astronauts or having people live on such bases, and that is food.

### 1.3 Space-Grown Plants for Long-Term Missions

People are some of the most valuable and capable assets for any space mission and are key to establishing colonies on other planets or our moon. Sending people on long-duration space missions or to colonize celestial bodies has been discussed in fiction and at conferences for some time, frequently revealing the challenges that such vessels must overcome. The primary limitation that we continue to toil against is how to provide the astronauts on these vessels with adequate sustenance. As additional astronauts and days are added to a mission, the quantity of food and water must increase, thereby increasing the payload weight of any spaceship\textsuperscript{20}. An increased payload weight necessarily increases the required fuel to launch a given spaceship, further increasing required weight; adding non-fuel weight will necessarily decrease the maximal distance a rocket can travel. As such, it is crucial to keep vessel weight as low as possible, which places a ceiling on the amount that can be sent into space\textsuperscript{21}. Therefore, proposals for providing lighter weight sustenance than traditional prepackaged meals are in demand.
One proposed method is to provide astronauts on these long flights or colonization missions the capability to grow their own food. Furthermore, space-grown edible plants present multiple benefits to the astronauts. Such organisms can participate in absorbing CO₂ and emitting O₂ to supplement existing cycling systems and help to recyle water and nutrients through transpiration and absorbance from human waste\textsuperscript{22,23}. However, plants are not typically suited to the conditions of spaceflight. Regardless, space-grown edible plants could open the doors for longer-term space missions and enable the acquisition of high-criticality elements. As such, research into making space-grown plants a reality would advance our space exploration efforts.
CHAPTER 2
BACKGROUND

2.1 Challenges for Space-Grown Plants

Plant architecture largely relies on standard gravity, thus making the microgravity experienced in spaceflight highly stressful to withstand\(^2^4\). We currently have an incomplete understanding of how a lack of gravity affects plants due to Earth’s gravity being largely inalienable; testing the response of plants to low gravity has frequently been conducted using approximations via random-positioning machines and clinostats\(^2^5\). In terms of physical response to these analogs and true microgravity, plants have shown altered growth patterns, increased heat-shock protein expression, and diminished immune response when placed on such clinostats or into spaceflight environments\(^2^6\)–\(^3^0\).

Furthermore, while many plants are well-acclimated to managing the effects of UV exposure, they are not as well-prepared to handle the unique components of cosmic radiation that a typical space vessel is exposed to which includes alpha, beta, as well as gamma ray particles\(^3^1\). These particles can damage a plant’s DNA, requiring the plant to spend large quantities of energy to repair and thereby lowering the pool of energy the plant can use for healthy growth\(^3^2\). Lastly, plants rely on the principle of convection to manage heat buildup\(^3^3\). Their metabolic processes generate a small amount of heat which is typically carried away by convective airflow, but a lack of gravity within sealed growth chambers may result in reduced convection. In these situations, plants are at risk of
accumulating unhealthy amounts of heat and triggering heat stress on top of other stresses of spaceflight. This particular stressor has been called HSRC (heat stress due to reduced convection). Though these problems are many, we can begin to understand plant responses to these stressors and perhaps mitigate them through the study of plant genomics, transcriptomics, and epigenomics.

2.2 Space Genomics using *Arabidopsis Thaliana*

2.2.1 *A. thaliana* as a Model Plant

The most common organism utilized for both broad plant genetic studies and its spaceflight counterparts is *Arabidopsis thaliana*, also known as thale cress. Friedrich Laibach was the first to propose *A. thaliana* as a model organism in 1945, but the plant gained proper attention as a model in the 1980s alongside maize, petunia, and tobacco\(^34\). Eventually, in 1986, *A. thaliana* had a breakthrough as a model that drove it into the center of attention: T-DNA-mediated transformation and the first successful cloning of an *A. thaliana* gene were described\(^{35,36}\). *A. thaliana*’s genome is relatively small, composed of about 157 megabase pairs on five diploid chromosomes, making sequencing and genetic mapping relatively simple\(^37\). It was the first plant to have its full genome sequenced, which was done in 2000 by the Arabidopsis Genome Initiative\(^38\).

Furthermore, *A. thaliana* can self-pollinate making it relatively easy to grow, store, and clone. It takes only 6 weeks to go from germination to mature seed\(^39\). Thus, many different labs across the world use *A. thaliana* as a model system for plant genetics. *A. thaliana*’s expansive genome annotation, accessible sequencing methods, and ease of growth make it a prime option for the study of the effects of spaceflight stressors on plants in general.
2.2.2 **Known Effects of Microgravity as a Stressor**

Plants such as *A. thaliana* have all adapted to a consistent gravitational value of 1 × g across Earth’s surface. This force is therefore important for plant growth and tropism\(^{40}\). Such tropisms are especially associated with directed growth and plant architecture in that shoot growth exhibits tropism away from the direction of gravity’s pull and towards light, while roots are opposite in following the gravity vector and avoiding light\(^{41,42}\). In addition to the main root/shoot’s tropisms, lateral roots and shoots maintain their own angles for growth based on another system partly controlled by gravity sensation\(^{43}\). A recent theory of graviperception in plants, dubbed the graviproprioceptive drive theory (GPD), proposes that stem straightness utilizes statoliths – membrane-bound starch grains within gravisensing cells previously described in the starch-statolith hypothesis – alongside mechanoreception by gravity-based tissue bending\(^{44–47}\).

As mentioned previously, plants grown in simulated microgravity display a number of maladaptive phenotypes such as altered growth, upregulation of heat-shock protein expression, and inhibited immune response\(^{26–28,30}\). These alterations are likely due to shifts in the plants’ proteome and transcriptome, as these mechanisms are responsible for managing the form and function of organisms. As such, it would follow that plant genetic research is a direct pathway to better understand these responses.

2.2.3 **Known Effects of Cosmic Radiation as a Stressor**

Radiation emitted by stars such as our Sun can travel through space without interference from magnetic barriers such as those found enveloping Earth\(^{48}\). Plants have
developed under the cover of our atmosphere, which protects them from more damaging alpha, beta, and gamma rays that the sun produces \(^{31}\).

There are two main types of cosmic radiation, delineated as Galactic Cosmic Rays (GCR) and Solar Particle Events (SPE), which both occur in slightly different fashion \(^{49}\). GCR are constant background radiation that can be observed throughout space, consisting of approximately 85% protons, 14% helium ions, and 1% heavier ions, while SPE are non-predictable bursts of radiation multiple times stronger than GCR. Both types of radiation produce secondary radiation when entering a shield such as the outer hull of a space station or spacecraft. This secondary radiation has been shown to influence space-grown plants by inducing mutations in their genomes \(^{50,51}\), such as a 30.2% genetic polymorphism rate observed in space grown rice \(^{52}\). The radiation can further cause complications by damaging single nucleotides, or creating single- and double-stranded breaks \(^{53}\). Though DNA repair mechanisms can typically repair such damage, \(\alpha\)-rays impair the ability of homologous recombination mechanisms to function, resulting in cellular death \(^{54}\). Space radiation causes chromosomal abnormalities in \(A.\ thaliana\) which reduces fertility, seed germination capability, and survival rate \(^{55}\).

However, cosmic radiation can be a useful catalyst in the creation of genetically-modified crops. For example, Chinese space breeding programs have produced 66 varieties of cotton, rice, tomato, and wheat \(^{56}\). Studying radiation as a method of generating new strains of plants for growth on Earth as well as in spaceflight is therefore valuable. The pathways observed in this research could be used as guides to manipulate aspects of \(A.\ thaliana\) and similar plants.
2.2.4 Convection and the Potential for Heat Stress in Airtight Containers

Convection is a type of movement that is dependent on gravity and, typically, temperature; gravity pulls more strongly upon denser substances such as cold air, leading to the well-known effect that heat rises while cool air falls\textsuperscript{57,58}. This flow does not occur in spaceflight conditions to any appreciable degree and has been shown in previous experiments to negatively impact plant growth, particularly due to effects of ethylene buildup and heat stress\textsuperscript{59}.

Ethylene is a plant signaling molecule responsible for inhibiting root elongation alongside the more commonly known molecule auxin\textsuperscript{60,61}. In addition, ethylene plays a key role in restriction of root growth in dense soil\textsuperscript{62}. Under normal conditions, diffusive processes evacuate ethylene from plants due in part to convection, as the heat generated by the plant will cause small amounts of convective airflow away from the plant even under windless conditions\textsuperscript{63,64}. In spaceflight, however, microgravity results in an artificial and unintended buildup of ethylene that thusly stunts root growth\textsuperscript{65}.

In addition to concerns regarding ethylene buildup, a lack of convection may result in simple heat buildup. Without gravity convection cannot drive hot air away and heat may remain on the plant’s exterior surface. Though responsible for other stresses, heat shock proteins (HSPs) were notably first discovered as chaperone molecules that aid in resisting heat-induced protein denaturation\textsuperscript{66–68}. In studies such as Choi et al.’s work in 2019, HSPs have been strongly induced in spaceflight when grown in hardware that provided little-to-no gas circulation for the subject plants\textsuperscript{69}. It is thus worth investigating experiments where restricted convection may be affecting pathways associated with heat stress.
2.2.5 The Need for Broad Analysis

Characterizing plant spaceflight stress at the genomic level is crucial to fully understanding the responses that are produced in such conditions, as well as accurately addressing those responses that may be harmful to the plant. Significant research has already been done to sequence and map pathways of differential expression among plants in space versus ground controls, but it is also important to itemize a list of what genes and pathways govern which responses. A study that looks exclusively at one stressor at the exclusion of all others is rare, so it becomes necessary to cross-reference multiple differential expression datasets to pinpoint what genes and pathways are common to different stressors. Moreover, pathways are particularly key to illustrate the systemic changes wrought by individual genetic changes, bridging the gap from more conceptual genetic function to demonstrable phenotype. With accurate lists of associated pathways, it may then be possible to perform follow-up research to identify plants that are amenable to typical spaceflight stressors or transgenically introduce adaptive genes to staple food items. Furthermore, especially in the case of radiation and heat stress, pinpointing responsible pathways and genes could allow future researchers to identify what stressors a plant faces in various growth chambers and centrifuges. For instance, if a plant grown in a particular growth chamber exhibits upregulation in pathways associated with heat stress, but no significant changes in radiation response pathways, it would be a reasonable conclusion that such a chamber is well-shielded from radiation but may require airflow management. Generation of data such as these is ongoing, but early analyses can guide future work. Broad analyses such as this can take together the breadth of multiple individual experiments to identify key common regulatory changes.
The full array of transcriptomic and physiological changes in plants during spaceflight has been broad and unwieldy to research, resulting in efforts to design tools and databases to aid researchers in examining, processing, and drawing conclusions about transcriptomic changes in the model plant *A. thaliana*\textsuperscript{70}. These efforts have focused upon providing open-source data for individuals to analyze, group, and draw conclusions upon. These databases and tools, while fantastic resources, still require analysis to produce a usable data product. In this study, I conducted a meta-analysis using 21 total datasets from 8 different experiments from NASA’s GeneLab repository to characterize what gene pathways were commonly found to correlate with stresses *A. thaliana* commonly encounters in spaceflight.
CHAPTER 3

METHODS

3.1 Data Collection

To conduct this study, I first sourced data from experiments hosted on NASA’s GeneLab website at genelab.nasa.gov. This was an appropriate source to utilize due to the ease of data acquisition, plentiful annotation and notes on the experiments, and quantity of individual experiments hosted. Furthermore, all data at genelab.nasa.gov follows a pipeline of data preparation and quality control outlined in the NASA GeneLab RNA-Seq Consensus Pipeline\textsuperscript{74}. I selected datasets from GeneLab that met three primary requirements: (1) the data set came from an experiment that included spaceflight trials or trials that otherwise exposed the subject organism to cosmic radiation, HSRC, or microgravity and included a ground control, in which the organisms were studied on Earth in equitable conditions; (2) the study organism was \emph{Arabidopsis thaliana}; and (3) the experimental data was available as microarray CEL files or normalized RNA-Seq counts files. Nineteen individual experiments met these criteria and were used for further analysis.

3.2 Data Filtering and Selection

I manually reviewed these selected experiments to evaluate the stressors \emph{A. Thaliana} were exposed to and culled any datasets that may have been affected by the use
of knockout or other mutants. I grouped the datasets by the stressors they were exposed to from the following list: microgravity, radiation, or HSRC. I organized the datasets into the microgravity group based on metadata provided by GeneLab descriptions for experiments. Datasets were determined to be exposed to radiation based on their spaceflight status; those organisms that were grown or exposed to conditions beyond Earth’s atmosphere were considered to undergo exposure to cosmic radiation. Finally, datasets were determined to undergo stress due to HSRC depending on the technology used to grow them. Sealed plates and BRIC units, for example, do not provide any airflow and therefore were classified into this category. A total of 8 datasets were available for use in the analysis after curation, including a total of 21 distinct spaceflight datasets that were stratified into their stressor groups (Figure 3-1).
3.3 Data Processing Protocol

I then analyzed the acquired data using R Studio version 2022.07.2+576 running R Version 4.2.1. A full list of utilized packages with their version numbers and citations is provided in Table B-1. I carried out the analysis in three stages: analysis and pathway generation, data cleanup, and data comparison. Analysis and pathway generation varied depending on if the input data was sourced from a CEL file or RNA-seq counts.
3.3.1 Microarray Processing

For CEL files derived from microarray experiments, I manually created a comma-separated values file (CSV) to identify which CEL file corresponded to spaceflight or ground experimental groups. I then applied this matrix to the CEL data after it was read into R Studio and converted into an ExpressionSet object using the affy package. This conversion process was done via robust multi-array average expression measure according to affy defaults and utilized the Affymetrix-provided ath1121501.db file to name probes. I then fitted linear models to the data and calculated empirical Bayes-moderated t-statistics, the F-statistic, and log-odds of differential expression using functions of the limma package. I adjusted for false discovery rate using in-built functionality of limma’s topTable function. I then wrote the resulting analyzed differential expression data to its own CSV file. This process was repeated for all experiments whose data was provided as CEL files as a result of microarray methodology.

3.3.2 RNA-Seq Processing

For RNA-seq counts, I created a matrix from the base data by rounding imported datasets, then manually creating a CSV file to assign spaceflight or ground status to each sample. Then, I used the DESeq package to create a DESeq Data Set object from the resulting matrix of the aforementioned files. I fed the DESeq Data Set object to DESeq’s primary wrapper function to calculate log fold changes and p-value significance for each gene. Log fold changes were calculated as log2 of ground over spaceflight expression, and the p-value was calculated by Wald test. I ordered the resulting table by p-value and wrote the data to a CSV.
3.3.3 Data Cleanup and Preparation for Rendering

I then cleaned the data using a script that removed repeat gene entries, reduced the data to only relevant columns, filtered by an adjusted p-value < 0.1, and separated the files into upregulated and downregulated genes & pathways. I combined the cleaned datasets into one data frame in R using the dplyr package to full join by gene name, then sorted by commonality. Commonality for this analysis is an arbitrary score defined as the number of datasets that share significant upregulation or downregulation of a given pathway or gene out of the total number of datasets that share a stressor. I selected a threshold to cull genes and pathways only significant in a small proportion of the experiments. This threshold varied by stressor group and upregulation/downregulation, but maintained an approximate ratio of 1 significant dataset with the named pathway to 3 datasets without it (Table B-3). The outlined statistical methods were used for all datasets (Figure 3-2).
**Figure 3-2:** Data processing pipeline before pathway analysis.
3.4 Data Graphing and Rendering

Figures generated for this analysis were made using the pathview and pathfindR packages\textsuperscript{83,84}. I generated pathview diagrams using pathways that were not culled by the threshold method stated previously using package defaults. I generated pathfindR graphics using all analyzed and annotated differential expression data from both RNA-Seq and microarray experiments. KEGG gene sets were used to run the pathfindR wrapper function with the method set to fuzzy\textsuperscript{84,85}. For pathfindR, R version 4.2.2\textsuperscript{77} and R Studio version 2022.12.0+353 was used alongside other updated packages as listed separately in Table B-2. The output of this function was used to generate enrichment diagrams as a secondary analysis to compare the pathview diagrams with. All pathways were annotated using org.At.tair.db\textsuperscript{86}. 
CHAPTER 4

RESULTS

4.1 HSRC Pathways

Multiple pathways of interest were identified as significant in a given stressor’s group, seen below (Figure 4-1).

<table>
<thead>
<tr>
<th>HSRC Upregulated</th>
<th>HSRC Downregulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribosome (ath03010)</td>
<td>Ribosome (ath03010)</td>
</tr>
<tr>
<td>Phenylpropanoid biosynthesis (ath0940)</td>
<td>Starch and sucrose metabolism (ath00500)</td>
</tr>
<tr>
<td>Photosynthesis (ath00195)</td>
<td>Phenylpropanoid biosynthesis (ath00940)</td>
</tr>
<tr>
<td>Proteasome (ath03050)</td>
<td>Endocytosis (ath04144)</td>
</tr>
</tbody>
</table>

Figure 4-1: Summary figure of all significant pathways of altered regulation by group. Pathway names are given, with pathway codes parenthetically noted.

In those plants which were subjected to the postulated HSRC stressor, two notable, unique pathways were found in highest commonality. Primarily, the starch and sucrose metabolism pathway (ath00500) was identified as significantly downregulated in 2 out of 3 HSRC datasets that contained significant downregulation (Table 4-1, 4-2). Among all datasets contained within the HSRC set, significant downregulation of this pathway was found in 2 out of 8. Additionally, the phenylpropanoid biosynthesis
pathway (ath00940) was found to be upregulated in 2 out of 5 HSRC datasets with significant upregulation and upregulated in 2 out of 8 sets among all included in the HSRC group. These pathways are related to structural support of the plant, as well as stress\textsuperscript{87}.

The only other common significant pathway found in this grouping was significantly upregulated in some studies, but significantly downregulated in others: ath03010, the ribosome pathway. More specifically, downregulation of this pathway was observed in 2 out of 8 HSRC datasets while it was also found to be upregulated in 3 out of 8 separate datasets (Table 4-1, 4-2). Note the Commonality column in Tables A-1 through A-6; these describe how many datasets exhibited up/downregulation out of the number of datasets with any significant results. The Commonality (Total) column describes how many datasets exhibited up/downregulation out of the total number of datasets categorized into the given stressor group.

In the secondary pathfindR analysis, HSRC datasets notably exhibited significant enrichment in the starch and sucrose metabolism pathway as well as the phenylalanine and tyrosine metabolism terms (Figure 4-2). The starch and sucrose pathway is directly supported as a result, but the phenylpropanoid pathway previously identified in the pathview analysis was indirectly supported by the phenylalanine and tyrosine pathway enrichment displayed as they are precursors to phenylpropanoids\textsuperscript{88}. Also notable is that the ribosome pathway shows highly significant, but low-magnitude fold enrichment. This aligns with the strongly observed contradictory regulation patterns of the ribosome pathway. Importantly, the pathfindR analysis functions only measure deviation from a standard value to calculate the fold enrichment value\textsuperscript{89}. In other words, the secondary
analysis confirms that there is significant altered regulation, but does not discriminate between up- or downregulation.

**Table 4-1**

Upregulated Pathways from the HSRC Stressor Group

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Mean Q-Value</th>
<th>Commonality (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribosome (ath03010)</td>
<td>0.0291</td>
<td>3/8</td>
</tr>
<tr>
<td>Phenylpropanoid biosynthesis (ath00940)</td>
<td>0.0237</td>
<td>2/8</td>
</tr>
</tbody>
</table>

**Table 4-2**

Downregulated Pathways from the HSRC Stressor Group

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Mean Q-Value</th>
<th>Commonality (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribosome (ath03010)</td>
<td>0.0129</td>
<td>2/8</td>
</tr>
<tr>
<td>starch and sucrose metabolism (ath00940)</td>
<td>0.0369</td>
<td>2/8</td>
</tr>
</tbody>
</table>
In datasets grouped on the basis of microgravity exposure, three total pathways were identified as commonly significantly upregulated. These pathways include photosynthesis (ath00195), ribosome biogenesis in eukaryotes (ath03008), and proteasome (ath03050) (Table 4-3, 4-4). The photosynthesis pathway was found in 7 out of 15 significant datasets whose subjects were exposed to microgravity, and the ribosome biogenesis in eukaryotes and proteasome pathways were found in 6 out of 15 significant datasets of the same group. In total, including non-significant datasets, these same numbers are instead counted out of 20. Conversely, three other pathways were identified as significantly downregulated. The plant-pathogen interaction pathway (ath04626),

**Figure 4-2:** PathfindR enrichment diagram for HSRC.

### 4.2 Microgravity Pathways

In datasets grouped on the basis of microgravity exposure, three total pathways were identified as commonly significantly upregulated. These pathways include photosynthesis (ath00195), ribosome biogenesis in eukaryotes (ath03008), and proteasome (ath03050) (Table 4-3, 4-4). The photosynthesis pathway was found in 7 out of 15 significant datasets whose subjects were exposed to microgravity, and the ribosome biogenesis in eukaryotes and proteasome pathways were found in 6 out of 15 significant datasets of the same group. In total, including non-significant datasets, these same numbers are instead counted out of 20. Conversely, three other pathways were identified as significantly downregulated. The plant-pathogen interaction pathway (ath04626),
phenylpropanoid biosynthesis pathway (ath00940), and endocytosis pathway (ath04144)
were each significant in 4 out of 12 significant datasets, or 4 out of 20 total microgravity
datasets.

Notably, the ribosome pathway was once again found to be significantly
upregulated in some studies while downregulated in others in the microgravity group,
being found to be upregulated in 9 out of 15 datasets but downregulated in 5 out of 15
datasets (Table 4-3, Table 4-4).

The proteasome pathway reappears in this secondary analysis and therefore is
supported more strongly (Figure 4-3). Oxidative phosphorylation also appears strongly in
this analysis Both the ribosome and photosynthesis pathways appear in this analysis as
well, corresponding with their high appearance in the primary analysis, though they
exhibit very different significance/enrichment relationships.

Table 4-3

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Mean Q-Value</th>
<th>Commonality (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribosome (ath03010)</td>
<td>0.0199</td>
<td>9/15</td>
</tr>
<tr>
<td>Ribosome biogenesis in eukaryotes (ath03008)</td>
<td>0.0164</td>
<td>7/15</td>
</tr>
<tr>
<td>Photosynthesis (ath00195)</td>
<td>0.0187</td>
<td>6/15</td>
</tr>
<tr>
<td>Proteasome (ath03050)</td>
<td>0.0205</td>
<td>6/15</td>
</tr>
</tbody>
</table>
### Table 4-4

Downregulated Pathways from the Microgravity Stressor Group

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Mean Q-Value</th>
<th>Commonality</th>
<th>Commonality (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribosome (ath03010)</td>
<td>0.000894</td>
<td>5/12</td>
<td>5/20</td>
</tr>
<tr>
<td>Plant-pathogen interaction (ath04626)</td>
<td>0.0139</td>
<td>4/12</td>
<td>4/20</td>
</tr>
<tr>
<td>Phenylpropanoid biosynthesis (ath00940)</td>
<td>0.0266</td>
<td>4/12</td>
<td>4/20</td>
</tr>
<tr>
<td>Endocytosis (ath04144)</td>
<td>0.0603</td>
<td>4/12</td>
<td>4/20</td>
</tr>
</tbody>
</table>
For datasets whose subject plants were exposed to radiation, the resulting significant pathways were almost entirely identical, featuring the plant-pathogen interaction pathway, phenylpropanoid biosynthesis pathway, and endocytosis pathway as significant in 4 out of the 16 total radiation datasets with significant pathways (Table 4-5, 4-6). Furthermore, photosynthesis pathway was found in 7 out of 16 datasets, while the proteasome and ribosome biogenesis in eukaryotes pathways share the same value of 6 out of a total 16 datasets. Including all datasets categorized into the radiation group, these counts are instead out of 20 total datasets.
The primary observable difference in significant pathways among these data is that the contradictory regulation pattern of the ribosome pathway was adjusted; instead of 9 instances of upregulation, there are 10 instances revealed in the radiation group. The number of downregulated instances remains the same at 5 instances (Table 4-5, Table 4-6).

**Table 4-5**

Upregulated Pathways from the Radiation Stressor Group

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Mean Q-Value</th>
<th>Commonality</th>
<th>Commonality (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribosome (ath03010)</td>
<td>0.0238</td>
<td>10/16</td>
<td>10/20</td>
</tr>
<tr>
<td>Photosynthesis (ath00195)</td>
<td>0.0164</td>
<td>7/16</td>
<td>7/20</td>
</tr>
<tr>
<td>Ribosome biogenesis in eukaryotes (ath03008)</td>
<td>0.0187</td>
<td>6/16</td>
<td>6/20</td>
</tr>
<tr>
<td>Proteasome (ath03050)</td>
<td>0.0205</td>
<td>6/16</td>
<td>6/20</td>
</tr>
</tbody>
</table>
### Table 4-6

Downregulated Pathways from the Radiation Stressor Group

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Mean Q-Value</th>
<th>Commonality</th>
<th>Commonality (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribosome (ath03010)</td>
<td>0.000691</td>
<td>4/12</td>
<td>4/20</td>
</tr>
<tr>
<td>Photosynthesis (ath00195)</td>
<td>0.0000263</td>
<td>4/12</td>
<td>4/20</td>
</tr>
<tr>
<td>Plant-pathogen interaction (ath04626)</td>
<td>0.0139</td>
<td>4/12</td>
<td>4/20</td>
</tr>
<tr>
<td>Phenylpropanoid biosynthesis (ath00940)</td>
<td>0.0266</td>
<td>4/12</td>
<td>4/20</td>
</tr>
<tr>
<td>Endocytosis (ath04144)</td>
<td>0.0603</td>
<td>4/12</td>
<td>4/20</td>
</tr>
</tbody>
</table>
Figure 4-4: PathfindR enrichment diagram of radiation datasets.

The pathfindR secondary analysis did not reveal many novel connections due to the radiation group’s overall similarity to the microgravity group (Figure 4-4). However, the results of the pathfindR analysis do serve as a form of bootstrapping to better support the significance of the proteasome pathway among the set of pathways of altered regulation.
5.1 Phonylpropanoids and Lignin Biosynthesis

As this analysis demonstrated, the phenylpropanoid biosynthesis pathway in *A. thaliana* is one of the most commonly altered in terms of expression. Furthermore, this pathway is mostly consistent in its downregulation among datasets tagged for microgravity and radiation. The phenylpropanoid pathway is a precursor to several important compounds in plants, including prenylated isoflavonoids, stilbenes, psoralens, coumarins, 3-deoxyanthocyanidins, flavonols, auronespterocarpans, and isoflavans, which exhibit broad-spectrum antimicrobial activity and are thus classified as ‘phytoalexins’\(^{90-92}\). Given many phenylpropanoids’ antimicrobial and immune-like functions in the plant, this could also potentially be tied to the observed downregulation of plant-pathogen response. The protocols for sanitation and pathogen control aboard the International Space Station are very strict and allow for little error\(^{93}\). The phenylpropanoid biosynthesis pathway is likewise tied to the immune response of plants via the phytoalexins and phytoanticipins, thus there is a potential link between the downregulation seen in both pathways. If plants spend time within such sterile environments, it would then follow that less transcriptive support would be given to immune response and its related systems. To further support this conclusion, plants that were sealed within airtight units, such as BRIC units or in sealed plates, did not exhibit
this downregulation. Many of these airtight growth chambers were pre-loaded in a lab before being sent to the International Space Station and may have been exposed to an environment more open to pathogenic contamination.

In addition, the phenylpropanoid pathway branches to monolignols: \( p \)-coumaryl, conferyl, and sinapyl alcohols, which are the monomeric precursors of the lignin polymer molecule in the respective units \( p \)-hydroxyphenyl, guaiacyl, and syringyl\(^\text{94} \). Lignin is more broadly a collection of polymer molecules responsible for structural elements in plants and can form in response to stressors\(^\text{92} \). These factors taken together form a particularly interesting picture when it is considered that the phenylpropanoid pathway was downregulated primarily in experiments where microgravity and radiation were present, but instead upregulated when airflow was restricted (thus being part of the HSRC stressor group). It is likely that such a relationship could arise from conflicting responses to heat generated by HSRC technology and the microgravity environment.

Graviperception is already known to affect the natural growth cycles and circumnutation of plants, thus it is likely that structural pathways are downregulated in microgravity as a result of no gravity being perceived to guide growth\(^\text{95} \). Furthermore, the GPD theory poses that sensing of tissue bending is an important component to plant stem straightness and circumnutation\(^\text{44} \). Furthermore, studies have revealed circumnutation decreased in microgravity, providing greater support to the action of microgravity on \( A. \) thaliana growth control\(^\text{96} \). Lacking any significant mechanical stimulation from gravity may stifle the plant’s ability to grow.

On the contrary, those plants that may have experienced HSRC saw an upregulation of this pathway. It is already understood that the phenylpropanoid
biosynthesis pathway is called upon to address stress response to low temperatures, among other stresses. Furthermore, in loquat it has been demonstrated that heat-shock factors – particularly EjHSF3 – do mediate plant lignin production in a dualistic fashion with EjHSF1 which responds instead to cold stress. Given this, the upregulation of phenylpropanoid biosynthesis in only the studies where HSRC was possible implies that these plants may indeed be undergoing unintended heat stress due to a lack of cooling via convection.

5.1.1 Evidence for Heat Stress via Starch Metabolism Alterations

The potential for heat stress in airtight or sealed growth chambers is highlighted by the observed upregulation in the starch and sucrose pathway in the HSRC group. As demonstrated by Min et al., (2014) high-temperatures in sensitive cotton plants induced overexpression of casein kinase I (GhCK1), inhibiting starch synthases and disrupting glucose homeostasis. In addition, soluble sugar has been found to function as a regulator of auxin biosynthesis through phytochrome interacting factors (PIFs) – factors which have been shown to function in high-temperature conditions – pointing to the importance of starch and sucrose metabolism regulation as a possible indicator of heat stress conditions. Taking this together, there are considerable links between airtight growth chambers and heat stress-associated responses. This reveals a notable concern that such airtight units may indeed overheat the plants within due to a lack of convection.

5.2 Proteasome Upregulation in Non-Airtight Growth Containers

In datasets that were exposed primarily to microgravity and radiation, a strong upregulation is noticeable in the proteasome pathway of A. thaliana. Notably, this pathway may be present in all spaceflight studies as the data for the HSRC stressor was
more difficult to find and in smaller numbers, but for this analysis it only appeared in studies whose technology was unsealed or used atmospheric controls to enable airflow. Furthermore, the microgravity and radiation groups share many datasets, with only 1 dataset included in the Radiation group that was not found in the microgravity group. As such, they share many of the observed upregulated and downregulated pathways, though the proteasome pathway was most notably not found in the data set from GLDS-213 that was exposed to Earth gravity. Tesei et al. (2021) provide some support to this idea, as they observed a notable downregulation – between 1.8 to 1.5-fold – in some proteasome complexes of black fungus *Knufia chersonesos* in normal gravity conditions as compared to simulated microgravity\textsuperscript{101}. This possible connection begs further research as a direct upregulation of proteasome activity in microgravity could broadly change much of a *A. thaliana*’s biological functions in such conditions.

The proteasome is a large protein complex responsible for degradation of ubiquitinated proteins, therefore its broad upregulation would point to higher rates of protein degradation\textsuperscript{102–104}. Upregulation was primarily seen in both the 19S Regulatory Particle and 20S Core Particle; both elements of the complex directly responsible for ubiquitin-mediated proteolysis (Figure A-9, A-10, A-11, A-12, A-13, A-14). Furthermore, small amounts of upregulation were observable in POMP (referred to as UMP1 analogs more commonly in plants) and P131. UMP1 aids in the maturation of the proteasome by binding along the \( \beta \)-ring in sequence for each \( \beta \) component\textsuperscript{105}. In addition, little upregulation is seen in the alternative product complex of the pathway – the immunoproteasome – which is the complex that degrades foreign proteins of pathogens. This lack of upregulation coincided with the observed downregulation of the plant-
pathogen interaction pathway and further highlights that the proteasome is exhibiting heightened transcription when exposed to microgravity. This single UMP1 analog, however, does not tell the whole story. Hu et al. (2021) identified three distinct homologs of UMP1 within common rice, which each had differing expression patterns on exposure to a fungal pathogen\(^{106}\). If \textit{A. thaliana} similarly possesses multiple UMP1 analogs of different function, then the expression patterns observed in this study may not fully explain the altered immune system gene expression especially due to a lack of downregulation in immunoproteasome pathways (Figure A-9, A-10, A-11, A-12, A-13, A-14).

Previous experiments have shown that plants grown in simulated microgravity on Earth exhibit muted immune responses, prompting further discussion and study. It is possible that the upregulation of the main proteasome pathway is essentially pulling greater quantities of material away from immunoproteasome construction, resulting in the knock-on effect of a net decrease in immunoproteasome production and ability to respond to pathogens. It is important to study this further by monitoring the proteasome pathway of clinostat-grown plants on Earth alongside exposure to various pathogens.

5.3 Contra-Regulation Observed in Ribosome and Photosynthesis Pathways

The final significant pathways of concern are the ribosome and photosynthesis pathways. Both pathways exhibited a pattern of regulation in which close-to-equal numbers of individual datasets show upregulation and downregulation. This typically can indicate potential errors or oversights in the data, but in this case, it appears to be a situation caused by confounding variables within the different experiments as well as factors that lie beyond the scope of this study’s categorization methods.
5.3.1 Photosynthetic Activity

To begin, the Photosynthesis pathway appears to be a direct result of varying conditions found in the experimental protocols under which the samples were exposed. In the case of the experiment using GeneLab ID GLDS-7, for example, one sample plant was placed directly under a fluorescent excitation LED for GFP imaging. In fact, the samples not directly under this LED were not kept wholly separate, so it is very probable that light from this LED could have contributed to increased photosynthetic activity. In the case of GLDS-314, upregulation of the Photosynthesis pathway was observed in the plants that were exposed to a 16-hour light regime followed by constant red light, also likely resulting in higher-than-typical photosynthetic activity. Finally, GLDS-469 utilized the highest light levels of any experiment analyzed – 120 μmol m\(^{-2}\) s\(^{-1}\) – which may have induced higher photosynthetic activity in the sample leaves since most experiments utilize a fraction of that light level, usually approximately 30-40 μmol m\(^{-2}\) s\(^{-1}\). In addition to these upregulated examples, however, there is also a significant quantity of data found to have downregulated photosynthetic activity.

In the case of GLDS-213, both datasets had significantly downregulated photosynthetic activity, though this was probably a result of the sample type which was a callus culture. Plant callus culture is a mass of unorganized dedifferentiated parenchyma cells – cells that typically serve numerous different functions throughout the plant, but primarily make up the non-vascular and non-dermal regions as well as the tissue that forms over plant wounds. As such, while some forms of parenchyma are capable of photosynthesis in leaves as mesophyll, most parenchyma are not specialized for this and, especially in a callus, do not have strong photosynthetic capability, thus
pointing to a reason as to why these samples showed a notable downregulation in both zero gravity and Earth gravity\textsuperscript{110}. GLDS-38, meanwhile, lists no light exposure and may have been grown in dark conditions\textsuperscript{111}. This would cause cells to downregulate photosynthesis since such machinery is useless without light. Aside from these two datasets, only the whole-plant sample from GLDS-7 exhibited downregulation of the Photosynthesis pathway, which was likely an error as two other categories of GLDS-7 samples – Hypocotyl and Shoot – exhibited upregulation.

Finally, an overall notable observation about the contradictory regulation patterns is that plants with upregulated photosynthesis generally grew for longer periods. Except for GLDS-314’s red light sample, all datasets with upregulated photosynthesis were grown for at least 11 days before collection\textsuperscript{27,76,107,112}. Meanwhile, except for GLDS-7’s whole-plant sample, samples exhibiting downregulation were grown for 5 days or less\textsuperscript{76,108,111}. This could potentially aid in explaining the observed upregulation, as longer-term exposure to growth-favoring conditions and high light can upregulate photosynthetic gene expression\textsuperscript{113}.

5.3.2 Ribosome Activity

The datasets that exhibit varying ribosomal regulation patterns, unlike the photosynthesis discussed prior, do not have a clear pattern between the variables in the experiments nor the three categories of space stressors (microgravity, radiation, HSRC). The primary notable correlation was that all datasets that exhibited downregulated ribosomal activity were samples grown from seedlings, rather than adult plants or other types of plants. Seedling-derived samples appeared amongst the datasets exhibiting upregulated ribosomal activity, however, pointing to a different cause. When considering
the quantity of contraregulated datasets and the lack of clear correlation instead likely points to factors beyond the scope of this study. As such, further research could be focused on ribosomal regulation in these datasets alongside more thorough investigation of metadata associated therein.

5.4 Other Pathways

Ribosome biogenesis in eukaryotes and the endocytosis pathways also exhibited significant regulatory change. In particular, the ribosome biogenesis in eukaryotes pathway is likely tied to the contradictory regulatory patterns previously described, especially due to the slightly higher tendency for the ribosome pathway to be upregulated compared to downregulated. Upregulation of the ribosome pathway would naturally coincide with increased ribosomal pathway regulation. In addition, this pathway is overrepresented by its proportional appearance in GLDS-469. All four datasets from GLDS-469 exhibit upregulation of this pathway, with only one dataset from GLDS-120 accompanying it. Coinciding with this, the endocytosis pathway was also found to be downregulated in all four datasets provided by GLDS-469 and only one dataset provided by GLDS-314. This is true in both the microgravity and radiation stressor groups. As such, with high representation from only one experiment it stands to reason that the altered regulation of these pathways is potentially a direct result of experimental protocol, and therefore less reliable for drawing conclusions upon. These pathways still must be noted, however, as their emergence could regardless point to significant spaceflight stress to be evaluated in future studies.
CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

Over the course of this study, a combination analysis using the pathview and pathfindR packages revealed several pathways of significant altered expression in *A. thaliana* spaceflight experiments. Among these pathways, the most significant were the phenylpropanoid biosynthesis, starch and sucrose metabolism, and proteasome pathways as they each are associated with the stressor group(s) they appeared in.

6.1.1 Phenylpropanoids and Lignin

The phenylpropanoid biosynthesis pathway was significantly downregulated in many studies in which subject plants were exposed to microgravity and radiation yet found significantly upregulated in the HSRC group. Phenylpropanoid biosynthesis includes the production of numerous compounds that exhibit antimicrobial properties, known as phytoalexins, and the production of precursors to lignin, a very important element of plant structure\(^{90-92,94}\). Given the strict sanitation protocols undergone for spaceflight experiments, downregulation of the phenylpropanoid biosynthesis pathway would follow as a plant has little need for antimicrobial compounds when exposed to little or no pathogens. This coincides with additionally observed – though less common – downregulation of the plant-pathogen interaction pathway. Plants grown in sealed technology such as BRIC units, thereby not subjected to stringent sanitation protocols,
did not show downregulation of immune-related pathways, thus lending credence to the possibility that sanitation differences cause these alterations.

In terms of the pathway’s other functions in production of lignin, plants have already shown that microgravity reduces circumnutation. This process is an important structural maneuver for plant growth and stem straightness; thus, it is possible that plants under microgravity reduce their growth in situations where gravity is unclear. It is possible that without clear graviperception, a plant will not engage in growth as actively, resulting in downregulation of associated growth pathways. Plants of the HSRC group, however, exhibited upregulation in this pathway likely due to previously shown associations between heat-shock factors and temperature stress. The emergence of upregulation in this pathway, despite otherwise observed downregulation, implies that HSRC may indeed be occurring in BRIC units and similar technology.

6.1.2 The Proteasome

The proteasome pathway is coincidentally involved in several internal processes, including an immune role as well as primarily control over the ubiquitination and degradation of proteins. According to my study and a one by Tesei et al., (2021) the upregulation observed appeared to be connected to exposure to microgravity. Primarily, the pathview analysis revealed that upregulation mostly occurred in structures directly responsible for ubiquitin-mediated proteolysis, but a small amount of upregulation was also observable in POMP, a UMP1 analog in plants. POMP is a factor that aids in maturation of the proteasome, implying that its upregulation represents a focus upon production of more proteasomes and increased protein degradation. It is also possible that this upregulation may be part of a larger picture; Hu et al. identified three
distinct UMP1 homologs in common rice with different expression patterns on exposure to a pathogen, so it is therefore not impossible that undiscovered UMP1 homologs within *A. thaliana* could be altering expression in non-direct ways. Also notable is that upregulation is hardly seen in the genes associated with the immunoproteasome, supporting earlier conclusions on the effects of sanitation on immune pathway regulation. Alternatively, it is instead possible that upregulation of alternative functions of the proteasome pathway is instead pulling resources away from the immunoproteasome production branch of the pathway, indirectly weakening the plant’s already reduced immune capability. Both of these possibilities are key to investigate further, as disease could be destructive to spaceflight plants based on these results.

6.1.3 **Starch and Sucrose Metabolism under Heat Stress**

Lastly, but not insignificantly, the starch and sucrose metabolism pathway showed downregulation in studies of the HSRC group. High temperatures have previously been shown to induce overexpression of starch synthase-inhibiting *casein kinase I* in cotton, thus disrupting typical glucose homeostasis. Furthermore, soluble sugar plays a role as a regulator of auxin biosynthesis through *PIFs*. Auxin is an important signaling molecule for plant stress, and *PIFs* function at high temperatures, implying their importance to mediating plant functions under heat stress. Slowed starch and sucrose metabolism may serve to increase soluble sugar levels within the plant, enabling more efficient signaling to address heat stress. If true, then it is likely that technologies such as BRIC cause undue heat stress in plants.
6.1.4 Other Pathways

Aside from the more conclusive identified pathways, it must be mentioned that the ribosome, ribosome biogenesis in eukaryotes, photosynthesis, and endocytosis pathways exhibited significant regulatory changes. The ribosome and photosynthesis pathways demonstrated largely contradictory regulatory patterns, showing both significant upregulation and downregulation depending on the dataset. Ribosome biogenesis in eukaryotes and endocytosis, on the other hand, were upregulated and downregulated respectively. Though the photosynthesis and endocytosis pathways may indeed be a result of confounding variables or overrepresentation in the data, the ribosome and ribosome biogenesis in eukaryotes pathways point to a larger effect on plants during spaceflight that falls out of the scope of this study. Although this study provides less direct targets for focusing upon ribosome activity, it is still worth pointing to and emphasizing the need for further study, as ribosomes are indeed a very key core feature of any organism.

In conclusion, each one of these pathways serve as a target to better aim future spaceflight experiments and contextualize much of the results attained in said experiments. Through these data and similar studies, more focused spaceflight studies can be devised to better characterize and understand the pathway alterations described.
6.2  Future Work

This study aimed to provide a number of pathways of significant interest for future research to expand upon in additional spaceflight studies. However, it must also be noted that despite the quality results produced, this study highlights the relatively incomplete nature of open data available for public analysis. Furthermore, this study should additionally serve as a call to improve open data in terms of metadata, accessibility, and format. Primarily, however, further work on the discovered pathways must be discussed.

6.2.1  Investigations into Graviperception

In particular, the phenylpropanoid pathway’s altered regulation demands further investigation to better characterize the extent to which the changes can be attributed to altered gravity or failed graviperception. To this end, it could be highly beneficial to carry out similar experiments upon plants that grow upon or underneath the surface of water. Water’s buoyant effects likely result in altered ways to sense gravity in such plants, and characterizing their graviperception abilities would potentially expand how we approach future investigations into terrestrial plant graviperception. It is also even possible that new genetic crosses could be engineered using water-based plants if their growth is less-dependent on graviperception and more upon factors such as tissue bending.

In a similar vein, it is also paramount to carry out studies focusing on graviperception-related pathways for plants under simulated microgravity. Given that the phenylpropanoid biosynthesis pathway was upregulated in both the microgravity and radiation stressor groups, it is impossible to discern if the observed regulatory changes are a direct consequence of microgravity, radiation, or an additive process between both
stressors. Research to this end would be required to reinforce the conclusions found in this paper.

6.2.2  **Plant Immune Function in Sterile Environments**

Many of the observed pathways have close ties to immune function, which was largely incorporated under the sterile aspects of spaceflight experiments versus ground controls. Though this presents a good case, such propositions demand verification; there is yet a causative link shown between the downregulated phenylpropanoid biosynthesis and proteasome immune-adjacent genes and sterile environments. As such, experiments should in future be carried out to better analyze regulation changes between a ‘typical’ lab environment and an extremely sterile environment mimicking that of a spaceflight experiment.

6.2.3  **Effects of Proteasome Upregulation on Proteomics**

It is additionally important to investigate not only the mechanisms responsible for proteasome pathway upregulation, but to also investigate its effects. Upregulation of protein degradation should have systemic effects observable in the subject plant, and such protein changes are important to characterize as proteins are the acting component of an organism. Proteomic investigation of the microgravity-caused upregulation alongside mutant overexpression of the proteasome pathway should give the broadest picture, allowing for strong conclusions on the effects of microgravity-induced proteasome upregulation and a greater understanding of this pathway’s regulation repercussions.

6.2.4  **Starch and Sucrose Metabolism and Heat Stress**

To then verify results regarding the starch and sucrose metabolism pathway’s regulation, it should follow that experiments should start by evaluating how auxin
signaling, heat stress, and glucose homeostasis are affected in *A. thaliana*, both via mutation and induction via high temperature. The conclusion that downregulated starch and sucrose metabolism increases auxin signaling efficiency is but one possible conclusion to be drawn from the data, and requires further support to stand on its own.

In addition, detailed testing of BRIC and similar hardware is an important step in assessing if such technologies’ lack of airflow in space results in heat stress as a result. Though simple in theory, testing this properly would require an actual spaceflight experiment as simulated microgravity would not alter convection as true microgravity does. Optimally, this research serves to inspire researchers to push for such an experiment to be incorporated into the next flight mission available within reason. Flawed hardware can cause many misinterpretations of data and if BRIC hardware is causing such confounding issues, it is paramount to address it.

One alternative way to assess the possibility of reduced convection’s heat effects is to introduce plants to a near-vacuum, where there is little air to circulate heat in the first place. This, in combination with sensitive temperature detection technology, could allow for actual validation of whether or not reduced convection indeed causes the proposed heat stress.

### 6.2.5 Data Availability, Formatting, and Metadata

A final component of future work that demands mentioning is the nature by which open data is available. NASA’s GeneLab repository is a fantastic resource, however the data available was swiftly reduced as many datasets had data of alternate formats, utilized nonfunctional methods, or were missing appropriate metadata. Steps are being taken to address this even as this thesis is being written, but polishing data should nonetheless
demand further attention. For example, many available experiments used in this study had incomplete information on the cell types that were sequenced from *A. thaliana*, had hard-to-find light exposure levels and protocols, and sometimes obfuscated what type of technology was used to grow the subject plants. These factors are all crucial to meta-analysis and their lack reduces future researchers’ abilities to draw solid conclusions upon the hard-wrought data.

Furthermore, open-source data must be remembered in efforts to archive and make compatible older methods of analysis. Some data that otherwise would have been usable in this study was instead culled due to incompatible microarray files with modern R scripts and packages. Compatibility issues such as these extend the time it takes to make complete analyses or necessitate ignoring such data. Neither alternative is preferable, so compatibility is another area worthy of concern for open data.

Overall, these hurdles add additional time to research efforts using open data, which can degrade the usefulness of studies produced from them. In conclusion, open data needs better curation to improve the scientific world’s capability to easily draw broader conclusions to support future focused endeavors, especially in the realm of space plant biology.

6.2.6 **Closing Remarks**

Overall, these pathways should serve as directions and targets for future research in addressing a food solution for astronauts of the future. Via careful and in-depth further analysis of those pathways presented here, it may be possible to resolve many of the challenges for space agriculture. With such a food source, resources such as rare metals once thought to be severely limited could be within our reach through space travel.
Furthermore, keen understanding of long-term space agriculture could even enable far-flung theories and postulations for space habitation over long terms and allow manned vessels to travel further into the stars than ever before. From these more capable spaceflight journeys, it is possible to access resources previously considered highly rare and limited, potentially expanding the quality of life for people on Earth and enabling more to enjoy many of the modern comforts of today.
APPENDIX A

SUPPLEMENTAL FIGURES
A.1 Phenylpropanoid Biosynthesis Pathview Diagrams

In this appendix section, the figures generated for datasets with significant regulatory changes to the phenylpropanoid biosynthesis pathway can be found for detailed analysis and interpretation (Figure A-1, A-2, A-3, A-4, A-5, A-6). These pathways were generated using KEGG pathway information. Green-labelled genes were found to be downregulated, and red-labelled genes were found to be upregulated. All pathways in this section were determined to be downregulated overall.

Figure A-1: Pathview diagram of GLDS-38 phenylpropanoid biosynthesis pathway.
Figure A-2: Pathview diagram of GLDS-120 (PhyD mutant, dark growth) phenylpropanoid biosynthesis pathway.
Figure A-3: Pathview diagram of GLDS-314 (red light) phenylpropanoid biosynthesis pathway.
Figure A-4: Pathview diagram of GLDS-469 (transgenic plants, 8-hour light period) phenylpropanoid biosynthesis pathway.
Figure A-5: Pathview diagram of GLDS-469 (wild-type, 8-hour light period) phenylpropanoid biosynthesis pathway.
**Figure A-6:** Pathview diagram of GLDS-469 (wild-type, 16-hour light period) phenylpropanoid biosynthesis pathway.
A.2 Starch and Sucrose Metabolism Pathview Diagrams

In this appendix section, the figures generated for datasets with significant regulatory changes to the starch and sucrose pathway can be found for detailed analysis and interpretation (Figure A-7, A-8). These pathways were generated using KEGG pathway information. Green-labelled genes were found to be downregulated, and red-labelled genes were found to be upregulated. All pathways in this section were determined to be downregulated overall.

Figure A-7: Pathview diagram of GLDS-38 starch and sucrose metabolism pathway.
Figure A-8: Pathview diagram of GLDS-120 (Col-0, light growth) starch and sucrose metabolism pathway.
A.3 Proteasome Pathview Diagrams

In this appendix section, the figures generated for datasets with significant regulatory changes to the proteasome pathway can be found for detailed analysis and interpretation (Figure A-9, A-10, A-11, A-12, A-13, A-14). These pathways were generated using KEGG pathway information. Green-labelled genes were found to be downregulated, and red-labelled genes were found to be upregulated. All pathways in this section were determined to be upregulated overall.
Figure A-9: Pathview diagram of GLDS-7 (hypocotyl) proteasome pathway.
Figure A-10: Pathview diagram of GLDS-7 (shoot) proteasome pathway.
Figure A-11: Pathview diagram of GLDS-213 (0G) proteasome pathway.
Figure A-12: Pathview diagram of GLDS-469 (transgenic plants, 16-hour light period) proteasome pathway.
**Figure A-13:** Pathview diagram of GLDS-469 (wild-type, 8-hour light period) proteasome pathway.
**Figure A-14**: Pathview diagram of GLDS-469 (wild-type, 16-hour light period) proteasome pathway.
A.4 Plant-pathogen Interaction Pathview Diagrams

In this appendix section, the figures generated for datasets with significant regulatory changes to the plant-pathogen interaction pathway can be found for detailed analysis and interpretation (Figure A-15, A-16, A-17, A-18). These pathways were generated using KEGG pathway information. Green-labelled genes were found to be downregulated, and red-labelled genes were found to be upregulated. All pathways in this section were determined to be downregulated overall.
Figure A-15: Pathview diagram of GLDS-120 (Col-0, light growth) plant-pathogen interaction pathway.
Figure A-16: Pathview diagram of GLDS-469 (transgenic plants, 16-hour light period) plant-pathogen interaction pathway.
Figure A-17: Pathview diagram of GLDS-469 (wild-type, 8-hour light period) plant-pathogen interaction pathway.
Figure A-18: Pathview diagram of GLDS-469 (wild-type, 16-hour light period) plant-pathogen interaction pathway.
A.5 Photosynthesis Pathview Diagrams

In this appendix section, the figures generated for datasets with significant regulatory changes to the proteasome pathway can be found for detailed analysis and interpretation (Figure A-19, A-20, A-21, A-22, A-23, A-24, A-25, A-26, A-27, A-28, A-29). These pathways were generated using KEGG pathway information. Green-labelled genes were found to be downregulated, and red-labelled genes were found to be upregulated. The pathway diagrams for datasets in figures A-19 through A-22 were determined to be downregulated overall. The pathway diagrams for datasets in figures A-23 through A-29 were determined to be upregulated overall.

Figure A-19: Pathview diagram of GLDS-7 (whole-plant) photosynthesis pathway.
**Figure A-20:** Pathview diagram of GLDS-38 photosynthesis pathway.
Figure A-21: Pathview diagram of GLDS-213 (0G) photosynthesis pathway.
Figure A-22: Pathview diagram of GLDS-213 (1G) photosynthesis pathway.
Figure A-23: Pathview diagram of GLDS-7 (hypocotyl) photosynthesis pathway.
Figure A-24: Pathview diagram of GLDS-7 (shoot) photosynthesis pathway.
Figure A-25: Pathview diagram of GLDS-120 (Col-0, light growth) photosynthesis pathway.
Figure A-26: Pathview diagram of GLDS-314 (Red light) photosynthesis pathway.
Figure A-27: Pathview diagram of GLDS-469 (transgenic plants, 8-hour light period) photosynthesis pathway.
Figure A-28: Pathview diagram of GLDS-469 (transgenic plants, 16-hour light period) photosynthesis pathway.
Figure A-29: Pathview diagram of GLDS-469 (wild-type, 16-hour light period) photosynthesis pathway.
A.6 Ribosome Pathview Diagrams

In this appendix section, the figures generated for datasets with significant regulatory changes to the proteasome pathway can be found for detailed analysis and interpretation (Figure A-30, A-31, A-32, A-33, A-34, A-35, A-36, A-37, A-38, A-39, A-40, A-41, A-42, A-43, A-44, A-45). These pathways were generated using KEGG pathway information. Green-labelled genes were found to be downregulated, and red-labelled genes were found to be upregulated. The pathway diagrams for datasets in figures A-30 through A-35 were determined to be downregulated overall. The pathway diagrams for datasets in figures A-36 through A-45 were determined to be upregulated overall.
Figure A-30: Pathview diagram of GLDS-7 (whole-plant) ribosome pathway.
Figure A-31: Pathview diagram of GLDS-38 ribosome pathway.
Figure A-32: Pathview diagram of GLDS-120 (WS, light grown) ribosome pathway.
Figure A-33: Pathview diagram of GLDS-251 ribosome pathway.
Figure A-34: Pathview diagram of GLDS-314 (dark) ribosome pathway.
Figure A-35: Pathview diagram of GLDS-314 (dark) ribosome pathway.
Figure A-36: Pathview diagram of GLDS-7 (hypocotyl) ribosome pathway.
Figure A-37: Pathview diagram of GLDS-7 (shoot) ribosome pathway.
**Figure A-38:** Pathview diagram of GLDS-120 (Col-0, light grown) ribosome pathway.
Figure A-39: Pathview diagram of GLDS-120 (WS, dark grown) ribosome pathway.
Figure A-40: Pathview diagram of GLDS-205 ribosome pathway.
Figure A-41: Pathview diagram of GLDS-213 (1G) ribosome pathway.
Figure A-42: Pathview diagram of GLDS-469 (transgenic plants, 8-hour light period) ribosome pathway.
Figure A-43: Pathview diagram of GLDS-469 (transgenic plants, 16-hour light period) ribosome pathway.
**Figure A-44:** Pathview diagram of GLDS-469 (wild-type, 8-hour light period) ribosome pathway.
Figure A-45: Pathview diagram of GLDS-469 (wild-type, 16-hour light period) ribosome pathway.
APPENDIX B
SUPPLEMENTAL DATA
B.1 R Programs and Packages

This appendix section outlines the packages utilized for the first analysis using the pathview package and the second analysis using the pathfindR package. The first table shows all packages used in scripts serving the pathview analysis, including version numbers and citations as they were provided in the R package itself (Table B-1). The second table of this section shows all packages used in scripts serving the pathfindR analysis, including version numbers and citations as they were provided in the R package itself (Table B-2).

Table B-1

Packages Used in Pathview Analysis

<table>
<thead>
<tr>
<th>Package</th>
<th>Citation</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>affy</td>
<td>78</td>
<td>1.74.0</td>
</tr>
<tr>
<td>annotate</td>
<td>114</td>
<td>1.74.0</td>
</tr>
<tr>
<td>AnnotationDbi</td>
<td>115</td>
<td>1.58.0</td>
</tr>
<tr>
<td>apeglm</td>
<td>116</td>
<td>1.18.0</td>
</tr>
<tr>
<td>ath1121501.db</td>
<td>79</td>
<td>3.13.0</td>
</tr>
<tr>
<td>ath1121501cdf</td>
<td>117</td>
<td>2.18.0</td>
</tr>
<tr>
<td>DESeq2</td>
<td>81</td>
<td>1.36.0</td>
</tr>
<tr>
<td>dplyr</td>
<td>82</td>
<td>1.0.10</td>
</tr>
<tr>
<td>gage</td>
<td>118</td>
<td>2.46.1</td>
</tr>
<tr>
<td>gageData</td>
<td>119</td>
<td>2.34.0</td>
</tr>
<tr>
<td>limma</td>
<td>80</td>
<td>3.52.4</td>
</tr>
<tr>
<td>Package</td>
<td>Version</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>oligo\textsuperscript{120}</td>
<td>1.60.0</td>
<td></td>
</tr>
<tr>
<td>org.At.tair.db\textsuperscript{86}</td>
<td>3.15.1</td>
<td></td>
</tr>
<tr>
<td>pathview\textsuperscript{83}</td>
<td>1.36.1</td>
<td></td>
</tr>
<tr>
<td>purrr\textsuperscript{121}</td>
<td>0.3.5</td>
<td></td>
</tr>
<tr>
<td>reshape\textsuperscript{122}</td>
<td>0.8.9</td>
<td></td>
</tr>
<tr>
<td>stringr\textsuperscript{123}</td>
<td>1.4.1</td>
<td></td>
</tr>
<tr>
<td>tidyverse\textsuperscript{124}</td>
<td>1.3.2</td>
<td></td>
</tr>
</tbody>
</table>

Table B-2

Packages Used in PathfindR Analysis

<table>
<thead>
<tr>
<th>Package</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>org.At.tair.db\textsuperscript{86}</td>
<td>3.15.1</td>
</tr>
<tr>
<td>pathfindR</td>
<td>1.6.4.9002</td>
</tr>
<tr>
<td>readxl</td>
<td>1.4.1</td>
</tr>
<tr>
<td>tidyverse</td>
<td>1.3.2</td>
</tr>
</tbody>
</table>
B.2 Additional Analysis Information

B.2.1 Commonality Thresholds

As previously described, an arbitrary threshold was determined to filter out pathways with low numbers of significant datasets. These thresholds – termed commonality thresholds – were approximately equal to one third of the total number of datasets, thereby requiring a dataset to be present in around one third of the datasets with any significant pathways. In the first two tables below, the commonality thresholds are listed for upregulated pathways and downregulated pathways respectively (Table B-3, B-4). Finally, this section includes the list of datasets used in this analysis, noting the GeneLab ID, Study Title, and number of individual datasets used from the experiment along with its relevant GeneLab citation (Table B-5).

Table B-3

Commonality Thresholds for Filtering Upregulated Pathways

<table>
<thead>
<tr>
<th>Stressor</th>
<th>Minimum Commonality</th>
<th>Minimum Commonality (Total)</th>
<th>Maximum Non-Significant Datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSRC</td>
<td>2/5</td>
<td>2/8</td>
<td>6</td>
</tr>
<tr>
<td>Microgravity</td>
<td>6/15</td>
<td>6/20</td>
<td>14</td>
</tr>
<tr>
<td>Radiation</td>
<td>6/16</td>
<td>6/20</td>
<td>14</td>
</tr>
</tbody>
</table>
### Table B-4

Commonality Thresholds for Filtering Downregulated Pathways

<table>
<thead>
<tr>
<th>Stressor</th>
<th>Minimum Commonality</th>
<th>Minimum Commonality (Total)</th>
<th>Maximum Non-Significant Datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSRC</td>
<td>2/3</td>
<td>2/8</td>
<td>6</td>
</tr>
<tr>
<td>Microgravity</td>
<td>4/12</td>
<td>4/20</td>
<td>16</td>
</tr>
<tr>
<td>Radiation</td>
<td>4/12</td>
<td>4/20</td>
<td>16</td>
</tr>
</tbody>
</table>

### B.2.2 GeneLab Data Sources

### Table B-5

Number of Utilized Datasets by Study

<table>
<thead>
<tr>
<th>GeneLab ID</th>
<th>Study Title^citation^</th>
<th>No. of Individual Datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLDS-7</td>
<td>The <em>Arabidopsis</em> spaceflight transcriptome: a comparison of whole plants to discrete root, hypocotyl and shoot responses to the orbital environment^125^</td>
<td>4</td>
</tr>
<tr>
<td>GLDS-38</td>
<td>Proteomics and Transcriptomics analysis of <em>Arabidopsis</em> Seedlings in Microgravity^126^</td>
<td>1</td>
</tr>
<tr>
<td>GLDS-120</td>
<td>Genetic dissection of the <em>Arabidopsis</em> spaceflight transcriptome: Are some responses dispensable for the physiological adaptation of plants to spaceflight?^112^</td>
<td>6</td>
</tr>
<tr>
<td>GLDS-205</td>
<td>HSFA2 functions in the physiological adaptation of undifferentiated plant cells to spaceflight microgravity environment^127^</td>
<td>1</td>
</tr>
<tr>
<td>GLDS-213</td>
<td>A whole-genome microarray study of <em>Arabidopsis</em> thaliana cell cultures exposed to microgravity for 5 days on board of Shenzhou 8^128^</td>
<td>2</td>
</tr>
<tr>
<td>GeneLab ID</td>
<td>Study Title</td>
<td>No. of Individual Datasets</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>GLDS-251</td>
<td>RNAseq analysis of the response of <em>Arabidopsis thaliana</em> to fractional gravity under blue-light stimulation during spaceflight$^{129}$</td>
<td>1</td>
</tr>
<tr>
<td>GLDS-314</td>
<td>Adaptive response of <em>Arabidopsis</em> seedlings in microgravity and Mars reduced gravity environment is enhanced by red light photostimulation$^{130}$</td>
<td>2</td>
</tr>
<tr>
<td>GLDS-469</td>
<td>Transcriptomic analysis of the interaction between FLOWERING LOCUS T induction and photoperiodic signaling in response to spaceflight$^{131}$</td>
<td>4</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


   doi:1988LPI....19..831N.


37. Bennett, M. D. *et al.* Comparisons with Caenorhabditis (~100 Mb) and Drosophila (~175 Mb) Using Flow Cytometry Show Genome Size in Arabidopsis to be ~157 Mb and thus ~25\% Larger than the Arabidopsis Genome Initiative Estimate of ~125 Mb. *Ann Bot* **91**, 547–557 (2003).


