Accelerating Space Life Sciences via Astronaut Omics Collection and Integrated Analyses

Henry Cope

10109029

Thesis submitted to the University of Nottingham for the degree of *Doctor of Philosophy,* April 2024

Supervised by:

Dr. Daniel Wilkinson

Dr. Michael Pound

Prof Nathaniel J Szewczyk

Signed:

HENRY STEPHEN COPE, 27/09/2024

Abstract

There are ambitious plans for human spaceflight on the horizon, including an increasingly diverse spacefaring population associated with a variety of commercial ventures, a return to the Moon, and deep-space voyages with the eventual goal of colonising Mars. Past missions - including continuous occupation of the International Space Station since the turn of the millennium - have highlighted stressors associated with spaceflight, such as altered gravity (including "microgravity", which refers to the near-weightlessness experienced due to low gravitational forces, such as during continuous freefall within orbiting spacecraft), cosmic radiation, isolation, and confinement within a closed spacecraft environment. These spaceflight stressors incite a distinct set of molecular effects and associated physiological and psychological changes that elevate the risk of certain health issues in humans and other Earth-based life. Health risks include loss of performance due to degradation of bone and muscle, skin rashes, vision problems, and cognitive and psychiatric disorders. These changes occur dynamically throughout the course of a mission and manifest variably based on differences between individuals and environmental parameters. The changing landscape of human spaceflight necessitates investigation into the precise mechanisms behind these detrimental biological changes, so that the interplay of risk factors can be comprehensively understood and countermeasures can be developed to increase safety and the likelihood of success across missions.

On Earth 'omics', including genomics and transcriptomics, have emerged as powerful 'Big data' tools for elucidating the molecular patterns behind diseases, and in many cases for improving diagnosis, monitoring, and treatment capabilities. Similar approaches are now being explored in the context of space life sciences. There is a need to optimise the generation of

new omics data, and the analysis of existing omics data, to maximise the scientific value of rare and expensive spaceflight missions and experiments.

As such, in this thesis I investigated gaps in data generation and data analysis capabilities in space life sciences. Regarding data generation, I identified a scarcity of multi-omic data from astronauts as a gap, investigated the topic through discussions with the international space life sciences community, and drafted recommendations for policy in the UK national, European, and the international context. This included investigation into the ethics of personal data in the context of biological and health data collected from astronauts. Regarding data analysis, I identified a need for strategies for maximising the utility of data, and led and contributed to analysis projects as a means of investigating this. In particular I focussed on transcriptomics, the most widespread omic type for space biology, and looked at integrated omics analyses, including pooling data from multiple missions, and analysing omics alongside phenotypic data to uncover mechanisms of spaceflight maladaptation; such approaches could be adapted to other spaceflight datasets, and could help to inform analysis strategies for the proposed multi-omic and multi-modal data generated from astronauts.

Acknowledgements

I would like to express my utmost gratitude to my supervisors Dr. Daniel Wilkinson, Dr. Michael Pound, and Prof Nathaniel Szewczyk. In particular I wish to acknowledge Prof Nathaniel Szewczyk for introducing me to this exciting field, frequently explaining new concepts to me, and presenting me with regular opportunities to evolve my research skills. I would also like to thank my family for supporting me throughout this process. I am grateful to the international space life sciences community, particularly the NASA OSDR team, the NASA OSDR AWGs, the ESA Space Omics Topical Team, the ISSOP consortium, and the Blue Marble Space Institute of Science. This community has fostered a supportive and opportunity-rich environment for early-career academics. Needless to say, without the open-science contributions of the data generators and curators, the majority of my research would not have been possible. I would like to thank the Horizon CDT team and my cohort for creating a friendly and supportive environment. I am thankful for NottSpace - the University of Nottingham space research team - for enabling me to contribute to exciting spacecraft engineering projects, and providing me with teaching and supervision opportunities.

Table of contents

List of Figures and Tables

Please note that supplementary figures, tables, and data are linked or included in the Appendix. Figure and table analysis, design, and final formatting are the sole work of the thesis author unless otherwise stated.

Figures

personalised medicine approaches, such as individualised medication prescriptions based on pharmacogenomics, or increased monitoring for early detection of specific health issues in predisposed individuals. **C)** Genomic data could improve fundamental space biology research, such as to detect genetic markers of spaceflight health issues, including spaceflight associated neuro-ocular syndrome (SANS).

- **Figure 4.2. Next steps to develop a routine health data collection programme for UK astronauts and spaceflight participants. A)** To design a world-leading programme, experts and stakeholders must achieve a consensus on which omics to collect, which physiological endpoints and environmental data to collect, and how best to implement policy pertaining to the governance of personal data from human subjects in this context. **B)** Participants, whether they be European Space Agency (ESA) astronauts or UK participants in commercial spaceflight missions, should be recruited for the programme and meaningful informed consent should be obtained. **C)** A pilot study should be performed, in which the health data collection programme is trialled in spaceflight, and opportunities for clinical actionability and research are documented.

- **Figure 5.1. Graphical summary of the key findings from the micro-16 analysis.**

- **Figure 5.2. Effect of spaceflight on body diameter and length of wild-type (wt) and dystrophin (*dys-1*) worms. A)** body diameter. **B)** body length. The diameters of both the strains grown at the International Space Station (ISS) are significantly different compared to ground controls, whereas there is no difference in the length of the worms. Sample size: n=30 for wt ground and flight, n=25 for *dys-1* ground, and n=29 for *dys-1* flight. All the data pass the normality test. Two-way ANOVA (Tukey multiple testing) was used for calculating significant differences; $p < 0.001$ is for ***. This figure was provided by Dr. Soni Purushottam.

- **Figure 5.3. Effect of spaceflight on muscle strength of wt and *dys-1* worms.** There is no difference in muscle strength between the strains on the ground, and the muscle strength of space-grown worms decreased by 16.6% and 33.4% for wt and *dys-1*, respectively. Sample size: n=30 for wt ground and flight both, n=25 for *dys-1* ground, and n=29 for *dys-1* flight. All the data pass the normality test. Two-way ANOVA (Tukey multiple testing) was used for calculating significant differences, * for $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$. This figure was provided by Dr. Soni Purushottam.

- **Figure 5.4. Global trends in spaceflight gene expression. A)** Principal component analysis (PCA) clustering of samples based on top 500 most variable genes. **B)** PCA loadings of top 500 most variable genes. **C)** Volcano plot for wt flight vs. wt ground differential expression analysis, annotated genes are those ranked in the top 20 upregulated/downregulated based on test statistic. **D)** Volcano plot for dys-1 flight vs. dys-1 ground differential expression analysis, annotated genes are those ranked in the top 20 upregulated/downregulated based on test statistic. Analysis and formatting of Panel C were done by Dr. Craig Willis.

- **Figure 5.5. Overlay of wt and *dys-1* transcriptome responses to spaceflight. A)** Overlay of genes upregulated by spaceflight in wt and/or *dys-1* worms. Venn diagram illustrates degree of commonality/uniqueness in spaceflight-upregulated genes between the two strains, while heatmap depicts representative Gene Ontology (GO) terms for common/uniquely upregulated genes. Up in flight. Venn shows the commonality and differential changes in wt vs *dys-1*. Table displays common and differential GO expression in wt vs *dys-1*. **B)** As per panel A, but for genes downregulated by spaceflight. This figure was provided by Dr. Craig Willis.

- **Figure 5.6. Cluster analysis to identify co-expressed gene modules. A)** Activity of gene co-expression modules in flight vs ground control comparison. **B)** Activity of gene co-expression modules in *dys-1* vs wt comparisons. **C)** Table of module annotations and summarised activation between pairwise comparisons. Panel C was designed by the thesis author but the final formatting of Panel C was done by Dr. Craig Willis

- **Figure 5.7.** Heatmap of potentially relevant chemical and biologic drug targets. Targets are predicted to be significantly activated or inhibited for significantly upregulated (UR) and downregulated (DR) genes from the wt flight vs. wt ground condition, highlighting potential for alteration in therapeutic potential in the context of spaceflight-induced changes. This figure was designed by the thesis author but the final formatting of this figure was done by Dr. Craig Willis.

- **Figure 5.8.** Detailed temperature profile of the culture bags during shipping from Texas Tech University (TTU) to Eastern Virginia Medical School (EVMS) and Launch. Bags were stored at 20ºC after receiving them at the ISS Space Automated Bioproduct Laboratory (SABL) and during transfer to the TTU Lab. This figure was provided by Dr. Soni Purushottam.

- **Figure 6.1. Global data overview. A)** Breakdown of the rodent datasets used in this study. **B)** Clustering of the most variable genes within the rodent datasets with functional annotation. Panel B was provided by Mr. Jonas Elsborg.

- **Figure 6.2. Differentially expressed genes shared between datasets. A)** An upset plot showing the number of significant (False discovery rate (FDR) ≤ 0.1) differentially expressed genes (DEGs) in spaceflight versus ground data subsets and the number of overlapping significant DEGs between these data subsets. The coloured annotation bar on the left of the plot shows how the original datasets divide into 10 data subsets based on all unique conditions including diet, biological sex and strain. The bar plot on the right of the upset plot shows the number of significant DEGs in each of the 10 data subsets. The bar plot on the top shows the number of intersecting DEGs between combinations of the data subsets, as indicated by the connected dots within the body of the upset plot. Black connecting lines indicate combinations spanning across multiple missions, and other connecting lines are coloured according to the annotation bar, based on their shared mission. **B)** A bar plot of significant Gene Ontology Biological Process (GOBP) pathways from significant DEGs shared by any subsets within the MHU-2 mission (i.e., DEGs from the blue bars in panel A).

- **Figure 6.3. All cross-mission genes involved in rodent skin spaceflight response.** A heatmap showing regulatory changes of all cross-mission genes (genes that are significantly (FDR ≤ 0.1) differentially expressed between flight and ground control across multiple missions) within each rodent data subset.

- **Figure 6.4. Cross-mission genes involved in rodent skin spaceflight response. A)** A heatmap showing regulatory changes of select cross-mission genes (genes that are significantly (FDR ≤ 0.1) differentially expressed between flight and ground control across multiple missions) within each rodent data subset **B)** A barchart showing the GOBP pathways that the full list of cross-mission genes are involved in. **C)** A decision boundary plot of the strongest cross-mission QLattice model using only the expression of two genes as input features to predict the spaceflight status of the all mice in the data. **D)** A decision boundary plot of the second-strongest model across all missions. Panels C and D were provided by Mr. Jonas Elsborg.

- **Figure 6.5. The profile of the cross-mission genes in astronauts. A)** Heatmap showing average expression in orthologs of the rodent skin cross-mission genes in

astronaut skin data from the Inspiration4 mission, for different skin layers. **B)** Heatmap showing $\log_2$ fold-change in orthologs of the rodent skin cross-mission genes in hair follicle samples at different time points from the JAXA HAIR astronaut study. Panel A was formatted by Dr. Jiwoon Park.

- **Figure 6.6. The profile of all the cross-mission genes in astronauts. A)** Heatmap showing t-score in orthologs of the rodent skin cross-mission genes in astronaut blood samples from sorted cell fractions from the NASA Twin Study. **B)** Heatmap showing $\log_2$ fold-change in orthologs of the rodent skin cross-mission genes in plasma samples at different time points from the JAXA CFE astronaut study. **C)** Heatmap showing average scaled expression in orthologs of the rodent skin cross-mission genes in astronaut PBMC data from the Inspiration4 mission at different timepoints. Panel A was formatted by Dr. Afshin Beheshti, and Panel C was formatted by Dr. Jiwoon Park.

- **Figure 6.7. Behaviour of the specific genes associated with skin health in astronauts and rodents. A)** The orange (suppressed) and green (enriched) heatmap shows the normalised enrichment score (NES) of curated skin health pathways (full list **Table S6.1**) that are significant (FDR ≤ 0.05) in at least one rodent subset. The red (upregulated) and blue (downregulated) heatmap shows the t-score for leading edge genes from the significant pathways that are significant (FDR ≤ 0.1) in at least two data subsets. **B)** Heatmap showing average expression in orthologs of the genes from Panel A in astronaut skin data from the Inspiration4 mission, for different skin layers. **C)** Heatmap showing $\log_2$ fold-change in orthologs of the genes from Panel A in hair follicle samples at different time points from the JAXA HAIR astronaut study. Panel B was formatted by Dr. Jiwoon Park.

- **Figure 6.8. The profile of skin health genes in astronauts. A)** Heatmap showing t-score in orthologs of the rodent skin, skin health genes (from **Figure 6.7**) in astronaut blood samples from sorted cell fractions from the NASA Twin Study. **B)** Heatmap showing $\log_2$ fold-change in orthologs of the rodent skin, skin health genes (from **Figure 6.7**) in plasma samples at different time points from the JAXA CFE astronaut study. **C)** Heatmap showing average scaled expression in orthologs of the rodent skin, skin health genes (from **Figure 6.7**) in astronaut PBMC data from the Inspiration4 mission at different timepoints. Panel A was formatted by Dr. Afshin Beheshti, and Panel C was formatted by Dr. Jiwoon Park.

- **Figure 6.9. DNA damage and repair pathways being regulated in rodents flown to space.** Heatmap of pathways relating to DNA damage response and repair mechanisms (full list **Table S6.1**), significant (FDR ≤ 0.05) in at least one data subset.

- **Figure 6.10. Mitochondrial specific analysis on rodent spaceflight skin tissue. A)** Heatmap of pathways relating to the mitochondria, significant (FDR ≤ 0.05) in at least one data subset. **B)** Decision boundary for the 2-gene model related to mitochondrial changes. The model indicates an increased removal of the toxic D2H compound in mitochondria through upregulation of *D2HGDH*, which is less pronounced when *PPP1R3B* expression is suppressed. Panel B was provided by Mr. Jonas Elsborg.

- **Figure 6.11. Astronaut physiological markers compiled from up to 50 astronauts. A)** Specific blood markers which contain data points for pre-launch (L-), flight (FD), and return to Earth (R+). The numbers on the x-axis indicate the number of days for each group. Interim Resistive Exercise Device (iRED) is shown in blue and Advanced Resistive Exercise Device (ARED) is shown in red. **B)** Specific blood markers which contain data points for pre-launch (L-) and return to Earth (R+).

C) Specific urine markers which contain data points for pre-launch (L-), flight (FD), and return to Earth (R+). The statistics on the data are *** (shown as *) $p < 0.001$ for significantly different from L-45 and ** (shown as ‡) $p< 0.01$ significantly different from ARED. This figure was provided by Dr. Afshin Beheshti.

- **Figure 6.12. Predicted potential countermeasures for mitigating spaceflight response to the skin.** Predicted drug signatures using the cross-mission genes across each dataset represented by a hierarchically clustered heat map. A positive (red) activation state implies cross-mission gene expression changes are consistent with expression changes observed with the indicated drug from curated causal gene expression relationship studies. A negative (blue) activation state implies the cross-mission gene expression changes are opposite to changes observed with the indicated drug.

Tables

Publications

- **Cope H***, Elsborg J*, Demharter S, McDonald JT, Wernecke C, Parthasarathy H, Unadkat H, Chatrathi M, Claudio J, Reinsch S, et al. Transcriptomics analysis reveals molecular alterations underpinning spaceflight dermatology. Communications Medicine. 2024;4(1), p.106.

  *These authors contributed equally to this work

- Rutter LA, MacKay MJ, **Cope H**, Szewczyk NJ, Kim J, Overbey E, Tierney BT, Muratani M, Lamm B, Bezdan D, Paul AM, et al. Protective alleles and precision healthcare in crewed spaceflight. Nature Communications. 2024 Jul 22;15(1):6158.

- Rutter LA*, **Cope H***, MacKay MJ*, Herranz R, Das S, Ponomarev SA, Costes SV, Paul AM, Barker R, Taylor DM, Bezdan D, et al. Astronaut omics and the impact of space on the human body at scale. Nature Communications. 2024 Jun 11;15(1):4952.

  *These authors contributed equally to this work

- Mathyk BA, Tabetah M, Karim R, Zaksas V, Kim J, Anu RI, Muratani M, Tasoula A, Singh RS, Chen YK, … **Cope H**, et al. Spaceflight induces changes in gene expression profiles linked to insulin and estrogen. Communications Biology. 2024 Jun 11;7(1):692.

- Soni P*, Edwards H*, Anupom T, Rahman M, Lesanpezeshki L, Blawzdziewicz J, **Cope H***, Gharahdaghi N, Scott D, Toh LS, et al. Spaceflight induces strength decline in Caenorhabditis elegans. Cells. 2023;12:2470.

  *These authors contributed equally to this work (ordered by affiliation).

- Manzano A, Weging S, Bezdan D, Borg J, Cahill T, Carnero-Diaz E, **Cope H**, Deane CS, Etheridge T, Giacomello S, et al. Enhancing European capabilities for application of multi-omics studies in biology and biomedicine space research. iScience. 2023 Sep 15;26(9).

- **Cope H**, Deane CS, Szewczyk NJ, Etheridge T, Williams PM, Willis CRG. A Data Collection Programme for Improving Healthcare in UK Human Spaceflight Ventures. Journal of the British Interplanetary Society. 2023;76:213-20.

- **Cope H**, Willis CRG, MacKay MJ, Rutter LA, Toh LS, Williams PM, Herranz R, Borg J, Bezdan D, Giacomello S, et al. Routine omics collection is a golden opportunity for European human research in space and analog environments. Patterns (N Y). 2022 Jul 30;3(10):100550.

- Overbey EG, Das S, **Cope H**, Madrigal P, Andrusivova Z, Frapard S, Klotz R, Bezdan D, Gupta A, Scott RT, et al. Challenges and considerations for single-cell and spatially resolved transcriptomics sample collection during spaceflight. Cell Reports Methods. 2022 Nov 21;2(11).

- Cahill T, **Cope H**, Bass JJ, Overbey EG, Gilbert R, da Silveira WA, Paul AM, Mishra T, Herranz R, Reinsch SS, et al. Mammalian and invertebrate models as complementary tools for gaining mechanistic insight on muscle responses to spaceflight. International Journal of Molecular Sciences. 2021 Aug 31;22(17):9470.

- Rutter L, Barker R, Bezdan D, **Cope H**, Costes SV, Degoricija L, Fisch KM, Gabitto MI, Gebre S, Giacomello S, et al. A new era for space life science: international standards for space omics processing. Patterns. 2020 Dec 11;1(9).

Chapter I: Background

1.1. From Last Universal Common Ancestor (LUCA) to Lunar: Human Adaptation to Space

For the majority of human history, evolution served as the primary mechanism for human adaptation and survival, guiding our development and responses to changing environments on Earth. From the Last Universal Common Ancestor (LUCA), a single-celled organism born billions of years ago (Schopf et al., 2002), life evolved through Darwinian mechanisms into a tapestry of forms, including multicellular organisms in the Cambrian explosion around half a billion years ago (Erwin et al., 2011). This led to the rise of primates within the last 100 million years and culminated in the emergence of **'**modern humans**'** (*Homo sapiens*) within the last few hundred thousands years, characterised by advanced cognition and dexterity (Bergström et al., 2021).

Remarkably, these abilities endowed by evolution have served in some part to supersede it. The process of evolution is inherently slow, constrained by the generational transmission of genetic material, a narrow bottleneck on the rate of adaptation. While fundamental evolutionary mechanisms like natural selection, mutation, genetic drift, and gene flow have shaped our species, they operate over extended timescales (Lenormand, 2002; Lynch et al., 2016). This gradual adaptation has failed to keep pace with humanity's inherent curiosity and ambition. As such, technological advancements - made possible through *Homo sapiens'* enhanced cognitive abilities and dexterity - have increasingly supplemented evolution. Indeed, at least within the last 10,000 years, societal technology in areas such as agriculture and medicine has begun to mitigate the selective pressures that would otherwise drive biological evolution. Under these extraordinary circumstances, technology and culture predominantly shape human adaptation to thrive in

new environments while biological evolution takes on a more subdued role (Richerson and Boyd, 2008). Indeed this process often manifests as humans altering natural environments or creating enclosed artificial environments within natural environments.

There is perhaps no example of this effect more hyperbolic than space travel. Space is an extreme environment, a near-perfect low-pressure vacuum, where bodies experience weightlessness, exposure to extreme temperatures, and cosmic radiation. All life as we know it has originated and evolved on Earth, and the only animal substantiated to survive in the space environment is the microscopic tardigrade, a so called ‘extremophile’, which through extraordinary mechanisms of desiccation and radiation-resistance can survive short periods of space exposure and in some cases even be recovered (Jönsson et al., 2008). Humans, not having these same biological mechanisms in place, have instead been able to somewhat adapt to space by engineering spacecraft and spacesuits - artificial **‘built’** environments for travelling beyond Earth and sheltering life within space. Tangentially, in the sole case of the Apollo program, this also involved supporting life beyond low earth orbit (LEO), including the lunar surface. This approach, however, has proven to be an imperfect solution. The built space environment (i.e. the spaceflight environment) remains a hostile habitat for life, characterised by stressors including “microgravity” (near-weightlessness experienced due to reduced gravitational forces, such as during the continuous freefall of orbiting spacecraft), cosmic radiation, and confinement within a noisy, closed, isolated, and pressurised vessel with elevated carbon dioxide levels (Afshinnekoo et al., 2020). Consequently, humans in space experience a litany of undesirable physiological and psychological effects with associated health risks (Afshinnekoo et al., 2020).

Some of these physiological changes driven by spaceflight showcase the human body’s remarkable malleability, and have been mitigated to

recoverable levels with contemporary countermeasures in mission formats of up to 1-year on the International Space Station (ISS) (Petersen et al., 2017). For example, while other factors appear to play a part, degradation of muscle and bone are primarily driven by the mechanical unloading induced by weightlessness (Shackelford, 2019). Aptly named 'anti-gravitational' muscles including the *soleus* and *gastrocnemius* in the calf are highly active during everyday tasks on Earth like standing and walking. In the spaceflight microgravity environment, the load on these muscles decreases dramatically, as they are no longer required to support the human frame against the force of weight. Providing adequate nutrition and resistive exercise in space is well established as a solution for mitigating the atrophy of muscles and bone (Petersen et al., 2016; Smith, MA Heer, et al., 2012), helping to maintain the whole-body health benefits of the musculoskeletal system, and easing readaptation to the 1G environment on Earth upon landing.

However, there are other physiological changes which do not appear to be sufficiently mitigated with existing countermeasures for current mission formats. Spaceflight Associated Neuro-ocular Syndrome (SANS) for example is an intriguing condition distinct to spaceflight. SANS is characterised by structural changes in the eye such as swelling of the optic disc and flattening of the posterior globe (Ong et al., 2023). These structural alterations are associated with symptoms including vision complications and headaches, and there are concerns surrounding the persistent and potentially permanent effects of SANS on vision (Mader et al., 2011). The precise causes and risk factors of SANS are yet to be established, with current hypotheses including microgravity-driven fluid shifts, nutritional deficiencies in B-vitamins, genetic predispositions in 1-carbon metabolism, and spaceflight-induced mitochondrial dysregulation (Ong et al., 2023; Waisberg et al., 2024; Zwart and Smith, 2022). SANS is not anomalous in this sense; several physiological changes and potential health risks

associated with the spaceflight environment require further investigation, including dermatological issues like rashes, spaceflight motion sickness, decreased cognitive performance, risk of deep venous thrombosis (DVT, i.e. blood clots), and cancers (Afshinnekoo et al., 2020; Heer and Paloski, 2006; Limper et al., 2021; Nguyen and Urquieta, 2023; Patel et al., 2020; Reynolds et al., 2021). Once the pathology of these issues is understood, risk stratification and mission planning can be improved and - in cases where opportunities for clinical actionability are identified - countermeasures can be designed and deployed to help mitigate these issues.

The motivation to understand and mitigate these issues goes beyond simply increasing health and safety during current governmental space missions to the International Space Station (ISS). International space agencies have set forth ambitious plans to travel beyond Low Earth Orbit (LEO), with goals to return to and establish a base on the lunar surface, such as via the Artemis program, and to eventually land the first humans on another planet, Mars. These lofty ambitions will require a huge feat of engineering, part of which includes preparing for the human body and mind to be placed under unprecedented stressors in challenging situations where options for communication, evacuation, and resupply are severely limited (Patel et al., 2020). Furthermore, while governmental agencies focus on the aforementioned goals, space access has entered a period of democratisation with increased involvement from commercial entities. There is now a trajectory of increasing diversity amongst humans entering space, in terms of factors like sex, age, physical fitness levels, and ethnicities. Just as the safety of newly engineered commercial spacecraft needs to be ensured, the biological impact of the full variety of commercial mission formats and health statuses needs to be assessed to also ensure health and safety in these burgeoning ventures.

1.2. Omics: The Biological Big Data Revolution on Earth

As technology has progressed on Earth, humanity has developed tools for querying and comprehending changes within biological organisms at an increasingly fundamental level. 'Omics' has emerged as the catch-all term for capturing and studying digital information derived from biological systems; indeed, omics has become synonymous with biological 'Big data' (Perakakis et al., 2018).

1.2.1 A Plethora of Omics

As visualised in **Figure 1.1**, the principal omics data types can be conceptualised as a layered cascade of information associated with different biomolecules. Genomics, the foundational layer of omics, involves the study of genomes, which are the complete set of an individual organism's DNA, organised into chromosomes capped by telomeres. DNA sequences contain genes and regulatory elements critical for development and function. The majority of genomic information is static and inherited, such as germline genetic variants which can directly introduce diseases or traits like eye colour, or may have indirect influences, such as elevating susceptibility to conditions like heart disease (Zhang et al., 2014). Genomic information can also be dynamic. For example, somatic genetic variants accumulate over a lifetime due to faulty DNA replication or environmental influences like DNA damage via ionising radiation (Behjati et al., 2016). If the automatic DNA damage repair mechanisms fail to correct somatic mutations, cells with these variants can proliferate, sometimes leading to diseases like cancer (Behjati et al., 2016). Epigenomics, the study of chemical modifications on DNA or DNA's histone 'packaging', follows genomics. While these epigenetic **'**marks**'** do not directly edit DNA sequences, they can alter the behaviour of genes for varying degrees of permanence and heritability. For example DNA methylation is a chemical modification which can turn genes off in response to environmental factors, such as to provide protective

effects (Tiffon, 2018). Transcriptomics examines the full set of RNA transcripts within cells and follows on from epigenomics via the process of transcription. During transcription, RNA polymerase attaches to DNA sequences, using the DNA as a template to synthesise RNA transcripts. A range of different transcript types can be produced through transcription (Perakakis et al., 2018), including messenger RNA (mRNA) which contain the instructions needed to produce proteins via a process called translation. This end-to-end process of DNA to a functional gene product via transcription and translation is termed gene expression; importantly, transcripts are sometimes also subject to further post-transcriptional modifications which can alter their stability, localisation, and functionality (Corbett, 2018). Following on from transcriptomics via the translation process, proteomics studies all proteins and their roles in cellular structure and functions. These proteins, which are often further modified via post-translational modifications (PTMs), play critical roles within cells including catalysing metabolic reactions and providing structural support (Pagel et al., 2015). Finally, metabolomics involves the analysis of the full range of metabolites, small molecular end-products of the cellular processes initiated by proteins. These metabolites include essential functional products like lipids, sugars, and hormones (Perakakis et al., 2018).

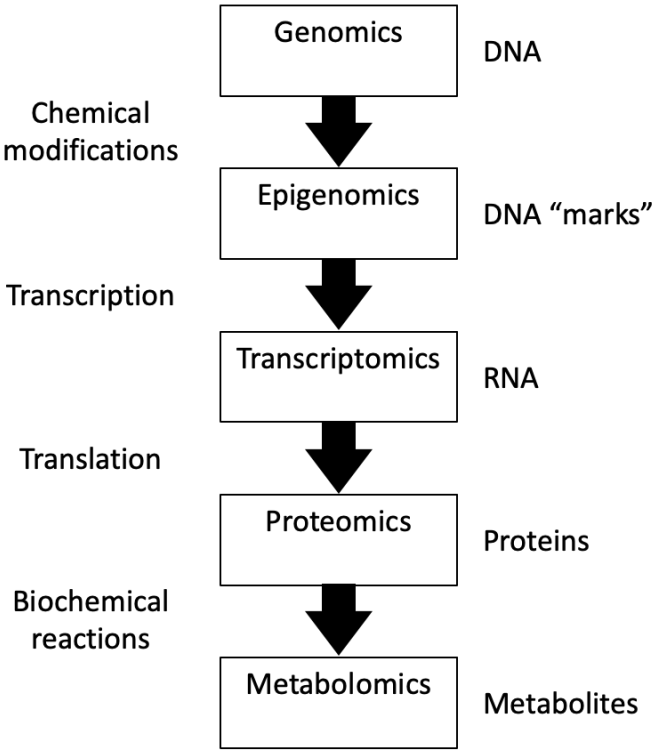

Figure 1.1. The omics cascade. Core omic types with their associated biomolecules (right side) and primary transitions to other omic layers (left side).

The flow of information between and within these layers is not always linear and there continues to be an explosion of additional sub-disciplines of omics with varying degrees of maturity. For example, microbiomics is rapidly revealing the impact of the collective genomes of dynamic microorganism populations (e.g. in the gut and on the skin) on host health (Kinross et al., 2011), and connectomics is an emergent field aiming to model the neuronal circuitry within the brain and wider nervous system, which could help to elucidate mechanisms governing mental health and cognition (Tozzi et al., 2020). In an effort to unite the plethora of omic sub-disciplines, 'multi-omics' has gained popularity as a term for integrating data from multiple omic data types to provide holistic perspectives of biological systems, such as by combining genomics and transcriptomics to investigate the impact of genetic variants on gene expression (Huang et al., 2014). Multi-omics presents an effort to combine multiple layers of biological data in order to

model the larger picture; thus making it a characteristic example of ‘systems biology’ (Hasin et al., 2017).

1.2.2 The Promise of Integrating Omics with Other Data

Beyond omics, phenotypic data encodes information pertaining to an organism’s phenotype, the full set of their observable traits, including clinical symptoms. As technology and scientific knowledge have evolved the definition of ‘observable’ has been continually stretched, incorporating a broader scope of traits with enhanced precision. For example, prior to the invention of X-ray imaging in 1895, assessment of a broken bone relied on manual examination and subjective descriptions of symptoms. Advancements in medical imaging now enable healthcare professionals to objectively visualise the extent and specifics of fractures, informing treatment. Beyond the boon of medical imaging and other morphological assessment technologies, phenotypic data encapsulates a huge spectrum of measures, such as behavioural logs for sleep, diet, mood, cognition, and exercise performance, and wearable biosensors for passively monitoring physiological metrics like heart rate, blood pressure, and more (Dunn et al., 2021; Zahedani et al., 2023).

Multi-modal approaches combine omics with phenotypic data and even environmental exposure data (which are sometimes even categorised as ‘phenomics’ and ‘exposomics’ respectively (Donovan et al., 2019)). The aim of multi-modal approaches is to describe the end-to-end chain of events leading to a phenotypic outcome. As a thought experiment, one might imagine constantly collecting a comprehensive set of multi-omics, phenotypic, and environmental exposure data from a comprehensive population of humans as they go throughout their lives (Chen et al., 2012; Gao et al., 2022). From this holistic longitudinal dataset of different data types it would in theory be possible to fully capture the cause and effect

patterns behind all health changes. Given this knowledge, it could be possible to use an individual's personal data to make tailored recommendations for enhanced performance, or even to spot a determinantal signature before the full chain of events has transpired, enabling opportunities for early intervention via countermeasures like targeted drugs. This is the ultimate fantasy of 'personalised/precision medicine', to be able to maximally optimise an individual's treatment using their personal data through processes like pharmacogenomics, where drug prescriptions are optimised based on genomic variants (Ginsburg and Phillips, 2018; Goetz and Schork, 2018).

1.2.3 Challenges of Integrated Omics

The reality is that this dream of a fully longitudinal and end-to-end approach centred around omics is currently infeasible at several stages. Firstly, sampling for omics requires extraction of biological material (i.e. biospecimens), which often entails invasive procedures by trained professionals (Hirahata et al., 2022); unlike certain phenotypic measures like heart rate, this cannot be performed passively at a constant frequency. Secondly, there is the matter of converting biological data into digital information, a process that often necessitates laborious and expert-driven lab procedures, the use of expensive consumables and machinery, and mass storage solutions for the resultant data (Rehm et al., 2013). The conversion of biological signals to digital signals tends to result in reduced precision and loss of information, despite the generation of often massive volumes of highly-dimensional data from each measurement. Thirdly, this digital data then needs to be further processed, including complex integration in the case of multi-modal and multi-omics, so that patterns can be identified; this requires significant computational resources alongside advanced algorithms, followed by expert manual interpretation and reporting. In a procedure that is notoriously vigorous and often lengthy,

targeted follow-up studies and comparisons to large cohorts are often needed to move beyond weak correlation and confirm any novel causal relationships between omic, phenotypic, and environment data. Fourthly, the extent to which these data can be collected, analysed, utilised, and shared is heavily dependent on societal perceptions and often slow-to-develop policy frameworks which are rooted in the latent ethical and cultural impacts of these technologies. Finally, artefacts can be introduced at all stages in this process, such as via software bugs or human-errors, and data generated at different times and by different labs may not be directly comparable to other datasets due to even slight differences in methodology, often linked to a lack of standardisation. As such, the fantasies of multi-omics and multi-modal data for biological discovery and personalised medicine are only fulfilled in part at present.

1.2.4 Omics in Current Practice

In spite of these considerable limitations, omics-centred approaches have become ubiquitous in contemporary biology and medicine, playing a pivotal role in both research and clinical workflows. As these techniques continue to advance, so too will their influence on research and practice.

One area of significant improvement has been the advent of next-generation sequencing (NGS) technologies, which are high-throughput devices capable of accurately and precisely reading DNA and RNA sequences (Satam et al., 2023). NGS technologies enable a broad spectrum of genomic analyses, from single-gene and small-panel sequencing to comprehensive assessments like whole exome sequencing (WES) - which targets the protein-coding regions of the genome - and whole genome sequencing (WGS) (Satam et al., 2023) - which targets all regions of the genome. Beyond genomics, NGS is also used for other omics, including transcriptomics and epigenomics. Notably, RNA-Seq has emerged as a

prominent technique in transcriptomics for measuring the abundance of transcripts (i.e. quantifying gene expression) and investigating post-transcriptional events (Satam et al., 2023). Just as RNA-Seq has largely supplanted microarrays for transcriptomics, newer single-cell and spatially-resolved techniques are quickly gaining traction due to their offer of unprecedented resolution for transcriptomics. Single-cell sequencing enables quantification of gene expression for individual cells, and spatially-resolved transcriptomics is a multi-modal technique which quantifies gene expression for cell populations across a given tissue sample while also providing the spatial context by linking gene activity to specific regions within the histology image of the sample (Andersson et al., 2020). For epigenomics, techniques like long-read sequencing can even simultaneously identify epigenetic marks during genomic sequencing (Satam et al., 2023). With the widespread adoption of NGS technology, the price of sequencing samples has continued to drop, enabling large-scale public health NGS projects. For instance, the 100,000 genomes project led by Genomics England generated a huge WGS dataset from UK National Health Service (NHS) patients for research and clinical purposes (Turnbull et al., 2018), and the NHS has also integrated routine WGS for unwell children with likely undiagnosed genetic conditions (R14 pathway) and WGS in the event of prenatal anomalies with potential genetic origins (R21 pathway) (Emms et al., 2022).

Another area of improvement to highlight has been the exponential increase in data sharing and the adoption of large-scale data and knowledge-base repositories. While data-sharing in omics has an established history, such as via the infamous Human Genome Project (Lander et al., 2001), there has been an explosion in the number of initiatives and repositories which house knowledge-bases, omics data, and even multi-modal data. In addition to the aforementioned 100,000 genomes project, initiatives include the Gene Expression Omnibus (GEO) database which contains over 4000 gene

expression profiling and/or RNA methylation profiling datasets (Clough and Barrett, 2016), the UK Biobank containing de-identified genomic and phenotypic data from half a million UK participants (Halldorsson et al., 2022; Sudlow et al., 2015), and ClinVar which hosts reports with supporting evidence describing the relationship between specific genetic variants and phenotypes (Landrum et al., 2016).

As omics initiatives have grown, policy has followed suit, with the governance of genomic data being regulated through acts such as the Genetic Nondiscrimination Act (GINA, 2008) in the United States, and the General Data Protection Regulation (GDPR, 2016) within the European Union and European Economic Area. These legal acts instil ethical principles described by international policies like the Declaration of Helsinki (1964) (World Medical Association, 2013), ensuring that omics research today includes mechanisms like informed consent, ethical review boards for oversight, and clear policies regarding follow-up counselling and handling of incidental findings. These policies and legal instruments also aim to ensure that personally identifiable information is carefully managed so as to not harm the patient through risks like discrimination or data breaches - an example of which is inappropriate disclosure of personal data, breaching individuals' privacy.

Finally, the availability and capability of analysis tools and computational resources has greatly improved. Bioinformatics is the interdisciplinary field of developing and applying software tools to the digital data extracted from biological investigations. To keep pace with the wealth of newly generated omic datasets, new analysis tools are constantly developed. These tools - often released as publicly accessible software libraries - are integrated into data analysis workflows, including standardised workflows known as 'pipelines'. These pipelines automate the processing of data through a series of steps; samples are run individually through the same pipeline, often

making use of parallel processing, to ensure efficient and consistent analysis. Machine learning, which develops computational models for mapping sets of inputs to outputs through the automated discovery and encapsulation of patterns in the data, has long been embedded within bioinformatics workflows. Unsupervised clustering and dimensionality reduction approaches including k-means, hierarchical clustering, and principal component analysis (PCA) are particularly well established in data exploration and visualisation tasks (Dalmaijer et al., 2022), and 'classical' supervised learning algorithms like support vector machines (SVMs), random forests, and regression approaches have been extensively studied in contexts like disease classification from omics data (Ahsan et al., 2022). With their success in other domains, and with the increased availability of omics datasets and computational resources (including popularisation of the cloud computing paradigm), deep learning approaches, which use artificial neural networks with many layers and tunable parameters, have seen recent emergence in bioinformatics. The scale of these deep learning models enables them to capture complex patterns from high-dimensional data, including omics. For example, a deep learning approach was recently applied to long-read genomic data for rapid classification of central nervous system (CNS) tumours during surgery, which can help to inform the surgeon on how best to operate on the tumour (Vermeulen et al., 2023). Another novel application of deep learning in bioinformatics is the development of the DeepVariant tool for calling genetic variants (Poplin et al., 2018). DeepVariant converts genomic sequence data into an image format in order to solve the problem of detecting and classifying germline variants as an image classification task via Convolutional Neural Networks (CNNs) (Poplin et al., 2018). CNNs are a specialised deep learning architecture which use filters to automatically learn meaningful features from image data; CNNs have enabled state-of-the-art performance in many imaging tasks including segmentation of medical imaging data (Ronneberger et al., 2015). Finally, the transformer deep learning architecture with its 'self-attention'

mechanism has triggered a recent AI boom, particularly in natural language processing (NLP) tasks where the architecture's ability to capture long-range contextual relationships between words (often subdivided into multiple tokens) enables it to achieve state-of-the-art performance across tasks like language translation, question answering, and text summarisation. In the context of omics the transformer architecture has started to be implemented as part of tools like AlphaFold2 and ESMFold to accurately predict 3D protein structure from protein sequence data (Jumper et al., 2021; Lin et al., 2023), and AlphaMissense and Primate-AI-3D which use the protein structural information and masked protein sequence prediction for the downstream task of predicting whether genetic variants are likely benign or pathogenic (i.e. disease-causing) (Cheng et al., 2023; Gao et al., 2023).

1.3. Space Omics: Omics meets Space Life Sciences

Space biology is the study of how Earth-originated life adapts to the environmental conditions of space, including the biological impact of spaceflight-associated stressors like microgravity and cosmic radiation. While related, space biology is distinct from the field of astrobiology, which investigates the origins and evolution of life throughout the universe, including the search for extraterrestrial life (Chyba and Hand, 2005). Moreover, space biology intersects with the field of space medicine, which is concerned with safeguarding human health, safety, and performance in spaceflight ventures (Hodkinson et al., 2017). By investigating the pathology of spaceflight-associated health risks, space biology provides crucial insights that can translate to occupational health strategies for humans in space. All of these fields can be grouped under the broader term, space life sciences.

While omic technologies have been revolutionising terrestrial life sciences, their adoption within space life sciences (i.e. 'Space Omics') has lagged comparably due to the significant challenges in designing and conducting

space biology experiments. Retrospective data analysis of the literature indicates an approximate 10-year lag on adoption of emergent omics technologies with space biology compared to Earth, including the adoption of microarrays and subsequent shift to RNA-Seq within space biology (Gupta et al., 2024). Nonetheless, as shown in **Figure 1.2**, as omics technologies have matured and gained popularity on Earth, omics have emerged as a crucial tool within space biology research, and hold transformative potential for space medicine (Rutter et al., 2020). Through the inclusion of omics in space biology investigations, several molecular hallmarks of spaceflight on biology have been revealed, including elevated oxidative stress, increased rates of DNA damage, mitochondrial dysregulation, epigenetic and gene expression changes, telomere length alterations, and shifts in host-microbiome interactions (Afshinnekoo et al., 2020).

Furthermore, while the development and validation of new biological techniques and technologies predominantly occurs terrestrially before being translated into space biology investigations, there are occasionally cases where technological advancements made in the *outside-the-box* niches of space life sciences research provide translational value for research and patients on Earth (Afshinnekoo et al., 2020; Shelhamer et al., 2020; Shirah et al., 2023).

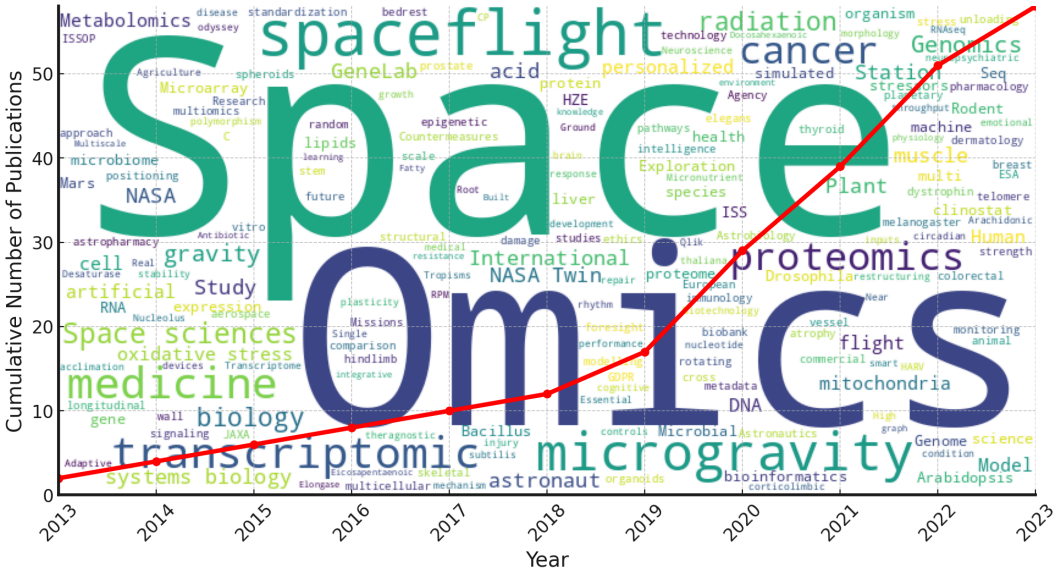

Figure 1.2. Growth of space omics literature. Growth in the number of publications on PubMed over the last decade for the search query ‘space’ AND ‘omics’, and commonality of words from the keywords of the returned literature. Data collected up to 19th November 2023.

1.3.1 Platforms for Space Omics

A variety of platforms have been developed for conducting space biology experiments, and can be broadly divided into four categories. The first two pertain to genuine spaceflight exposure: 'Human Spaceflight' where human subjects are studied *in vivo* in space, and 'Model Organism Spaceflight' where experiments are conducted with model organisms *in vivo* or *in vitro* in space. The final two categories are based on analogue environments that simulate aspects of spaceflight exposure: 'Human Analogue' for *in vivo* study of human subjects on Earth and 'Model Organism Analogue' for *in vivo* or *in vitro* experiments with model organisms on Earth. The specifications and availability of platforms has changed throughout the history of space life sciences, and will continue to evolve with the development of new facilities and technologies on Earth, and the changing

landscape of built environments in space. Each platform has advantages and disadvantages for conducting space biology studies, as summarised in **Figure 1.3**, and studies will sometimes make use of multiple experiments across platform types to provide robust insights.

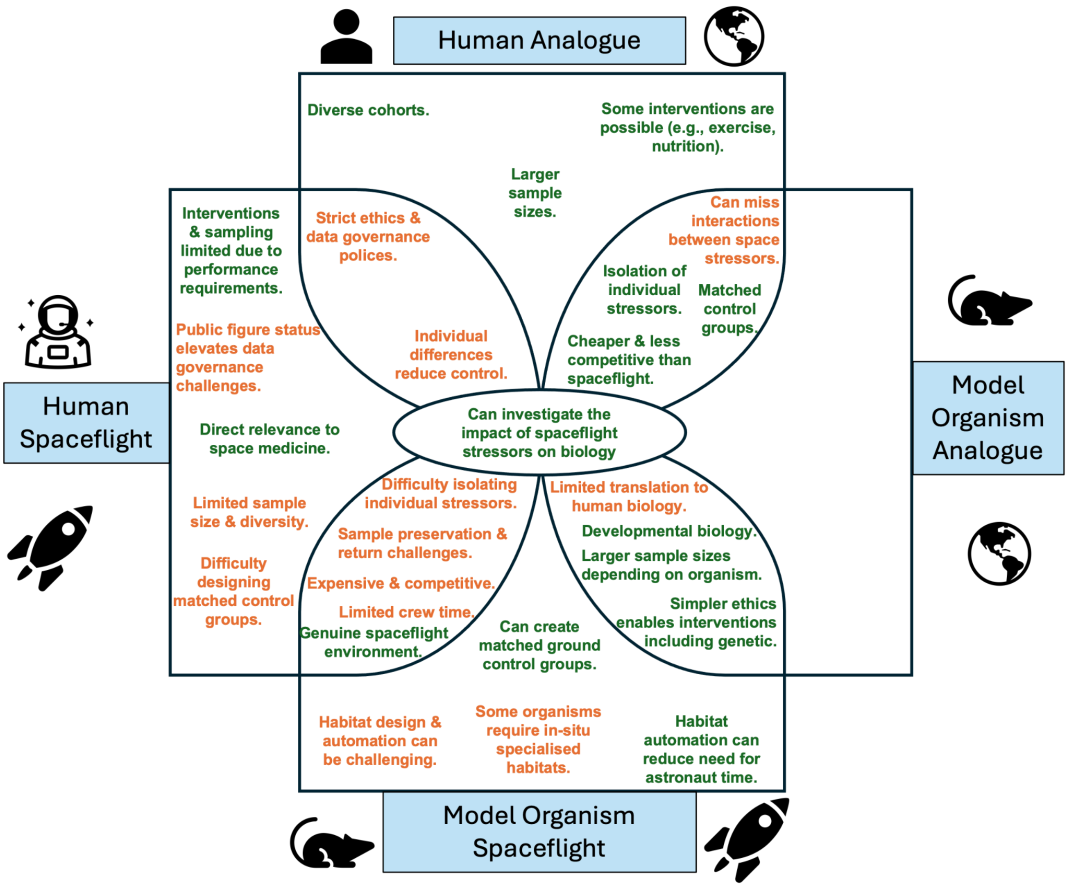

Figure 1.3. Platforms for space biology. Pros (green text) and cons (red text) of the four main platform types for space biology experiments: ‘Human spaceflight’, ‘Human analogue’, ‘Model organism spaceflight’, and ‘Model organism analogue’.

1.3.1.1 Human Spaceflight

The first platform category, human spaceflight, has a rich history. While the advancement and adoption of rocketry technologies, including in aircraft, were increasing prior to the war, the use of rockets for spaceflight were primarily catalysed by the development of missiles and aircraft throughout World War II and into the subsequent Cold War ‘Space Race’ between world

superpowers, the Soviet Union and the US. In spite of prior testing with model organisms and training in analogue environments, when the first human subject entered space in 1961 there was an uncertainty over whether the journey would even be survivable. Systems for monitoring heart rate, blood pressure, body temperature were put in place, and during the successful 108 minute mission, Yuri Gagarin is reported to have conveyed that he was feeling well over the radio several times. Although the Cold War motivated a lack of public disclosure by the Soviet Union, no health issues were reported, even with Gargarin reportedly experiencing 8g forces during re-entry (Bruhns, 2012). Following this maiden flight, human subjects have been studied within a wide variety of spacecraft (**Table 1.1**) primarily via government programmes with a gradual trend towards international collaboration and increased commercialisation (Griko et al., 2022).

Table 1.1. Crewed spaceflight programs preceding the ISS.

Agency	Program	Years Active	Duration Range	Crew Size	Spacecraft
Soviet/ Roscosmos	Vostok	1961-1963	Minutes to 1 day	1	Capsule
	Voskhod	1964-1966	1 to 2 days	2-3	Modified Vostok Capsule
	Almaz	1964-1978	Days to months	2-3	Military Space Station
	Soyuz	1967-	Crewed for hours to a few days (months if docked)	1-3	Capsule
	Salyut	1971-1986	Days to 8 months	1-5	Space Station
	Mir	1986-2001	Days to over 1 year	2-6*	Modular Space Station (*without shuttle)
NASA	Mercury	1958-1963	15 minutes to 1.5 days	1	Capsule
	Gemini	1961-1966	5 hours to 14 days	2	Capsule
	Apollo	1961-1972	10-13 days	2-3	Command/Service Module, Lunar Module
	Skylab	1973-1974	24-84 days	3	Space station

	Space Shuttle	1981-2011	2 to 17.5 days	2-8	Spaceplane, also housed ESA Space Lab
Soviet/NASA collaboration	Apollo-Soyuz Test Project	1975	9 days	5	Command/Service Module, Capsule
Roscosmos/ NASA collaboration	Shuttle-Mir	1994-1998	Days to over 1 year	5-13	Spaceplane, Space Station, Capsule

With its continuous occupation since 2000, the ISS has emerged as the premier platform for conducting human subjects research in space. During this period only a few rare human subject studies onboard the ISS have incorporated omics, with the most notable example being the NASA Twins study (Garrett-Bakelman et al., 2019). The NASA Twins study was a groundbreaking multi-omics experiment conducted with two identical twins. One twin completed a long-duration mission onboard the ISS from March 2015 - March 2016, while the other twin remained on Earth as a control. Notably, the monozygotic nature of the twins - and the fact that they were raised together and both became career astronauts - meant that their genomes and lifestyle background were remarkably similar at the study inception, theoretically helping to isolate the signal of stressors associated with the space mission. Throughout the full mission lifecycle, including pre-flight, in-flight, and post-flight timepoints, multiple biospecimens were collected from each subject in parallel, with subsequent multi-omics data including genomics, epigenomics, transcriptomics, proteomics, metabolomics and microbiomics being generated and analysed by select groups of approved researchers (Garrett-Bakelman et al., 2019). NASA runs a standard measures programme for participating U.S. and ISS International Partner crew members; thus, in addition to the unprecedentedly comprehensive set of longitudinal omics captured by the Twins study, phenotypic and environmental measures were also collected, including cognitive and sleep performance.

The clear advantage of studying human subjects in space is the direct relevance of the experiments to both space biology and space medicine. Stated plainly, the environment in which the experiments are conducted is a genuine spaceflight environment, and the subjects themselves are humans, so findings do not need to be translated from another organism. However, studying humans directly in the spaceflight environment does present significant challenges.

To start with, there are some general challenges associated with conducting biological experiments within the spaceflight environment. The capacity and frequency of spacecraft launching to and returning from space is extremely limited and the associated costs are significant, so slots for flying even compact payloads - such as handheld ultrasound scanners - onboard spacecraft are highly competitive and require considerable funding (Rutter et al., 2020). Secondly, astronaut time is also a valuable commodity, with astronauts spending their time in space working to full schedules of spacecraft maintenance tasks, outreach, workouts, sleep, and scientific experiments (Rutter et al., 2020). Thus, scheduling time for astronaut involvement in experiments in space is highly competitive, often with only very short windows granted and limited options for rescheduling in the case of unplanned changes to experiments, such as needing to repeat a measure. These scheduling difficulties extend to pre-flight and post-flight measures from the astronauts, albeit to a lesser degree, so the number of sampling timepoints is limited. Thirdly, biological studies often entail specialist training and equipment. Astronauts are highly trained but may lack the time for thorough training in the specialised skills needed to deliver specific space biology experiments (Rutter et al., 2020). Equipment and protocols for biological studies may also require time consuming and costly adaptation to spaceflight due to factors like altered fluid dynamics in microgravity, spacecraft size constraints, and stringent requirements for mitigating cross-contamination risks, including the spread of microbial populations

within the spacecraft. Fourthly, there are significant challenges in preserving and returning biological samples and data to Earth (Overbey et al., 2022; Rutter et al., 2020). For instance, some biomolecules like RNA degrade relatively quickly(hours) following biospecimen collection, which can impact gene expression findings (Gallego Romero et al., 2014). Thus, samples should be frozen for preservation as quickly as possible. While common in labs on Earth, liquid nitrogen snap-freezing is too dangerous an option for spacecraft, so samples on the ISS have typically been placed into freezers where they freeze to -80°C at a slower rate. Recently, rapid freezing technology to quickly bring samples to -185°C was introduced on the ISS for minimising these effects (Overbey et al., 2022). Additionally, media can be added to samples prior to freezing - such as RNALater - to help mitigate degradation (Lai Polo et al., 2020). While the ability to analyse samples directly in space is being explored for certain use cases - including analysing microbial swabs from crew members (Castro-Wallace et al., 2017; Stahl-Rommel et al., 2021) - samples are typically returned to Earth and transported to a terrestrial lab for analysis, creating situations where further degradation of samples can occur via factors like partial thawing, vibrations, and hypergravity (Garrett-Bakelman et al., 2019). This procedure of sample preservation and return can therefore lead to artefacts in omics data generated from in-flight samples.

There are also challenges of working with astronauts as human subjects. Astronaut performance must be maintained, so interventions or procedures like invasive tissue biopsies which could potentially affect performance, such as via soreness or infection, are risky and challenging to gain approval for despite common use during human subjects research on Earth. Furthermore, astronauts are a small population of public figures, so data governance challenges like protecting privacy have motivated strict policies pertaining to which data can be collected, and how it can be shared and utilised (Antonsen and Reed, 2019; Reed and Antonsen, 2018). While

diversity has increased, astronauts are historically a relatively homogeneous cohort when sampled against the worldwide population of humans on Earth, with the typical first-time astronaut being a physically-fit American male in their forties with a military background and at least a Master's level education (Smith et al., 2020). When coupled with the already low sample size, the statistical power of astronaut studies and their ability to generalise findings to heterogeneous human biology is limited. In spite of the relative similarities between astronauts, all humans are individuals with varied genomic and lifestyle backgrounds, and different habits in areas like diet, sleep, exercise, and adherence to study protocols. Thus, it is challenging to perform controlled studies with human subjects in general, and it is also difficult to match the spaceflight group with suitable control groups on Earth to enable statistical comparison. Finally, while human subjects in space being exposed to the full array of spaceflight stressors is beneficial for studying global signatures of the spaceflight environment on human biology, it does make it challenging to decipher the roles that individual stressors play in the biological impacts of spaceflight.

1.3.1.2 Model Organism Spaceflight

The use of model organisms in spaceflight predates human spaceflight, with initial model organisms sent to space as an often tragic means of gauging survivability. Subsequently, model organisms have been used as an essential means to study the effects of spaceflight on biological systems. A huge range of model organisms have been flown to space, including cell cultures, plants, invertebrates like tardigrades, worms, fruit flies and spiders, amphibians and aquatic animals including frogs and fish, reptiles including tortoises and geckos, and mammals including rodents, rabbits, dogs, monkeys, and apes (Burgess and Dubbs, 2007; Rutter et al., 2020).

The range of organisms used in model organism studies has mostly normalised in recent decades due to several factors (Deane et al., 2022; Rutter et al., 2020). Firstly, societal perceptions and policy regarding the ethics of animal research has changed over time, so at least for the ISS, space experiments with dogs and primates would likely not be conducted today unless no other organism was appropriate (Katahira, 2001). Secondly, the type of model organisms that can be sent to space is dependent on the enclosed habitats that are available to them onboard the spacecraft (Rutter et al., 2020). While this factor is a challenge for model organisms requiring larger specialised habitats - such as aquatic tanks and rodent cages - for smaller model organisms this factor is less of an issue as compact habitat infrastructure is often included within the payload sent to space (Rutter et al., 2020). The third factor for normalisation is that the space biology community, and indeed the wider life sciences community on Earth, has converged on several key model organisms for their scientific value and practicality (Bult and Sternberg, 2023). While less common model organisms can be useful for answering specific research questions - such as spiders for the impact of spaceflight on web building behaviour (Zschokke et al., 2020) - the standardisation of model organisms is useful for enabling comparability and collaboration across studies, and for supporting the development of shared analysis approaches and tools (Bult and Sternberg, 2023).

Rodents, especially rats and mice, are key model organisms in space biology. Forefront initiatives include NASA's Rodent Research (RR) and Japan Aerospace Exploration Agency's (JAXA) Mouse Habitat Unit (MHU) projects (Choi et al., 2020; Matsuda et al., 2019; Shiba et al., 2017). Although Europe and Russia have engaged in rodent research, such as via the Mouse Drawer System (MDS) payload and Bion M1 unmanned satellite experiment respectively, rodent models are not currently a programmatic focus in these regions (Cancedda et al., 2012; Deane et al., 2022; Radugina et al., 2018). The MHU and RR programmes have developed and installed

their own separate habitat infrastructures on the ISS for housing rodents, and both projects have conducted a series of experiments with their hardware (Choi et al., 2020; Horie et al., 2019; Shiba et al., 2017). Although factors like mouse strain, sex, and age are variable across the series of experiments, the repeated use of the same hardware enhances comparability between studies.

While strictly regulated and logistically challenging due to animal care needs, performing research in rodents comes with some distinct advantages. Firstly, rodents share many genetic and physiological similarities to humans, so while care should be taken due to some unconserved differences between the species, they are long-established as a popular model for translational research (Perlman, 2016). Secondly, mice within the same inbred strains have almost identical genomic profiles, which allows for more controlled experiments. Thirdly, building upon this point, while animal research in space is carefully regulated, it is possible to test interventions that may impact performance, including gene knock-out or gene knock-in models and the use of strains with established susceptibilities (Cadena et al., 2019; Uruno et al., 2021). Fourthly, alongside controlling for genomic variation and background and lifestyle factors like lighting, diet, age, and biological sex, spaceflight rodent experiments can be paired with matched control groups on the ground. These ground control groups are typically kept within the same hardware, with efforts taken to mirror factors like scheduling and atmospheric conditions (Beheshti et al., 2018; Choi et al., 2020). To this point, the MHU hardware also allows for centrifuge control groups, including creating 1G conditions within the spaceflight habitat; this can help to isolate the signal of microgravity on biological changes (Shiba et al., 2017).

One disadvantage of rodent studies in space is that the sample size is still relatively low due to the animals' size and housing requirements; so

statistical relationships are often based on only three replicates per condition. Additionally, rodent research has highlighted some of the major challenges with animal care, sample preservation and return. Firstly, the death of multiple mice during the long-duration MDS study on the ISS highlights both the difficulty of automating model organism experiments and dealing with operational stressors throughout the entire mission duration. The death of one mouse from the MDS ISS study was linked to failure of an automatic food dispenser (Cancedda et al., 2012; Neutelings et al., 2015), a failure which similarly led to mice fatalities during the Bion M1 mission (Andreev-Andrievskiy et al., 2014); while automation of experiments can improve control and reduce the time needed for human intervention, systems can fail, compromising the experiment, potentially harming the model organisms, and potentially affecting nearby experiments such as via thermal or fluid leakages. The death of another mouse in the MDS ISS study was linked to a spinal injury potentially caused by launch forces (Cancedda et al., 2012; Neutelings et al., 2015); scheduling delays, temperature fluctuations, and hypergravity events during launch and landing can have a significant impact on biological systems so must be considered during planning and interpretation of results (Choi et al., 2020; Lai Polo et al., 2020). For example, during the RR program, different sample dissection and return protocols have been trialled, concluding that Live Air Return (LAR) of animals prior to dissection is confounded by landing forces and recovery time, while euthanasia injections in space followed by freezing of tissues dissected on-orbit or whole carcasses can help to preserve the effects of microgravity exposure, subject to potential degradation due to the previously discussed freezing and preservation strategies (Choi et al., 2020; Lai Polo et al., 2020; Overbey et al., 2022).

With fruit flies being the first living organisms intentionally sent into space in 1947 (Burgess and Dubbs, 2007), small invertebrates have continued their legacy as key model organisms for spaceflight, with a particular focus on *C.*

elegans worms, *Drosophila melanogaster* fruit flies, and tardigrades. These invertebrates offer particular advantages over larger model organisms when it comes to cost, sample sizes, ease of handling and automation of compact habitats to minimise required astronaut time and payload volume, simplified animal research policies and thus opportunities for testing genetic and pharmaceutical interventions, and the ability to study different developmental stages of organisms, including sustaining multiple generations in space. Additionally, while not quite as similar to humans as rodents, invertebrates do share many evolutionary conserved biological mechanisms with humans, enabling translational research (Markaki and Tavernarakis, 2010; Yamamoto et al., 2023). For instance, *C. elegans* are microscopic nematodes with short lifespans (~3 weeks) which have seen extensive use as a model organism both on Earth and in space to study metabolic, developmental, reproductive, neurological, and neuromuscular processes (Cook et al., 2019; Markaki and Tavernarakis, 2010; A Scott et al., 2023).

Aside from rodents and small invertebrates, studies with cell cultures and plants remain popular within space biology. Plant studies tend to focus on the impact of spaceflight on plant growth, such as the gravitropism impact of microgravity on roots, rather than translational research for human biology. One of the key aims of plant space biology investigations is to demonstrate and optimise the growth of plants - including crops - in spaceflight environments, as demonstrated by a recent study testing *Arabidopsis thaliana* growth in lunar soil samples from Apollo missions (Paul et al., 2022). Cell culture studies are also a popular model for space biology as they lend themselves to automation and compact payloads. For example, tissue-on-a-chip models are emerging as a practical choice for studying the impact of spaceflight on specific populations of human cells (Parafati et al., 2023). Finally, in addition to being collected from humans, microbial samples can be swabbed from the built environment within spacecraft, such as to

monitor the spread of potentially harmful microorganisms (Avila-Herrera et al., 2020; Stahl-Rommel et al., 2021).

1.3.1.3 Human Analogue & Model Organism Analogue

Analogue studies are investigations that simulate some aspects of the spaceflight environment and thus tend to serve as a cheaper, less competitive, and more controlled methodology for conducting space biology studies or even testing experiments in preparation for spaceflight deployment (Loehr et al., 2011). Crucially, the ability to generalise findings from analogue studies to spaceflight is limited as simulated stressors may hold varying degrees of fidelity with genuine spaceflight stressors, the analogue study may introduce new stressors, and the removal of individual stressors from the complete set experienced during spaceflight may overlook important interactions between stressors (Robson and Cappelletti, 2022). Nonetheless, analogue studies can be an excellent tool for dissecting the effects of spaceflight stressors on biological systems and for developing countermeasures.

In human subjects research, analogues to spaceflight include immobilisation studies like head down tilt bed-rest experiments which can mimic the mechanical unloading and the cephalic fluid shift of microgravity (Jost, 2008; McGregor et al., 2021), trips to remote research stations which mimic the isolation and confinement of spaceflight (Van Ombergen et al., 2021), and longitudinal studies of workers in hazardous environments which can mimic stressors like prolonged radiation exposure (Boscolo and Durante, 2022). Sometimes these analogue environments are even included within astronaut training, such as caving to assess and develop behavioural skills (Sauro et al., 2021). Aside from isolating the effect of specific spaceflight-associated stressors, a key advantage of analogue studies is the opportunity for larger sample sizes and more diverse cohorts compared to spaceflight studies.

Additionally, for human subjects research, terrestrial analogues allow for easier matching of control groups and greater opportunity for invasive sampling and interventions that may affect performance, such as difference exercise and dietary regimes (Arc-Chagnaud et al., 2020; Loehr et al., 2011). Ethical approval for human subject studies on Earth is a rigorous process, but policies are well-established and procedures such as preserving privacy through data pseudo-anonymisation are typically straightforward when compared to the additional considerations entailed by working with public figures like astronauts (Antonsen and Reed, 2019; Reed and Antonsen, 2018).

In recent decades, space agencies have conducted several human analogue campaigns, with the European Space Agency (ESA) including immobilisation studies as a particular programmatic focus (Deane et al., 2022). While these campaigns often centre on unique objectives and interventions - such as to validate a new analogue methodology, or assess specific nutritional, artificial gravity, or exercise interventions (Arc-Chagnaud et al., 2020; Clément, Rittweger, et al., 2022; Robin et al., 2023) - they have increasingly undergone standardisation of sampling protocols to support comparability between studies. This standardisation has resulted in the emergence of a set of core measurements to be collected across ESA human analogue studies, and the emergence of a set of International Standard Measures (ISM) for bed rest studies analogous to spaceflight, including NASA-funded bed rest studies (Clément, Crucian, et al., 2022). The ISM measures and ESA core measures include phenotypic data, such as DEXA scan body composition, bone mineral density, and exercise performance including vertical jump and treadmill tests (Clément, Crucian, et al., 2022; Clément, Rittweger, et al., 2022; Sundblad et al., 2016).

Analogue studies also extend to investigations with model organisms. For instance hindlimb unloading is used to mimic microgravity in rodents, and

can even be paired with radiation beam exposure to mimic cosmic radiation (Cekanaviciute et al., 2018). Additionally, model organisms can be studied within drop towers, parabolic flights, centrifuges, clinostats, and random positioning machines (RPMs) to investigate biological impact across gravity spectrums, including near-microgravity and hypergravity conditions (Robson and Cappelletti, 2022; Rutter et al., 2020; Willis et al., 2020).

1.3.2 The New Era of Open-Science for Space Biology

In 2011, following the explosion of omics research on Earth and early demonstrations within space biology investigations (Däpp et al., 2004; Selch et al., 2008; Ward et al., 2006; Wilson et al., 2007), an expert panel from the US National Academy encouraged NASA to improve informatics capabilities as a means to support space biology researchers interested in omics and associated system biology approaches. To address this need, they suggested the creation of a specialised database for housing space omics data, including data derived from model organism and analogue studies (Scott et al., 2020). This suggestion quickly materialised into funding for the NASA GeneLab project, which was launched in 2015 (Ray et al., 2019). GeneLab has quickly grown into the *de facto* public data repository for space omics, providing access to over 700 space omics datasets, often from international data generators. A considerable amount of work has been performed by data generators and the GeneLab team to curate the datasets and ensure standardisation of metadata between them. Notably, these metadata include fields that are uniquely important for space biology investigations such as whether the dataset is a spaceflight study or an analogue study; in the case of the latter, additional fields denote which spaceflight-associated stressors were modelled. Many of the datasets have also been run through standardised data processing pipelines by the NASA GeneLab team (Overbey et al., 2021). These curation and standardisation efforts aim to improve comparability across datasets as part of an overarching goal to maximise the utility of rare space biology investigations with often limited sample sizes (Ray et al., 2019). As shown in **Figure 1.4A**, while there are now many omic types within the GeneLab data repository, transcriptomic assays have emerged as the most popular thus far, so much of the standardisation effort has been focussed on transcriptomics, and particularly RNA-Seq due to its increasing traction within the field (**Figure**

1.4B and 1.4C) (Deane et al., 2022; Lai Polo et al., 2020; Overbey et al., 2021, 2022; Rutter et al., 2020).

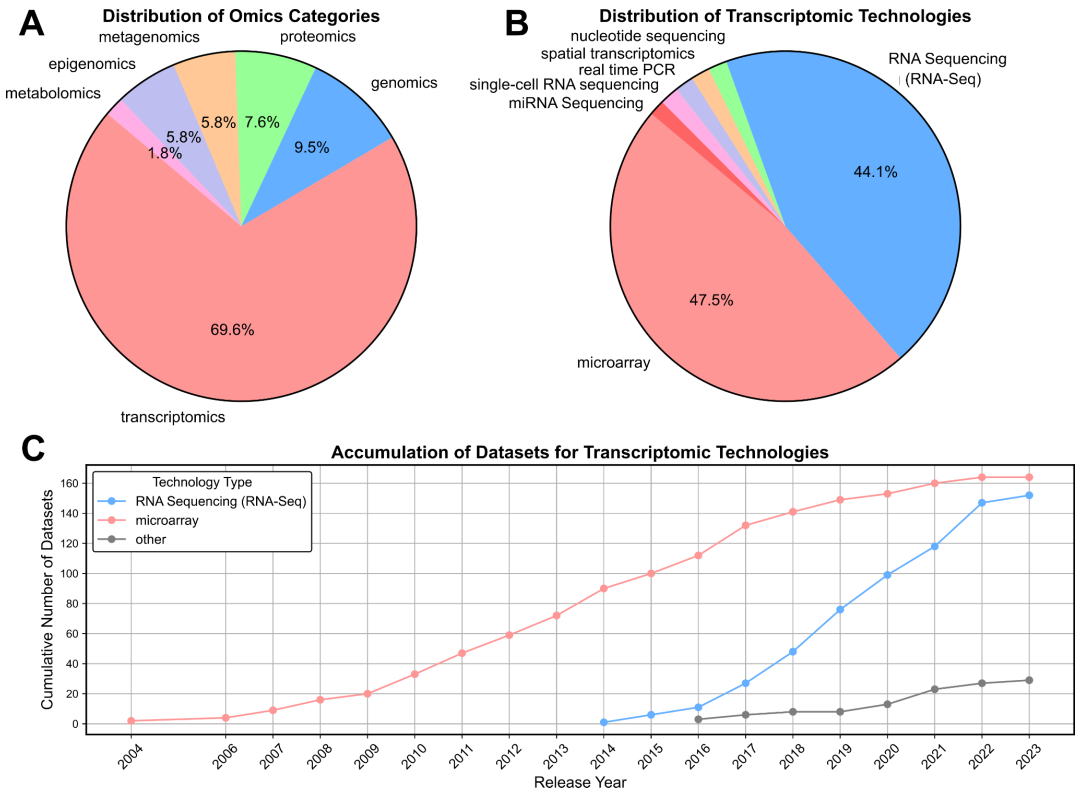

Figure 1.4. Omic data trends within the NASA Open Science Data Repository (OSDR). Generated from a dataset mined from OSDR by Dr. Pedro Madrigal (Gupta et al., 2024), with data up to 4th October 2023. Release year can predate OSDR due to transfers from other repositories. **A)** Distribution of omic categories across OSDR datasets. **B)** Distribution of transcriptomic technologies across OSDR transcriptomic datasets. **C)** Adoption of microarray and RNA-Seq transcriptomics datasets in OSDR over time.

Alongside the formation of the NASA GeneLab data repository, analysis working groups (AWGs) have been established. The AWGs aim to build sub-communities for researchers with common interests within space biology to collaborate and increase scientific return of data housed in the GeneLab database. The Animal AWG has focussed on how best to obtain space biology insights using vertebrate and invertebrate model organism

studies (Cahill et al., 2021; Paul et al., 2021), while the Plant AWG has focussed on space biology investigations with plants (Barker et al., 2023), and the Microbe AWG has focussed on studying the response of microbial populations to spaceflight stressors (Brereton et al., 2021). Furthermore, the multi-omics and artificial intelligence/machine learning (AI/ML) AWGs have respectively focussed on leading investigations into the multi-omic integration of space biology data (Afshinnekoo et al., 2020; da Silveira et al., 2020), and AI/ML applications within space biology (Sanders et al., 2023; RT Scott et al., 2023). Further community groups have also formed, such as the International Standards for Space Omics Processing (ISSOP) consortium for investigating how best to standardise space biology investigations across institutions internationally (Rutter et al., 2020), and the ESA Space Omics Topical team for investigating how best to organise European efforts for space omics (Herranz et al., 2022; Madrigal et al., 2020). While each group focuses on specific challenges and opportunities, there is significant overlap and cross-talk between all of these sub-groups due to the shared goals of the space biology field and the aforementioned challenges of space biology experiment platforms.

In a monumental effort to move towards the establishment of stronger statistical associations, NASA GeneLab has recently been folded into a wider concept termed the NASA Open Science Data Repository (OSDR). Far from just a name change, this rebrand has involved the integration of NASA GeneLab with other data sources, including the NASA Ames Life Science Data Archive (ALSDA) (Scott et al., 2020). The ALSDA contains phenotypic space biology data, such as videos and images from spaceflight experiments, and other standardised assays such as calcium uptake and bone density measurements. Much of the effort in curating the data within ALSDA and ensuring compatibility with data in NASA GeneLab has been led by the ALSDA AWG (Scott et al., 2020). Additionally, through ALSDA, the OSDR houses environmental measures, including radiation exposure data

for many studies; the NASA Biological Institutional Scientific Collection (NBISC) biospecimen repository has now also been included under the OSDR umbrella. Crucially for omics, biobanked samples - such as the tissues collected during spaceflight studies and stored in NBISC - can be used to generate new omics retrospectively.

1.4. Research Gaps

The field of space biology has made significant progress in recent years when it comes to harnessing data from model organisms and analogue studies to elucidate mechanisms behind spaceflight adaptation and to inform new space medicine countermeasures for alleviating health concerns during current mission formats. However, space biology is a relatively niche field which faces significant barriers to conducting controlled experiments, such as high costs, mass constraints, competitive availability for experiment slots, and limited hands-on astronaut time. Additionally, the field is confronted with emergent challenges, including the need to understand the effects of spaceflight on an increasingly diverse sparefaring population, largely during commercial missions, and understanding the impact of new mission formats and spacecraft, such as for long-duration missions beyond LEO. Due to these challenges, many gaps remain in the understanding of how life adapts to space, including the pathology of health risks associated with spaceflight and how best to manage them. To address these unknowns, new strategies for generating data with maximal utility (i.e. advanced data generation), and new approaches to maximise the utility of existing data are warranted (i.e. advanced data analysis).

1.4.1. Optimised Generation of Human Space Omics Data Internationally

The NASA Twins study in 2015/2016 was a groundbreaking initiative, demonstrating the feasibility of collecting multi-omics from an astronaut and a control subject throughout the entire course of a long-duration spaceflight

mission (Garrett-Bakelman et al., 2019). Data from the NASA Twins study has generated new insights into how the human body adapts to long-duration spaceflight and the health risks associated with this (Garrett-Bakelman et al., 2019). Subsequently, additional international omic studies have recently been performed on humans venturing into space, such as during the SpaceX Inspiration4 mission (Jones et al., 2024), and the JAXA CFE mission. However, there is yet to be a similar study performed at a European level, such as with ESA astronauts.

Additionally, while these standalone studies generate useful data for the scientific community, the datasets are still anomalous, only representing a handful of subjects. Omics is yet to be integrated into a routine programme for spacefarers, akin to NASA's standard measures from ISS astronauts and ESA's core measures during bed rest studies. Statements have been recently made that NASA will integrate omics into the standard measures programme in future, and the recent SOMA initiative from Christopher Mason's lab at Weill Cornell Medicine (https://www.masonlab.net/) is taking steps towards collecting standardised omics during commercial missions (Overbey et al., 2024). However, it remains unclear how or whether ESA and other space agencies will integrate omics into a routine programme.

Adopting routine omics would help to generate new datasets for understanding how the human body adapts to spaceflight, and how best to mitigate health issues. With the limited number of humans entering space, collecting omics routinely alongside other standard measures will help to establish a more substantial dataset, enabling robust conclusions. Additionally, collecting omics from a more diverse population of individuals via commercial missions, and during new mission formats, such the Artemis programme, will help to shed light on the role of individual differences and environmental differences on the human body in space. While collecting human space omics will be important for fundamental space biology

research, it could also help to inform risk modelling and occupational healthcare in space, such as via personalised medicine approaches (Schmidt and Goodwin, 2013). In alignment with calls to improve European capabilities in space (High-Level Advisory Group on Human and Robotic Space Exploration for Europe, 2023), it is of paramount importance for ESA to collect omics data from their own astronauts, so that this data can be used for these individualised approaches; additionally, due to stringent data protection laws in Europe, such as the GDPR, there is also a need to develop clear policy around the collection and usage of this data. This need also presents an opportunity to take a leading role on the international stage when it comes to policy development pertaining to data governance and ethical conduct for human omics studies in space (Antonsen and Reed, 2019; Rahimzadeh et al., 2023). ESA has historically focussed on flying non-mammalian model organisms and performing human analogue studies on Earth for space biology research; thus, there is a salient opportunity for ESA to incorporate routine omics for Europeans participating in space missions and align these with their human analogue studies for comparison. This would maximise the use of astronauts as existing test subjects, and reduce the need to develop and fly model organisms and housing. To this point, humans will continue to fly to space after the decommissioning of the ISS, but the future of model organism studies in space is uncertain, particularly for mammalian model organisms which require larger habitats. While an ESA initiative in human space omics is a clear gap warranting further investigation, the importance of international collaboration and standardisation cannot be stressed enough, as comparability across international datasets will yield greater insights through improved statistical power.

1.4.2. Data Science Approaches for Maximising the Utility of Space Omics

The launch of NASA GeneLab (Ray et al., 2019) in 2015 was a pivotal action for the field of space biology. The curation and storage of omics data generated from model organism studies, and the establishment of international AWGs has fostered an ecosystem conducive of international data sharing and collaboration. This has led to the extraction of a wealth of new insights, often from data that had been archived after initial analysis and publication.

Given the scarcity of spacelight datasets and the monetary and time expense of generating this data, maximising existing data is an invaluable initiative. A multi-omics analysis published in 2020 identifying mitochondrial dysregulation as a central hub for spaceflight adaptation across species is perhaps the most poignant showcase of these data sharing efforts to date (da Silveira et al., 2020). However, there remain challenges in the analysis and interpretation of these existing omic datasets, such as how best to compare changes between species and between missions. This includes investigating how best to deal with the significant metadata differences (e.g. duration, strain, age) between different missions, when performing cross-mission meta-analyses.

The recent conglomeration of phenotypic data via NASA ALSDA and space omics via NASA GeneLab into the NASA OSDR presents a shift in the goals for the space biology field (Scott et al., 2020). By housing omic and phenotypic data together, so that omic data and associated phenotypic data can be viewed and downloaded concurrently, NASA OSDR has set the stage for analyses that integrate omic and phenotypic data. As such, there is now a clear gap in how best to integrate omic and phenotypic data to move beyond weak correlational relationships and towards stronger correlational, or even causal, relationships between omic changes and phenotypic

changes. This will include addressing how best to deal with the low sample counts often inherent to spaceflight datasets, and how best to deal with the high-dimensionality associated with omics data. Additionally, there is a challenge in maximising interpretability of any approaches used for analysis. Results from techniques such as machine learning approaches can sometimes be complex and difficult for humans to trace and understand, yet interpretability is important in the biological context where clear mechanistic insights are crucial for the development of interventions.

Chapter II: General Overview and Methods

Two distinct gaps were identified during the initial literature review; thus, this thesis is divided into two parts, with Chapters III and IV discussing the opportunity for human space omics. Chapter III discusses the opportunity from the European perspective while Chapter IV considers a UK national focus. Chapters V and VI are case studies reporting results from the analysis of space biology experiments. These chapters arise from two primary goals within the thesis context; firstly, to utilise omics to expand fundamental knowledge on the effects of spaceflight-associated stressors on biological systems, and secondly, to gain insights into the challenges and opportunities for maximising the utility of space biology datasets via data science approaches. Chapter VII is a general discussion which summarises contributions from the chapters and provides perspectives on future directions for improving data generation capabilities and data science capabilities for space biology. The following chapter, Chapter II, outlines the general approach for how these two research gaps were explored.

The thesis author's active involvement as a NottsSpace team member for biomedical payload engineering projects at the University of Nottingham has also provided insights within the thesis context. While these projects are not directly in pursuit of either of the two research gaps, they have helped to inform discussion within Chapter I and VII, particularly related to the challenges of performing model organism spaceflight work. The thesis author was the software lead for VITA (Visualising In-space Tx-TI Astropharmaceuticals), a payload which was selected for ESA's Orbit Your Thesis program, with a goal of trialling on-orbit synthesis of proteins for pharmaceuticals (Green et al., 2022), and WormSail, a small satellite containing multiple payloads, including a habitat for studying *C. elegans* (Robson et al., 2021). The thesis author's involvement as a collaborator in

COVID-19 International Research Team (COV-IRT) projects is outside the scope of this thesis (Guarnieri et al., 2023).

2.1. Optimised Generation of Human Space Omics Data Internationally

Following the identification of a lack of a European initiative to generate human space omics data during an initial literature review, a community-driven approach was adopted to investigate the gap.

Firstly, the ethical and legal aspects of the astronaut personal data were explored, to identify challenges. This investigation was motivated by the lack of human space omics datasets in the NASA GeneLab data repository. This work was performed by joining and engaging with the International Standards for Space Omics Processing (ISSOP) consortium in the early stages of the consortium's formation in 2020. The goal of which was to engage with the international community of expert space scientists to gain perspectives on human biology research in space, and to review approaches that have worked in the terrestrial context internationally. The plan was to create a publication summarising the current state of space omics internationally and future directions in standardisation within the field. As part of this paper, a discussion on human space omics was intended to be included. Due to the scale of the paper, it was decided that the human data component would be cut out, with the manuscript published with a focus on model organism work (Rutter et al., 2020). Several authors from this paper, including the thesis author, continued to work on the human data component as a subgroup. A literature search was conducted by the thesis author to identify work related to the ethics of human omics studies in spaceflight, and this was used to structure the section initially. Following identification of key challenges, such as privacy, discrimination, and incidental findings, the literature was searched to identify differences in laws related to these challenges internationally, as well as discussions of the topics within the general omics literature. This led to the discovery of the

Global Alliance for Genomic Health (GA4GH) as an important organisation driving international standardisation of omics. The ISSOP subgroup, including the thesis author, met with members of the GA4GH and then remotely attended their 2021 (28th-29th September 2021), which helped to shape ideas for the section. Following this, the section was written up by the thesis author, but given the scale of other sections within the paper, the section was written in a rather condensed format and included within the publication (Rutter et al., 2024).

During this ethical and legal literature search, the scope of the GDPR in this context was investigated. The thesis author joined the ESA-funded Space Omics Topical Team to pursue the targeted research gap by focussing on the opportunity for a European human space omics initiative. These regular online discussions with European space science experts and subsequent research performed by the thesis author and collaborators led to the development of a paper focused on the European opportunity for human space omics (Cope et al., 2022). This paper makes up Chapter III of this thesis. An extended abstract for this paper was also requested by the Committee on Space Research (COSPAR) Space Research Today publication, and was included within Volume 215 (*Space Research Today*, 2022). During two in-person meetings of the Space Omics TT (12th-13th December 2022, Centro de Investigaciones Biológicas (CSIC), Madrid, Spain; 18th April 2023, European Space Research and Technology Centre (ESTEC), Noordwijk, Netherlands), the thesis author led discussions based on the paper to gain input from the topical team. These discussions fed into the formation of the final topical team recommendations, with a section on human research recommendations written by the thesis author. These final recommendations were submitted directly in a report to ESA and published as a paper (Manzano et al., 2023). Meanwhile, after receiving a request via the UK Space Agency for inputs from the space science and exploration community, the author worked in a smaller subgroup of UK researchers to

consider the gap from the UK national perspective and to present this work at the British Interplanetary Society (BIS) “Future of UK Human Spaceflight Policy Symposium”. Due to the thesis author being in the US at the time of the symposium, the presentation was given by Dr. Craig Willis (26th October 2022, British Interplanetary Society, London, UK). The paper arising from this work was also published in the society journal and sent to the UK Space Agency via the British Interplanetary Society, although it is unclear whether it made it into parliamentary record. This paper makes up Chapter IV of this thesis (Cope et al., 2023).

2.2. Data Science Approaches for Maximising the Utility of Space Omics

Similarly, from an initial literature review, it was clear that further work was needed to fully capitalise on archived space omics data. The strategy adopted for exploring this gap was to identify and collaborate on analysis projects, so that data science approaches could be investigated in a practical and applied setting. For this thesis, two particular case studies have been focussed on where the thesis author played a leading role in the analysis.

Firstly, the thesis author joined a collaboration between University of Nottingham researchers and Texas Tech University researchers, to assist with the data analysis from a spaceflight model organism experiment. The experiment was micro-16 which was launched to the ISS on February 20th, 2021 with the aim of investigating spaceflight-induced losses in muscle performance through the use of *C. elegans* worms housed within a microfluidic device (Soni et al., 2022). The scope of this project changed significantly throughout the thesis. Initially the intention was for the thesis author to complete the phenotypic analysis (via computer vision techniques), the transcriptomic analysis, and work on connecting these via omic-phenotypic analysis. Due to modifications to the engineering and scientific specifications of the payload, changes in the roles of other

personnel, and challenges in data sharing, the analysis scope for the thesis author was reduced to parts of the transcriptomic analysis only. The analysis performed by the author is included within a publication which is incorporated as Chapter V of this thesis (Soni et al., 2023).

Secondly, as the original scope of the thesis was focussed specifically on loss of muscle performance in spaceflight, it included a planned transcriptomic analysis and omic-phenotypic analysis of a human immobilisation study at the Royal Derby Hospital, to study muscle atrophy in an environment analogous to spaceflight microgravity (Bass et al., 2021). Due to the 2020 Covid-19 pandemic, this study was paused indefinitely, and timelines for the availability of omics data were unclear. Since the data from micro-16 data was also unavailable at this time, the thesis author looked for alternate data sources to work with. Tangentially, the thesis author did not have any involvement with the study once it was resumed, but did submit two unsuccessful funding applications in 2021 with the study's PI to try and acquire funding for a portable ultrasound scanner to perform an additional project within the scope of this thesis for developing computer vision models to automatically extract muscle architectural measures in near real-time, based on foundational work from other institutions (Cronin et al., 2020). Through the thesis author's involvement in NASA open science analysis working groups (AWGs), particularly the NASA GeneLab Animal Analysis Working Group, it became apparent that there were a number of unanalysed transcriptomic skin datasets available from mice flown across multiple missions. As such it was decided that the thesis author would lead a project analysing these skin datasets. Part of this project was completed during a remote internship with NASA GeneLab in 2021, and the scope of this project continued to increase via opportunities to include additional data and methods. Ultimately the paper ended up analysing model organism spaceflight, and human spaceflight data across a variety of missions. This paper was published and is incorporated as Chapter VI of this thesis (Henry

Cope et al., 2024). Work from this paper was also presented at the in-person Royal Aeronautical Society Aerospace Medicine Symposium (8th December 2021, Royal Aeronautical Society, London, UK) and online 2022 AWG annual symposium (30th June 2022) by the thesis author, and the online 2023 AWG annual symposium (16th March 2023) by Mr. Jonas Elsborg.

Additional analysis projects have taken place within the context of the thesis, including collaborations on transcriptomics analysis projects within the NASA open science AWGs (Cahill et al., 2021; Mathyk et al., 2024). Additionally, the thesis author led a project analysing transcriptomic data and root architecture data, including omic-phenotypic analysis with machine learning methods, using archived plant data from the CARA spaceflight mission (Paul et al., 2017). This bulk of this work took place during an in-person internship with the NASA ALSDA in 2022, but the project was not completed and published due primarily to unsatisfactory results from the analysis approaches trialled. While these additional projects are not focussed on as chapters, they are used to help inform the discussion in Chapter VII.

Chapter III: Astronaut Omics - The European Opportunity

The contents from this chapter are published in (Cope et al., 2022):

***Cope, H.**, Willis, C.R., MacKay, M.J., Rutter, L.A., Toh, L.S., Williams, P.M., Herranz, R., Borg, J., Bezdan, D., Giacomello, S. and Muratani, M., et al., 2022. Routine omics collection is a golden opportunity for European human research in space and analog environments. Patterns, 3(10).*

All work is that of the author of this thesis, other than "3.3.6. Sampling Methods for Spaceflight Omics Collection", which was co-written with Dr. Craig Willis. Dr. Mathew MacKay assisted with the figures by providing initial code for mining the data for **Figure 3.2**, and providing an astronaut icon for **Figure 3.4**.

This work is incorporated into this thesis in accordance with the terms of the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0). Modifications have been made by the copyright holder for its inclusion within this thesis, which include adjustments to figure, table, and section numbering, as well as the conversion of the text to British English.

3.1. Summary

Widespread generation and analysis of omics data has revolutionised molecular medicine on Earth, yet its power to yield new mechanistic insights and improve occupational health during spaceflight is still to be fully realised in humans. Nevertheless, rapid technological advancements and ever-regular spaceflight programmes mean that longitudinal, standardised, and cost-effective collection of human space omics data is firmly within reach. Within this chapter the practicality and scientific return of different sampling methods and omic types are considered in the context of human spaceflight. Additionally, ethical and legal considerations pertinent to omics data derived from European astronauts and spaceflight participants (SFPs) are appraised. Ultimately, a routine omics collection programme in spaceflight and analogue environments presents a golden opportunity (**Figure 3.1**). Unlocking this bright future of artificial intelligence (AI)-driven analyses and personalised medicine approaches will require further investigation into best practices, including policy design and standardisation of omics data, metadata, and sampling methods.

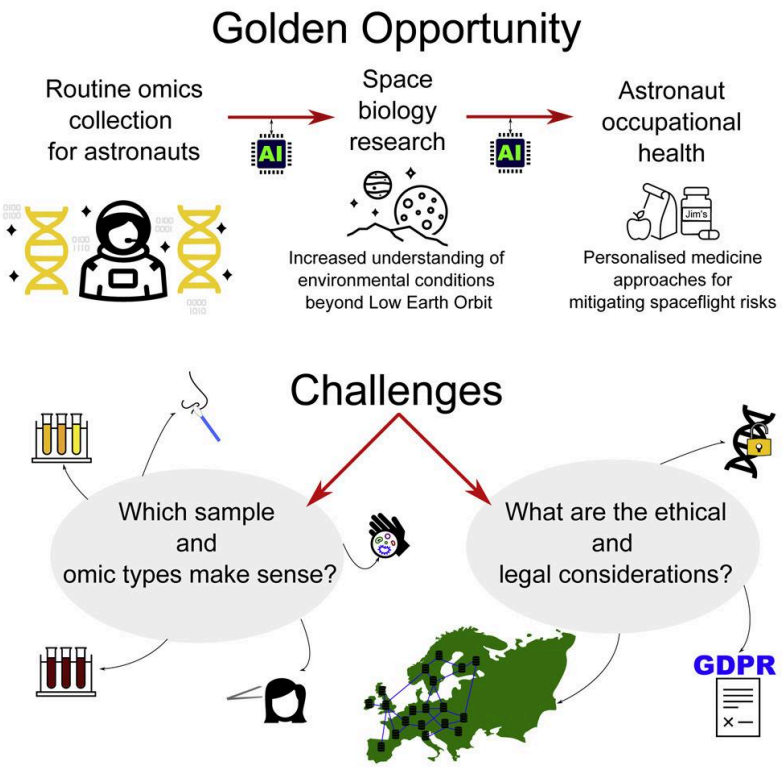

Figure 3.1. Graphical summary of the European opportunity and considerations for human space omics.

3.2. Introduction

In 1978 Vladimír Remek of the Czech Republic became the first human from the European Union to cross the boundary into space. Since then, astronauts from an array of European countries have followed suit (**Figure 3.2A**), with many of these trips occurring on board the International Space Station (ISS) since its first occupation at the turn of the millennium. Despite a steady increase in the number of European astronauts venturing beyond Earth (**Figure 3.2B**), astronauts remain a relatively scarce population. Each astronaut therefore presents a rare and unique opportunity to directly study the impact of spaceflight on the human body, mind and spirit. Additionally, with the emergence of commercial spaceflight, and European citizens already flying to space via these programmes (e.g. Axiom Mission 1, Blue

Origin NS-16), there is an emerging opportunity to expand the population of European human subjects that can be studied in space (Griko et al., 2022). The spaceflight environment is characterised by simultaneous exposure to stressors including cosmic radiation, altered gravity, isolation, and confinement within a closed environment (Afshinnekoo et al., 2020). Additionally, demand on the human body changes dynamically throughout different time points within a typical spaceflight mission, including pre-flight training, launch, extravehicular activity (EVA), landing, and post-flight rehabilitation.

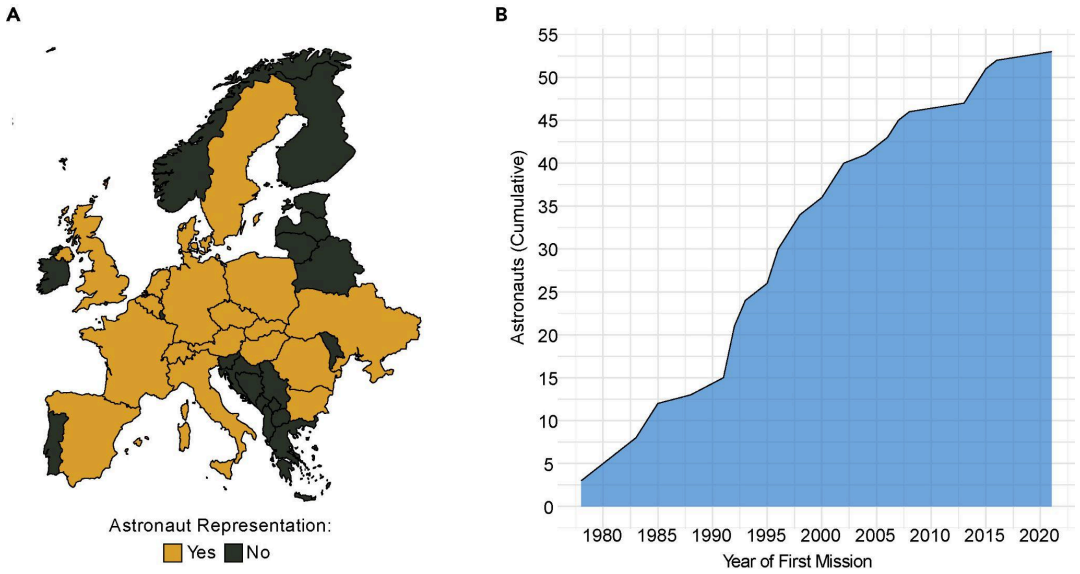

Figure 3.2. European astronaut population. The distribution and population of European astronauts who have crossed the Kármán line (as of 9/20/21). **A)** Geographically highlights the range of nationalities of European astronauts. **B)** Shows the cumulative population of European astronauts over time, based on astronauts' first mission into outer space.

To date, studies of astronauts have shown that long-duration spaceflight induces multisystem physiological deconditioning, such as degradation of muscle and bone (Gabel et al., 2022; Rittweger et al., 2018), and detrimental changes to the eye known as spaceflight associated neuro-ocular syndrome (SANS) (Lee et al., 2017). However, the precise mechanisms behind these responses are yet to be fully understood, largely due to the pace and practicalities of conducting comprehensive molecular studies in astronauts.

Biological changes and associated medical risks during spaceflight missions, including in the commercial spaceflight setting, have recently been reviewed elsewhere (Afshinnekoo et al., 2020; Griko et al., 2022; Stepanek et al., 2019). The European Space Agency (ESA) and other space agencies have historically taken two complementary approaches to accelerate the understanding of molecular causes of biological changes in astronauts: translational research and space analogue research. Translational research utilises model organisms, including rodents, for molecular studies in space (Rutter et al., 2020) with the aim of understanding and countering detrimental physiological changes in astronauts. Space analogue research utilises facilities on Earth which enable one to study individual aspects of the spaceflight environment without having to conduct studies in space. In the context of human research, this can include studying inactivity via bedrest studies and isolation via use of remote research stations, such as Concordia (Van Ombergen et al., 2021).

At the 2022 European Space Summit, the European Association of Space Explorers, called for the development of European vehicles for transporting humans into space. Irrespective of whether or not this becomes a near to medium term reality, not only will ESA astronauts continue to regularly fly to the ISS, they will also fly in future exploration missions with international partners. One such example is the Artemis 2 mission to the Moon, which could occur as early as 2025. The new generation of spacecraft used for these missions are likely to have highly limited cargo capacity, therefore countermeasures such as the Advanced Resistive Exercise Device (ARED) used on board the ISS to suppress muscle and bone loss, may not be suitable (Stokes et al., 2016). Additionally, during these exploration missions, astronauts will venture beyond Low Earth Orbit (LEO) for the first time since the end of the Apollo program in 1972. Within LEO, the Earth's magnetosphere provides partial protection against radiation; thus, humans venturing beyond this will experience greater exposure to solar particle event

radiation and ionising galactic cosmic radiation (Afshinnekoo et al., 2020; Zeitlin et al., 2013; Zhang et al., 2020). Whilst there is a lack of data on the molecular effects of the beyond LEO environment, an increase in radiation-associated health risks is almost a certainty (Restier-Verlet et al., 2021; Simonsen et al., 2020; Walsh et al., 2021). Further considerations of future missions beyond LEO include health risks associated with unprecedentedly long-duration missions, lunar dust (Linnarsson et al., 2012), and EVA (Afshinnekoo et al., 2020). To the latter point, in the case of a Mars mission, the crew would likely need to safely perform EVA tasks in partial gravity (0.38g) following long-duration spaceflight, without the extensive rehabilitation support available on Earth (Gaffney et al., 2020; Petersen et al., 2017).

With these new spacecraft and mission plans, it will be crucial to assess compatible countermeasures, while also gaining greater insight into the new challenges to human health posed by environments where biological and molecular adaptations remain uncharted. A key challenge for translational research from a European standpoint is that ESA does not currently have rodent research facilities in space, and it seems unlikely that such facilities will be rapidly developed for future exploration vehicles. Thus, new approaches that do not depend upon rodent research facilities should be considered by ESA. In this context, a salient opportunity for space analogue research is to harmonise measures and endpoints between research in analogue environments and in space, increasing fidelity. For example, the efficacy of new countermeasures could be evaluated in both environments, and the efficacy of rehabilitation following long-duration spaceflight (Petersen et al., 2017) could be studied and improved by testing novel rehabilitation techniques following long-duration bedrest. For these reasons, it could prove cost-effective for ESA to pursue an approach that focuses on improved capabilities for research on astronauts themselves, and with

increased harmonisation of data and endpoint measures obtained in space and analogue environments.

On Earth, medicine has undergone a revolution through the widespread generation and analysis of multiple omic data types, including genomics, transcriptomics, and proteomics. The strength of omic-based, data-driven approaches is their power to discover unanticipated effects, elucidate molecular mechanisms, and generate novel hypotheses to guide follow-up targeted studies (Hasin et al., 2017; Schmidt and Goodwin, 2013). In scientific studies on Earth, omic-derived molecular changes are regularly correlated with performance metrics to uncover molecular drivers of physiological change (Willis et al., 2021). Omics have also become a valuable tool for clinical practice. For example, the National Health Service (NHS) in the UK: (1) Routinely offers all newborn babies blood spot test screening for several genomic-based disorders, including sickle cell disease, cystic fibrosis, congenital hypothyroidism, and six inherited metabolic diseases (van Campen et al., 2019) (2) Recently launched a pilot project with Genomics plc to evaluate incorporating polygenic risk factor data into the clinical risk assessment for cardiovascular disease (Fuat et al., 2024) (3) Recently approved use of personalised medications for treating cystic fibrosis based on the nature of the mutation(s) present in the patient's CFTR encoding gene (Granger et al., 2022). Furthermore, the COVID-19 pandemic has highlighted how omics can rapidly be employed in combination with clinical data to determine risk of severe disease course, identify potential underlying mechanisms of susceptibility, and identify potential personalised treatments (Kousathanas et al., 2022).

Whilst omics have become a significant part of scientific research and standard clinical practice on Earth, for humans in space, omics are a relatively new and untapped technology. Between March 2015 and March 2016, NASA conducted the seminal Twins study. The study monitored an

astronaut during a one-year mission in space, while the astronaut's identical twin remained on Earth as a control. Both subjects were comprehensively profiled using multi-omic measures across several time points pre-flight, in-flight, and post-flight. Spaceflight-induced changes were detected across multiple systems, with most of these alterations returning to baseline levels post-flight (Garrett-Bakelman et al., 2019). The longitudinal profiling of the NASA Twins study, and similar profiling on the recent short-duration SpaceX Inspiration4 commercial mission, demonstrates that multi-omics profiling of astronauts is feasible. Accordingly, NASA has recently announced that they have elected to add omics to the Standard Measures programme for the ISS, which ESA astronauts can opt into. NASA's Standard Measures programme also routinely collects select physiological measures and select biomarkers from blood and urine; commercial spaceflight missions are similarly exploring these possibilities. For example, the Translational Research Institute for Space Health (TRISH), a partner of NASA's Human Research Program, recently created a database for commercial spaceflight health data, including crew omics data from the Inspiration4 mission (Urquieta et al., 2022). It is therefore important that ESA also explores these new opportunities, leveraging existing expertise in European spaceflight omics (Deane et al., 2022) and facilities for human subjects research. To that end, the key question is how best to develop a routine longitudinal omics programme for European astronauts and/or European spaceflight participants (SFPs).

A routine omics collection programme could be viewed as an initiative to develop a high-value data resource which can be analysed with ever-evolving big data approaches, thereby improving ESA's Artificial Intelligence (AI) readiness. Access to sensitive data collected during the programme would need to be controlled so that scientific research is supported, while protecting the rights of the subjects. Such a programme could help to elucidate key molecular drivers behind spaceflight health

issues which could lead to improved risk quantification and approaches to spaceflight occupational health. One example of how this could be used is for the development of evidence-based recommendations on training and medical standards for spaceflight; these standards will become especially important as the spacefaring population increases in heterogeneity and begins to include humans with pre-existing conditions (Griko et al., 2022; Stepanek et al., 2019). Furthermore, a routine omics programme could be used as a basis for personalised medicine approaches. As an example, researchers have previously found that genetic variations in one-carbon metabolism genes may contribute to susceptibility to SANS in astronauts (Zwart et al., 2016), and other genomic studies have shown individual variability associated with spaceflight-relevant factors, including bone mineral density and radiation sensitivity (Griko et al., 2022). Pre-flight genomic screening of astronauts therefore holds potential in developing personalised medicine approaches to risk assessment and countermeasures, including pharmacogenomics and nutrigenomics (Lulli et al., 2018; Pavez et al., 2021; Schmidt and Goodwin, 2013). Pharmacogenomics aims to optimise medication therapy, with respect to the patient's genomic data, to ensure maximum efficacy with minimal adverse effects. For example, a study of medication on board the ISS found that the metabolism of approximately a third of the drugs in the ISS repository was affected by polymorphic metabolising enzymes (Stingl et al., 2015). Through further data collection, cargo could potentially be optimised to only include the best available medication for the individual crewmembers, reducing costs and adverse drug reaction/therapy failure risk (Sawyers et al., 2022). Furthermore, omics could prove valuable for longitudinal health monitoring during spaceflight. For example, TRISH recently suggested that omics could be used alongside other datasets obtained via point-of-care (POC) devices to diagnose approximately one third of NASA’s list of medical conditions that are of concern for deep-space missions (Strangman et al., 2019). Finally, discoveries obtained using human

research-derived omics data could benefit health on Earth. For example, the World Health Organization has listed physical inactivity as the 4th leading risk factor for global mortality, estimated to have been responsible for 9% of global premature mortality in 2008 (Lee et al., 2012). Mechanical unloading of the musculoskeletal system due to microgravity during spaceflight, or due to immobilisation or reduced step count in analogue environments, induces similar effects to physical inactivity, such as muscle atrophy and altered glucose handling; thus, discoveries and countermeasures developed for astronauts and/or SFPs could prove useful for terrestrial efforts against physical inactivity.

3.3. Considerations for a European Space Research Routine Omics Programme

Designing an appropriate routine omics collection programme for European astronauts and/or SFPs requires careful consideration of a number of factors. The following section entails a discussion of some of the key considerations including: collecting data with AI in mind, which sampling methods and omics data types should be collected based on practicality and perceived scientific return, and which ethical and legal challenges may need to be addressed.

3.3.1. AI-readiness of Omics Data & Metadata

Collecting the same standardised measures routinely from astronauts and/or SFPs across multiple missions would greatly improve the AI-readiness of the generated datasets, especially when combined with accompanying standardised metadata, including spacecraft environmental measures (e.g. radiation dosimetry) (Rutter et al., 2020; Scott et al., 2020). One important practical consideration here will be digital storage requirements, as omics experiments generate large volumes of data. Similarly, harmonising approaches with other space agencies will be important. In this context

there is an opportunity for ESA to expand collaboration with the Frontier Development Lab, which it currently partners with to assist with AI approaches to Earth Observation, and who is also currently partnered with NASA to assist with AI approaches across the entire Science Mission Directorate. In pursuit of AI-readiness, it will be important to adhere to guidelines, such as well-implemented FAIR (findable, accessible, interoperable, and reusable) principles where possible (Jacobsen et al., 2020; Wilkinson et al., 2016), which could include machine-interpretable file formats such as standardised processed omic data (Hasin et al., 2017; Overbey et al., 2021) alongside raw data and metadata. In order to further enhance the quantity of data, identical measures could be collected from well-designed ground-based analogue studies, to facilitate comparison and to unlock powerful machine learning (ML) techniques based on transfer learning. With the ever-increasing adoption and utility of AI-based approaches for healthcare and life sciences research (Bohr and Memarzadeh, 2020), the importance of designing data collection and curation with AI capabilities in mind cannot be overstated. For example, it may eventually be possible to develop omic-based digital twin approaches to be used in applications such as mission planning and countermeasure design. These approaches could model the physiological impact that mission parameters such as duration and radiation exposure may have on individual astronauts (Björnsson et al., 2019; RT Scott et al., 2023). Whilst there are a few interesting exceptions, at present most AI approaches require large volumes of high-quality and domain-relevant training data (Cohain et al., 2017; Ma et al., 2021; Mostavi et al., 2021). Notably, robust population-based analyses will be particularly challenging to achieve for humans in space, given their rarity. It is therefore important to initiate collection and curation of standardised astronaut and/or SFP omics data and metadata as soon as possible, thereby unlocking a future of powerful AI-based omic analyses for personalised medicine approaches and mechanistic discovery.

3.3.2. Omic Data Types

A key consideration for implementing a routine omics programme will be selecting a panel of omic assays to capture from human subjects. Omics is increasingly being used as shorthand for big data approaches in biology and medicine, with numerous different omic data types and technologies now available at decreasing costs. Different omics data types and technologies offer both overlapping and unique insights into molecular changes within the body (Hasin et al., 2017). Integrating multiple omic types (i.e. “multi-omics”) via systems biology approaches can lead to a more holistic picture of spaceflight-associated molecular patterns (da Silveira et al., 2020). Initial multi-omics studies during the NASA Twins mission and SpaceX Inspiration4 mission have opted for a relatively broad approach to assay selection (Garrett-Bakelman et al., 2019). It makes sense to use findings from these missions to optimise a panel of assays for routine omics. This harmonised approach would also support comparability between data captured via a European routine omics programme, and these initial international studies. To further support comparability, it also seems logical to collect data types that are prevalent in large-scale terrestrial omic initiatives, such as The Cancer Genome Atlas and the European Genome-phenome Archive. The ability to easily compare astronaut data to terrestrial disease cohorts could allow for the identification of similar patterns, which could aid processes such as countermeasure design (e.g. omic-based drug repurposing) (Pulley et al., 2020). While collecting the same measures over a long period of time is desirable, some flexibility is to be expected due to the continued evolution of omic technologies.

Although coverage of all potential omic types and technologies is beyond the scope of this manuscript, there are a few clear candidates worth noting. The most immediately actionable and cost-effective measure in terms of healthcare will likely be pre-flight genomic screening (Stark et al., 2019),

such as whole-genome-sequencing (WGS). For example, a recent proof-of-concept study by the Mayo Clinic found that pre-emptive use of basic pharmacogenomic testing, alongside participants' medication history, enabled pharmacists to offer medication improvement opportunities in 56% of participants (Matey et al., 2022). Some of the medications involved in the study (e.g. ondansetron, metoprolol, aripiprazole, sertraline and phenytoin), have been documented as being part of the ISS medication kit (Stingl et al., 2015). Thus, pre-flight genomic screening could potentially be used to inform medication management and personalised medicine approaches to spaceflight countermeasures (Pavez et al., 2021). Similarly, microbiomic approaches such as studying the genomic makeup of microbial communities via metagenomics are likely to have high clinical relevance for humans in space. Metagenomic analyses of crew samples can be used to study microbiome shifts (Afshinnekoo et al., 2020) and potential health consequences such as viral (e.g. herpes) reactivation and skin rashes (Urbaniak et al., 2020; Voorhies et al., 2019); the addition of environmental samples to study microbial interactions between crew members and the spacecraft environment would also be useful for monitoring crew health (e.g. infectious disease management) (Avila-Herrera et al., 2020; Stahl-Rommel et al., 2021). On board the ISS, at least some coverage of metagenomics would likely fall under NASA's microbial monitoring programme. Other candidate omic data types for studying regulatory molecular changes, discovering biomarkers, and designing interventions for spaceflight-associated health risks include: transcriptomics, epigenomics, proteomics, and metabolomics (Hasin et al., 2017) (**Figure 3.3**).

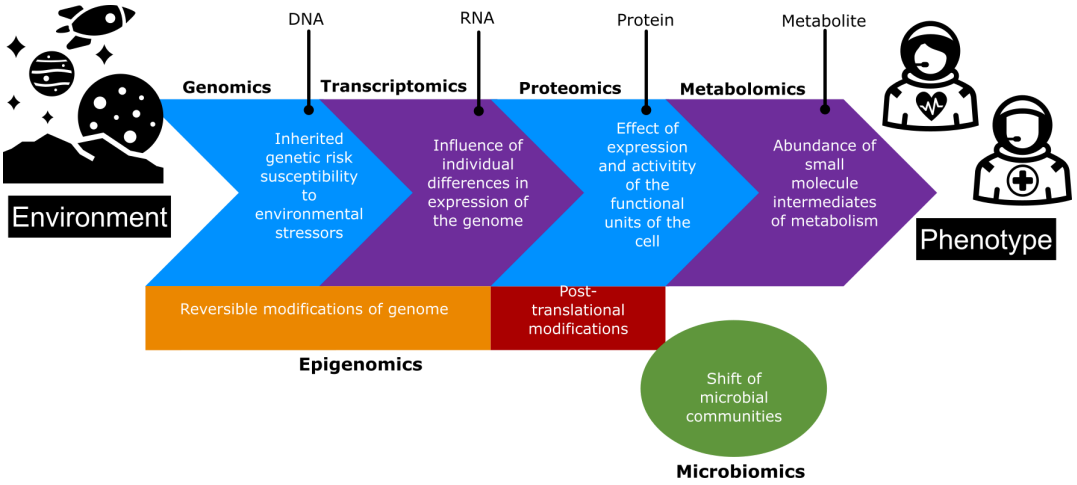

Figure 3.3. Incorporating omics into astronaut studies. Overview of how candidate omic data types for studying and counteracting astronaut health risks interact with one another, the environment, and phenotype. Epigenomics is shown to span several omic data types to demonstrate that environmental exposures and activity occurring at the genomic, transcriptomic, and proteomic levels influence epigenomic status. Adapted from Yu et al (Yu et al., 2022). under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

3.3.3. Sampling Time Points

Due to the dynamic nature of physiological changes throughout the course of a typical spaceflight mission, it is scientifically meaningful to acquire omic measures across an array of time points. At a minimum, a sampling point just before flight and a sampling point taken shortly after the mission would prove useful for quantifying the molecular impact of the spaceflight mission. However, averaging multiple pre-flight measures would be valuable for establishing a robust molecular baseline, and multiple post-flight measures would prove useful for studying recovery and long-term health. To the latter point, long-term (i.e. “lifetime”) post-flight timepoints will be particularly valuable for missions beyond LEO, to study whether the human body fully recovers from spaceflight environmental factors, such as DNA damage response following the increased radiation dosage. To that point, it has

recently been reported, albeit in a single pair of monozygotic twin astronauts, that onset of clonal haematopoiesis may be influenced by spaceflight (Mencia-Trinchant et al., 2021). Clonal haematopoiesis results in lifelong changes in the blood that are associated with increased risk of cardiovascular disease and blood cancers.

The addition of in-flight measures, particularly for long-duration missions, would provide greater temporal resolution, allowing for spaceflight-induced molecular changes to be monitored over time. As a starting point, it may make sense to focus timepoints around expected points of significant physiological change, especially when clinically relevant. Significant physiological changes can occur rapidly, so early in-flight time points could be useful for studying their onset. For example, the first in-flight time point for the NASA Twins study was at 14 days into the flight (Garrett-Bakelman et al., 2019), whereas spaceflight research has indicated that significant physiological changes can occur before this point and thus, earlier time points may be valuable. For instance, a recent study investigating anaemia in astronauts collected blood and exhaled air samples at 5 ± 1 days into the 6-month flights and results already indicated significantly increased breakdown of red blood cells (haemolysis). Markers for increased haemolysis persisted throughout the in-flight timepoints (11 ± 1 days, 64 ± 5 days, 157 ± 31 days), with five of 13 astronauts reaching clinical levels of anaemia when measured 4 ± 1 days after landing (Trudel et al., 2022). The potential implication that the longer the duration in space, the worse certain health risks including anaemia, suggests that in-flight timepoints across the full duration of spaceflight missions may be valuable. For example, other useful in-flight time points may include shortly before the end of the flight, so that this can be compared to the first post-flight time point to measure the impact of landing and initial readaptation, and shortly before and after EVA to investigate the molecular impact of such events.

The sampling schedule also has implications for occupational health. For example, pre-flight omic measures could be used as a baseline for personalised medicine approaches. Measures taken in-flight across different time points could then eventually be compared against the baseline data with ML approaches to predict health outcomes and suggest preventative interventions (Mias et al., 2021; Scott et al., 2021). Insights related to the molecular impact of mission events such as EVA could also be used to inform mission design, or indeed day-to-day task scheduling to manage stress responses.

While frequent sample time points are desirable, it will be important to balance the sampling schedule with regard to practical considerations, including cost, use of crew time, sample storage space, and subject recovery time (dependent on sample invasiveness). Additionally, if in-flight time points are included in the sampling schedule, it will be important to thoroughly train crew members for successful sample collection.

3.3.4. Accompanying Functional Measures

It is beneficial to routinely collect performance measures in tandem with omics, such as self-assessments (e.g. cognition, nutrition, sleep quality), exercise outputs, and physiological measures collected from wearables and POC devices (Roda et al., 2018; Strangman et al., 2019). When combined, omic and functional measures can be used to infer causal molecular-phenotypic links which can then be used to inform and test interventions and discover novel biomarkers; if adopted, this approach would be a major advance in research capabilities, enabling widespread hypothesis-driven molecular mechanistic research to be conducted in astronauts. While outside the scope of this omic-focussed manuscript, the choice over which performance measures to collect must also be considered via an evidence-based approach, balancing scientific and clinical

value with considerations regarding practicality, such as levels of invasiveness. Importantly, NASA's standard measures programme, which ESA astronauts can opt in to, currently captures a curated selection of standardised functional measures from astronauts on the ISS.

3.3.5. Sample Processing

One practical consideration regarding routine omics collection is the logistics of processing the samples. Currently, human biological samples collected on the ISS are typically preserved (e.g. via rapid freezing) and stored before being returned to Earth for sequencing and analysis. The multi-omics profiling conducted for the NASA Twins study shows that this model is feasible for longitudinal human studies on the ISS (Garrett-Bakelman et al., 2019), yet the approach does come with some disadvantages, including sample degradation, delays in receiving results, and dependence on freezer space. Thus, while this approach with well-designed procedures for sample processing (Rutter et al., 2020) is likely to be appropriate for routine omics on the ISS, future commercial space stations, and short-duration missions, it may not be appropriate for all mission types. Therefore, at least for long-duration exploration missions, it may make sense to move towards an approach with reduced reliance on sample return to Earth.

One step towards enabling this in-flight autonomy is improving in-flight equipment capabilities, which has been demonstrated recently via usage of the portable Oxford Nanopore MinION sequencer on board the ISS for sequencing DNA (Castro-Wallace et al., 2017) and even human-derived RNA directly (Personal communication, Sarah Castro-Wallace). Using this equipment, environmental swabs have been collected in-flight and then processed and sequenced on board (swab-to-sequence) (Burton et al., 2020; Stahl-Rommel et al., 2021). As the technology becomes more

compact and lightweight, autonomy of the sample processing may improve further still through robotic automation of spaceflight laboratories (Sanders et al., 2023). One consideration is that in-flight generation of omics data presents challenges regarding data storage and transfer back to Earth. Importantly, downlink challenges compound with the communication delays introduced by increasing distance from Earth. In-flight computational processing of raw omics data can mitigate this issue by significantly reducing file size, and full analysis pipelines have even been executed on-orbit, eradicating the need to downlink data to obtain results (Personal communication, Sarah Castro-Wallace).

Importantly, in-flight processing of omics data has implications for occupational health in space. For example, in the context of NASA's microbial monitoring programme, in-flight processing can enable crew health to be monitored (e.g. infectious disease diagnosis) in near real-time (Burton et al., 2020; Stahl-Rommel et al., 2021). As another use case, results obtained in-flight could eventually be inputted directly into AI models, such as Clinical Decision Support Systems (CDSS) to inform the decisions of onboard medical officers in the event that the healthcare support teams on Earth cannot be contacted.

3.3.6. Sampling Methods for Spaceflight Omics Collection

Within the following subsection biospecimens and sampling methods for omics data collection are discussed (**Figure 3.4**), focussing on approaches that have spaceflight heritage.

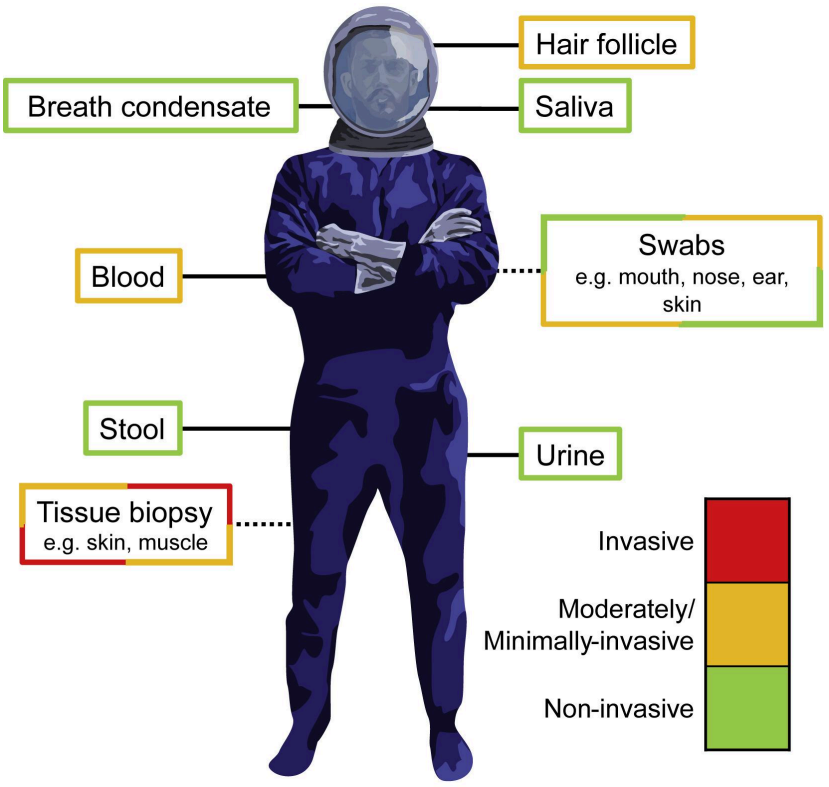

Figure 3.4. Omic sampling from astronauts. The medical invasiveness of different sampling methods for spaceflight omics collection from astronauts or spaceflight participants (SFPs). The astronaut icon used was provided by Dr. Mathew MacKay.

3.3.6.1. Tissue Biopsy

Biopsies involve the removal of a piece of solid tissue or a sample of cells from the body. Thus, while biopsies present the gold standard for molecular profiling of human physiology, such procedures are invasive in nature, require specialised expertise, are often characterised by localised pain, and

naturally carry an increased risk of infection. It is no surprise, therefore, that routine tissue biopsy sampling of astronauts in-flight is lacking, and omic studies in this sense remain constrained to those of cultured human cells flown to space – examples include transcriptomics and/or proteomics of human induced pluripotent stem cell-derived cardiomyocytes (Wnorowski et al., 2019) and human intestinal epithelial cells (Barrila et al., 2021). Given the inherent challenges of human tissue biopsy sampling in space, routine omics data collection from astronauts via this tool is most immediately likely through regular sampling programmes pre- and post-flight, particularly for more readily extractable tissue types such as skin and skeletal muscle. Indeed, in the case of the latter, muscle biopsies have previously been obtained from astronauts post- *versus* pre-flight to infer human skeletal muscle functional, morphological, biochemical and, more recently, proteomic changes in space (Fitts et al., 2010, 2013; Rittweger et al., 2018; Trappe et al., 2009). With the application of various omic techniques (transcriptomics, metabolomics, proteomics, epigenomics etc.) to biopsied skeletal muscle now being commonplace in terrestrial human physiology research, future multi-omic analyses of astronaut muscle (and outwardly skin) obtained pre- and post-flight are entirely plausible. The existence of less invasive ('microbiopsy') techniques for sampling skeletal muscle (Corvelyn et al., 2020) and skin (Berekméri et al., 2019) could make routine collection of these tissue types from astronauts during flight a realistic future prospect. In terms of a more immediate scientific return, omics collection via tissue biopsy thus lends itself as a promising tool for better understanding molecular determinants of skeletal muscle and skin responses to the spaceflight environment.

3.3.6.2. Blood

As a minimally invasive procedure, blood sampling has long been used as a primary alternative to tissue biopsies (i.e. "liquid biopsy) for molecular characterisation of human health on Earth. Given the clear technical and

operational advantages of blood *versus* tissue sampling in the spaceflight environment, blood sampling currently appears to be more viable for longitudinal monitoring of astronaut responses at the molecular level, particularly in-flight. Blood samples are thus often collected from astronauts for analyses on Earth, and targeted molecular profiling of astronaut blood obtained before, during and/or after spaceflight has subsequently been performed on numerous occasions, with particular emphasis on spaceflight-induced immune system dysregulation (Buchheim et al., 2020; Crucian et al., 2013, 2014, 2015). Additionally, comprehensive analyses of astronaut blood samples from the NASA Twins study have yielded several new and important insights. For instance, initial multi-omic (epigenomic, metabolomic, transcriptomic, proteomic) analyses revealed wholesale gene expression changes in peripheral blood mononuclear cells and metabolite changes in plasma during spaceflight that were largely normalised upon return to Earth (Garrett-Bakelman et al., 2019). Additional plasma proteomic profiling also uncovered patient-specific exosome protein signatures following the mission (Bezdan et al., 2020), with these blood samples also proving useful towards confirming human relevance of findings derived from omic analysis in model organisms (Malkani et al., 2020). Other independent studies have also explored plasma metabolomic or proteomic changes in astronauts following long-term spaceflight exposure (Brzhozovskiy et al., 2019; Pastushkova et al., 2021; Rittweger et al., 2018). Routine longitudinal multi-omics of astronaut blood is thus firmly within reach. As the number of available samples across astronauts and time points continues to grow, blood omics therefore holds accelerated potential to expedite molecular understanding of human spaceflight responses, in areas such as immune system dysregulation.

3.3.6.3. Urine & Stool

Both normal human waste products that can be cheaply, frequently and non-invasively sampled without specialised equipment or skill, urine and

stool (faeces) naturally present themselves as strong candidate sample types for routine, longitudinal astronaut omics collection. Indeed, alongside blood, urine is a common form of liquid biopsy for terrestrial biomonitoring of humans, and can be used to screen for a variety of different (patho)physiological biomarkers (renal, cardiac, prostate, bone etc) (Brzhozovskiy et al., 2017; Kuo and Chen, 2017; Swensen et al., 2021). However, unlike blood, urine is absent of a homeostatic mechanism and can therefore retain/ accommodate more substances that reflect changes produced *in vivo*, while its low complexity lends itself to straightforward detection of low-abundance proteins (Jing and Gao, 2018). As such, urine may be considered as better suited for early-stage biomonitoring, particularly in terms of the kidneys (Gao, 2015; Jing and Gao, 2018). Astronaut urine omics is consequently starting to grow, with several recent studies demonstrating the multi-omic (namely metabolomic and proteomic) potential of this sample type for gaining new information on human body adaptations following spaceflight (Brzhozovskiy et al., 2017; Garrett-Bakelman et al., 2019; Pastushkova, Kashirina, et al., 2019; Pastushkova, Rusanov, et al., 2019). For instance, metabolomic/proteomic analyses of urine from the NASA Twins Study revealed urinary excretion of COL1A1 and COL3A1 as potential biomarkers of muscle, tendon, or bone changes during spaceflight (Garrett-Bakelman et al., 2019). Cardiovascular-related urine proteome changes, including some associated with autonomic regulation of heart rate, have also been delineated, offering several promising urine biomarkers for screening cardiovascular responses to spaceflight (Pastushkova, Kashirina, et al., 2019; Pastushkova, Rusanov, et al., 2019).

On Earth, stool has become a primary sample choice for metagenomic profiling of the gut microbiome, which plays a critical role in the internal environment and concomitantly (patho)physiological state of an individual (Kinross et al., 2011; Tang et al., 2020; Valdes et al., 2018). Recognition is

now also growing for a potential role of the gut microbiome towards physiological maladaptation of astronauts during spaceflight (Turroni et al., 2020). Metagenomic studies of astronaut stool have thus gained traction, demonstrating spaceflight-induced alterations in both the composition and function of an astronaut's gut microbiome (Z Liu et al., 2020; Voorhies et al., 2019), as well as uncovering genetic and immunological evidence for transfer of environmental strains to an astronaut's gut microbiome on board the International Space Station (ISS) (Danko et al., 2020). Thus, astronaut stool (and outwardly other samples; see below) metagenomics can serve as a useful tool not only for guiding continual health monitoring, but also mission planning and habitat design (Danko et al., 2020). Future omic studies of astronaut stool would perhaps benefit from complementary metabolomics in order to improve functional readout of gut microbiome activity (Zierer et al., 2018).

In closing, it appears clear that stool and urine analyses could prove practical for routine astronaut omics collection, with strong potential for rapid scientific return.

3.3.6.4. Saliva & Body Swabs

While a potent contributor, the human microbiome is not merely defined by the gut, but rather the full complement of microbes (and their genes) that reside on or within the host. These microbes interact together alongside host genetics and the environment to contribute towards an individual's health state (Blum, 2017; Fan and Pedersen, 2021; Gilbert et al., 2018; Salvucci, 2019; Ursell et al., 2012). Wider-ranging microbiome understanding in space is thus essential for biomonitoring astronaut health and, as noted above, other factors such as habitat design. Indeed, the ISS is a unique closed environment with rich microbial diversity (Be et al., 2017; Checinska Sielaff et al., 2019; Singh et al., 2018; Stabler et al., 2017) which, combined with possible influx and outflow of microorganisms at times of spacecraft

exchange (Avila-Herrera et al., 2020), may hold potential to have an impact on astronaut health. In terms of spaceflight, saliva, nose and mouth/buccal samples present alongside stool as useful for providing insight on astronaut health and physiological conditions, whilst skin and the ear are prime sampling sites to explore interactions between an astronaut and the spaceflight environment (Morrison et al., 2021). Importantly, routine, longitudinal metagenomics collection from all these locations is a highly feasible prospect in the context of spaceflight, since all are readily accessible for frequent non-invasive sampling via swab (or 'spit' sample in the case of saliva) at little cost or time expense and without the need for specialised skill. Demonstrating the associated scientific potential are several other recent metagenomic studies of astronauts, revealing new information on host- and site-(in)dependent microbiome responses across these various locations during spaceflight, as well as microbial transmission between the astronauts and the ISS surface and *vice versa (Avila-Herrera et al., 2020; Lee et al., 2021; Morrison et al., 2021; Urbaniak et al., 2020; Voorhies et al., 2019)*. Thus, as complementary to stool, metagenomics collection from saliva and body swabs should serve as a strong tool for better understanding molecular determinants of astronaut health in the immediate future.

3.3.6.5. Hair

While perhaps an intriguing sample choice to note on face value, hair as a source for omics collection may represent a useful complementary source for routine biomonitoring of human physiological responses to spaceflight. Indeed, hair is readily available and easily sampled via extraction of hair follicles, requiring no specific personal expertise nor complex hardware (Salama et al., 1999; Terada et al., 2013). Moreover, hair cells can reflect host physical conditions, including human peripheral circadian clock and organismal metabolic responses to environmental changes (Terada et al., 2013). Omic studies of astronaut hair are thus beginning to emerge. Notably,

a seminal study by the Japan Aerospace Exploration Agency (JAXA) conducted transcriptomics on astronaut hair follicles collected before, during and after 6-months of spaceflight (Terada et al., 2016). While several markers of hair growth were unsurprisingly dysregulated during flight, interestingly, hair *PCDH8* was derived as a potential biomarker for astronaut neurochemical changes in the space environment (Terada et al., 2016). Terrestrial studies of human hair also span other omic lines (e.g. proteomics and metabolomics) (Adav et al., 2018; Chen et al., 2021), highlighting the potential for multi-omic profiling of astronaut hair samples. Nevertheless, the extent to which various physiological responses to spaceflight can be tracked by biomarkers in hair warrants a better understanding, before the true scientific potential of hair omics towards biomonitoring astronaut health can be established.

3.3.6.6. Breath

A final sample type worthy of note when it comes to astronaut omics collection is breath. Indeed, molecular analysis of exhaled breath is rapidly emerging as a useful tool on Earth for biomonitoring of an individual in health and disease ("breath-print"), both at the level of the lung as well as other vital organs via systemic circulation (e.g. the heart) (Cikach and Dweik, 2012; Ibrahim et al., 2021). Exhaled breath analysis is plausibly suited to the spaceflight environment, as samples can be acquired frequently in a non-invasive manner, with relative ease and at little expense (Cikach and Dweik, 2012). Moreover, for certain omic types, such as proteomics, exhaled breath condensate may be better suited for analysis in short time *versus* other forms of liquid biopsy such as blood and urine (Kononikhin et al., 2019). Nevertheless, omic studies of astronaut breath remain very much in their infancy. In one of the only studies to date (Kononikhin et al., 2019), proteomic analysis of exhaled breath condensate obtained from cosmonauts before and after long-term spaceflight revealed strong enrichment for proteins involved in keratinisation and pathogenic *E. coli* infection, offering

support to the potential utility of breath analysis for non-invasive monitoring of astronaut health and respiratory tract pathology (Kononikhin et al., 2019). On Earth, metabolomic study of human exhaled air ("breathomics") has also proven to be highly feasible and consequently gained great traction (Hauschild et al., 2015), though transcriptomic studies are currently lacking, perhaps owing to poor knowledge of present messenger RNAs due to high fragmentation of RNA isolated from exhaled breath condensate (Campanella et al., 2019). Nonetheless, sample collection advantages coupled with multi-omic potential (Campanella et al., 2019) make exhaled breath a highly promising complementary source for molecular profiling of astronaut health, particularly with respect to the respiratory system.

3.3.7. Ethical & Legal Considerations for Human Space Omics

In this section, the ethical challenges and legislative landscape within Europe will be considered in the context of astronaut and SFP derived omics data. Focus is given to harmonious laws across Europe, but it should be acknowledged that national laws and the involvement of non-European law, via multinational entities (e.g. commercial spaceflight companies) and international collaborations, can introduce further complexities. Ultimately, legislation aims to find a balance between supporting scientific research and healthcare while protecting participants from harm; data should be as open as possible in order to facilitate further analysis and as protected as necessary for the participants donating data.

3.3.7.1. When is Data Personal?

In May 2018 the General Data Protection Regulation 2016/679 (GDPR) came into force in the European Union. The regulation aims to enhance individuals' control over their data. The impact of the GDPR on omics in healthcare and research has recently been reviewed in a detailed report by the PHG Foundation (Mitchell et al., 2020).

Article 3(1) of the GDPR states that “This Regulation applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not”. Therefore, a routine omics programme for ESA would most likely fall within the territorial scope of the GDPR, but importantly this is only the case if the data is deemed to be ‘personal data’.

From the definition of ‘personal data’ in Article 4(1) and Recital 26 of the GDPR, it would seem apparent that omics data could be considered to be personal data if it is deemed ‘reasonably likely’ to be either directly or indirectly identifiable. Hypothetically, if a database contains solely omics with all other identifiers removed, additional information would be needed to indirectly identify the natural person from which the data is derived. This process is sometimes referred to as a linkage attack.

Holding some similarity to rare disease cohorts (Haendel et al., 2020; Nguengang Wakap et al., 2020; Nguyen et al., 2019), astronauts and SFPs are currently a small population of public figures. Consequently, information including mission status and phenotypic information such as hair colour, height, and sex pertaining to individuals is readily available to the public. It could therefore be possible to link astronauts’ or SFPs’ omic data to this additional information in order to identify individuals (Shabani and Marelli, 2019). However, an important caveat is that different omics data types have different levels of identifiability when combined with phenotypic information, meaning that some will qualify as personal data while others do not.

Omics describing the unique DNA sequence of individuals, such as single nucleotide polymorphisms (SNPs), are generally regarded as having the highest level of identifiability. Traits including eye colour, skin colour and sex are considered simple to accurately predict from data sources such as

whole genome sequencing (Lippert et al., 2017). Thus, in the unique context of spaceflight cohorts, such data types would likely be classed as personal data. However, the identifiability of other omics types, such as transcriptomics and proteomics, is also an emerging area of research (Bandeira et al., 2021; Elhaik et al., 2021; Gürsoy et al., 2021; Harmanci and Gerstein, 2016). Furthermore, the generation of multi-omics datasets adds an additional layer of complexity because there could be relationships between the different omic types that increase the overall likelihood of identification (Dupras and Bunnik, 2021). Additionally, raw data holds increased potential for identification compared to processed data, so file formats are an important consideration for appropriate data dissemination. Finally, aggregated data which combines datasets from a population of individuals (e.g. to produce an average value) can reduce identifiability. However, aggregated approaches can significantly reduce utility for research into individual differences associated with spaceflight-induced biological changes, and therefore utility in personalised medicine approaches.

It is also worth mentioning that the familial element of some omic data types means that if it is deemed 'reasonably likely' for someone to identify a family member from the astronaut or SFP's personal omics data, then it would also be classified as that family member's personal data, and so their rights may also be brought into consideration.

3.3.7.2. Approval for Data Collection & Processing

When it comes to navigating the ethics of conducting human research studies in Europe, there are several prominent non-legal binding instruments, including the Declaration of Helsinki which was first formulated in 1964 with several subsequent revisions (World Medical Association, 2013). These instruments have introduced key concepts that have become encoded within national laws and regulations across Europe. Two of the

primary concepts are independent ethical review committees and informed consent.

Firstly, all research collecting omics data from human subjects must gain approval from an independent ethical review committee before the research can commence. The committee consists of a panel of experts who provide oversight for human research, ensuring that the research is conducted in agreement with internationally and locally accepted ethical guidelines, and in compliance with the law. In the case of data derived from ESA astronauts flying on missions with international collaborators, approval will likely need to be granted by international ethics committees in addition to European committees. For example, the Human Research Multilateral Review Board (HRMRB) exists for overseeing human research on the ISS, with board members from the international partners.

As part of addressing the ethical review committees, investigators would be expected to demonstrate collection of informed consent from all study participants. Prior to providing such consent, participants should be clearly informed of the research purpose, experimental and data handling procedures, the potential risks, and their right to withdraw from the study at any time. As mentioned in the previous section, utility of omics is ever-increasing with the true capability and by consequence, identifiability, of omics yet to be unlocked. This landscape of unknowns makes informed consent difficult to meaningfully define, particularly when research is not necessarily hypothesis-driven, as is the case with the longitudinal omics collection programme proposed in this manuscript. In this context, broad consent has emerged in law (e.g. GDPR Recital 33) and practice as a potential means of obtaining consent for collecting omics for use in unspecified future research projects, including biobanking (Hallinan, 2020). Notably - although in the United States - broad consent was used in the recent multi-omics profiling study onboard the commercial SpaceX

Inspiration4 mission, to gain consent from the crew for the archiving and future use of their samples and data for space health research (Urquieta et al., 2022).

In order for processing to be permitted, European space omics data that is deemed to be personal data must satisfy a legal basis pursuant to Article 6 of the GDPR. In the context of space agencies or commercial companies, several of the reasons enumerated in Article 6 could prove relevant. To paraphrase, relevant bases may include (a) consent, (b) performance of a contract, and (e) public interest. However, omics personal data would also be considered special category data within the GDPR, either under the classification of 'genetic data' or 'data concerning health' in Article 4 and Recitals 34 and 35 (Shabani and Borry, 2018). Article 9 states that such data types cannot be processed unless at least one of ten exceptions applies. To summarise, potentially relevant exceptions include (a) explicit consent, (e) the data subject makes their data public, (h) medical purposes, and (j) archiving scientific data in the public interest. Ultimately, the restrictions on processing imposed by Article 6 and Article 9 of the GDPR require careful consideration in the context of a space agency or commercial company processing astronaut or SFP personal omics data. There are many interesting angles to consider based on space agencies' multifaceted roles as research institutions, government agencies, employers, and primary care providers for astronauts (Reed and Antonsen, 2018), and the societal role of spaceflight. To the latter point, the United Nations' Outer Space Treaty 1967 - considered to be the backbone of international space law - states that astronauts should be regarded as "envoys of mankind". The advent of commercial spaceflight missions adds additional considerations, especially due to the variety of mission types with varying degrees of scientific value.

3.3.7.3. Ethical Issues Pertaining to Data Usage

A key ethical challenge involved in omics research is the handling of incidental findings in consideration of duty of care. Analysis of astronaut or SFP derived omics could reveal pertinent findings such as gene mutations associated with increased SANS risk, however incidental findings that aren't necessarily relevant in the spaceflight context, such as increased risk of late-onset Alzheimer's disease, could also be unintentionally revealed (Antonsen and Reed, 2019). The ethical dilemma lies in the decision of whether to inform the participants, and indeed their family members, of these incidental findings which may not be medically actionable. Ultimately, the protocol for handling incidental findings should be clearly defined when obtaining informed consent. Protocol could include giving participants the choice to opt-in or opt-out of receiving these results as well as developing appropriate support services when needed.

A further ethical challenge surrounding the processing and usage of omics data is genetic discrimination, which is when an individual receives differential treatment based on their genetic information. In Europe, genetic discrimination is typically addressed in national laws, which are influenced by regional instruments including the 2000 Charter of Fundamental Rights of the European Union, and the 1997 Convention on Human Rights and Biomedicine (also known as the Oviedo Convention) (Joly et al., 2020). Chapter IV of the Oviedo Convention prohibits any form of discrimination based on genetic heritage, and also prohibits the use of tests to predict genetic predisposition to diseases, unless these tests are carried out for medical or scientific purposes, together with the provision of adequate genetic counselling. Thus, this would seem to permit the use of omics to study individual susceptibility to spaceflight risks, such as radiation-induced cancers, as long as genetic counselling is provided. However, the potential role of omics-based susceptibility and predisposition in processes such as

flight assignment nonetheless remains a contentious issue. In the United States, the 2008 Genetic Information Non-Discrimination Act (GINA) prevents NASA from making employment decisions, such as flight assignment, based on genetic data (Antonsen and Reed, 2019; Reed and Antonsen, 2018). From a European perspective, largely due to differences in national laws, the legality of using omics data in employment decisions such as flight assignment appears to be less clear. However, it seems likely that in flight assignment, the use of genomic susceptibility to disease, as opposed to manifestation, would be viewed as discriminatory. Additionally, ESA's recent Parastronaut Feasibility Project could be seen to suggest a more inclusive approach to flight assignment; thus, the focus of omics technologies should probably be to reduce risk and improve health outcomes via personalised medicine approaches (Pavez et al., 2021). However, some personalised medicine approaches are likely to be based on the processing of omics data deemed to be personal data under the GDPR. This could present a new challenge in operations with international partners. For example, if other space agencies such as NASA deploy personalised countermeasures but are unable to process data from ESA astronauts in compliance with the GDPR, either ESA astronauts will not be able to partake in the use of these countermeasures to improve occupational health or ESA will need to be able to deploy personalised countermeasure via an ESA programme; personalised medicine approaches have recently been considered by an ESA Topical Team (Pavez et al., 2021). Similarly, emergent challenges relating to GDPR compliance during processing of personal data from Europeans on commercial flights need to be considered in the context of initiatives such as the database recently established by TRISH (Urquieta et al., 2022).

3.3.7.4. Storing & Sharing Data

As an independent treaty organisation based in Europe, sharing omics data within ESA's network is of great importance. All 22 Member States,

Associate Members, and countries with Cooperation Agreements, including Canada which sits on the Council, appear to follow the GDPR or are considered, via adequacy decisions, 'secure' third countries due to similarities in data protection law. Thus, data can be transferred between these countries, as long as all parties first reach a consensus on issues such as the basis for lawful processing and what constitutes personal data. While this sounds simple, reaching a consensus can be challenging because the GDPR allows scope for variation via national laws (e.g. Article 9), which may lead to key differences in interpretation and application of the GDPR (Molnár-Gábor et al., 2021; Molnár-Gábor and Korbel, 2020). Furthermore, space biology research typically involves a great deal of international collaboration with other space agencies, posing a need for international sharing of data.

For astronaut and/or SFP omics data deemed not to be personal data, international sharing could be relatively simple. For example, the data could potentially be uploaded to a publicly accessible biobank, such as the NASA GeneLab data repository (Berrios et al., 2021). For context, NASA GeneLab has emerged as the primary database for spaceflight omics data, with significant European representation in GeneLab's Analysis Working Groups (AWGs) and experimental datasets (Madrigal et al., 2020). However, if the data is deemed to be personal data, transfers to countries without adequacy decisions, including the United States, become challenging and unclear at present due to the recent "Schrems II" judgement (Hallinan et al., 2021).

Due to challenges with ensuring GDPR-compliance during global sharing of sensitive healthcare and research data, alternative models for data sharing have been considered. For example, in congruence with a declaration signed by 21 European countries to transnationally share data on at least one million human genomes by the end of 2022, large-scale initiatives such as the European Genome-Phenome archive are shifting towards federated

approaches (Saunders et al., 2019). Federated approaches can invert the traditional data sharing paradigm of bringing the data to the analysis, by instead bringing the analysis to the data. Through federated approaches, researchers can analyse data across a distributed network of databases, and then combine the results, potentially avoiding the transfer of personal data across jurisdictions (Rehm et al., 2021; Thorogood et al., 2021). This model can also be applied to the training of AI models, in a process known as federated learning. For example, a successful model for predicting dyspnea, a common side effect after radiotherapy treatment of lung cancer, was trained on sensitive clinical data from several different hospitals across Europe, without the data ever having to leave the individual hospitals (Jochems et al., 2016). Federated approaches could potentially serve as an appropriate solution for enabling international access and analysis of sensitive astronaut and/or SFP healthcare or research data, while ensuring international legal compliance. Other potential approaches to ease data sharing include data sanitisation methods, to convert identifiable data into non-identifiable formats (Gürsoy et al., 2020). However, methods that modify the data should be carefully balanced against the potential loss of scientific utility.

With the specific security issues of data sharing strategies in mind (Rieke et al., 2020; Thorogood et al., 2021; Truong et al., 2021), data storage solutions such as biobanks housing identifiable astronaut and/or SFP omics data should meet high standards of security, to prevent issues such as data breaches. Chapter 4 of the GDPR requires that data controllers approach “data protection by design”, and establish safeguards to protect the privacy and security of data in a manner that is proportionate to the risks involved.

3.4. Conclusion

Humanity has laid ambitious plans to venture to Mars and establish settlements on other planetary bodies, and commercial spaceflight is

gaining rapid traction with the spacefaring population set to quickly increase both in terms of size and diversity. These new endeavours present new challenges for human health in space. Crucially, more data is needed, especially for environmental conditions beyond LEO where biological and molecular adaptations remain uncharted. Recently the first multi-omic profiling studies of astronauts have been performed in the NASA Twins and SpaceX Inspiration4 missions. With their feasibility for spaceflight now demonstrated, these biological big data approaches are primed to transform our understanding of health in space, as they have for health on Earth. Importantly, in addition to elucidating mechanistic understanding of spaceflight-induced physiological deconditioning, these approaches also hold strong potential to improve risk quantification and occupational health in space, such as through the design and deployment of personalised medicine approaches. Thus, routinely applying these same approaches to European human subjects, including participants in analogue environments and astronauts and/or SFPs in spaceflight missions, is a golden opportunity.

In order to achieve the full potential of these big data approaches several challenges need to be addressed in a cohesive European fashion. Firstly, how do we best store and ensure collected data is AI-ready? Secondly, which datasets should we collect and when? Thirdly, how do we enable open science approaches while protecting personal data? While the challenges and some opportuities have been discussed within this chapter, it is important to achieve European consensus on these questions (**Table 3.1**). Forming an ESA topical team and/or holding an ESA workshop presents a strategic path forward. The goal is to bring together European human omics/data science researchers and human subjects researchers to address these questions in coordination with ESA and stakeholders from the commercial spaceflight sector. Importantly, this consensus will need to consider the international context of human research on the ISS, the commercial aspects of data science, and the personal nature of the data.

These latter aspects could be achieved via collaboration with organisations such as International Standards for Space Omics Processing (ISSOP), HRMRB, and TRISH.

Table 3.1. Goals and potential topics of discussion for suggested European Space Agency (ESA) topical team/workshops.

Goal: To explore routine omics collection by deciding which omics data to collect routinely, from which samples, and at which time points.
Potential topics to be addressed
➔ Which omic types and technologies have the highest potential for scientific return and clinical actionability (with additional consideration to multi-omic combinations)?
➔ Which sample types are the most practical to collect (i.e., cost, sample processing procedures, and invasiveness)?
➔ What metadata would be the most useful and practical to standardize alongside omics collection (e.g., physiological, environmental, and lifestyle)?
➔ At which time points should omics be collected?
➔ How can the AI readiness of the generated datasets be maximized?
Goal: To help establish clear ESA policy on human omics research governance
Potential topics to be addressed
➔ What is the potential identifiably of different omic types, in the specific context of astronauts?
➔ How should ESA go about obtaining meaningful informed consent for omics research?
➔ Which legal bases should be used for processing astronaut and/or SFP omics data?
➔ How should ESA handle potential ethical issues, such as incidental findings and discrimination?
➔ How should data storage and data sharing between ESA's European and international network be handled?

Chapter IV: Astronaut Omics - The UK Opportunity

The contents from this chapter are published in (Cope et al., 2023):

***Cope, H.**, Deane, C.S., Szewczyk, N.J., Etheridge, T., Williams, P.M. and Willis, C.R.G., 2023. A Data Collection Programme for Improving Healthcare in UK Human Spaceflight Ventures. Journal of the British Interplanetary Society, 76, pp.213-220.*

All work is that of the author of this thesis, other than "4.5. Data beyond omics", which was co-written with Dr. Colleen Deane.

This work is incorporated into this thesis in accordance with the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0). Modifications have been made for its inclusion within this thesis, which include adjustments to figure, table, and section numbering.

4.1. Summary

Over the next decade the number of humans venturing beyond Earth is projected to rapidly increase in both quantity and diversity. Humans will regularly fly to the International Space Station until it is decommissioned by 2031, will return to the Moon by 2025 via the Artemis programme, and will fly to space via commercial ventures. Spaceflight presents a hazardous environment for human health. To understand spaceflight-associated health risks further and to increase safety via advanced healthcare approaches, including personalised medicine, more data must be collected. Importantly, this data must be derived from a diverse cohort of participants and a range of mission formats. An opportunity exists for the UK to start to consider all citizens venturing into space as potential participants from which health and biological data could be consensually collected. Importantly, this routine data collection programme should adopt a similar strategy to the UK National Health Service and the UK Biobank, by including "omics" data for scientific and healthcare purposes. This chapter looks at how a world-leading programme, kick-started via a pilot study, might be realised through appropriate policy design, including which measures to collect, when to collect them, and unique ethical considerations pertaining to the spacefaring population.

4.2. Introduction

Travelling into space is a hazardous venture. Firstly, in rare cases such as the Columbia disaster, design choices can prove fatal. Secondly, there are the stressors experienced by the body during launch, adaptation to space, landing, and post-flight adaptation to Earth or indeed other celestial bodies. This exposure to a unique array of environmental stressors, including microgravity and cosmic radiation, is associated with detrimental biological changes and subsequent health risks (Afshinnekoo et al., 2020; Patel et al., 2020). NASA (U.S National Aeronautics and Space Administration) maintains a register of the primary risks to human health in space (**Table 4.1**). In a recent review paper, authors from NASA enumerated the highest priority (i.e. “red”) human health risks for future missions to Mars as (i) space radiation health effects that include cancer, cardiovascular disease, and cognitive decrements (ii) Spaceflight-Associated Neuro-ocular Syndrome (SANS) (iii) behavioural health and performance decrements, and (iv) inadequate food and nutrition (Patel et al., 2020). Importantly, if a medical incident occurs beyond Earth, evacuation options and medical resources are limited, with a significant decline in the performance of an individual crew member potentially endangering other crew members and reducing the overall likelihood of mission success (Antonsen et al., 2022).

Table 4.1. Unordered subset of human health risks from NASA's risk register. Based on high priority (i.e. "red") risks for Mars missions (Patel et al., 2020). The full list of risks and connected evidence reports can be found online via the NASA Human Research Roadmap website (https://humanresearchroadmap.nasa.gov/, Accessed: 1st November 2022)

Risk Title	Risk Statement
Spaceflight Associated Neuro-ocular Syndrome (SANS)	Given that ocular and brain structural changes develop during spaceflight, there is a possibility that these changes could lead to short-term or long-term vision alterations, cognitive effects, or other deleterious health effects.
Adverse Cognitive or Behavioral Conditions and Psychiatric Disorders	Given that crews of future exploration missions will be exposed to extended durations of isolation and confinement, greater distances from Earth, as well as increased exposures to radiation and altered gravity, there is a possibility that these singular or combined hazards could lead to (a) adverse cognitive or behavioural conditions affecting crew health and performance during the mission; (b) development of psychiatric disorders if adverse behavioural health conditions are undetected or inadequately mitigated; and (c) long term health consequences, including late-emerging cognitive and behavioural changes.
Cardiovascular Adaptations Contributing to Adverse Mission Performance and Health Outcomes	Given exposure to the spaceflight environment can contribute to cardiovascular deconditioning, dysfunction, and remodelling, there is a possibility that crews will experience impaired performance and negative cardiovascular health outcomes during and after spaceflight and planetary operations.
Performance Decrement and Crew Illness Due to Inadequate Food and Nutrition	Given that there is a constrained spaceflight environment with a limited-source food supply and altered nutrient requirements, there is a possibility of inadequate nutritional delivery resulting in performance decrement, crew illness, and long term health effects.
Radiation Carcinogenesis	Given crew is exposed to radiation from the space environment, there is the possibility for increased cancer morbidity or mortality.

Within the last few decades humans have been striving to make the spaceflight experience safer, and indeed technological improvements - including new in-flight exercise devices to counteract microgravity-driven

muscle and bone degradation (Loehr et al., 2011)- have helped to minimise physiological deconditioning during regular ≈6-month missions onboard the International Space Station (ISS). However, there are two key challenges on the horizon:

A. New mission formats beyond Low Earth Orbit (LEO) - In pursuit of a lofty goal to land humans on Mars, NASA and the European Space Agency (ESA) are collaborating through the Artemis programme to return humans to the Moon by 2025. Importantly, this will reintroduce stressors associated with travel beyond LEO that have not been directly studied in astronauts since the Apollo programme, including (i) adaptation and extravehicular activity (EVA) in partial gravity (ii) exposure to irritant lunar dust (iii) increased distance from Earth (i.e. communication delays, limited resupply and evacuation options), and (iv) exposure to increased doses of radiation. Although challenging to quantify in the absence of high fidelity data, these factors contribute to a greater predicted risk of medical events for Mars and Lunar missions relative to current ISS missions (Antonsen et al., 2022). Improved healthcare approaches, suitable for long-duration deep space missions onboard compact spacecraft, will be required to decrease human health risks to acceptable levels during these upcoming missions.

B. An increasingly diverse spacefaring population - Knowledge of the impact of spaceflight on human biology has been garnered from a rare and unique population of less than 650 rigorously selected individuals, with high levels of fitness and significant training. Recently, the first commercial missions have been performed in space, including the 17-day Axiom Mission 1 (Ax-1) in which four commercial crew members visited the ISS in April 2022, the 3-day SpaceX Inspiration4 mission, in which four commercial crew

members entered LEO via a Crew Dragon spacecraft in September 2021, and the 10-minute Blue Origin NS-18 suborbital mission, in which four crew members ascended 107 km from the Earth within a New Shepard 4 vehicle in October 2021. One of the crew members on the Inspiration4 flight was a childhood cancer survivor with a prosthesis, and the NS-18 mission broke the record for the oldest human to cross the Kármán line and enter space, at 90 years of age. Similarly, the Artemis programme will see the first female astronauts to fly beyond LEO, and the ESA Parastronaut Feasibility Project looks to study the impact of certain disabilities on spaceflight performance, with a UK Paralympian recently selected. This trend of an increasingly diverse population of humans venturing beyond Earth necessitates further research into the role of individual differences in spaceflight response, including the effect of pre-existing medical conditions and lower levels of physical fitness. If risks can be more accurately quantified, they can be minimised to increase safety during human spaceflight, including commercial ventures.

To meet the healthcare needs of this changing landscape of human spaceflight, rapid advancements in space biology research and space healthcare technology are required. On Earth, the National Health Service (NHS) has been investigating advanced approaches to improve the standard of public healthcare for the UK population. As a result, the UK has become world-leading in the domain of genomic research and healthcare through initiatives like the UK Biobank and the Genomics England-led 100,000 Genomes Project. The UK Biobank is an NHS supported initiative to collect health data on over 500,000 participants for use in healthcare research, including whole genome sequencing (WGS) data. UK Biobank data has been used in a myriad of research projects, including identification of variants associated with Covid-19 mortality (Hu et al., 2021), early prediction of lung cancer (Muller et al., 2017), and understanding of genes associated with

muscular strength (Tikkanen et al., 2018). Similarly, the NHS has been implementing genomics and other omics (e.g. transcriptomics) within clinical care via approved molecular tests (**Table 4.2**). To develop this approved list, the NHS regularly performs trials to evaluate new technologies; for example, the NHS recently commenced a pilot study with Genomics plc to evaluate the inclusion of polygenic risk scores into routine assessment of cardiovascular disease risk. In their 2020 strategy paper entitled "Genome UK: the future of healthcare" the NHS set the target of being the first national healthcare system in the world to offer WGS as part of routine care, and highlighted clinical actionability of this data via approaches such as pharmacogenomics, where a patient's genomic information is used to optimise medication dosage and avoid adverse drug reactions. A classic example of a pharmacogenetic approach is NHS routine screening for HLA-B*57:01 in HIV patients before prescribing abacavir, since the drug can cause potentially life-threatening hypersensitivity reactions in patients with the risk allele (Martin and Kroetz, 2013).

Table 4.2. Unranked subset of genetic tests from the NHS's genomic test directory for cancer. Note that multiple methods with varying eligibility criteria may be available for a given clinical indication, with only one example method selected per clinical indication for inclusion into this table. The full array of tests for cancer and rare and inherited diseases can be found online via the NHS National genomic test directory webpage (https://www.england.nhs.uk/publication/national-genomic-test-directories/, Accessed: 01/11/22)

Clinical indication	Method
Solid tumour, Melanoma - Adult	Multi-target NGS panel - structural variant (NTRK1, NTRK2, NTRK3)
Possible mitochondrial disorder	Whole mitochondrial genome sequencing
Histiocytosis	Multi-target NGS panel - small variant (BRAF, MAP2K1, NRAS, KRAS, HRAS, ERBB3, ARAF, MAP3K1, PIK3CA, PIK3CD)
Solid tumour, Breast Cancer - Adult	Multi-target expression array (Oncotype DX)

Human spaceflight could adopt a similar strategy to the NHS, to accelerate space biology research, improve risk modelling, and deploy personalised healthcare approaches, including pharmacogenomics. The key to unlocking this strategy will be to start collecting comprehensive routine health and biological data - including omics data from astronauts and commercial spaceflight participants - and to upload this data into the UK Biobank or a repository akin to the UK Biobank. The feasibility of routine health data collection has been demonstrated by NASA's Spaceflight Standard Measures (SSM) programme, in which a consistent set of core health-related measurements are collected from US ISS crew (**Table 4.3**). While biospecimen samples including blood and urine are collected, the SSM programme does not currently include comprehensive untargeted omics, although statements have been made that additional omics will be added in the future. The feasibility of collecting omics for human spaceflight has previously been demonstrated in studies including (i) the NASA Twins study, where regular samples for omics processing were collected from an astronaut, pre-flight, post-flight and in-flight for a year-long mission in space, while his identical twin brother was studied in parallel on Earth as a control (Garrett-Bakelman et al., 2019) (ii) the Japan Aerospace Exploration Agency's (JAXA) recent Cell-Free Epigenome (CFE) liquid biopsy study, and (iii) a study conducted during the commercial Inspiration4 mission (Urquieta et al., 2022). The thesis author is not currently aware of any UK or European plans to conduct similar studies. The UK could collaborate with ESA and/or commercial spaceflight companies to develop a similar study, leveraging significant regional expertise in omics research and medicine. Such a study could act as a pilot for routine data collection for improving the standard of care in space for astronauts and commercial spaceflight participants. Ultimately, collecting this data can be seen as an investment in the future; the study could inspire the UK public to engage with similar health data initiatives on Earth (e.g. the NHS supported "Our Future Health" Programme), and commencing construction of a well-designed data

resource now means that the necessary quantity and quality of data will be available for future analysis capabilities and healthcare applications. Some practical considerations and recommendations for designing such a world-leading pilot study and eventual routine programme are introduced within the following subsections.

Table 4.3. Standard Spaceflight Measures (SSM) routinely collected from US ISS crew by NASA. Based on information from (https://www.nasa.gov/feature/spaceflight-standard-measures-characterizing-how-humans-adapt-in-space, Accessed: 01/11/22)

Standard Spaceflight Measures	Methodologies	Impact
Ability to stand	Sensorimotor measures	Body kinematics, heart rate, blood pressure and functional measurements are taken pre- and post-flight, providing information on the sensory motor system. These results will help identify counter-measures that maintain peak functional performance.
Biomarkers	Biochemical markers (e.g. proteins, lipids, amino acids)	Urine and blood samples provided before, during (blood only) and after spaceflight are used to identify health-related biomarkers. This information will help optimise nutritional and exercise interventions.
Heart health	Carotid intima-media thickness (cIMT)	The thickness of the carotid artery, measured via cIMT pre and post-flight, is used as an early indicator of oxidative stress which may lead to serious heart complications. These results will help monitor astronaut heart health and will aid the development of countermeasures against atherosclerosis.
Cellular profile/immune system	Cellular profile	Blood and saliva provide pre-, during and post-flight This information will help understand infections and how astronaut's immune system adapts to spaceflight, leading to the development of interventions to boost the immune system.
Isolation and confinement	Personality survey	Monthly surveys regarding stress levels, mood, perceived performance and group cohesion are collected. This information will help determine the optimal team characterisers for long duration spaceflight missions.
Microbiome	Microbiome	Body swabs, faecal samples and saliva samples, coupled with questionnaires, are collected to quantify the types and concentrations of microorganisms. These results help determine how spaceflight affects the microbiome and how microbial dynamism relates to health changes.
Sleep in space	Actigraphy & sleep survey	The wrist-worn wearable measures activity and ambient light sources, which combined with a short sleep survey regarding sleep patterns, provides information on sleep quality. This information will help improve alertness, reduce fatigue and improve sleep quality.
Space fog	Cognition	Pre, during and post-flight, 10 brief cognitive tests are performed to help devise interventions that optimise cognition in space.

4.3. Pre-flight genomic screening for spaceflight

Adopting routine pre-flight genomic sequencing of participants could provide results that are clinically actionable in the immediate term, and provide valuable data for advancing fundamental space biology research (**Figure 4.1**).

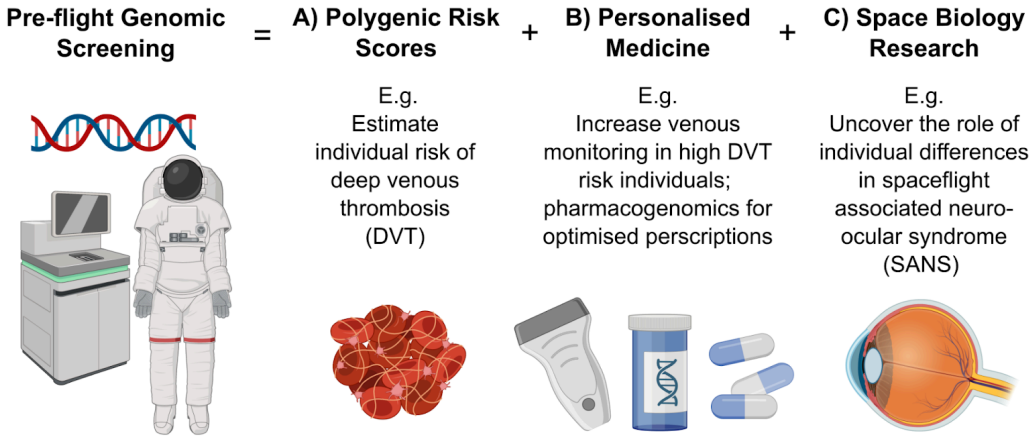

Figure 4.1. Potential benefits of pre-flight genomic screening of astronauts with examples. A) Genomic data could be used for spaceflight polygenic risk scores, to increase the accuracy of risk estimation for spaceflight health risks, including deep venous thrombosis (DVT). **B)** Genomic data could enable personalised medicine approaches, such as individualised medication prescriptions based on pharmacogenomics, or increased monitoring for early detection of specific health issues in predisposed individuals. **C)** Genomic data could improve fundamental space biology research, such as to detect genetic markers of spaceflight health issues, including spaceflight associated neuro-ocular syndrome (SANS).

Firstly, pre-flight genomic screening of astronauts would be useful for risk quantification, such as via polygenic risk scores (**Figure 4.1A**). For example, spaceflight may increase the risk of deep venous thrombosis (DVT) (Limper et al., 2021), with a case of internal jugular DVT recently reported to have occurred on the ISS (Auñón-Chancellor et al., 2020; Marshall-Goebel et al.,

2019). Genetic polymorphisms associated with DVT on Earth are well documented and this information could potentially be used in assessing risk of developing DVT in flight. Most commonly found in individuals of Northern European ancestry, the Factor V Leiden mutation associates with an approximate five to seven-fold increase in risk of developing DVT when present in its heterozygous form (Ornstein and Cushman, 2003). The rarer homozygous form of the Factor V Leiden mutation increases the risk of developing clots by approximately 25 to 50 fold (Ornstein and Cushman, 2003), and other genetic mutations such as the Prothrombin G20210A (Factor II Mutation) can also elevate risk of DVT on Earth (Varga and Moll, 2004). Additionally, it is worth noting that a plethora of other factors can increase individual DVT risk on Earth and potentially in spaceflight, such as the use of oral contraceptives (Zwart et al., 2022). Thus, genetics can be combined with other risk markers (Harris et al., 2022) as part of routine individual risk assessment for DVT.

Secondly, pre-flight genomic screening could be used to deploy personalised medicine strategies (**Figure 4.1B**) to ameliorate health risks in space. For example, if an individual is predicted to have a greater risk of developing DVT during a mission, this information could be used to increase occupational surveillance in measures relating to DVT. This approach could help to detect blood clots requiring treatment before adverse health outcomes occur, as highlighted by the detection of DVT on the ISS during an ultrasound monitoring study (Marshall-Goebel et al., 2019). Increased surveillance in high-risk individuals could also potentially detect markers of early onset of DVT, enabling the deployment of preventative countermeasures. The latter point was highlighted by NASA's follow-up occupational surveillance programme of 11 astronauts, which found abnormal flow characteristics in-flight to varying degrees between individuals, which may be predictive of DVT (Pavela et al., 2022). In addition to increased surveillance, if individuals with a higher DVT risk are part of the

crew, this information could be used during mission planning to ensure that the spacecraft pharmacy contains recommended medications for treating DVT (Mazzolai et al., 2018) from mission inception. Notably in the DVT event on the ISS, a resupply payload had to be delivered to the ISS containing an oral anticoagulant and a reversal agent, as the ISS pharmacy only contained an injectable anticoagulant when the DVT was detected (Auñón-Chancellor et al., 2020). In longer duration missions medication resupply may not be a feasible option, so launching with the appropriate pharmaceuticals will be of paramount importance. In support of this, genomics data could also be used to deploy pharmacogenomic and nutrigenomics approaches, optimising the medication and diet provided to participants based on their individual genomic differences. Previous research identified that 31% of pharmaceuticals onboard the ISS are metabolised by polymorphic enzymes, which can significantly contribute to individual variability in drug pharmacokinetics, efficacy, and safety [5]. A recent study by the Mayo Clinic demonstrated that basic pharmacogenomic screening of just nine genes improved prescriptions in 46 out of 82 participants, with several of the drugs involved in this study also overlapping with drugs previously reported to be in the ISS pharmacy [5,6]. Thus, genomic data from individual spacefaring participants could be used to individualise drug prescriptions, improving medication outcomes and optimising pharmaceutical supply payloads.

Thirdly, pre-flight genomic screening would also be useful for advancing fundamental space biology research (**Figure 4.1C**), such as by unveiling the role of individual genomic differences in spaceflight associated physiological deconditioning. For example, SANS is a distinct condition causing structural changes to the eyes of astronauts (**Table 4.1**). The mechanistic cause of SANS is yet to be established, but the phenotypic response appears to vary between individuals. Previously, researchers have hypothesised that genetic factors, such as differences pertaining to one-carbon metabolism genes, may be associated with SANS (Lee et al., 2020). Routine pre-flight genomic

screening of astronauts could be used to establish whether individual variation in these genes or other genes can be linked to the development of SANS. An improved mechanistic understanding of SANS could then be used to develop molecularly rationalised countermeasures. Similarly, some physiological changes that occur in space have well-established genomic markers on Earth, but it is often not known how well markers for the Earth phenotype correlate with markers for the spaceflight phenotype in humans. Uncovering this relationship could help to provide data on the validity of spaceflight analogues, such as bedrest for muscle atrophy.

Practically, pre-flight genomic screening of astronauts and/or spaceflight participants could be performed via standard WGS protocols used in initiatives including the UK Biobank. To support the use of genomics data in mission planning, and for convenience regarding busy astronaut training schedules, samples could be collected and sequenced well in advance of the flights. The cost of next generation sequencing technologies has continued to drop well below £1000 per human genome, which is low relative to human spaceflight costs. Pre-flight genomic screening of spacefarers would certainly be cost-effective if it can reduce risk and improve human health and performance in space, such as via optimised pharmaceutical prescriptions.

4.4. Omics beyond pre-flight genomic screening

In addition to pre-flight genomics, collecting other omic measures (e.g. transcriptomics, proteomics) longitudinally would prove valuable (Metwally et al., 2022). Pre-flight measures at frequent intervals could be used to establish a robust molecular baseline for an individual and to study the impact of training exercises, in-flight measures could be used to monitor biological changes during adaptation to spaceflight and to study the impact of events (e.g. infections, EVA), and post-flight measures could be used to monitor recovery. Spaceflight impacts human biology rapidly, so early

in-flight timepoints are essential. For example, space motion sickness and hemolysis associated with anaemia both occur within the first few days of spaceflight (Heer and Paloski, 2006; Trudel et al., 2022).

A variety of biospecimen collection methods have been used in human spaceflight studies to date (Cope et al., 2022). For frequent sampling, particularly in-flight, non-invasive or minimally invasive sample collection methods (e.g. urine, saliva, stool, blood) could prove to be an efficient use of expensive astronaut time while also providing valuable biological insights. Tissue biopsy and micro-biopsy techniques are also worth investigating for less frequent sample collection, such as muscle biopsies for revealing the mechanisms behind spaceflight muscle atrophy or monitoring muscle/metabolic health. Of note, skin punch biopsies were explored as a routine monitoring technique on the recent Inspiration4 mission.

As part of NASA's SSM, microbiome data is collected routinely from astronauts (**Table 4.3**). This data is particularly useful for monitoring the microbiome of the built environment within the ISS, which can be used to detect potentially dangerous pathogens. Due to the closed environment of the ISS, microorganisms can proliferate and pass between surfaces and crew members (Avila-Herrera et al., 2020). Microbiome monitoring work on the ISS has also helped to develop capabilities for in-situ analysis of microbes, including on-orbit sequencing of DNA and RNA via the Oxford Nanopore Technologies MinION sequencer (Stahl-Rommel et al., 2021). These tools could be used for in-situ processing and analysis of human data (Castro-Wallace et al., 2017). In-situ processing and analysis capabilities will be invaluable for missions beyond LEO, where sample return becomes increasingly challenging.

4.5. Data beyond omics

In addition to genomic and other omics data, it will also be crucial to routinely collect physiological endpoints and environmental measures throughout missions. These could then be correlated with omics data to identify omic-phenotypic relationships, such as omic variations related to decline of muscle performance in space (Tikkanen et al., 2018).

Physiological endpoints should be routinely collected to monitor the phenotypic changes induced by spaceflight. A range of physiological endpoints are included in NASA's SSM programme (**Table 4.3**), to characterise the behavioural health and performance, cellular profile/immunology, microbiology, biochemistry, sensorimotor, and cardiovascular effects of living and working in space. These SSM enable high-level understanding of the biological effects of spaceflight, permit the monitoring of countermeasure effectiveness and inform on future hypothesis-driven, mission-enabling research. Thus, adopting a similar set of feasibly obtainable core measurements consistently collected pre-flight, in-flight and post-flight in future spacefarers and missions, including during the Artemis programme, is of prime importance. Astronauts from international partner agencies can opt into the SSM programme on the ISS.

Importantly, these SSM represent easily obtainable measures that do not often require high levels of experience, expertise or training. For example, astronauts currently don wrist-worn wearables, akin to a smartwatch that monitors their activity levels and ambient light sources. This, in conjunction with a sleep pattern questionnaire, provides information on the quality of their sleep (**Table 4.3**). Indeed, the potential of wearables as an easy-to-implement SSM in spaceflight goes beyond collecting sleep-related data, with the potential for collecting a multitude of lifestyle-related data to provide personalised predictions and recommendations for optimising

health in space. For example, on Earth, a wrist-worn wearable was able to detect aberrant physiological (heart rate) and activity signals (steps) associated with the onset of early infection, providing a real-time alerting wearable that can be used for early detection of infection and other stressors (Alavi et al., 2022). This could have significant application to spaceflight by facilitating the prevention and detection of infections during microgravity stays. Along similar lines, wearable-monitored vital signs, namely continuously monitored heart rate, body temperature, electrodermal activity and movement, combined with machine learning models (e.g. random forest) were able to predict clinical lab measurements: Electrodermal activity was a strong predictor of haemoglobin, for example (Dunn et al., 2021). In the spaceflight setting, these models may also be able to identify context specific risks such as individuals suffering from dehydration. Of particular note, wearables provided a more consistent depiction of heart rate when compared to clinically obtained heart rate measurements (Dunn et al., 2021), mostly likely due to the amount of time monitored (i.e. wearable is continuous, clinical is a snapshot). This implies that not only will wearables provide an easy way to obtain a wealth of lifestyle data during spaceflight, but they will also provide consistent and robust results that have predictive clinical value. Similarly, wearables on Earth can identify inflammatory responses and distinguish physiological differences between insulin-sensitive and -resistant individuals (Li et al., 2017), both of which have relevance to spaceflight-induced biological changes. Indeed, the Canadian Space Agency has developed a wearable shirt for assessing physiologic parameters in the context of spaceflight health monitoring (Mastrandrea et al., 2021).

Environmental measures could include data collected passively via sensors throughout the spacecraft, including oxygen levels, carbon dioxide levels, pressure levels, and radiation dosimetry data. These measures are collected on the ISS and will certainly be collected in future spacecraft too, since

monitoring these parameters is essential for maintaining a safe habitat for crew. For example, dosimetry data can be used to estimate radiation exposure in crew members and monitor lifetime radiation exposure; lifetime radiation exposure limits are currently limited to 1 Sievert (Sv) for ESA astronauts, yet limits vary between international space agencies (Shavers et al., 2024). Accuracy of radiation dosage estimated will be dependent on how localised sensor data can be to the studied individual. A recent pilot study of a single participant on Earth performed a combined analysis of a plethora of longitudinal environmental parameters (Gao et al., 2022); such an approach would be immensely useful in linking spaceflight physiologic changes to environmental parameters.

4.6. Ethical considerations

Personal health data is sensitive information, as such, initiatives that collect and/or use this data are subject to legislation and policy. The overarching goal of policy in this context is to protect individuals and their families from exposure to ethical harms while enabling their personal data to be used in appropriate use cases. Procedures including ethical review boards, genetic counselling for discussing incidental findings, and informed consent are implemented in human subject research involving omics and health data on Earth. Similarly, these same procedures are required for human spaceflight research (Cope et al., 2022). Notably, broad consent has emerged as a solution to obtain consent for studies like the proposed spaceflight routine data collection programme, where the future uses of the data cannot be granularly defined. Broad consent has been used to collect and utilise data from commercial spaceflight participants in the recently established EXPAND database by the Translational Research Institute for Space Health (TRISH) (Urquieta et al., 2022). However, studies involving governmental astronauts, including the NASA Twins study, have traditionally used continuous/ongoing consent in order to provide continuous protection for their astronaut's privacy concerns.

Privacy preservation for human spaceflight studies is especially challenging due to the current rarity and public nature of spacefaring humans. Individuals could potentially be identified by linking health or biological data with phenotypic information from public media. Identifiability risk is trivially significant with certain genomic information, such as WGS data (e.g. eight SNPs can be used to predict eye colour (Hart et al., 2013)), yet other health and omic data types could also prove identifiable, particularly in combination (Dupras and Bunnik, 2021). While this presents a challenge, it is not insurmountable, and shares similarity with rare disease and elite athlete cohorts on Earth. Given the sensitivity of certain health and biological data types, approaches including access levels where only individuals from approved research projects can access sensitive data, and federated or distributed approaches where only summary statistics or machine learning model weights are seen by the accessors, should be considered. Federated and distributed approaches have been used in initiatives on Earth such as multi-site clinical studies where individual health providers are not permitted, or do not wish, to share patient data outside of the primary site of care (Jochems et al., 2016). As more astronauts and commercial spaceflight participants are studied, the reasonable likelihood of identification can be reduced providing that metadata such as names, dates and mission tag are withheld. It should however be noted that the withdrawal of metadata, such as age, can reduce the scientific utility of the data, so attempts at pseudonymisation of the data should be considered carefully. Under the General Data Protection Regulation (GDPR) and equivalent regulations in countries including the UK, data deemed to be non-personal could likely even be made publicly accessible, such as via the NASA Open Science Data Repository (Berrios et al., 2021; Cope et al., 2022).

Discrimination is also a challenging topic in the context of human spaceflight research. Space agencies have a mixed role as both the healthcare provider

and the employer of astronauts. This raises concerns over the use of personal health information in employment decisions and flight assignment decisions. The current astronaut selection process at ESA is rigorous and applicants must meet medical standards, such as the European Part-MED, Class 2 (this includes vision, hearing, respiratory, cardiovascular, metabolic, and muscular health checks). Similarly, medical standards are being implemented for commercial spaceflight in the UK via the UK Space Industry Act 2018, and indeed insights from the proposed routine data collection programme could help to inform medical standards in the future. Genetic discrimination is a particularly sensitive topic; to the thesis author's knowledge genomic data is not currently used during selection or flight assignment, although genetically-influenced traits such as height and family medical history may be used. Ultimately, discrimination should be avoided where possible. Informing consenting participants of their predicted spaceflight health risk levels would enable greater decision autonomy; if they still wish to fly, their data should be used to decrease risk levels (e.g. via personalised medicine approaches), and in cases where risk is deemed to be too high, these risk factors should be integrated into spaceflight medical standards.

4.7. Conclusion

Human spaceflight has reached an exciting and pivotal point, with ambitious plans for missions beyond LEO, and an increasingly diverse population of individuals venturing into space, including via commercial programmes. To increase safety in these ventures, health and biological data, including omics data, should be collected routinely via a standardised programme. The resultant data resource would improve risk estimation, enable advanced healthcare approaches, and aid fundamental space biology research. Increasing the size and richness of this data resource will increase statistical power and unlock new analysis methods, including machine learning

approaches; thus it is important to start to build this resource today in anticipation of future insights.

As a world leader in genomics, the UK should make steps towards developing this routine health data collection programme for astronauts and spaceflight participants, following a similar strategy to that adopted by the NHS. Firstly, experts in human research should achieve a consensus on which measures to collect and at which time points (**Figure 4.2A**). This should include pre-flight genomic screening of spacefarers in addition to further omic measures, including at regular time points throughout the flight. Additionally, this should include environmental data and physiological endpoints, collected passively via sensors and wearables. Where appropriate, these efforts could be coordinated with existing programmes including NASA's SSM. This consensus on which data to collect should also include policymaking. Policy should be developed to gain meaningful informed consent from participants and to support data sharing for appropriate research and healthcare uses. This policy should take due consideration of the unique nature of the astronaut population, to protect participants from ethical harms such as privacy breaches and discrimination.

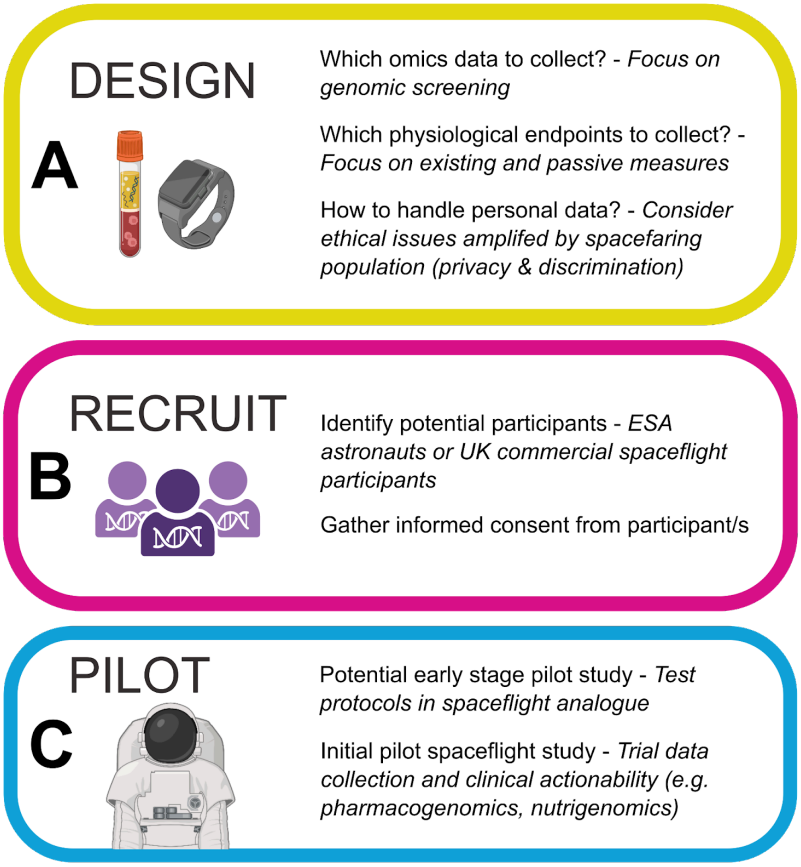

Figure 4.2. Next steps to develop a routine health data collection programme for UK astronauts and spaceflight participants. A) To design a world-leading programme, experts and stakeholders must achieve a consensus on which omics to collect, which physiological endpoints and environmental data to collect, and how best to implement policy pertaining to the governance of personal data from human subjects in this context. **B)** Participants, whether they be European Space Agency (ESA) astronauts or UK participants in commercial spaceflight missions, should be recruited for the programme and meaningful informed consent should be obtained. **C)** A pilot study should be performed, in which the health data collection programme is trialled in spaceflight, and opportunities for clinical actionability and research are documented.

Following study and policy design, participants should be recruited (**Figure 4.2B**), whether they be ESA astronauts, or UK spaceflight participants flying via commercial missions. Once participants have been recruited, an initial pilot study can be conducted in the first cohort of participants (**Figure 4.2C**).

In this pilot study, the routine data should be collected from the participants and any opportunities for clinical actionability of this data should be documented and potentially pursued. Prior to a spaceflight pilot study, it may make sense to validate the procedures and protocols of the programme in a spaceflight analogue environment, such as via a cohort of UK citizens in an isolation study. After the spaceflight pilot study, the programme should be assessed and adapted as needed for future routine adoption as a new standard of care.

Chapter V: Data Analysis Case Study - Transcriptomic Analysis of Worms

The contents from this chapter are published in (Soni et al., 2023):

Purushottam Soni, Hunter Edwards*, Taslim Anupom, Mizanur Rahman, Leila Lesanpezeshki, Jerzy Blawzdziewicz,* ***Henry Cope****, *Nima Gharahdaghi, Daniel Scott, Li Shean Toh, Philip M. Williams, Timothy Etheridge, Nathaniel Szewczyk, Craig R.G. Willis, and Siva A. Vanapalli. Spaceflight induces strength decline in Caenorhabditis elegans. Cells. 2023.*

*Equal contribution. Author list was ordered by institution.

Figure 5.1 (i.e. the graphical abstract) was developed by the thesis author. Principal component analysis and associated Figures 5.4A and 5.4B were done by the thesis author, gene module analysis and associated Figures 5.6 was done by the thesis author (but the final version of 5.6C was formatted by Dr. Craig Willis), drug target analysis was done by the thesis author but the final version of Figure 5.7 was formatted by Dr. Craig Willis. The thesis author was responsible for the paper formatting, and significant work on the writing and interpretation of results.

This work is incorporated into this thesis in accordance with the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). Modifications have been made for its inclusion within this thesis, which include adjustments to figure, table, and section numbering, as well as the conversion of the text to British English.

5.1. Summary

Understanding and countering the well-established negative health consequences of spaceflight remains a primary challenge preventing safe deep space exploration. Targeted/personalised therapeutics are at the forefront of space medicine strategies, and cross-species molecular signatures now define the 'typical' spaceflight response. However, a lack of direct genotype-phenotype associations currently limits the robustness and, therefore, therapeutic utility of putative mechanisms underpinning pathological changes in flight. As a validated model of space biology, the worm *Caenorhabditis elegans* was used in the following study (A Scott et al., 2023) Additionally, 'NemaFlex-S' microfluidic devices were used for assessing animal strength production. Wild-type and *dys-1*(BZ33) strains (a Duchenne muscular dystrophy (DMD) model for comparing pre-disposed muscle weak animals) were cultured on the International Space Station in chemically defined media before loading second generation gravid adults into NemaFlex-S devices to assess individual animal strength. These same cultures were then frozen on orbit before return to Earth for next generation sequencing transcriptomic analysis. As summarised in **Figure 5.1**, this study observed that neuromuscular strength was lower in flight *versus* ground controls (16.6% decline, $p < 0.05$), with *dys-1* significantly more (23% less strength, $p < 0.01$) affected than wild-types. Furthermore, the transcriptional gene ontology signatures characterising both strains of weaker animals in flight strongly corroborate previous results across species, enriched for upregulated stress response pathways and downregulated mitochondrial and cytoskeletal processes. Functional gene cluster analysis extended this to implicate decreased neuronal function, including abnormal calcium handling and acetylcholine signalling, in space-induced strength declines, under the predicted control of UNC-89 and DAF-19 transcription factors. Finally, gene modules specifically altered in *dys-1* animals in flight again cluster to neuronal/neuromuscular pathways, suggesting strength loss in

DMD comprises a strong neuronal component that predisposes these animals to exacerbated strength loss in space. In conclusion, highly reproducible gene signatures strongly associate with space-induced neuromuscular strength loss across species, and neuronal changes in calcium/acetylcholine signalling require further study. These results promote targeted medical efforts towards, and provide an *in vivo* model for, safely sending animals and people into deep space in the near future.

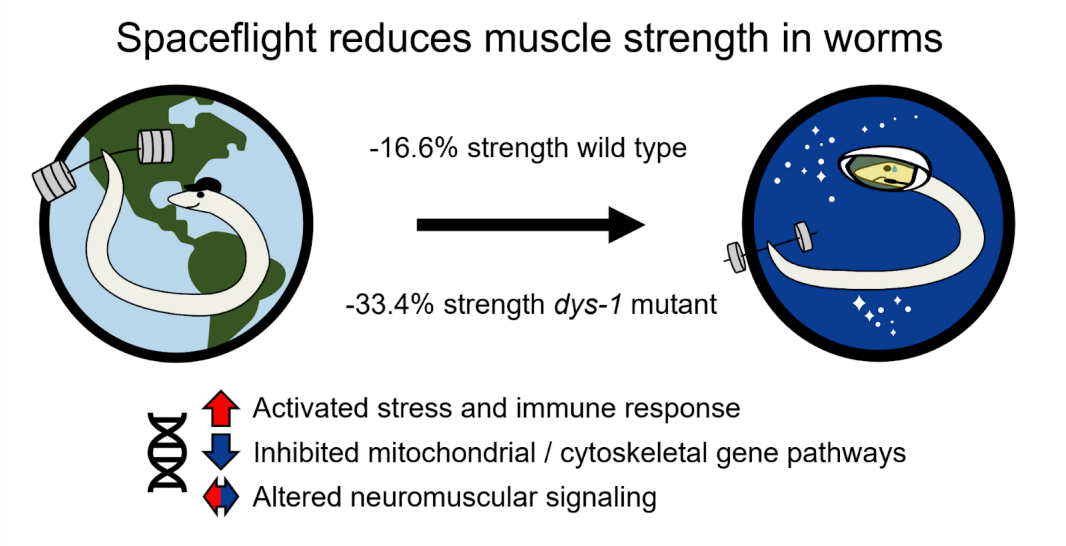

Figure 5.1. Graphical summary of the key findings from the micro-16 analysis. Strength loss refers to reduction of strength when comparing 2nd generation adult spaceflight *C. elegans* with worms on the ground.

5.2. Introduction

As a species, we have been working and living in space for more than 60 years, with continuous occupation for the last 23 years via the International Space Station (ISS) (Douglas et al., 2020). During this time, we have discovered that spaceflight induces numerous physiologic alterations including but not limited to muscle and bone loss, structural changes in the eyes and brain, altered cardiovascular function, and altered cognitive function (Patel et al., 2020). In some cases, we have developed effective countermeasures for these undesirable outcomes, but in other cases

understanding and countering the causes of these outcomes remain active areas of investigation. For example, adequate exercise and nutrition are sufficient to limit bone and muscle loss in astronauts onboard the ISS (Lane et al., 2013). In contrast, causes of and countermeasures for alterations in the brain remain active areas of research (Roy-O'Reilly et al., 2021). As research advances, it has become increasingly clear that there are some molecular hallmarks of spaceflight (Afshinnekoo et al., 2020), which include mitochondrial dysregulation, oxidative stress, DNA damage, and epigenetic changes. Current challenges of space life sciences research include linking these molecular alterations to the physiologic ones as well as to the known spaceflight hazards. These hazards include distance from Earth, confinement, being in a hostile and closed environment, altered gravity, and increased and altered radiation exposure (Afshinnekoo et al., 2020).

With our increased presence in space, our ability to conduct space medicine continues to improve. Currently, there is an increased interest in personalised approaches for optimising astronaut health (Pavez et al., 2021; RT Scott et al., 2023; Stroud et al., 2022). For example, exercise and nutrition can be individually tailored to prevent muscle loss in flight (Garrett-Bakelman et al., 2019; Lane et al., 2013). The success of these approaches on Earth combined with the increased use of -omics data in terrestrial medicine has also led to using -omics approaches in astronauts (Cope et al., 2022; Garrett-Bakelman et al., 2019; Stroud et al., 2022; Zheng et al., 2023). While still in its infancy, it seems clear that combining -omics data from space-flown model organisms such as *C. elegans*, *Drosophila*, and rodents with astronaut data can accelerate the discovery process, in part by compensating for the current rarity of human space -omics datasets (Cope et al., 2022; Rutter et al., 2020). For example, *C. elegans*, rodents, and humans all display alterations in insulin linked gene expression in response to spaceflight (Mathyk et al., 2024). Given the central role of insulin in human health and longevity on Earth, it is highly likely this system is equally

important to maintain in space. Indeed, recent data from space-flown rodents confirm that the liver-muscle axis is central to regulating both global metabolic health and muscle health functions the same in space as on Earth, and that its perturbation may contribute to both liver and muscle pathology in flight (Vitry et al., 2022).

While our advances in fundamental space life sciences and space medicine have been substantial, these are largely limited to habitation in Low Earth Orbit (LEO). With planned government and commercial missions beyond LEO (BLEO), understanding fundamental changes to biological systems and countermeasures to detrimental alterations at new destinations such as the Moon and Mars are emerging challenges. For example, current exercise countermeasures for maintaining muscle on the ISS are not feasible on currently planned spacecraft due to size constraints (Laws et al., 2020). Similarly, storing adequate nutrition further away from Earth is operationally challenging (Tang et al., 2021). Thus, despite not being a current priority for research on the ISS, understanding molecular mechanisms underpinning muscle atrophy in space remains important for these new exploration class missions.

Currently, designing space biology experiments for BLEO means that the experiments must be small and autonomous (Everroad et al., 2021). For example, using a CubeSat such as Biosentinel (Massaro Tieze et al., 2023), as a passive or mostly passive payload on Artemis, or an autonomous instrument on a Commercial Lunar Payload Services mission. Thus, using standard flight genomic model organisms (Rutter et al., 2020) that are small such as bacteria, yeast, *C. elegans*, or *Drosophila* is feasible. In terms of muscle strength, a microfluidic device has been developed for assessing strength in *C. elegans* (Rahman et al., 2018) and this has been successfully used to both demonstrate decreased strength in mutants with defective muscle structure (Etheridge et al., 2015; Hewitt et al., 2018; Lesanpezeshki

et al., 2021) and to identify drugs to improve strength both in muscle mutants (Ellwood et al., 2021) and with age (Vintila et al., 2023). Thus, the use of *C. elegans* for studying and countering muscle strength decline on the Moon is technically feasible; albeit challenging.

The use of *C. elegans* in space biology has recently been reviewed (A Scott et al., 2023; Soni et al., 2022). Notably, past studies have demonstrated conserved gene expression changes in response to spaceflight between *C. elegans*, rodents, and humans. These include declines in muscle contractile genes (Higashibata et al., 2006), mitochondrial genes (Afshinnekoo et al., 2020; Higashibata et al., 2016), and insulin signalling (Honda et al., 2012; Mathyk et al., 2024)
(Honda et al., 2012; Mathyk et al., 2024). Knockdown of these genes on Earth is sufficient to induce strength decline (Ellwood et al., 2021; Etheridge et al., 2015). Therefore, the hypothesis of the study was that worms would be weaker in space and the aim was to directly test this on the micro-16 ISS payload, flown starting in February 2021 (NG-15 launch). The results from this experiment pave the way for testing pharmaceutical and nutrient countermeasures for strength decline in space-flown *C. elegans* as well as for testing *C. elegans* muscle strength at various BLEO destinations.

5.3. Results

5.3.1 Growth

Observation of ground control worms indicated that worm growth was as expected based upon results from the experiment verification test (Soni et al., 2022). The crew assessed flight growth by eye, based on perceived density (e.g. small or medium) of worms within Fluorinated Ethylene Propylene (FEP) bags. Growth was adequate for ability to load worms into NemaFlex-S devices as indicated by loading of 37 *wt* including 15 adults and 25 *dys-1* including 13 adults on the crew's training time point session.

For the experimental time point, 41 *wt* including 30 adults and 49 *dys-1* including 29 adults were loaded into the devices.

5.3.2. Body diameter and length

Force estimation using NemaFlex devices is dependent on adjustment for body diameter (Rahman et al., 2018). Therefore, body diameter and length of both *wt* and *dys-1* adult worms were measured from the recorded movies. The diameter of both strains cultured on ISS was significantly lower than both strains cultured on Earth (**Figure 5.2A**, **Table 5.1**). The body diameter of *wt* worms was 5.5% less and in *dys-1* worms it was 7.9% less. The smaller diameter might be due to altered metabolism as this has previously been reported for space-flown *C. elegans* (Higashibata et al., 2016). Indeed, the gene expression data in the study outlined within this thesis chapter are consistent with previously reported gene expression data for space-flown *C. elegans* (Higashibata et al., 2006, 2016; Honda et al., 2012; A Scott et al., 2023; Selch et al., 2008; Sudevan et al., 2022). The length of both strains was not significantly different in flight than on the ground. Note that this contrasts a previous flight where length decreased significantly by 5.5% (Higashibata et al., 2016). This discrepancy could be because mixed populations of worms were used in the current study, whereas (Harada et al., 2016; Higashibata et al., 2016) used age-synchronised worms or because the current study employed a different diet which is known to profoundly impact *C. elegans* morphology and life history (Szewczyk et al., 2006).

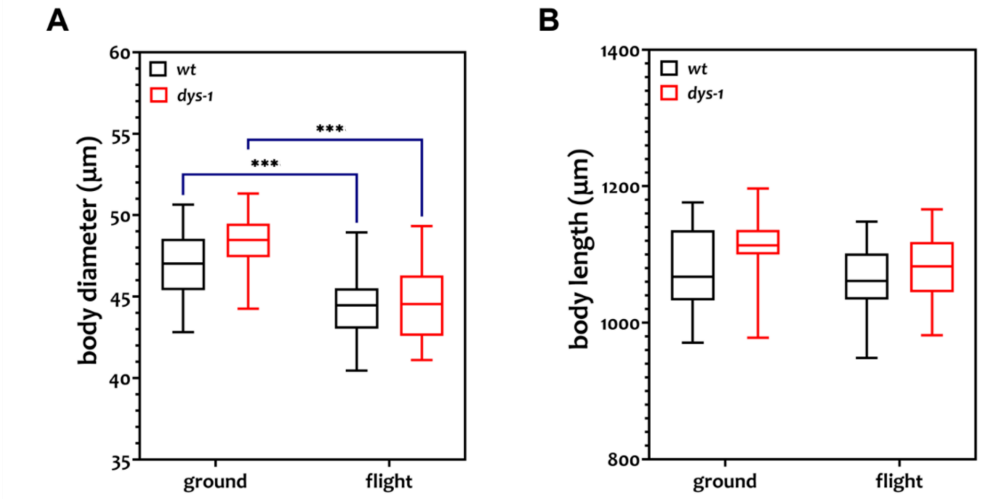

Figure 5.2. Effect of spaceflight on body diameter and length of wt and dystrophin (*dys-1*) worms. A) body diameter. **B)** body length. The diameters of both the strains grown at the International Space Station (ISS) are significantly different compared to ground controls, whereas there is no difference in the length of the worms. Sample size: n=30 for wt ground and flight, n=25 for *dys-1* ground, and n=29 for *dys-1* flight. All the data pass the normality test. Two-way ANOVA (Tukey multiple testing) was used for calculating significant differences; $p < 0.001$ is for ***. This figure was provided by Dr. Soni Purushottam.

Table 5.1. The effect of space flight on body diameter, length, and muscle strength for adult wt and *dys-1* worms. This table was provided by Dr. Soni Purushottam.

Ground				
Strain	**Sample Size**	**Diameter (µm)**	**Length (µm)**	**f_{95} (µN)**
wt	30	46.97 ± 1.92	1077 ± 60	22.34 ± 5.67
dys-1	25	48.24 ± 1.87	1110 ± 43	21.38 ± 5.39
Flight				
Strain	**Sample Size**	**Diameter (µm)**	**Length (µm)**	**f_{95} (µN)**
wt	30	44.35 ± 1.88	1062 ± 42	18.62 ± 4.05
dys-1	29	44.41 ± 2.29	1080 ± 45	14.23 ± 3.87

5.3.3. Muscle strength

Strength was measured in the second generation in flight to reflect strength in worms wholly developed onboard ISS and with in flight loading methods and video quality that had been tested on first generation from Earth prior to use with the experimental time point. To measure adult worms' muscle strength, a 60-second-long video of crawling worms in NemaFlex-S chambers was recorded. The recorded videos were analysed, and the pillar with the maximal deflection was identified in each image and generated a cumulative probability distribution with all the maximal deflections, as previously described (Rahman et al., 2018). he 95th percentile of this maximal force distribution, referred to as f_{95}, as a measure of muscle strength. As shown in **Figure 5.3**, on Earth, gravid young adult *dys-1* worms are not significantly weaker than *wt*. This is consistent with previous reports where *dys-1* mutants only display strength deficits vs. *wt* post day 1 of adulthood (Hewitt et al., 2018).

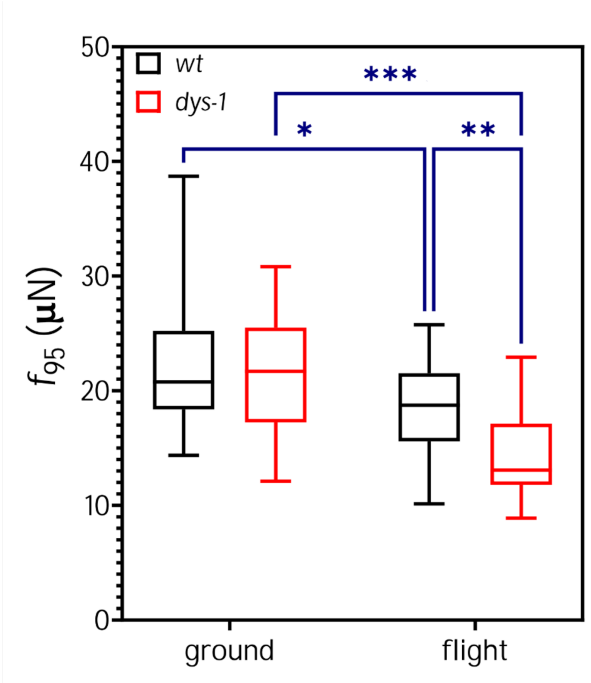

Figure 5.3. Effect of spaceflight on muscle strength of wt and *dys-1* worms. There is no difference in muscle strength between the strains on the ground, and the muscle strength of space-grown worms decreased by 16.6% and 33.4% for wt and *dys-1*, respectively. Sample size: n=30 for wt ground and flight both, n=25 for *dys-1* ground, and n=29 for *dys-1* flight. All the data pass the normality test. Two-way ANOVA (Tukey multiple testing) was used for calculating significant differences, * for $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$. This figure was provided by Dr. Soni Purushottam.

The muscle strength of the *wt* worms cultured on-board the ISS was significantly less than on Earth (ground: $f_{95} = 22.34 \pm 5.67$, flight $f_{95} = 18.62 \pm 4.05$, $n = 30$ per group, $P \leq 0.05$; **Figure 5.3**). Notably, the strength deficit (16%) is roughly similar to what is observed in Astronauts' muscles (6-14%, (English et al., 2020)) and similar to the previously published quantitative changes in muscle contractile protein during flight (7-10%, (Higashibata et al., 2006)). In flight, *dys-1* mutants were also weaker than on Earth (ground: $f_{95} = 21.38 \pm 5.39$, $n = 25$; flight

$f_{95} = 14.23 \pm 3.87,\ n = 29,\ P \leq 0.001$; **Figure 5.3**). It should be noted that the decrease in body diameter of both strains is similar in flight (see **Figure 5.2** and **Table 5.1**), indicating the strength decrement is not due to difference in the measurement technique. The strength in the *dys-1* mutants was 23% less than *wt* worms on-board the ISS. This may reflect the small sample size or a difference in the response to spaceflight in *dys-1* mutants vs. wild-type, as previously reported (Xu et al., 2018). It could be that the altered neuromuscular health of *dys-1* worms predisposes them to the negative effects of spaceflight on the neuromuscular system.

5.3.4. Gene Expression Analysis

Past gene expression analysis of space-flown *C. elegans* has revealed decreased expression of muscle cytoskeletal genes and mitochondrial genes (Higashibata et al., 2006, 2016; A Scott et al., 2023; Selch et al., 2008). It was previously shown that mutation of some of these muscle genes results in decreased strength (Etheridge et al., 2015) and that declines in mitochondrial function also result in decreased strength (Ellwood et al., 2021; Vintila et al., 2023), both in *C. elegans* on Earth. Therefore, to confirm that decreased expression of cytoskeletal genes or mitochondrial genes could be contributing to strength decline in flight, we measured gene expression was measured in the same cultures of worms that strength was measured in, as well as two additional cultures.

Unlike most past flights of *C. elegans* (A Scott et al., 2023), we employed an unbiased approach to the analysis of gene expression was pursued (e.g. not focussing solely on specific subsets of genes). As shown in **Figure 5.4A**, Principal Component Analysis (PCA) of the most variable genes showed a distinct clustering of samples in line with their experimental conditions. The tight clustering of samples within conditions indicates that the cultures without strength measures have similar gene expression profiles to the cultures from which strength measures were obtained. The first principal component (PC1), which accounted for 72.9% of the variance, separated the samples based on environment (i.e., flight vs. ground control). Meanwhile, the second principal component (PC2) differentiated the samples based on genetic background, distinguishing between *wt* and *dys-1* mutant, and explaining 13.5% of the variance. These results suggest that spaceflight is the main driver of the bulk of gene expression changes while the *dys-1* mutation has a lesser but significant effect. This is consistent with a past report demonstrating that *dys-1* mutation has an impact on the transcriptional response to spaceflight (Xu et al., 2018). The outlier genes

driving the variation in the PCA are shown in **Figure 5.4B**, the descriptions of these genes match the descriptions of the clusters of genes identified in the module analysis discussed later in the chapter).

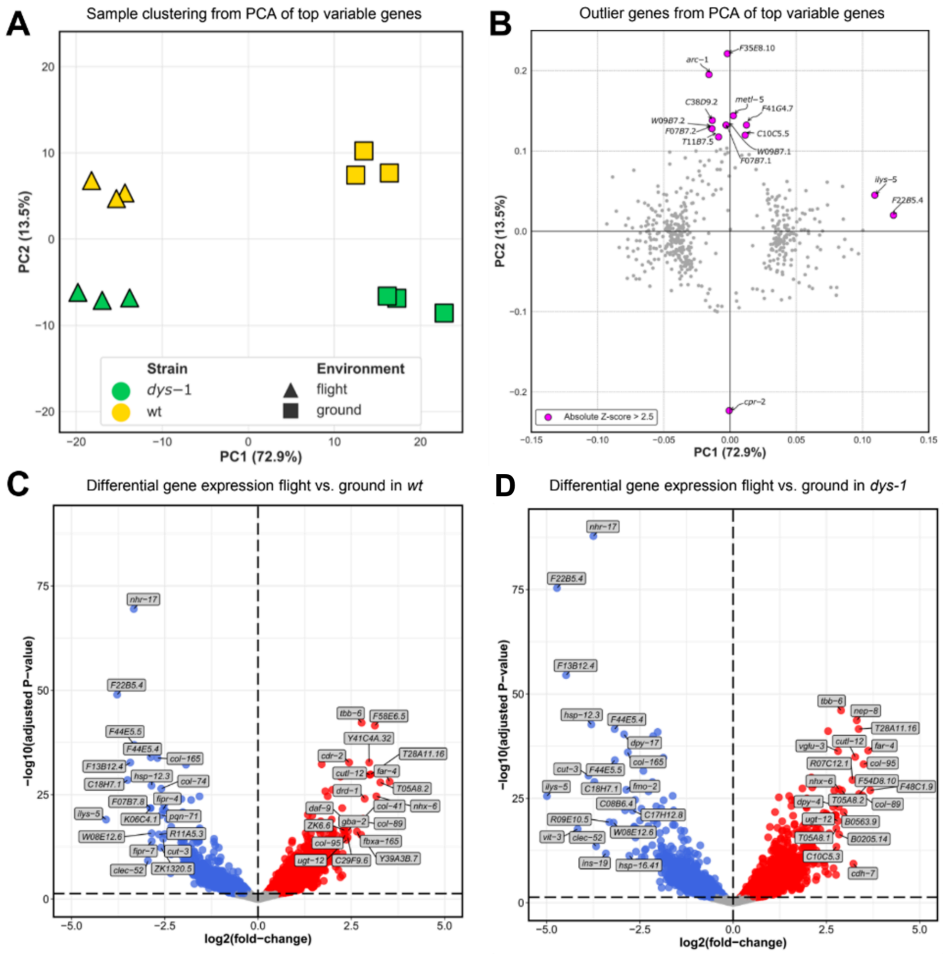

Figure 5.4. Global trends in spaceflight gene expression. A) Principal component analysis (PCA) clustering of samples based on top 500 most variable genes. **B)** PCA loadings of top 500 most variable genes. **C)** Volcano plot for wt flight vs. wt ground differential expression analysis, annotated genes are those ranked in the top 20 upregulated/downregulated based on test statistic. **D)** Volcano plot for *dys-1* flight vs. *dys-1* ground differential expression analysis, annotated genes are those ranked in the top 20 upregulated/downregulated based on test statistic. Analysis and formatting of Panel C were done by Dr. Craig Willis.

In terms of Differential Expressed Genes (DEGs) with spaceflight, the distribution of expression changes is largely similar between the two strains

as shown in **Figure 5.4C** and **Figure 5.4D**. While it is tempting to examine individual DEGs as meaningful, biological systems are complex and genes do not act in isolation (Kitano, 2002). Therefore, the full set of DEGs is provided (**Spreadsheet S1**) but the discussion focuses on sets of genes with changing expression rather than individual genes.In terms of genome level scale of changes with spaceflight (e.g. each strain's flight response normalised against each strain's ground control), 4726 genes displayed significant changes in expression in response to spaceflight with 409 uniquely up in *wt*, 394 uniquely down in *wt*, 1069 uniquely up in *dys-1*, 765 uniquely down in *dys-1*, and 1295 up in both strains and 819 down in both strains (**Figure 5.5A**).

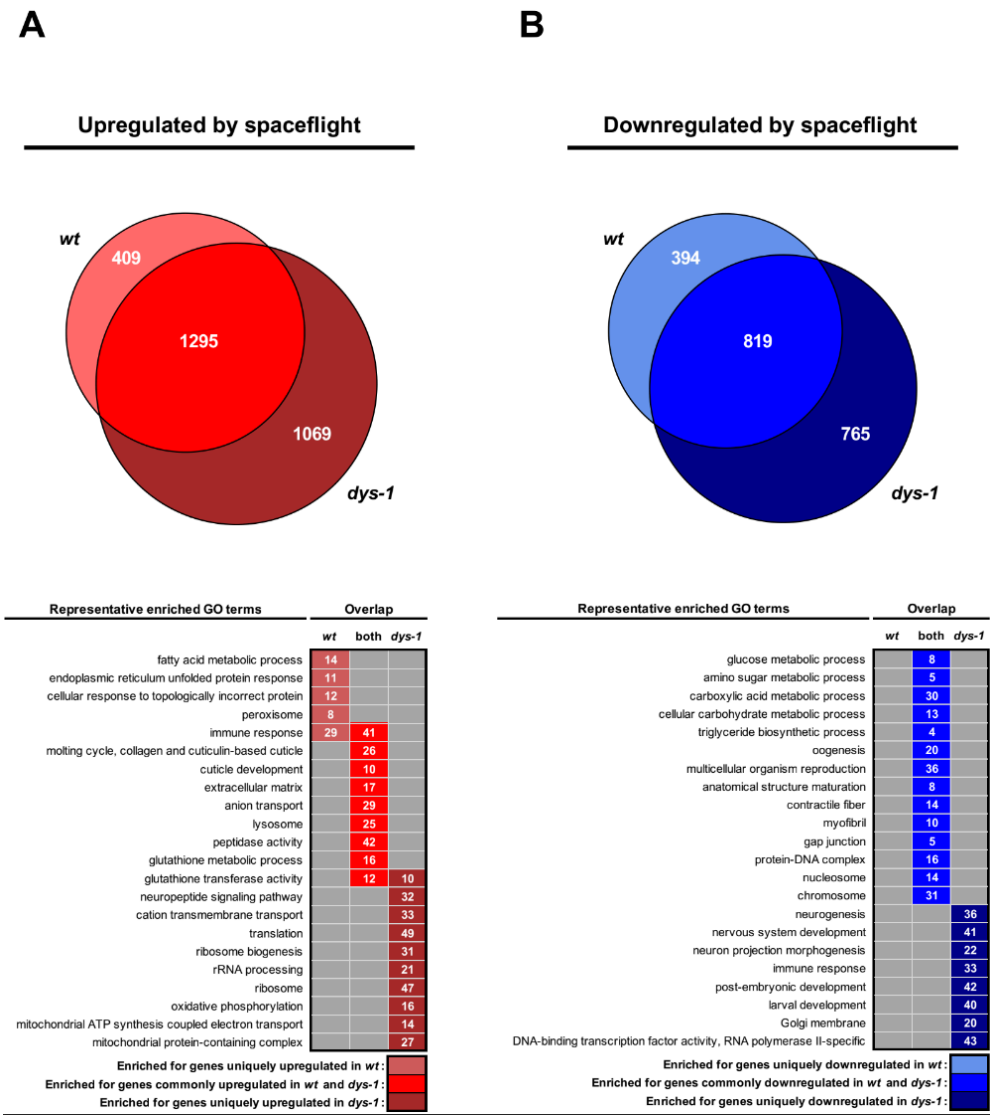

Figure 5.5. Overlay of wt and *dys-1* transcriptome responses to spaceflight. A) Overlay of genes upregulated by spaceflight in wt and/or *dys-1* worms. Venn diagram illustrates degree of commonality/uniqueness in spaceflight-upregulated genes between the two strains, while heatmap depicts representative Gene Ontology (GO) terms for common/uniquely upregulated genes. Up in flight. Venn shows the commonality and differential changes in wt vs *dys-1*. Table displays common and differential GO expression in wt vs *dys-1*. **B)** As per panel A, but for genes downregulated by spaceflight. This figure was provided by Dr. Craig Willis.

As shown in **Figure 5.5B**, gene ontology analysis of DEGs reveals changes in innate immune response as up in both strains, stress response as up in both strains, and metabolism and cytoskeleton as down in both. These changes are broadly similar to past analyses of space-flown *C. elegans* where stress response is up and metabolic and cytoskeletal genes are down (Higashibata et al., 2006, 2016; Honda et al., 2012; A Scott et al., 2023; Selch et al., 2008). These gene expression changes could underlie the strength decline in flight as strength goes down in both strains. These gene expression changes could also potentially underlie the decrease in muscle size that has recently been reported for other space-flown *C. elegans (Kim et al., 2023)*. Changes not likely to underlie strength decline in both strains in flight include upregulation of protein synthesis in *dys-1* in flight, and decreased neural development/function in *dys-1* in flight.

C. elegans was the first multicellular animal for which a transcriptional co-regulation map was generated (Kim et al., 2001). This map has recently been updated, employing advances in both DEG identification and analysis (Cary et al., 2020). When *wt* and *dys-1* DEGs are analysed for the effect of spaceflight vs. ground controls, three modules are inhibited in response to space flight and four modules are activated in both *wt* and *dys-1* (**Figure 5.6A**). Only one module is uniquely activated in *wt* whereas three are uniquely activated in *dys-1* (**Figure 5.6A**). These observations are consistent with spaceflight being the major driver of gene expression changes (**Figure 5.4A**). The modules inhibited in both include identifications (**Figure 5.6B**) consistent with past spaceflight experiments (A Scott et al., 2023). Module 65 was previously (Cary et al., 2020) specifically identified as responsive to spaceflight and other environmental conditions and may be controlled by the transcription factor DAF-16, amongst others, where DAF-16 has been suggested to be a controller of a *C. elegans* stress to spaceflight (Higashibata et al., 2016; Honda et al., 2012; Selch et al., 2008; Willis et al., 2020). Module 169 is a HIF-1 responsive module (Cary et al., 2020) and is

associated with mitochondrial function, this is consistent with both past *C. elegans* gene expression changes and the recent observation that decreased mitochondrial gene expression is a fundamental feature of biological alterations in response to spaceflight (da Silveira et al., 2020). Notably, this module may also be regulated by SKN-1 which has also previously been suggested to be a transcription factor regulating the response to spaceflight (Honda et al., 2012). Module 185 is associated with calcium handling in neurons and may be regulated by UNC-89 and DAF-19 (Cary et al., 2020). This module is particularly interesting following recent observations that neuronal morphology (Laranjeiro et al., 2021) and neurotransmitter production (Sudevan et al., 2022) are altered in *C. elegans* in flight. Similarly altered calcium handling in ageing *C. elegans* muscle has recently been shown to negatively impact mitochondrial health (Higashitani et al., 2023) and therefore might be an alternative mechanism by which mitochondrial gene expression declines in response to spaceflight. In terms of commonly activated modules, modules 47 and 118 are both responsive to mitochondrial stress (Cary et al., 2020) which is consistent with modules 65 and 169 being inhibited. Potential transcriptional regulators of these modules are displayed in (**Figure 5.6B**). Module 61 is associated with response to inhibited Acetylcholinesterase (Cary et al., 2020) which is consistent with Acetylcholinesterase gene expression being decreased in response to spaceflight in *C. elegans* (Honda et al., 2012). Module 151 is associated with desmosomal cell adhesion and calcium metabolism in the hypodermis (Cary et al., 2020). This result contrasts decreased cell adhesion in *C. elegans* (Higashibata et al., 2006, 2016) and human muscle (Murgia et al., 2022) in response to spaceflight. It could be that this response is unique to *C. elegans* as they have a hydrostatic skeleton that would be predicted to be altered by spaceflight, it could be related to alterations in global calcium homeostasis in *C. elegans* and the potential regulation of mitochondria via both calcium and cadherins in *C. elegan*s (Higashitani et al., 2023; Vintila et al., 2023), or it could be an over compensation of adhesome structures as a

response to failure as recently suggested for ageing human muscle (Deane et al., 2023). As with the DEG GO analysis, the module analysis identifies decreased expression of mitochondrial metabolism genes and increased expression of mitochondrial/other stress response genes as a potential cause of the decreased strength in flight. Unlike the DEG GO analysis, the module analysis also suggests decreased neuronal function and specifically altered acetylcholine signal from nerve to muscle (and consequent post-synaptic remodelling of excitation contraction coupling) as another potential contributor to strength decline in flight.

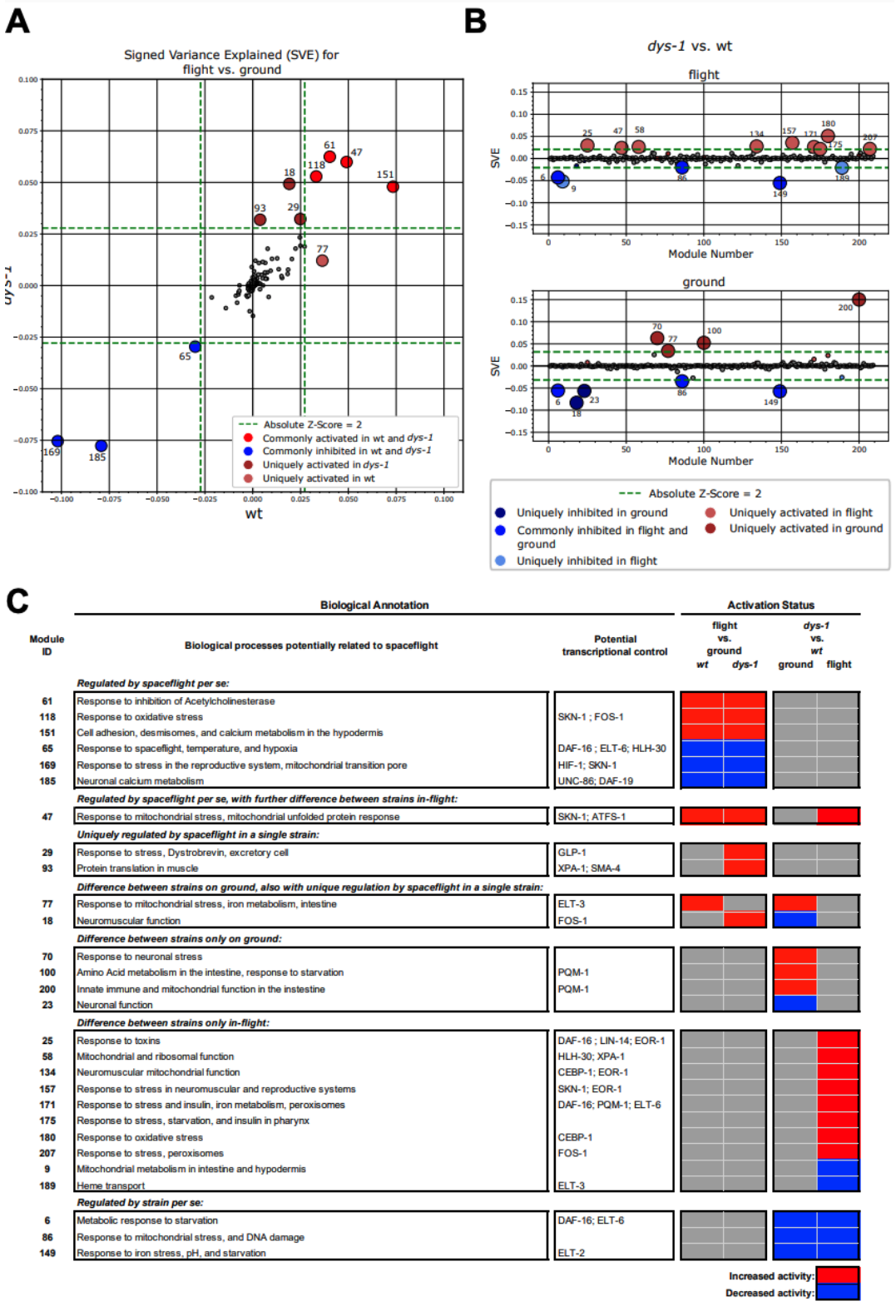

Figure 5.6. Cluster analysis to identify co-expressed gene modules. A) Activity of gene co-expression modules in flight vs ground control comparison. **B)** Activity of gene co-expression modules in *dys-1* vs wt comparisons. **C)** Table of module annotations and summarised activation

between pairwise comparisons. Panel C was designed by the thesis author but the final formatting of Panel C was done by Dr. Craig Willis.

Consistent with the past report that *dys-1* modulates the DEG in response to spaceflight (Xu et al., 2018) the DEG (**Figure 5.5**) and module analysis (**Figure 5.6**) confirm *dys-1* modulates the DEG response to spaceflight. As shown in Figure 6A, module 77 is uniquely activated in *wt* and is responsive to mitochondrial stress (Cary et al., 2020). Interestingly, module 77 is significantly activated in *dys-1* vs. *wt* on the ground (**Figure 5.6C**), potentially explaining the lack of further significant activation in *dys-1* in flight. Activation of this module at baseline is not surprising given the impact of *dys-1* on mitochondrial function (Ellwood et al., 2021, 2022; Hewitt et al., 2018) and gene expression (Hrach et al., 2020). Consistent with a baseline mitochondria stress response in *dys-1*, the modules uniquely activated in *dys-1* in response to flight include a stress responsive module that includes the dystrophin associated protein Dystrobrevin (29), a neuromuscular function module (18), and a protein translation in muscle module (93) (**Figure 5.6**). The activation of these modules suggests that the baseline neuromuscular dysfunction in *dys-1* is exacerbated by the added negative effect(s) of spaceflight on the neuromuscular system. Interestingly, three modules that are inhibited at baseline in *dys-1* do not alter in response to spaceflight (**Figure 5.6B & 5.6C**). These are module 6, metabolic response to starvation, module 86, response to mitochondrial stress, and module 149, response to iron stress, pH, and starvation (Cary et al., 2020). This, again, suggests that the baseline neuromuscular dysfunction in *dys-1* is exacerbated by the added negative effect(s) of spaceflight on the neuromuscular system. This may explain why *dys-1* are significantly weaker in flight than *wt* (small sample size being the other obvious reason).

5.3.5. Genes predicted to be altered in space are predicted to be altered by drugs on Earth

With the increased interest in precision medicine in space, the commercialisation of space, and the rise of synthetic biology, astropharmacy has recently emerged as a recognized field of specialisation (Aziz et al., 2022; Sawyers et al., 2022). The goals of astropharmacy mirror those of pharmacy on Earth but with a specific focus on medication storage, access, use, and manufacturing beyond Earth (Aziz et al., 2022). Current challenges include understanding how detrimental physiologic changes in flight might be countered by existing medications, how medications may work differently in flight and how to ensure our explorers receive the best possible outcome while minimising side effects despite limited resources. Having found a number of gene expression changes that were conserved with past missions (Higashibata et al., 2006, 2016; Honda et al., 2012; A Scott et al., 2023; Selch et al., 2008; Sudevan et al., 2022). and profiles indicative of pathways that medications target (e.g. mitochondria, metabolism, neurotransmitters), the potential of whether drug target predictions could be used to identify compounds that could work better or worse in flight based upon underlying gene expression changes was explored. As shown in **Figure 5.7**, QIAGEN Ingenuity Pathway Analysis identifies 31 drugs as being regulators of genes, whose expression is either increased or decreased in flight. These results suggest, unsurprisingly, that pharmacological activity is likely to be affected resulting in alterations of medications efficacy and safety profile in flight, based upon underlying alterations in gene expression due to not only physiological changes in spaceflight but also further exacerbation by alterations in gene expressions due to medications such as on the list. Potential consequences include the need for dose adjustments, changes to classic first-line treatment options, and ultimately for drug selection to be optimised through pharmacogenomics and/or modelling. Clearly, this is an important area of

future research as we begin to live and work on other celestial bodies with a more diverse population with different comorbidities needing better understanding of astropharmacy.

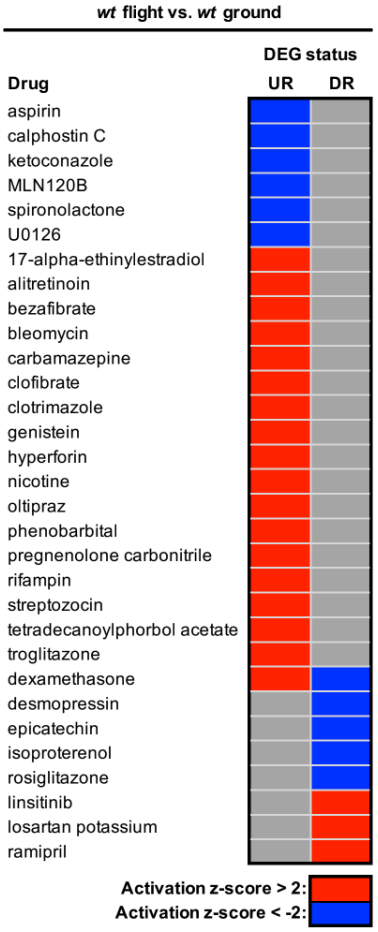

Figure 5.7. Heatmap of potentially relevant chemical and biologic drug targets. Targets are predicted to be significantly activated or inhibited for significantly upregulated (UR) and downregulated (DR) genes from the wt flight vs. wt ground condition, highlighting potential for alteration in therapeutic potential in the context of spaceflight-induced changes. This figure was designed by the thesis author but the final formatting of this figure was done by Dr. Craig Willis.

5.4. Discussion

Historically, space biology experiments have been constrained by mass, power, size, and crew time limitations (Ferranti et al., 2020). Now with a functional ISS, there is functional laboratory equipment, such as the

microscope used in this study and crew time for carrying out complex tasks, such as in this experiment or extracting and sequencing DNA on ISS (Stahl-Rommel et al., 2021). With these improved abilities to conduct science, this investigation was able to extend past studies using *C. elegans* in space to include capture not only of transcriptomic data (A Scott et al., 2023) but to study physiology more comprehensively than simple movement analysis (Soni et al., 2022). Using the NemaFlex-S worm strength was successfully measured both in flight and on the ground. Using these extended capabilities the hypothesis that worms, like people, are weaker in space was tested and confirmed. This demonstration extends our understanding of muscle response to spaceflight by demonstrating that not only are the molecular changes in muscle sarcomeric gene expression driven by altered MyoD expression conserved between worms and people (Higashibata et al., 2006) but so too are the functional consequences. These technical advances and increased knowledge may now be combined to test various nutrient and pharmaceutical interventions to improve muscle strength in flight as recently demonstrated for DMD and ageing muscle on Earth (Ellwood et al., 2021; Vintila et al., 2023). This experiment presents an example of how near-simultaneous in-flight collection of -omics alongside phenotypic measures can enable insights into the potential molecular mechanisms behind spaceflight-associated physiological deconditioning (Scott et al., 2020).

Response of gene expression to spaceflight has gained increased interest in the past few years, principally due to the implementation of NASA's GeneLab program (Ray et al., 2019). This has even resulted in introducing network biology into scientific roadmaps, for example ESA's Biology roadmap 9E (Angeloni et al., 2021). At the simplest level DEG analysis focuses on individual genes, while this may be useful it ignores the power of big data such as reproducibility across experiments and systems and may be biassed based upon differences in computational approach/pipeline

(Arora et al., 2020). For example, the gene identified as most down regulated in this study was previously identified as one of the most up regulated in the first unbiased whole genome response of *C. elegans* to flight (Selch et al., 2008); NB this difference could also be due to operational differences in the two missions. In contrast, the use of GO and network/module analysis has, once again (A Scott et al., 2023), revealed a consistent and reproducible change in gene expression in *C. elegans* in response to spaceflight. A key finding of this study is that strength decreases in flight. As discussed above, declines in mitochondrial gene expression are both repeatedly observed in response to spaceflight (A Scott et al., 2023) and are a treatable cause of strength decline on Earth (Ellwood et al., 2021; Vintila et al., 2023). Therefore, testing of mitochondrial interventions to reverse strength decline in space are now required to determine if the correlation between strength and mitochondrial decline is causal or not. Additionally, as our gene annotation and computational skills improve it is possible to get increased knowledge from gene expression data than in the past. For example, the module analysis reveals that certain gene expression changes can be attributed to specific tissues (for example cell adhesion in the hypodermis) whereas others are more universal (for example mitochondrial stress). Further, this new analysis reveals that calcium metabolism in neurons is also a potential cause of strength decline in space, more specifically by altering acetylcholine signalling, again suggesting an interventional study is required to (dis)prove casualty of this link. This suggestion, unsurprisingly, has also previously been made for human strength decline in flight (Lee et al., 2022).

As we confront the new age of commercial spaceflight it remains unclear how experiments such as this one will be possible on as yet unbuilt commercial space stations. The largest space agencies (the Japanese Space Exploration Agency, the European Space Agency, and the National Aeronautics and Space Administration) have not announced any plans for funding scientists to conduct research on such platforms. Given that the US

National Lab program relies on scientists to provide their own funding it seems current US policy may continue and there will be no funding for academic research on commercial space stations provided by the US government. Clearly this is an opportunity for smaller agencies such as the United Kingdom, the Australian, the Italian, and the United Arab Emirates Space Agencies to grow their research portfolio by continuing to purchase commercial access to these new space stations as they currently do for the ISS. However, for the larger agencies it is clear that the goal is moving BLEO (Everroad et al., 2021). For example, NASA's Thriving in Deep Space initiative, the Artemis program, and the commercial lunar payload services (CLPS) program. In this return to the Moon push the Artemis program is effectively a return to the Apollo program for space biology which means mass, size, power, and crew time will all be limiting to experiment design again. Similarly, the CLPS program is uncrewed and currently without sample return capabilities. For these reasons, autonomous experiments with data received by telemetry are essential, for example the recently selected Lunar Explorer Instrument for space biology Applications (LEIA) payload. Cube-sats provide an ideal example of how to achieve this with more than 1,500 having previously flown and Biosentinal having already demonstrated that biology experiments beyond low Earth orbit can be successfully conducted (Massaro Tieze et al., 2023). Conducting an experiment such as this one on the Moon or another BLEO destination is technically feasible. Worms have previously been autonomously cultured on ISS for six months with video data returned via telemetry (Oczypok et al., 2012). Thus the key challenges to testing worm strength on the Moon are designing culture chambers and strength measurement chambers to function in a cube-sat platform. Currently, funding for development of hardware for use on the Moon is a challenge via the largest space agencies. Additionally, late load for biology payloads is a key requirement that is currently not accommodated by CLPS so consideration of inert experiments that can be activated in flight is key. This is possible with *C. elegans* where

developmentally arrested larvae have previously been restored to normal development once on orbit (Higashitani et al., 2009). With new destinations for exploration come new challenges for space biology. However, cutting edge experiments such as conducted here with *C. elegans* on ISS are within technical reality for execution on the Moon or beyond.

5.5. Conclusion

This study extends the growing literature base solidifying the general molecular features underpinning spaceflight-related health decline. By directly linking neuromuscular strength loss with increased stress response and reduced mitochondrial / cytoskeletal gene pathways, this study demonstrates a robust framework in which to develop targeted therapeutics against a primary maladaptation to space habitation. The informatic pipeline, combined with use of *dys-1* muscle weakness mutants, further highlights perturbed calcium handling and acetylcholine signalling within neurons as primary candidates causing impaired neuromuscular strength in space. These findings provide the foundations for, and *in vivo* model of, space-induced strength loss to employ on near-term missions beyond low Earth orbit and to the Moon.

5.6. Materials and Methods

The following supporting information for Chapter V can be downloaded via the associated publication (Soni et al., 2023): Table S5.1 - RNA-Seq_analysis_data.

5.6.1. Worm preparation

Wild-type (*wt*) Bristol isolate (N2), and dystrophin (*dys-1*) mutant (BZ33) worms were maintained in *C. elegans* Maintenance Medium (CeMM) (Szewczyk et al., 2003), purchased on contract from Cell Guidance Systems,

Cambridge, UK. Cultures were established as previously described (Soni et al., 2022) and maintained at two separate sites (Texas Tech University, USA and University of Nottingham, UK). 10 days before the launch, approximately 1000 larvae from a stock culture were transferred into a Fluorinated Ethylene Propylene (FEP) bag (Saint Gobain Performance Plastics Corporation, Ohio, USA) containing 20 mL CeMM at Texas Tech University (TTU). This procedure was conducted by collecting three 100 µl aliquots from the stock culture, using a dissecting microscope to count the worms per aliquot, averaging this to get the number of worms per 100 µl, and then scaling up to transfer the volume needed to reach approximately 1000 larvae in the FEP. These primary culture bags and CeMM-Filled FEP bags were shipped to the Eastern Virginia Medical School (EVMS), Norfolk, on 10th February 2021. The culture bags were shipped with phase change material at a temperature of 16 ± 2°C. Upon arrival (following a lengthy delay of shipment in transit due to poor weather), culture bags were inspected for contamination and stored in the incubator at 22°C for 2 days (flight bags) and 4 days (ground bags). The flight bags were handed over to NASA cold stowage on 17th February 2021 at a temperature of 15 ± 1°C for the 20th February 2021 NG-15 launch. Ground control bags were sent to TTU, and the temperature profile of flight bags was replicated on a time delay of two days (Figure 1). Upon arrival at the ISS, the culture bags were incubated at 20°C, and the CeMM bags were stored at 4°C. Cultures were incubated for a week in microgravity before initiating the multigenerational culturing outlined in (Soni et al., 2022). All experiments used a starting mixed population of well-fed animals with force measurements only being made on gravid adults. Upon experiment completion, ground bags at TTU were frozen and stored at -80°C, while flight bags were frozen in the ISS Minus Eighty Degree Laboratory Freezer for ISS (MELFI) at -80°C until download to Earth via NASA cold stowage at -20°C, followed by dry ice transport to TTU for storage at -80°C.

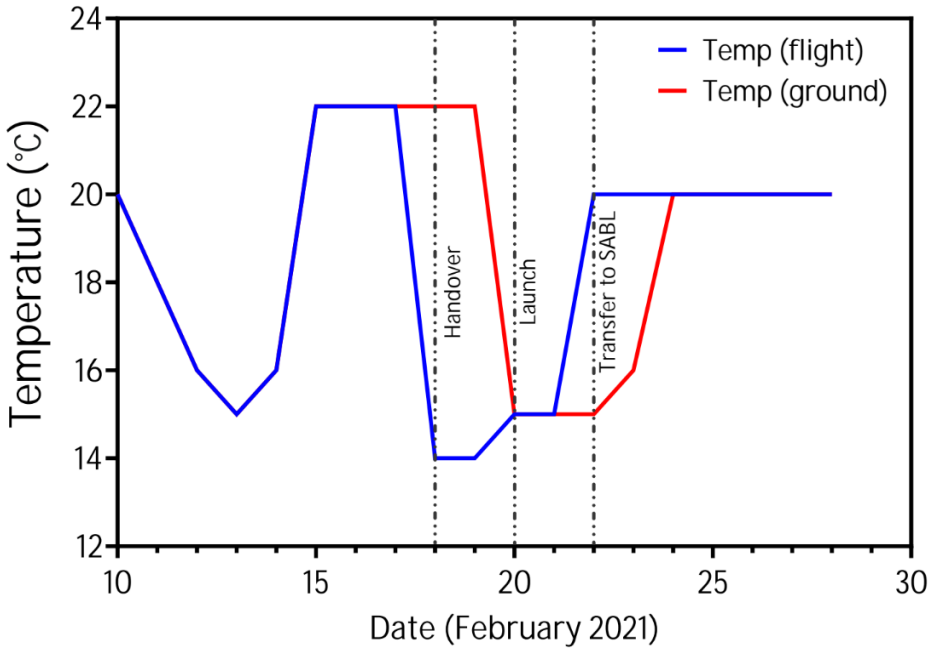

Figure 5.8. Detailed temperature profile of the culture bags during shipping from Texas Tech University (TTU) to Eastern Virginia Medical School (EVMS) and Launch. Bags were stored at 20°C after receiving them at the ISS Space Automated Bioproduct Laboratory (SABL) and during transfer to the TTU Lab. This figure was provided by Dr. Soni Purushottam.

5.6.2. Fabrication and assembly of a microfluidic device in Worm Loading Apparatus (WLA)

Devices were fabricated and assembled in WLA as previously described (Rahman et al., 2018; Soni et al., 2022). Fabrication of devices took place at TTU with devices shipped to EVMS. Fabrication of WLA took place at BioServe Space Technologies with WLA shipped to EVMS. Assembly took place at EVMS and was conducted by BioServe Space Technologies (Boulder, CO).

5.6.3. Image acquisition and processing

The worms were loaded into NemaFlex-S chambers using a previously described procedure (Soni et al., 2022). For on orbit loading, an additional manual centrifugation of worms toward the loading port of the FEP bags was achieved by having the crew member swing the bag in a circular motion with the injection port pointed away from the crew member. Additionally, loading volumes on orbit were calibrated based upon results of a training time point prior to the experimental time point and based upon crew member impressions of culture density for the experimental time point vs. the training time point. Once loaded, worms were allowed to habituate to the micropillar arena for approximately 10 minutes before imaging. One-minute-long videos of crawling worms were acquired with a Nikon inverted microscope (Eclipse TS 100) at 40X magnification with a camera resolution of 1920×1080 pixels recorded at 50 frames per second. All videos were recorded at a temperature of 22 ± 2°C. The recorded movies were processed offline using custom routines written in MATLAB (Mathworks, R2018b) for the quantification of pillar displacements as previously published (Rahman et al., 2018). Recorded videos were analysed manually using ImageJ 1.48v for measuring worm diameters at the mid-section and body length at the centerline of the worm. Videos were processed manually for quantification of coiling phenotyping. Adult worms were identified based on their body size and the presence of eggs.

5.6.4. RNA extraction, sequencing and data pre-processing

RNA was extracted from frozen worm samples using Direct-zol RNA Miniprep kit (Zymo Research cat# R2050). Two to three independent biological replicates were prepared for each treatment. RNA concentration was determined using Nanodrop 2000 Spectrophotometer (Thermo Fisher

Scientific). Library preparation and next generation sequencing was subsequently performed by the Beijing Genomics Institute (BGI, Hong Kong), with strand-specific (second strand cDNA synthesis with dUTP) 100 bp paired-end reads generated using the DNBseq platform. Cleaned reads (reads with adaptor sequences, contamination and low-quality reads removed via the SOAPnuke software developed by BGI (Chen et al., 2018)) obtained from BGI were deemed to be of good quality (no over-represented sequences or adapter sequences and median per base quality scores always > 30, as determined using FastQC; Babraham Bioinformatics) and transcript-level abundances consequently estimated via pseudo-alignment to the *C. elegans* reference transcriptome (Ensembl release 108) using Kallisto (version 0.48.0; (Bray et al., 2016)). Gene counts were then inferred via the tximport R package (version 1.28.0, (Soneson et al., 2015)) and lowly expressed genes filtered out (genes with a count < 10 in every sample) to leave 13,897 genes for downstream analyses.

5.6.5. Gene expression analysis

Differential gene expression analysis was performed via DESeq2 (version 1.40.2, (Love et al., 2014)) in R (version 4.3.1). Beforehand, principal component analysis (PCA) of the top 500 most variable genes (with variance stabilising transformed counts used as input) was undertaken for unsupervised clustering of samples. Wald tests were then used to test for differential gene expression, with pairwise comparisons made between flight and ground samples per strain, as well as between the ground samples of each strain and between the flight samples of each strain. Log fold-change shrinkage was performed using an adaptive shrinkage method (ashr) (Stephens, 2017) and the Benjamini-Hochberg procedure used to adjust *P* values to control for false discovery rate (FDR). Significant gene expression changes in each case were defined at the adjusted $P < 0.05$ level. Functional characteristics of differentially expressed gene lists were elucidated by

undertaking over-representation analysis of Gene Ontology (GO) terms using the clusterProfiler R package (Wu et al., 2021). In which case, each GO sub-category (biological process, cellular component, molecular function) was considered, with the corresponding background gene list being the genes input into differential expression testing. Enriched GO terms were defined as those with a Benjamini Hochberg corrected $P < 0.05$. Mapping of this gene expression data to established gene co-expression modules was conducted using the genemodules software for *C. elegans* (Cary et al., 2020). Tool 1 was used per pairwise comparison with $\log_2$ fold-change values of all genes subject to differential expression testing (i.e., the 13,897 genes that survived pre-filtering) as input in each case, to determine directional activity of the 209 defined transcriptional modules. Module descriptions potentially relevant to spaceflight and putative transcriptional regulators were subjectively assigned from Supplemental Table 2 (from (Cary et al., 2020)) and Tool 3. For drug target prediction, differentially expressed genes were first mapped to human orthologs, using OrthoList2 (Kim et al., 2018) with one-to-many mapping. Following this, the QIAGEN Ingenuity Pathway Analysis (IPA) (version 01-22-01) Upstream Regulator Analysis tool (Krämer et al., 2014) was used on the human orthologs of significant differential expressed genes (adjusted $P < 0.05$), with the human orthologs for all genes used as input into differential expression testing serving as the reference/background set. Targets were considered inhibited if the activation Z-score was < -2, and activated if the activation Z-score was > 2 (Krämer et al., 2014). Additionally, targets were only deemed significant if the Benjamini-Hochberg corrected P value was < 0.05.

Chapter VI: Data Analysis Case Study - Transcriptomic Analysis of Murine Skin

The contents from this chapter are published, in (Henry Cope et al., 2024):

Cope H*, *Elsborg J*, Demharter S, McDonald JT, Wernecke C, Parthasarathy H, Unadkat H, Chatrathi M, Claudio J, Reinsch S, et al. Transcriptomics analysis reveals molecular alterations underpinning spaceflight dermatology. Communications Medicine. 2024;4(1), p.106**.***

**These authors contributed equally to this work.*

All analysis and figures were done by the thesis author, with the exception of the clustering (**Figure 6.1B**) and machine learning models (**Figure 6.4C** and **Figure 6.4D**, **Figure 6.10B)** done by Mr. Jonas Elsborg, the astronaut physiological markers analysis and figure done by Dr. Afshin Beheshti (**Figure 6.11**), and the Inspiration4 data heatmaps done by Dr. Jiwoon Park (**Figure 6.5A, Figure 6.7B**, **Figure 6.6C**, and **Figure 6.8C**). In the case of the Inspiration4 heatmaps, due to data governance policies, gene lists were sent to Dr. Jiwoon Park by the thesis author, and heatmaps were produced and sent back. Similarly, for **Figures 6.6** and **Figure 6.8A**, the thesis author sent the code to produce heatmaps to Dr. Afshin Beheshti, who then executed the code to produce heatmaps from the data. The thesis author was responsible for formatting of the paper, and also responsible for a significant contribution to the writing and interpretation of results, which was shared with other collaborators (Particularly Dr. Pinar Avci and Mr. Jonas Elsborg).

This work is incorporated into this thesis in accordance with the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). Modifications have been made for its inclusion within this thesis, which include adjustments to figure, table, and section numbering, as well as the conversion of the text to British English.

6.1. Summary

Spaceflight poses a unique set of challenges to humans and the hostile spaceflight environment can induce a wide range of increased health risks, including dermatological issues. The biology driving the frequency of skin issues in astronauts is currently not well understood. To address this issue, a systems biology approach utilising NASA's Open Science Data Repository (OSDR) on space-flown murine transcriptomic datasets focused on the skin, biochemical profiles of 50 NASA astronauts and human transcriptomic datasets generated from blood and hair samples of Japan Aerospace Exploration Agency (JAXA) astronauts, as well as blood samples obtained from the NASA Twins Study, and skin and blood samples from the first civilian commercial mission, Inspiration4 was employed. Key biological changes related to skin health, DNA damage & repair, and mitochondrial dysregulation were identified as potential drivers for skin health risks during spaceflight. Additionally, a machine learning model was utilised to determine gene pairings associated with spaceflight response in the skin. While the findings appear to be indicative of spaceflight-induced dysregulation, such as alterations in genes associated with skin barrier function and collagen formation, the results also highlight the remarkable ability for organisms to re-adapt back to Earth via post-flight re-tuning of gene expression. The findings can guide future research on developing countermeasures for mitigating spaceflight-associated skin damage.

6.2. Introduction

Throughout the course of a spaceflight mission, the astronaut exposome includes altered gravity, elevated radiation, and confinement within a closed environment with unique hygiene procedures, light cycles, and ventilation (Afshinnekoo et al., 2020). These stressors perturb biological systems, inducing gene regulatory changes, mitochondrial dysregulation, microbiome shifts, and DNA damage (Afshinnekoo et al., 2020). Dermatological issues

are not typically regarded as a key risk to astronaut health and mission success, yet they are amongst the most common in-flight health issues reported by astronauts. During regular International Space Station (ISS) missions averaging 6-months, skin rashes have been identified as the most frequently reported in-flight clinical symptom, with 1.1 cases per flight year, accounting for 40% of all notable medical events and a 25-fold increase compared to the United States' general population (Crucian, Babiak-Vazquez, et al., 2016). An additional 0.3 cases per flight year include skin manifestations accompanied by symptoms of infection; skin lesions associated with viral reactivation were also reported and have been studied in-flight (Crucian, Babiak-Vazquez, et al., 2016; Mehta et al., 2022; Rooney et al., 2019). Notably, an in-flight skin rash was reported to have occurred during a 6-month ISS mission, worsening immediately after extravehicular activity (EVA) (Crucian, Johnston, et al., 2016), and a post-flight skin rash was reported following the 1-year-long NASA Twins spaceflight study (Law et al., 2020). These events arouse particular concern for future Moon and Mars missions, which will be longer, include high-levels of EVA, and also include risk of exposure to irritant dust, reported previously to induce skin issues in Apollo astronauts (Linnarsson et al., 2012). Elucidating the biological response of skin in space could aid development of new countermeasures to manage dermatological issues and optimise astronaut performance during these future missions.

As the human body's largest organ, the skin plays a crucial role in numerous essential health functions (McKnight et al., 2022). These include facilitating fluid diffusion, promoting wound healing, regulating body temperature, and enabling tactile sensation (McKnight et al., 2022). Skin is also the first line of defence against chemicals, allergens, infectious agents, and for spaceflight radiation (McKnight et al., 2022). Studies of the astronaut skin microbiome have identified microbial interchange between the skin and the ISS interior (Avila-Herrera et al., 2020), and have also hypothesised that abnormal

proliferation of certain types of opportunistic microorganisms on astronaut skin may stem from the unusual hygiene procedures on the ISS, where wipes are used as opposed to showering (Sugita et al., 2016; Xiao et al., 2019). While these microbiome shifts and their associated health effects require further studies, investigations into the molecular response of skin tissue in space are lacking. In murine skin, the 13-day STS-135 study and the 91-day Mouse Drawer System (MDS) study reported significant spaceflight-induced modulation of extracellular matrix (ECM) genes (Mao et al., 2014; Neutelings et al., 2015); the set of genes did not overlap, which could be due to significant differences in study design (e.g. duration). The MDS study also reported a 42% increase in synthesised procollagen coupled with a 15% reduction in dermal thickness likely indicative of increased collagen turnover, and an increase in hair follicles growing in the anagen stage accompanied by dysregulation of hair follicle genes (Neutelings et al., 2015). Corroborating gene regulatory changes associated with hair cycle were also reported in an analysis of hair follicle samples from 10 astronauts in a Japan Aerospace Exploration Agency (JAXA) study (Terada et al., 2016), yet reports of skin physiological changes in astronauts, including dermal atrophy, have been mixed (Braun et al., 2019; Heinrich, 2008). In a recent analysis of astronaut skin punch biopsies taken before and after the 3-day commercial Inspiration4 (i4) mission (Jones et al., 2024), gene expression changes within different layers of the skin were explored alongside microbiome changes, revealing signatures of increased inflammation and mitochondrial dysregulation across skin compartments, and gene regulatory changes associated with activation of DNA damage response (DDR) processes, T cell immunity, and increased barrier disruption in the outer epidermal layer of the skin (Dupras and Bunnik, 2021).

This study involved an analysis of five previously unreported murine skin RNA-Seq datasets from the JAXA Mouse Habitat Unit 2 (MHU-2) (Matsuda et al., 2019), NASA Rodent Research 5 (RR-5) and 7 (RR-7) experiments, to

identify global gene expression signatures in spaceflight-exposed skin and to investigate the effect of study design on biological signatures relating to skin health and common spaceflight themes including DNA damage and repair, and mitochondrial dysregulation (Afshinnekoo et al., 2020). In addition to examining gene regulatory patterns, an explainable artificial intelligence (AI) modelling approach was used to construct interpretable machine learning models to identify synergistic effects between pairs of genes, which can be interpreted as possible biological interactions, that reveal putative disequilibria of dependent processes. Gene regulatory changes observed in the rodent skin data and in astronaut data, including JAXA hair follicle data (Terada et al., 2016), skin data from the i4 mission, and blood gene expression data from the NASA Twins (Garrett-Bakelman et al., 2019), JAXA Cell-Free Epigenome (CFE), and i4 studies were also investigated. The chapter concludes with how these molecular signatures may eventually lead to follow-up studies and pharmaceutical interventions.

6.3. Results

6.3.1. Spaceflight transcriptome global changes in rodent skin reflect common biological hallmarks of spaceflight

To explore whether transcriptomic changes in the skin would occur during spaceflight, five RNA-Seq datasets from the NASA Open Science Data Repository (OSDR) (Ray et al., 2019), all representing skin tissue from space-flown mice and matching ground control replicates. These datasets are derived from three different spaceflight missions (**Figure 6.1A**). The most variable genes across the datasets cluster into functional groups related to established physiological risks of spaceflight (**Figure 6.1B**). For example, spaceflight is known to induce immune dysfunction (Afshinnekoo et al., 2020) and Cluster 1 involves highly-correlated genes associated with immune response, including genes linked to modulation of Immunoglobulin

G (IgG) levels. Microgravity is also well established to cause degradation of muscle (Afshinnekoo et al., 2020), and Cluster 9 contains a significant number of genes related to muscular morphology and muscle disorders. Ionising radiation is known to cause renal and pulmonary fibrosis (Klaus et al., 2021; Murray, 1994) and profibrosis-like alterations were previously observed in lung tissue of mice returned from Space Shuttle Endeavour (STS-118) (Tian et al., 2010). Renal and lung fibrosis associated genes found in Cluster 7, and genes encoding proteins involved in sclerotic diseases grouped in Cluster 8, further supports these reports.

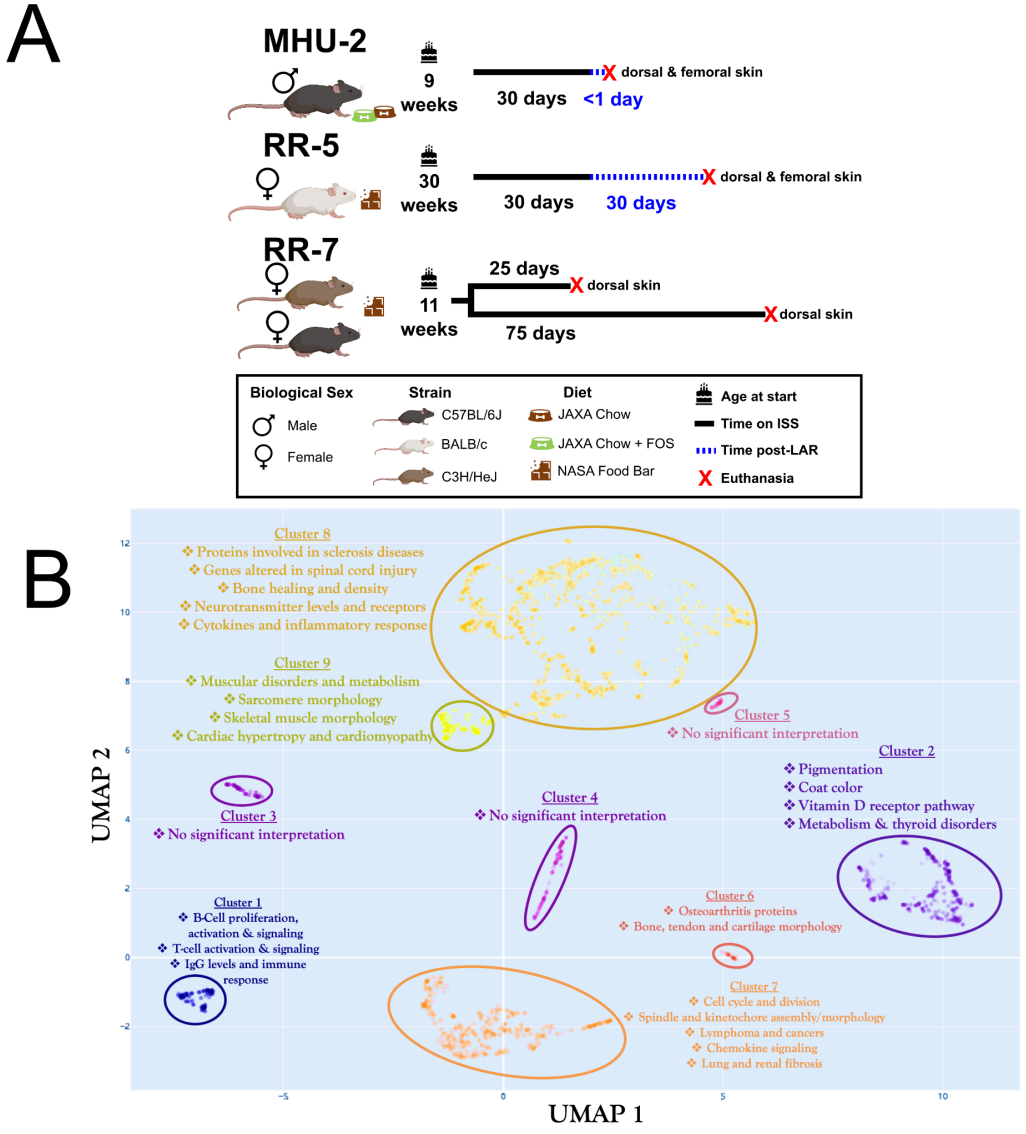

Figure 6.1. Global data overview. A) Breakdown of the rodent datasets used in this study. **B)** Clustering of the most variable genes within the rodent datasets with functional annotation. Panel B was provided by Mr. Jonas Elsborg.

Upon subsetting each dataset based on experimental groups (i.e. diet, strain, skin site) and performing differential gene expression (DGE) analysis for spaceflight samples versus ground control samples, resulting in 10 data subsets (**Figure 6.2A**). In total 476 unique differentially expressed genes (DEGs) are significant (False Discovery Rate (FDR) ≤ 0.1) in at least two out of the ten data subsets. One significant DEG shared across data subsets from the RR-7 study is *GLYCAM1* which is shared between both data subsets for the C57BL/6J strain. *GLYCAM1*, a ligand for L-selectin that has been proposed to modulate leukocyte transendothelial migration from blood (Howell et al., 2012), is strongly upregulated (Log_2 Fold Change (LFC) ≈ 18.81, FDR < 10^{-6}) at the 25 day time point and strongly downregulated (LFC ≈ -15.66, FDR < 0.002) at the 75 day time point, which could indicate a relationship between spaceflight duration and immune alterations for the C57BL/6J mice. Over representation analysis (ORA) reveals that a large proportion of the 244 significant DEGs shared between MHU-2 data subsets are associated with the organisation of collagen and the ECM (**Figure 6.2B**). Significant modulation of genes associated with ECM homeostasis was previously reported in analyses of the skin of space-flown mice (Mao et al., 2014; Neutelings et al., 2015). The MHU-2 subset from the dorsal skin of mice fed a standard JAXA chow diet with supplemental prebiotic fructooligosaccharides (FOS) contains the highest quantity of significant (FDR ≤ 0.1) DEGs by a substantial margin (**Figure 6.2A**). In ground-based rodent studies, FOS has been shown to improve gut microbiome balance, increase bone density, and affect the immune system through increased short chain fatty acid (SCFA) production (modifying interleukin production and natural killer cell activity), and modification of the immune system via gut-associated lymphoid tissue (Akiyama et al., 2020; Matsuda et al., 2019).

Due to this disparity between the number of significant DEGs in the data subsets, and the variety of conditions between data subsets, care must be taken when combining these data subsets to infer shared responses of rodent skin to spaceflight.

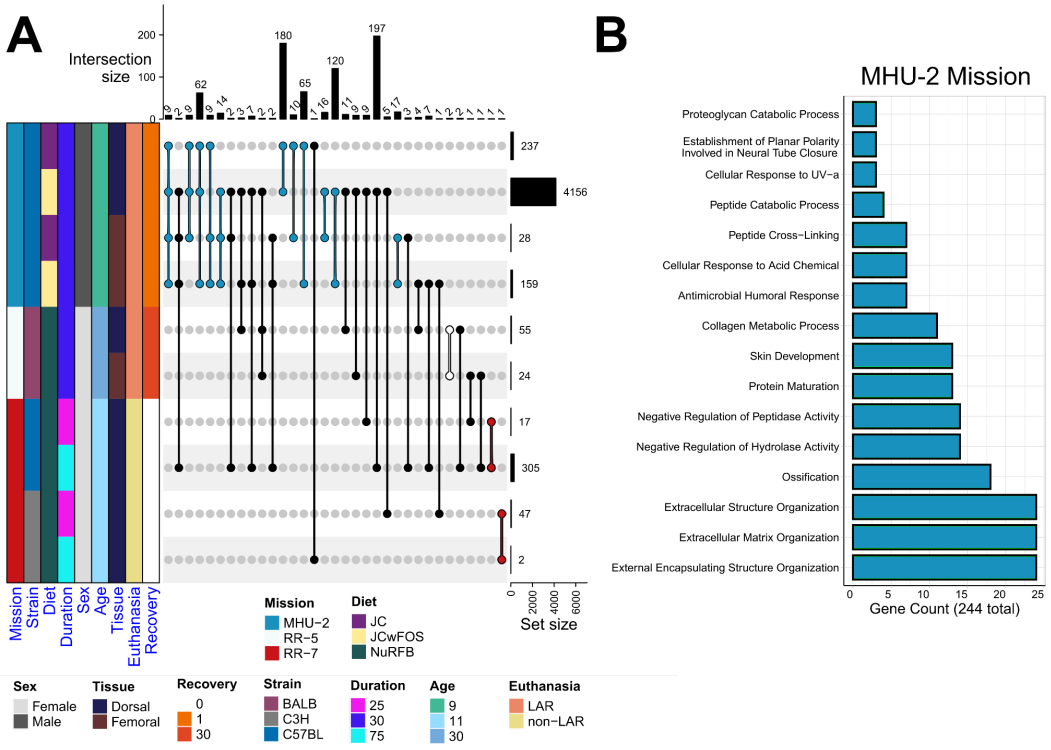

Figure 6.2. Differentially expressed genes shared between datasets. A) An upset plot showing the number of significant (False discovery rate (FDR) ≤ 0.1) differentially expressed genes (DEGs) in spaceflight versus ground data subsets and the number of overlapping significant DEGs between these data subsets. The coloured annotation bar on the left of the plot shows how the original datasets divide into 10 data subsets based on all unique conditions including diet, biological sex and strain. The bar plot on the right of the upset plot shows the number of significant DEGs in each of the 10 data subsets. The bar plot on the top shows the number of intersecting DEGs between combinations of the data subsets, as indicated by the connected dots within the body of the upset plot. Black connecting lines indicate combinations spanning across multiple missions, and other connecting lines are coloured according to the annotation bar, based on their shared mission. **B)** A bar plot of significant Gene Ontology Biological Process (GOBP) pathways from significant DEGs shared by any subsets within the MHU-2 mission (i.e., DEGs from the blue bars in panel A).

6.3.2. Genes significantly altered in rodent skin spaceflight response across multiple missions are associated with cell cycle processes

To address the previously mentioned bias in the MHU-2 dorsal skin data subset with a FOS-supplemented diet, and to delineate the common response of murine skin to spaceflight across multiple datasets, a list of 189 Differentially Expressed Genes (DEGs) was generated exhibiting significance (False Discovery Rate ≤ 0.1) in data subsets from at least two distinct missions. For the majority of the significant DEGs, there is a trend of downregulation in the MHU-2 dorsal skin dataset with the FOS-supplemented diet, and in the RR-7 mission C57BL/6J mice after 75 days of flight (**Figure 6.3)**. A smaller cluster of genes (e.g., *LAMA1, S1PR1, HMGCS2, TXNIP, IGFBP3, ADGRF5, CAR4, HSD17B11, C1QTNF9, KLF11*) are also significant in those same data subsets, but are significantly upregulated (**Figure 6.4A**).

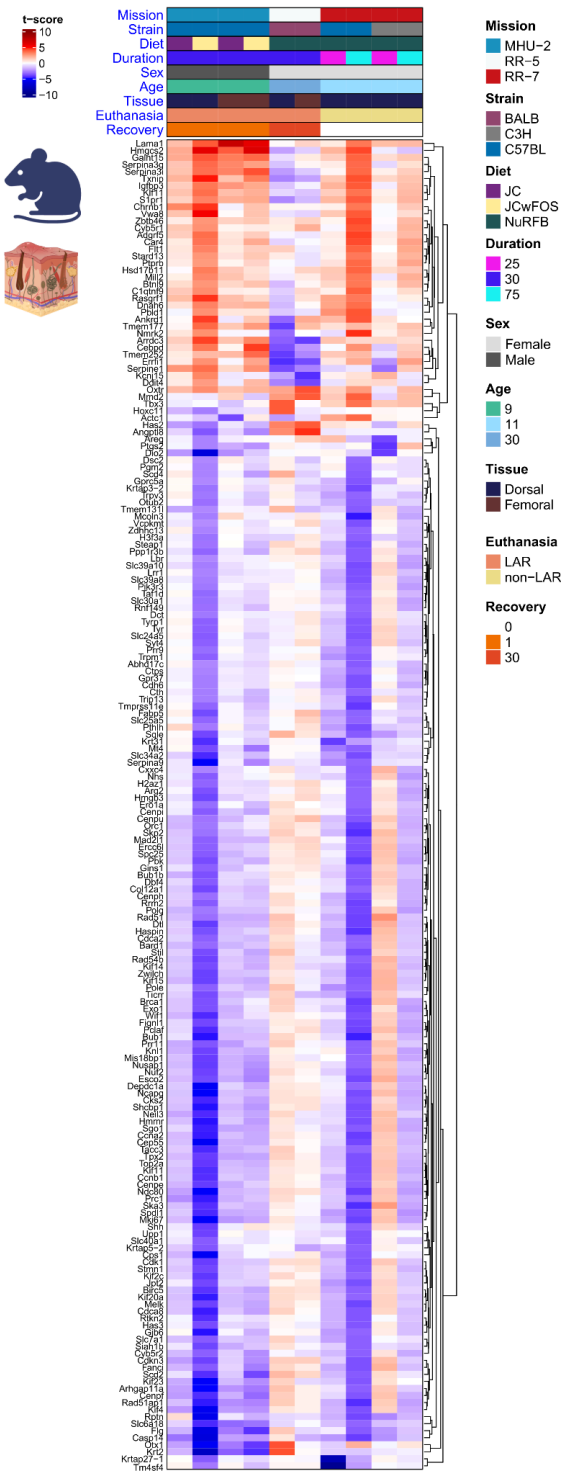

Figure 6.3. All cross-mission genes involved in rodent skin spaceflight response. A heatmap showing regulatory changes of all cross-mission genes (genes that are significantly (FDR ≤ 0.1) differentially expressed between flight and ground control across multiple missions) within each rodent data subset.

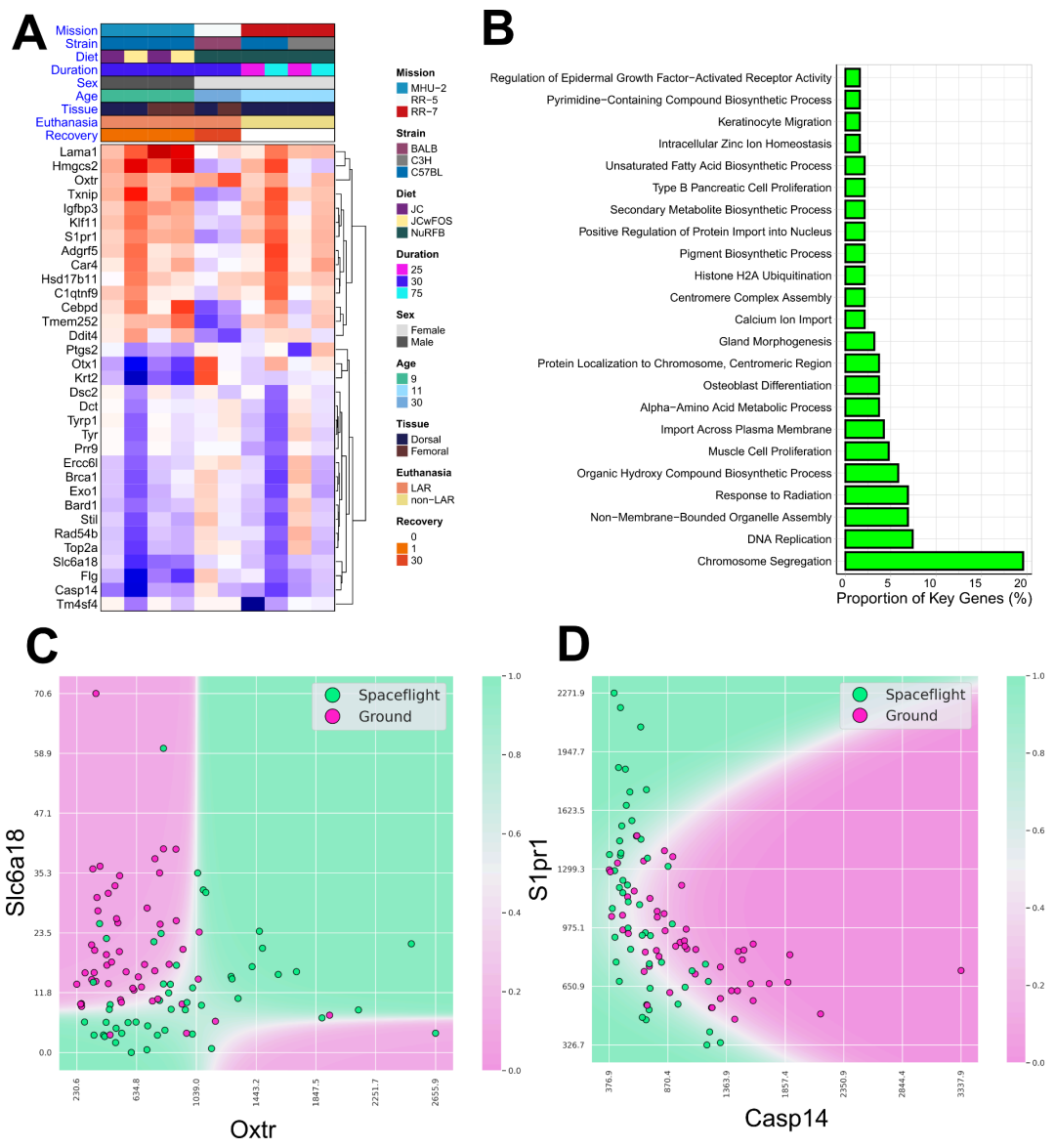

Figure 6.4. Cross-mission genes involved in rodent skin spaceflight response. A) A heatmap showing regulatory changes of select cross-mission genes (genes that are significantly (FDR ≤ 0.1) differentially expressed between flight and ground control across multiple missions) within each rodent data subset **B)** A barchart showing the GOBP pathways that the full list of cross-mission genes are involved in. **C)** A decision boundary plot of the strongest cross-mission QLattice model using only the expression of two genes as input features to predict the spaceflight status of the all mice in the data. **D)** A decision boundary plot of the second-strongest model across all missions. Panels C and D were provided by Mr. Jonas Elsborg.

To go beyond single-gene significance, QLattice modelling was used to construct models that predict the spaceflight status of all samples in the data (i.e. from all data subsets). These models used the expression of two genes at most to arrive at maximally relevant and minimally redundant models. In **Figure 6.4C** and **Figure 6.4D**, decision boundaries of the two strongest models found are visualised, with Area Under Curve (AUC) scores of 0.86 and 0.85, respectively. Both models show that the spaceflight status is well separated using only the expression of two genes as input, while also having plausible biological interpretations, as discussed in the following sections.

The 189 cross-mission DEGs are primarily involved in cell-cycle pathways (**Figure 6.4B**), with several of the genes (e.g. *LAMA1, C1QTNF9, HMGCS2, TXNIP, KLF11*) clustering with average trends of upregulation across the datasets (**Figure 6.4A**). Spaceflight is well-documented to perturb the cell-cycle (Afshinnekoo et al., 2020). These genes tend also to be involved in metabolic pathways, particularly those associated with diabetes; *HMGCS2* has been shown to regulate mitochondrial fatty acid oxidation (Vilà-Brau et al., 2011), while *LAMA1* and *KLF11* variations have been shown to be risk factors in type 2 diabetes (Neve et al., 2005; Perry et al., 2012). Aside from being a protector against oxidative stress, *TXNIP* is similarly implicated in metabolic diseases, and is typically upregulated in diabetic and prediabetic muscle tissue (Parikh et al., 2007). *C1QTNF9* (also known as *CTRP9*) has also emerged as a potentially important component of pathways involving lipid metabolism and adipose tissue, exemplified by the fact that *C1QTNF9* transgenic mice have been shown to resist diet-induced weight gain and related metabolic dysfunctions including insulin resistance (Peterson et al., 2013). It is interesting to note that upregulation of these metabolic pathway genes appears generally weaker for female mice. This might be linked to the same mechanisms offering higher protection for females against

diet-induced obesity compared to male mice, as has been noted in other studies (de Souza et al., 2022).

Sphingosine-1-phosphate (S1P)-Sphingosine-1-phosphate receptor (*S1PR*) axis have been implicated in pathogenesis of both inflammatory and autoimmune skin diseases and also plays a role in skin's immune response to viral infection (Skon et al., 2013). In both MHU-2 and RR-7 missions *S1PR1* is enriched, while this effect is reversed in RR-5 after 30-day recovery (**Figure 6.4A**). On the contrary *FLG* (Filaggrin), *CASP14*, *KRT2*, genes involved in cornified envelope and skin barrier formation(Jung et al., 2014; M Liu et al., 2020; Winge et al., 2011), generally showed an opposite trend in gene expression patterns (**Figure 6.4A**). *CASP14* is also the main driver of the second-most predictive 2-gene QLattice model obtained from the cross-mission key gene list, which involved a combined model of *CASP14* and *S1PR1* (**Figure 6.4D**).

OXTR, identified as a key driver in the most predictive 2-gene QLattice model derived from the cross-mission key gene list, along with *SLC6A18* (**Figure 6.4C**), is predominantly localised in the basal layer of the epidermis. Its primary expression is attributed to keratinocytes and dermal fibroblasts (Deing et al., 2013). *OXTR* together with oxytocin has been shown to mediate oxidative stress response in the skin (Deing et al., 2013). The analysis outlined within this thesis chapter revealed an increase in gene expression of *OXTR* across all missions and experimental conditions except for C3H/Hej mice in the RR-7 mission at the 25-day time point (**Figure 6.4A**). Both estrogens and androgens play an important role in skin and hair physiology. *HSD17B11,* which has a stronger expression in steroidogenic cells such as *sebaceous* glands*,* has been predicted to be involved in androgen metabolism during steroidogenesis, and conversion 5α-androstane-3α, 17β-diol to androsterone (Chai et al., 2003). Similar to the

pattern observed in *OXTR*, *HSD17B11* expression was increased in all murine datasets except for RR-5 dorsal skin group (**Figure 6.4A**).

6.3.3. Astronaut data correlates with select gene expression changes occurring across missions in the murine skin models

Having investigated a set of cross-mission genes regulated by spaceflight in rodent skin, the expression of human orthologs of these genes in astronaut studies was then investigated. Firstly, post-flight vs pre-flight gene expression data from skin samples collected during the i4 mission was looked at. Similar to the patterns observed in murine models (**Figure 6.4A**), *FLG* and *CASP14*, which play significant roles in skin barrier formation, were also downregulated in the epidermis of i4 astronauts (**Figure 6.5A**), where they are primarily expressed. In the literature, *FLG* is reported to mainly be expressed in stratum granulosum (Edqvist et al., 2015), with mutations associated with dry skin, atopic dermatitis, contact dermatitis as well as ichthyosis vulgaris (Lagrelius et al., 2020; Wallmeyer et al., 2017; Winge et al., 2011). *CASP14* on the other hand is known to be primarily expressed in differentiating and cornifying layers (Markiewicz et al., 2021) and decreased levels were detected in lesional skin biopsies of patients with atopic dermatitis and contact dermatitis (Hvid et al., 2011; Markiewicz et al., 2021). Likewise, *KRT2* which was downregulated across all subgroups in the MHU-2 and RR-7 missions (**Figure 6.4A**), was also downregulated in the skin of i4 astronauts, with strongest downregulation in outer epidermis where it is primarily expressed but also in inner epidermis and outer dermis (**Figure 6.5A**). Additionally, similarly to the trend of upregulation in the rodent data (**Figure 6.4A**), *OXTR* is upregulated across all compartments studied in the i4 skin (**Figure 6.5A**) and *HSD17B11* is upregulated in both compartments of the epidermis and in the outer dermis (**Figure 6.5A**).

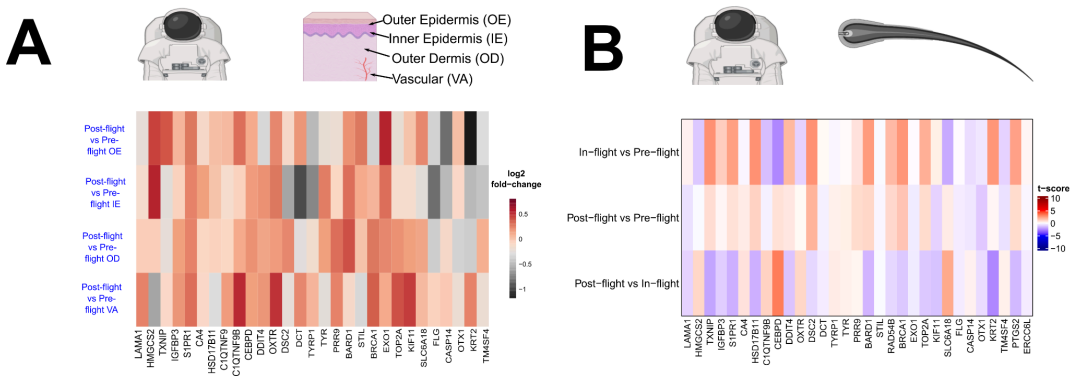

Figure 6.5. The profile of the cross-mission genes in astronauts. A) Heatmap showing average expression in orthologs of the rodent skin cross-mission genes in astronaut skin data from the Inspiration4 mission, for different skin layers. **B)** Heatmap showing $\log_2$ fold-change in orthologs of the rodent skin cross-mission genes in hair follicle samples at different time points from the JAXA HAIR astronaut study. Panel A was formatted by Dr. Jiwoon Park.

In some astronauts returning from long-term spaceflight, decrease in melanin content has been previously reported (Radstake et al., 2022). In the inner epidermis of the i4 skin data, where melanin synthesis takes place, there was downregulation of two genes associated with melanin biosynthesis, *TYRP1* and *DCT (Pavan and Sturm, 2019)* (**Figure 6.5A**). However, *TYR* which synthesises tyrosinase, the rate limiting enzyme of melanin synthesis was upregulated in the same compartment (Pavan and Sturm, 2019) (**Figure 6.5A**). *TYR, DCT,* and *TYRP1* clustered in the rodent data, with an average trend of downregulation, particularly in the MHU-2 dorsal skin dataset with the FOS-supplemented diet, and in the RR-7 mission C57BL/6J mice after 75 days of flight (**Figure 6.4A**).

Gene expression data from the blood of astronauts from the NASA Twins, JAXA CFE studies, and i4 studies was also investigated (**Figure 6.6**). In the NASA Twins Study, the most noteworthy results are in CD4 cell type in-flight vs pre-flight samples with significant downregulation in a group of genes (e.g. *DSC2*, *TMEM252*, *CEBPD*, *CA4*, and *PTGS2,* and *OTX1*) persisting

post-flight (**Figure 6.6A**). These genes display no significant changes in CD8 cells, which may indicate that their downregulation is a response to immune system stressors affecting the helper function and signalling of CD4 cells. *DDIT4* (DNA-damage-inducible transcript 4), *STIL* and *TOP2A* cell-cycle associated genes are amongst the most strongly upregulated genes in-flight compared to pre-flight in the CD8 cells, and exhibit downregulation post-flight to trend towards pre-flight levels (**Figure 6.6A**). In the CFE study, several genes were strongly upregulated in-flight and then downregulated post-flight (e.g., *OXTR*, *PRR9*, *ERCC6L*) in the blood samples from JAXA astronauts (**Figure 6.6B**). In the i4 mission blood samples the majority of genes show a pronounced suppression in the time point immediately before flight, which could be indicative of increased stress prior to launch. Several genes then show strong upregulation at the first post-flight time point (e.g., *BARD1, BRCA1, RAD54B, EXO1, OXTR,* and *HSD17B11)* (**Figure 6.6C**), suggesting upregulation due to spaceflight or landing stressors, followed by post-flight recovery towards baseline expression levels from second post-flight time point and beyond. However, some other genes appear to show a delayed upregulation following spaceflight, as they are upregulated in the second post-flight time point, two months after flight (e.g. *KLF11*, *CA4*, *OXT1*, *TM4SF4*) and return towards baseline at the third post-flight time point (**Figure 6.6C**). The gene expression alterations in these sets of genes are potentially due to a re-adaptation response.

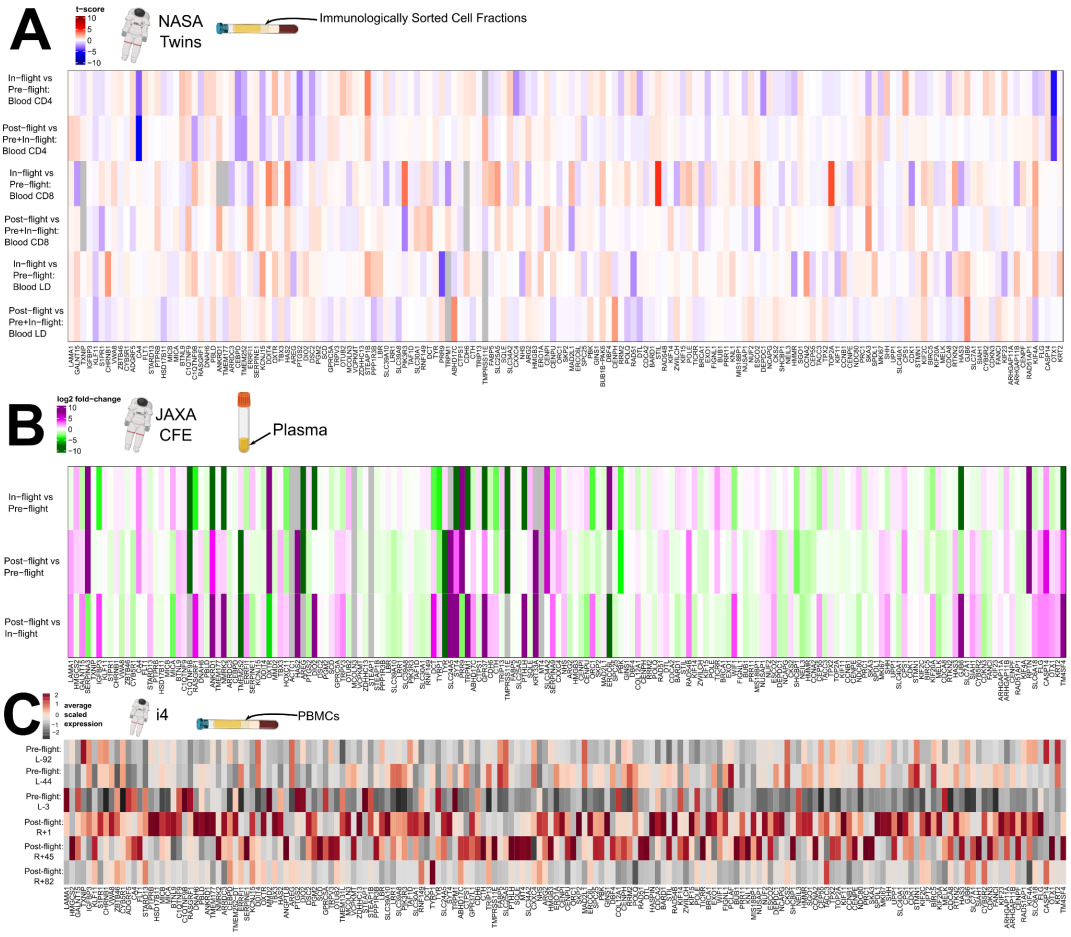

Figure 6.6. The profile of all the cross-mission genes in astronauts. A) Heatmap showing t-score in orthologs of the rodent skin cross-mission genes in astronaut blood samples from sorted cell fractions from the NASA Twin Study. **B)** Heatmap showing $\log_2$ fold-change in orthologs of the rodent skin cross-mission genes in plasma samples at different time points from the JAXA CFE astronaut study. **C)** Heatmap showing average scaled expression in orthologs of the rodent skin cross-mission genes in astronaut PBMC data from the Inspiration4 mission at different timepoints. Panel A was formatted by Dr. Afshin Beheshti, and Panel C was formatted by Dr. Jiwoon Park.

6.3.4. Specific pathways and genes related to skin health are altered in spaceflight

To determine direct relevance of spaceflight to skin health, a targeted analysis with a curated list of key pathways involved in skin health from The Molecular Signatures Database (MSigDB) was performed, identifying which

of these pathways and associated genes were significantly modulated in the murine skin datasets. A Collagen biosynthesis pathway and a set of collagen genes (e.g., *COL1A1, COL1A2, COL3A1, and COL5A1*) were significantly (FDR ≤ 0.1) suppressed/downregulated across all dataset subsets for the MHU-2 mission, while showing a trend of weak enrichment/upregulation across the two other missions (**Figure 6.7A**). Additionally, several collagen and/or ECM related genes (e.g., *MMP3*, *COL5A2*, *COL6A1*, *COL6A2*, *COL15A1*, *COL6A3, SPARC, PCOLCE,* and *PYCR1*) were significantly (FDR ≤ 0.1) downregulated in all of the MHU-2 data subsets except femoral skin with the standard diet, and a few collagen related genes were significantly downregulated only in the dorsal subsets from the MHU-2 mission (e.g., *P3H1*, *CRTAP*, *COL14A1,* and *FKBP14)* (**Figure 6.7A**). Overall, the distinct collagen/ECM downregulation signature (**Figure 6.7A**) suggests that a study design difference in the MHU-2 mission (compared to the RR-5 and RR-7 missions), where the use of young single-housed male mice or dissection shortly after the hypergravity event of Live Animal Return (LAR) may have suppressed collagen biosynthesis. Similarly, there was an overall trend of enrichment of pathways relating to thin skin and dermal atrophy in the RR-5 and RR-7 mice, with suppression occurring in the MHU-2 mice (**Figure 6.7A**), likely driven by regulatory patterns in the collagen-associated genes. While there is a general trend of upregulation, the RR-5 mission generally lacks any significant results for genes relating to skin health, which could be due to the 30 days of recovery post-flight. In addition to being reported as cross-mission genes in the previous section, *CASP14, FLG* and *KRT2* were also identified as key skin health genes (**Figure 6.7**).

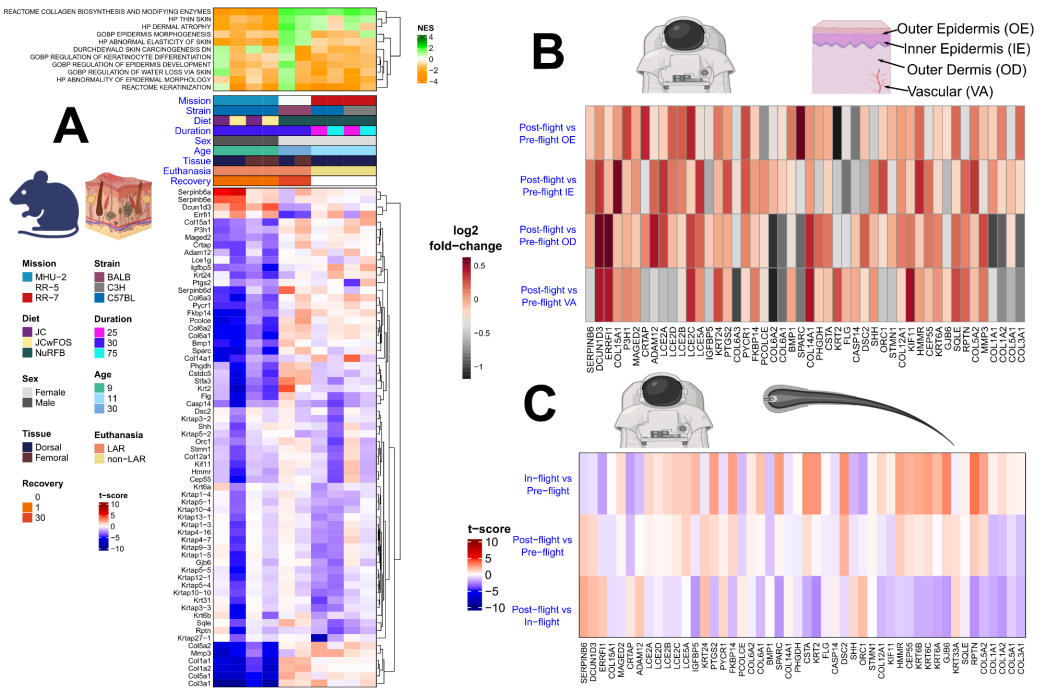

Figure 6.7. Behaviour of the specific genes associated with skin health in astronauts and rodents. A) The orange (suppressed) and green (enriched) heatmap shows the normalised enrichment score (NES) of curated skin health pathways (full list Table S6.1) that are significant (FDR ≤ 0.05) in at least one rodent subset. The red (upregulated) and blue (downregulated) heatmap shows the t-score for leading edge genes from the significant pathways that are significant (FDR ≤ 0.1) in at least two data subsets. **B)** Heatmap showing average expression in orthologs of the genes from Panel A in astronaut skin data from the Inspiration4 mission, for different skin layers. **C)** Heatmap showing $\log_2$ fold-change in orthologs of the genes from Panel A in hair follicle samples at different time points from the JAXA HAIR astronaut study. Panel B was formatted by Dr. Jiwoon Park.

In post-flight vs pre-flight skin samples from the i4 mission, several collagen/ECM genes (e.g., SPARC, *COL1A1*, *COL1A2*, *COL3A1*, *COL6A1*, *COL6A2*, and *COL6A3*) were downregulated across the whole dermis, whereas *COL5A1* was only downregulated in the vascular compartment of the dermis (**Figure 6.7B**). *MMP3*, which induces accelerated loss of skin collagen when there is inflammation, was upregulated across the whole dermal layer (Mirastschijski et al., 2019) (**Figure 6.7B**). Likewise, several

genes (e.g., *P3H1*, *COL14A1*, *COL15A1*, and *COLA12A1*) were upregulated in both compartments of the dermis (**Figure 6.7B**).

KRT24, involved in terminal differentiation of keratinocytes, is reported to be mainly confined to stratum spinosum with slight overexpression in senescent keratinocytes (Min et al., 2017) and has been shown to induce apoptosis of keratinocytes (Min et al., 2017). In the post-flight vs pre-flight i4 skin data *KRT24* is upregulated in all compartments of the skin with a more prominent increase in outer epidermis and vascular layers (Min et al., 2017) (**Figure 6.7B**). Late cornified envelope genes (LCE) are located in the epidermal differentiation complex on chromosome 1 and in the skin their expression is known to be mainly confined to the epidermis, specifically to upper stratum granulosum and premature cornified envelopes (Bergboer et al., 2011; Jackson et al., 2005). In both the inner epidermis and outer dermis of the i4 skin data, *LCE2A, LCE2B, LCE2C, LCE2D*, genes that respond to UV and calcium (Jackson et al., 2005) and *LCE5A* were upregulated post-flight vs pre-flight (**Figure 6.7B**).

In the NASA Twins study blood samples, *COL6A2* was upregulated in-flight and then downregulated post-flight in CD8 cells, with *PYCR1* following the same pattern in CD4 cells (**Figure 6.8A**). In the CFE study, blood samples from JAXA astronauts had strong in-flight upregulation of *RPTN* and *COL3A1*, and in particular *LCE2B*, where upregulation increased further post-flight (**Figure 6.8B**). On the other hand, *PYCR1* and *GJB6* showed the opposite trend in gene expression, with strong downregulation in-flight and almost complete reversal upon return (**Figure 6.8B**). *PYCR1,* which was upregulated post-flight vs pre-flight in all skin layers of i4 astronauts (**Figure 6.8C**) encodes pyrroline-5-carboxylate reductase 1, an enzyme that catalyses the last step in proline synthesis via utilisation of *NADH* or *NADPH*. Proline plays an important role in collagen synthesis and *PYCR1* deficiency results in collagen and elastin abnormalities (Kretz et al., 2011).

Conversely, *COL3A1* encodes alpha 1 chain of type III collagen, which is an EMP synthesised in cells as pre-collagen and its accumulation was shown to cause fibrosis in various organs (D'hondt et al., 2018; Kuivaniemi and Tromp, 2019). In case of *LCE2C* and *KRT33A*, weak in-flight downregulation was followed by strong post-flight upregulation (**Figure 6.8B**)**.** Although less apparent**,** this trend for *KRT33A* was also observed in the JAXA HAIR data (**Figure 6.7C**). *KRT33A*, one of the type 1 hair keratin genes, is known to be highly expressed in the cortex of hair follicles, and its expression is reported to decrease with age (Giesen et al., 2011). Lastly, genes such as *COL6A2* and *FKBP14,* a collagen/ECM related gene associated with diseases such as kyphoscoliotic Ehlers-Danlos syndrome that is characterised by joint hypermobility, hyperelastic skin, hearing impairment, muscle hypotonia and occasional vascular fragility like aortic rupture (Brady et al., 2017), were downregulated in-flight and their average expression persisted post-flight (**Figure 6.8B**).

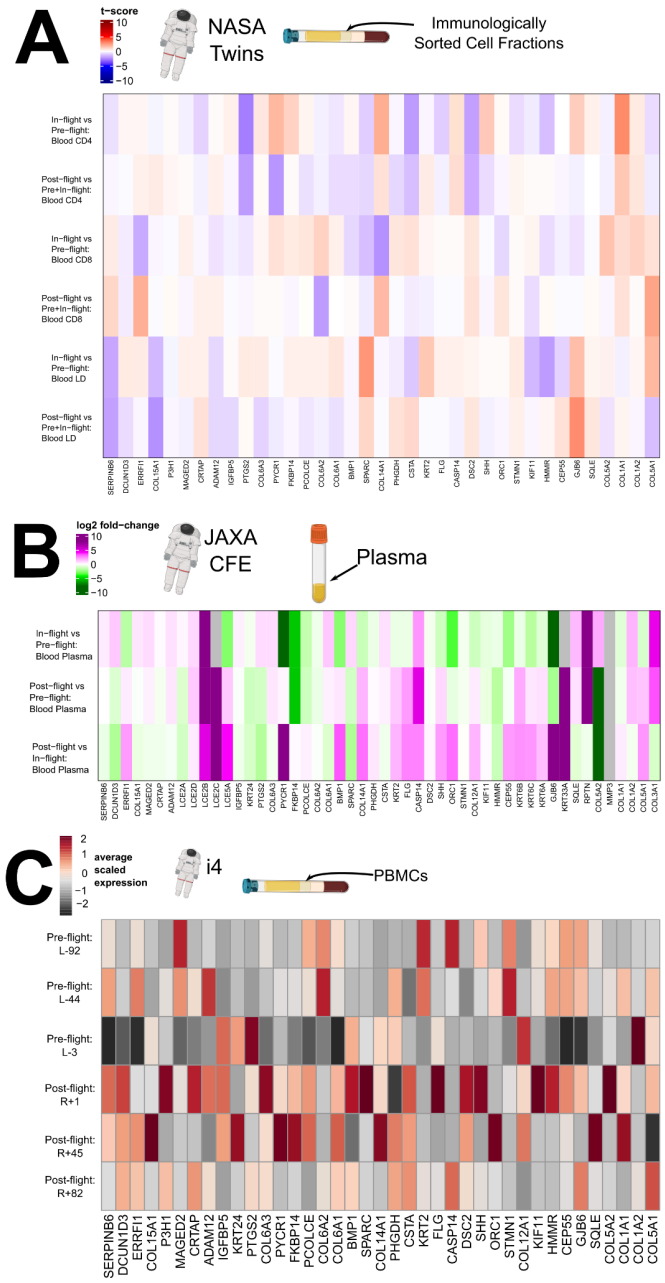

Figure 6.8. The profile of skin health genes in astronauts. A) Heatmap showing t-score in orthologs of the rodent skin, skin health genes (from **Figure 6.7**) in astronaut blood samples from sorted cell fractions from the NASA Twin Study. **B)** Heatmap showing $\log_2$ fold-change in orthologs of the rodent skin, skin health genes (from **Figure 6.7**) in plasma samples at different time points from the JAXA CFE astronaut study. **C)** Heatmap showing average scaled expression in orthologs of the rodent skin, skin health genes (from **Figure 6.7**) in astronaut PBMC data from the Inspiration4 mission at different timepoints. Panel A was formatted by Dr. Afshin Beheshti, and Panel C was formatted by Dr. Jiwoon Park.

6.3.5. Radiosensitivity of mouse strains impacts DNA damage response in the skin following spaceflight

As part of the targeted analysis, modulation of DNA damage and repair pathways in rodent skin was also investigated. DNA damage and repair is a well-established response to space radiation (Beheshti et al., 2021; Garrett-Bakelman et al., 2019) and little research is reported on the consequences this will have on the skin during spaceflight. When ionising radiation hits DNA molecules, single-strand breaks (SSBs) and double-strand breaks (DSBs) occur, with DDR mechanisms activated to repair these breaks (Beheshti et al., 2021). For the majority of DNA damage and repair pathways, both the dorsal and femoral skin RR-5 mission data subsets and the 25-day time point from the C3H/HeJ mice in the RR-7 mission show an opposing pattern compared to the other data subsets (**Figure 6.9**). In these three data subsets, pathways relating to DNA damage and repair are generally enriched, while being suppressed in other data subsets. BALB/c and C3H/HeJ mice, as used in these data subsets, have been shown to be more radiosensitive compared to C57BL/6J mice (Cekanaviciute et al., 2023; Mukherjee et al., 2014; Roderick, 1963), so repair mechanisms may be activated to mitigate increased radiation-induced DNA damage. In the C3H/HeJ mice there is a trend of decreasing activity for the DDR pathways from the 25-day to 75-day time points which could indicate either adaptation of the DNA repair pathways over time in space or dysregulation following extended-duration spaceflight (**Figure 6.9**). The enrichment of DDR pathways in RR-5 also follows 30-days of spaceflight and a post-flight recovery period of 30-days (**Figure 6.9**). However, DNA repair genes were reported to still be enriched compared to pre-flight levels when evaluated at 6-months post-flight during the NASA Twins study (Garrett-Bakelman et al., 2019), so persistent DDR activation during post-flight recovery in the mice could be indicative of a similar mechanism.

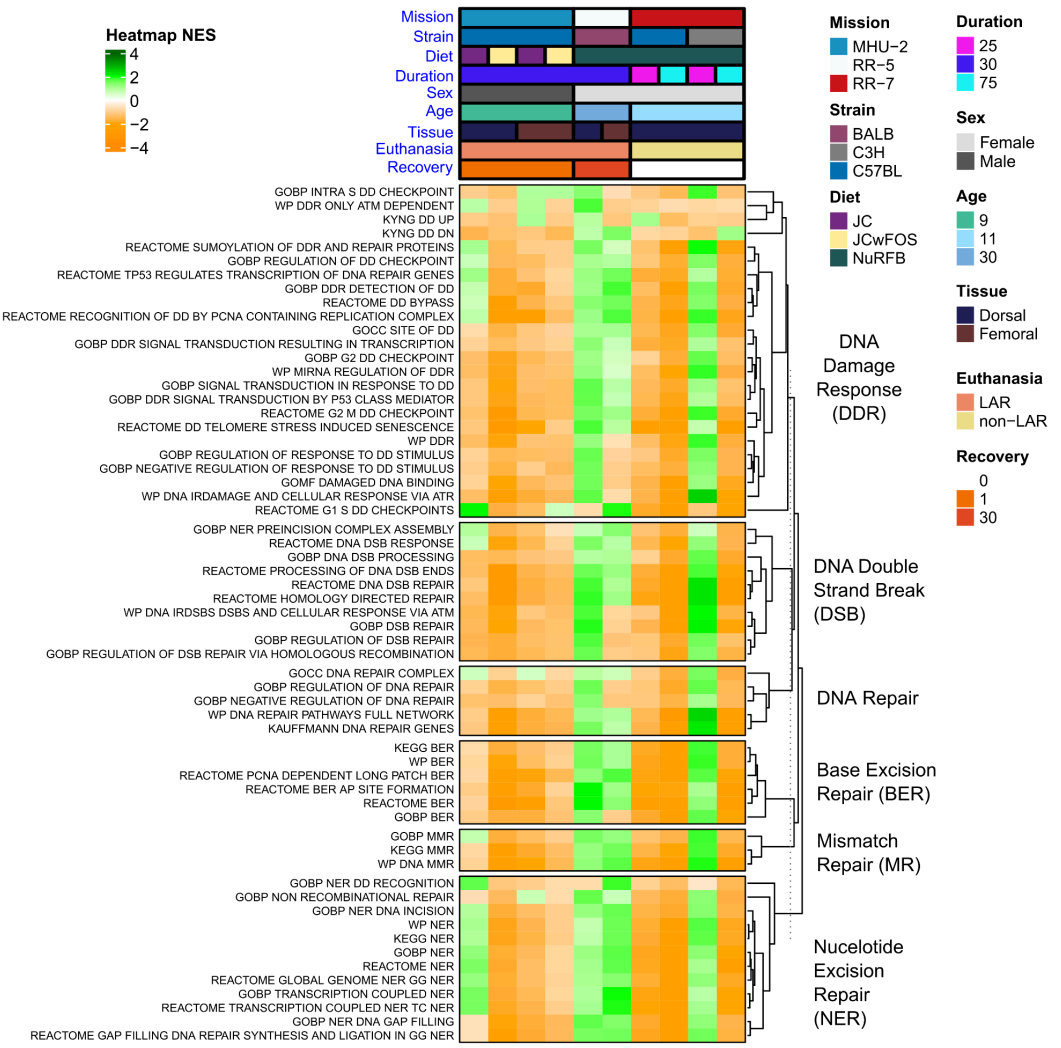

Figure 6.9. DNA damage and repair pathways being regulated in rodents flown to space. Heatmap of pathways relating to DNA damage response and repair mechanisms (full list Table S6.1), significant (FDR ≤ 0.05) in at least one data subset.

6.3.6. Mitochondrial dysregulation increased in the skin during spaceflight

Mitochondrial stress has been identified as a key hub for spaceflight response in multi-tissue analysis, yet skin was not included (da Silveira et al.,

2020). Skin is a tissue with high turnover and energy requirements (Sreedhar et al., 2020), motivating an investigation into spaceflight changes relating to mitochondrial stress in murine skin. It was observed from the data that spaceflight alters mitochondrial pathways in the skin (**Figure 6.10A**), as previously observed in other tissues and across species (da Silveira et al., 2020). Enrichment of an integrated stress response (ISR) pathway in all but two of the dataset subsets (**Figure 6.10A**) is consistent with a previous report of potentially activated ISR due to mitochondrial dysfunction in space-flown mice (da Silveira et al., 2020). Interestingly, the two data subsets where ISR appeared to instead be suppressed were the 25 day time points for both strains from the RR-7 mission (**Figure 6.10A**), implicating spaceflight duration as an important factor for modulating ISR. Additionally, the QLattice machine learning model was then run on the full gene set to obtain a set of models that were unbiased by feature selection due to e.g. variance filtering and the selection of cross-mission genes. One of the 10 strongest 2-feature models involved the expression of the genes *D2HGDH* (D-2-hydroxyglutarate dehydrogenase) (Achouri et al., 2004) and *RPLP0-PS1* (a gene coding for the ribosomal protein RPLP0). The model's decision boundary (**Figure 6.10B**) demonstrates how the simultaneous upregulation of *D2HGDH* and downregulation of *RPLP0-PS1* characterises the space-flown mice, with a resulting AUC score of 0.82 (Achouri et al., 2004). The upregulation of *D2HGDH* generally indicates an increased ability to break down the toxic D-2-hydroxyglutarate (D2H) compound in the mitochondria. In the case of the spaceflight data, this could be a compensatory mechanism due to larger build-up of the compound, and the model indicates that this is further dependent on the abundance of the RPLP0 protein.

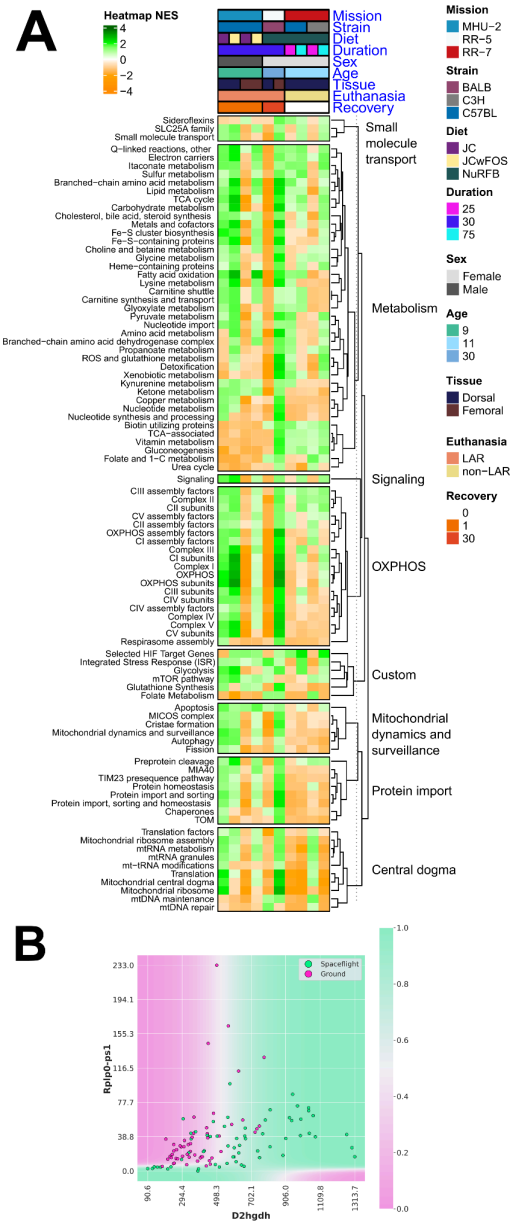

Figure 6.10. Mitochondrial specific analysis on rodent spaceflight skin tissue. A) Heatmap of pathways relating to the mitochondria, significant (FDR ≤ 0.05) in at least one data subset. **B)** Decision boundary for the 2-gene model related to mitochondrial changes. The model indicates an increased removal of the toxic D2H compound in mitochondria through upregulation of *D2HGDH*, which is less pronounced when *PPP1R3B* expression is suppressed. Panel B was provided by Mr. Jonas Elsborg.

6.3.7. Circulating physiological markers from astronauts indicate that exercise countermeasures may have improved skin health

The next stage of the investigation looked at standard physiological biomarkers collected from astronaut urine and blood (Smith et al., 2015; Smith, MA Heer, et al., 2012) to see whether average trends in these biomarkers connect to spaceflight gene regulatory changes occurring in the skin (**Figure 6.11**). In-flight increases in IGF-1, leptin and white blood cell levels may indicate altered stress response due to spaceflight. Following a hypothesis that improved countermeasures may improve skin physiological parameters in astronauts on the ISS (Braun et al., 2019; Heinrich, 2008), the data was split into astronauts that used the older Interim Resistive Exercise Device (iRED), and astronauts that used the newer Advanced Resistive Exercise Device (ARED). Notably, overlapping data to that shown in (**Figure 6.11**) have been used as evidence of improvements in parameters relating to bone mineral density (Smith, MA Heer, et al., 2012). Indeed, while vitamin D decreased in-flight and normalised upon return to Earth for both exercise devices, ARED appeared to reduce this drop. Vitamin D is well-established to be an important mediator of skin health (Mostafa and Hegazy, 2015). Notably, this decrease in vitamin D is observed despite consistent programmatic vitamin D supplementation since the collection of blood samples (Smith, MA Heer, et al., 2012).

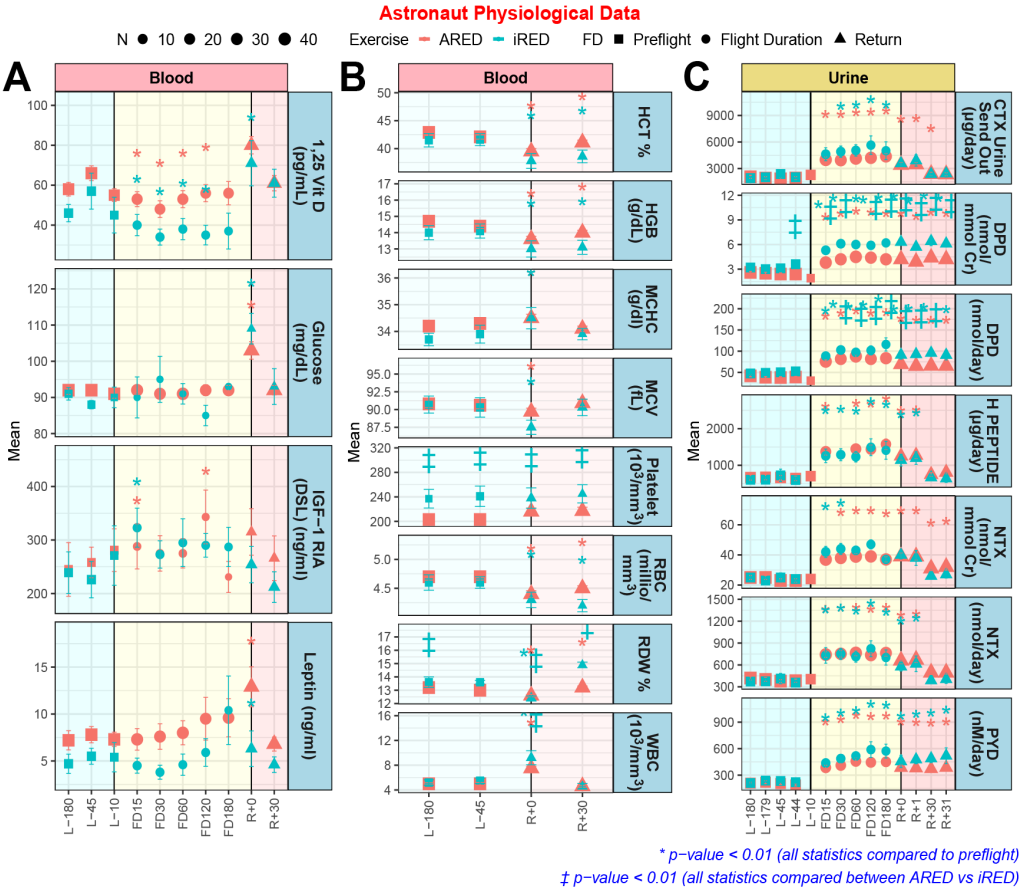

Figure 6.11. Astronaut physiological markers compiled from up to 50 astronauts. A) Specific blood markers which contain data points for pre-launch (L-), flight (FD), and return to Earth (R+). The numbers on the x-axis indicate the number of days for each group. Interim Resistive Exercise Device (iRED) is shown in blue and Advanced Resistive Exercise Device (ARED) is shown in red. **B)** Specific blood markers which contain data points for pre-launch (L-) and return to Earth (R+). **C)** Specific urine markers which contain data points for pre-launch (L-), flight (FD), and return to Earth (R+). The statistics on the data are *** (shown as *) $p < 0.001$ for significantly different from L-45 and ** (shown as ‡) $p< 0.01$ significantly different from ARED. This figure was provided by Dr. Afshin Beheshti.

6.3.8. Spaceflight-induced gene regulatory changes in the skin correlate with relevant drug responses

Having established a list of genes that changed in the rodent skin across multiple missions, potential drug targets and mechanisms for these genes were investigated via Ingenuity Pathway Analysis (IPA) (**Figure 6.12**). As with the DDR pathways (**Figure 6.9**), both the dorsal and femoral skin RR-5 mission data subsets and the 25-day time point from the C3H/HeJ mice in the RR-7 mission show an opposing pattern compared to the other data subsets for the predicted upstream regulators (**Figure 6.12**).

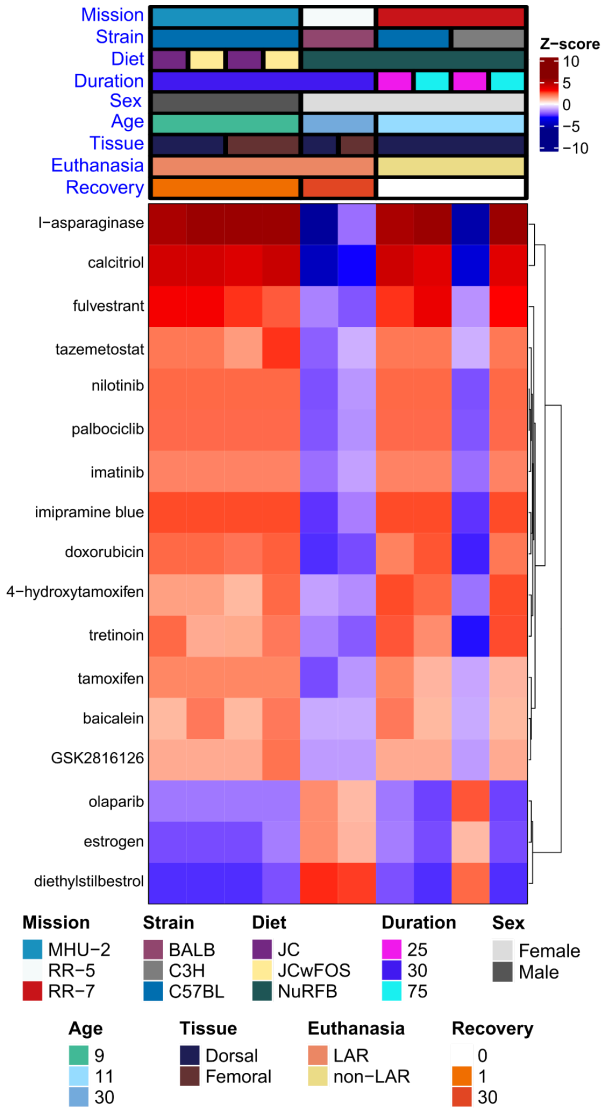

Figure 6.12. Predicted potential countermeasures for mitigating spaceflight response to the skin. Predicted drug signatures using the cross-mission genes across each dataset represented by a hierarchically clustered heat map. A positive (red) activation state implies cross-mission gene expression changes are consistent with expression changes observed with the indicated drug from curated causal gene expression relationship studies. A negative (blue) activation state implies the cross-mission gene expression changes are opposite to changes observed with the indicated drug.

Several drugs were predicted to exhibit significant positive correlation with gene expression changes seen in the space-flown mice, with the exception

of dorsal skin tissues obtained from C3H/HeJ mice that spent 25 days in flight for the RR-7 mission, and dorsal and femoral skin from the RR-5 mice (30-day recovery) (**Figure 6.12**). For example, calcitriol is the active form of Vitamin D, and has been observed to affect skin barrier formation and epidermal differentiation (Rid et al., 2013); it is used clinically to treat plaque psoriasis (Alia and Kerr, 2021). Calcitriol is also associated with immune system alterations in the skin, and topical application in mice has been shown to exacerbate atopic dermatitis (Li et al., 2006). On the other hand, in a murine model, epidermal permeability and antimicrobial skin barrier impaired by corticosteroid, was shown to be reduced via calcitriol (Hong et al., 2010). Alteration of gene regulatory patterns associated with calcitriol, in the mouse data subsets (**Figure 6.12**) could indicate activation of compensatory mechanisms for spaceflight-induced skin barrier dysfunction, epidermal differentiation, and immune system alterations. L-asparaginase showed the strongest correlation with gene regulatory changes in these spaceflight-exposed mice (**Figure 6.12**). L-asparaginase is used in treatment of acute lymphoblastic leukaemia and lymphoblastic lymphoma and exerts its action by depriving leukemic cells of circulating asparagine, resulting in DNA breaks, cell cycle arrest and cell death (Broome, 1981; Burke and Zalewska-Szewczyk, 2022). It is also an immunogenic protein and leads to skin rash, dry skin, and hypersensitivity reactions (Burke and Zalewska-Szewczyk, 2022). Similarly, there was strong positive correlation between gene regulatory changes in these spaceflight-exposed mice associated with use of the Vitamin A derivative tretinoin (all-trans retinoic acid) (Yoham and Casadesus, 2023), and two selective oestrogen receptor modulators, tamoxifen and 4-hydroxytamoxifen (Boyd and King, 1999; Gallo and Kaufman, 1997; Lee et al., 2014; Thornton, 2013).

On the contrary, oestrogen and diethylstilbestrol, a synthetic form of oestrogen, showed a reversed pattern compared to the aforementioned drug targets (**Figure 6.12**). These drugs showed positive correlation with the

expression changes in the RR-5 datasets, and the RR-7 25-day C3H/Hej data subset, and negative correlation in the other data subsets (**Figure 6.12**).

6.4. Discussion

Skin is well-established as an essential organ for health on Earth, yet despite the frequency of dermatological issues in astronauts (Crucian, Babiak-Vazquez, et al., 2016), the molecular response of skin to spaceflight is understudied. This chapter outlined a comprehensive study on the molecular impact of spaceflight on skin, with the aim of addressing gaps in the understanding of spaceflight associated skin health risks. The analysis indicated intriguing gene regulatory changes occurring in the skin of spaceflight vs ground control mice during the MHU-2, RR-5, and RR-7 spaceflight experiments, that also show some overlaps with molecular alterations observed in astronauts.

The genes exhibiting the highest variability across the datasets formed clusters associated with well-established physiological risks of spaceflight, notably immune dysfunction and muscle degradation (**Figure 6.1B**). Furthermore, the majority of genes showing significant differential expression across multiple missions in murine data were linked to cell cycle processes, particularly chromosome segregation (**Figure 6.4A and Figure 6.4B**). This suggests that skin mirrors common spaceflight signatures observed in other tissues, including dysregulation of cell cycle processes (Afshinnekoo et al., 2020; Garrett-Bakelman et al., 2019). Further to that point, there was an enrichment of DDR pathways in some data subsets from radiosensitive BALB/c and C3H/HeJ strains (Cekanaviciute et al., 2023; Mukherjee et al., 2014; Roderick, 1963) (**Figure 6.9**), which is also a well-established spaceflight response (Beheshti et al., 2021; Garrett-Bakelman et al., 2019). These repair mechanisms may be activated to mitigate radiation-induced DNA damage and appear to remain active during the post-flight recovery process, as seen in the RR-5 mission where samples were collected 30 days

after flight. This finding is consistent with reports of sustained post-flight DDR activation from the NASA Twins study (Garrett-Bakelman et al., 2019). Notably, DNA damage in the skin has important implications for skin cancers including melanoma, of which astronauts carry a higher risk, although the direct correlation between spaceflight and melanoma is still under investigation (Nguyen and Urquieta, 2023; Reynolds et al., 2021). Similarly, mitochondrial dysregulation is established as a key molecular signature of spaceflight (da Silveira et al., 2020), and there was enrichment of an ISR pathway in all but two of the murine skin data subsets (**Figure 6.10A**), which may be driven by increased oxidative stress or DNA damage. ISR activation appears to be impacted by spaceflight duration, as the 25 day time points for RR-7 show a suppression of the ISR pathway (Afshinnekoo et al., 2020; da Silveira et al., 2020). Use of an explainable machine learning approach also highlighted the counteractive mechanism played by *D2HGDH*, which was generally upregulated in space-flown mice (**Figure 6.10B**) indicating an increased need for breakdown of toxic D-2-hydroxyglutarate (D2H) molecules in mitochondria. In this model, the partner gene *RPLP0-PS1* codes for the ribosomal protein RPLP0 and is down-regulated in space-flown mice. RPLP0 has previously been implicated in the top significant network for the mitochondria-mediated cycle of Alzheimer's disease (Oka et al., 2016). Low or absent levels of RPLP0 seem to limit the compensatory mechanism in the model (**Figure 6.10B**). If this occurs, build-up of D2H and suppression of certain enzyme functions can result, causing DNA and histones to enter hypermethylated states and activate oncogenes and inhibit tumour suppressors (Xu et al., 2011). Build-up of D2H in the mitochondria may thus be related to cell changes induced by DNA damage from space radiation. Notably, gene regulatory changes associated with mitochondrial dysregulation and DDR were also reported in the skin data from the i4 mission (Dupras and Bunnik, 2021), corroborating these changes as key molecular hallmarks of spaceflight in the skin.

The downregulation of collagen/ECM genes and the suppression of collagen/ECM-related pathways in the MHU-2 mouse mission, where spaceflight samples were collected shortly after landing (**Figure 6.2B and Figure 6.7A**), mirror findings from previous studies where a reduction in collagen synthesis occurred under hypergravity conditions in cultured human fibroblasts (Seitzer et al., 1995). In contrast, RR-7 mice (frozen in space) generally exhibited an upregulation of the same collagen/ECM genes (**Figure 6.7A**), consistent with results from the aforementioned human fibroblast study, which reported a 143% increase in collagen synthesis during microgravity (Seitzer et al., 1995). Similar to RR-7, RR-5 also showed a trend of upregulation in collagen/ECM genes (**Figure 6.7A**). The opposing patterns between the MHU-2 mission and RR-5/RR-7 may arise from diverse factors owing to experimental design differences. Notably, the timing of sample collection introduces a critical variable, as seen in the MHU-2 mission where samples were collected less than one day post-flight, in contrast to RR-5, where collection occurred after a 30-day post-flight recovery, and RR-7, where samples were obtained directly in space. This temporal distinction implies that collagen suppression observed in MHU-2 mice may be influenced, at least in part, by the acute hypergravity response, a known factor in decreasing collagen synthesis (Seitzer et al., 1995), and the stress-induced release of glucocorticoids (Horii et al., 2017; Kahan et al., 2009; Stowe et al., 2001), rather than being solely attributed to the spaceflight environment. Moreover, differences in sex and housing arrangements further complicate the comparison; MHU-2 mice are singly-housed males, while RR-5 and RR-7 mice are multi-housed females, introducing additional variables that may contribute to the observed disparities (Horii et al., 2017; Kahan et al., 2009; Kim et al., 2024; Stowe et al., 2001).

Skin has an anatomically heterogeneous gene expression profile (Yan et al., 2023). For instance, fibroblasts exhibit distinct gene expression patterns based on their positional identity in relation to the major anatomical axis (Rinn et al., 2006). This phenomenon in part explains the predisposition of different sites to different skin conditions (Yan et al., 2023). Although several genes related to collagen synthesis and ECM were downregulated across all four MHU-2 experimental groups, some of these genes showed statistical significance in the dorsal group but not in the femoral group (**Figure 6.7A**). Likewise in the RR-5 mission, during terrestrial recovery, skin biopsies obtained from dorsal skin generally showed a higher reversal of gene expression pattern compared to those obtained from the femoral skin (**Figure 6.4A and Figure 6.7A**). These changes may be indicative of the variations in gene expression profile, their adaptation response to space flight and re-adaptation to terrestrial environment, at different anatomical sites.

Skin biopsies collected from i4 astronauts one day post-flight exhibited a marked downregulation of various collagen/ECM-associated genes in the dermal layer compared to pre-flight conditions (e.g., *SPARC, COL1A1, COL3A1, COL6A1*) (**Figure 6.7B**). However, certain genes, such as *SPARC* and *COL3A1*, showed an upregulation in one or both compartments of the epidermis (**Figure 6.7B**). Moreover, several other collagen/ECM associated genes (e.g. *COL14A1*, *COL15A1*, *MMP3*) were upregulated in both epidermis and the dermis (**Figure 6.7B**). As with the MHU-2 mice, the i4 astronauts' skin biopsies were obtained one-day post-flight, so some of these effects may have been induced by hypergravity. Nevertheless, blood data obtained from the same i4 astronauts at 1-, 45-, and 82-days post-flight suggests that most gene alterations observed upon landing revert back to normal levels within 82 days following flight, so one may expect a similar response in the skin tissue (**Figure 6.6C and Figure 6.8C**). Previously reported findings from the NASA Twins showed in-flight activation of

collagen formation pathways and an increase in COL1A1 and COL3A1 proteins detected in the urine, indicative of structural remodelling (Garrett-Bakelman et al., 2019). These changes, as with the majority of spaceflight-induced changes, were also reported to revert to baseline post-flight (Garrett-Bakelman et al., 2019).

Some of the changes reported by astronauts during and after spaceflight, such as dryness and contact dermatitis, point to an impaired barrier function (Dupras and Bunnik, 2021; Nguyen and Urquieta, 2023); the results within this thesis chapter support this hypothesis. *FLG* is a gene expressed in the stratum granulosum layer of the epidermis, during the epidermal differentiation process (Edqvist et al., 2015). It encodes profilaggrin, a major component of keratohyalin granules found in the keratinocytes that is dephosphorylated and cleaved into filaggrin (Resing et al., 1989; Rosso et al., 2016). During conversion of keratinocyte cell membrane to cornified cell envelope, filaggrin aggregates and aligns intermediate filaments such as keratin 1 and keratin 10, within the cytoskeleton of keratinocytes, and subsequently a cross-linking process is initiated by proteins such as involucrin and loricrin (Brown and Irvine, 2008; Nithya et al., 2015). This results in compactization and flattening of the corneocytes, one of the constituents of the skin barrier (Ishitsuka and Roop, 2022). *FLG* was identified as one of the cross-mission genes in the spaceflight vs ground control mice data with downregulation across the majority of data subsets (all except RR-5 dorsal skin) (**Figure 6.4A**). Additionally, post-flight vs pre-flight downregulation of *FLG* was observed in the epidermis of i4 astronauts (**Figure 6.7B**). Caspase-14, an enzyme encoded by the *CASP14* gene, cleaves filaggrin, breaking it down into free amino acids (Hoste et al., 2011). These amino acids serve as integral components of natural moisturising factors, contributing to the preservation of skin hydration and pH (Ye et al., 2014). Additionally, they play a crucial role in antimicrobial defence, further contributing to the multifaceted functions of the skin barrier

(Čepelak et al., 2019; Denecker et al., 2008; Markiewicz et al., 2021). Similar to *FLG*, *CASP14* demonstrated downregulation in the majority of murine skin subsets (excluding RR-5 dorsal skin) (**Figure 6.4A**), the epidermal layer of i4 astronaut skin tissue (**Figure 6.7B**), and JAXA hair follicles (**Figure 6.7C**). This observation is important considering the crucial role *CASP14* plays in skin barrier formation and its involvement in skin conditions such as contact dermatitis and atopic dermatitis (Hvid et al., 2011; Jung et al., 2014; Kirchmeier et al., 2017; Markiewicz et al., 2021). The link between filaggrin deficiency and dry skin, altered pH, atopic and contact dermatitis, and ichthyosis vulgaris is also well-known (Lagrelius et al., 2020; Palmer et al., 2006; Smith et al., 2006; Winge et al., 2011; Ye et al., 2013, 2014).

CASP14 was also the driver of one of the two main QLattice models, found to be maximally relevant and minimally redundant among the cross-mission genes, resulting in a very simple model involving only *CASP14* and *S1PR1* with an AUC of 0.85 (**Figure 6.4D**). *S1PR1* was upregulated in most murine skin subsets (**Figure 6.4A**), the in-flight vs pre-flight samples from the JAXA hair follicle study (**Figure 6.5B**), and in all compartments for the post-flight vs pre-flight i4 astronauts' skin samples (**Figure 6.5A**). In the RR-5 mission, after a 30-day recovery, in-flight effects may have been reversed as *S1PR1* expression was downregulated (**Figure 6.4A**), which is also consistent with downregulation in the post-flight vs pre-flight samples from the JAXA hair follicle study (**Figure 6.5B**). Downregulation of *S1PR1* has been reported to be necessary for activation of resident memory CD8+ T cells in skin (Skon et al., 2013) and has previously been observed in CD8+ T cells from murine skin during response to Herpes Simplex Virus Infection (Mackay et al., 2013). While the S1P-S1PR axis has inhibitory roles in keratinocyte proliferation, it was also shown to contribute to the process of chemotaxis (Masuda-Kuroki and Di Nardo, 2022). For instance, in Staphylococcus-aureus-stimulated keratinocytes; TNF alpha, IL-8 and IL-36γ release was induced via involvement of *S1PR1* and *S1PR2 (Igawa et al., 2019)*. Higher expression

levels of *S1PR1* and *S1PR2* observed in the skin of impetigo patients and changes in the distribution patterns of these receptors were also aligned with these findings (Igawa et al., 2019). Herpes Simplex Virus reactivation (Mehta et al., 2022; Nguyen and Urquieta, 2023; Rooney et al., 2019) and Staphylococcus colonisation in the skin (Dupras and Bunnik, 2021; Morrison et al., 2021; Xiao et al., 2019) (which could also be influenced by breakdown of filaggrin altering pH levels (Ogonowska et al., 2020)) have been widely reported in astronauts, therefore alterations in *S1PR1*, *CASP14,* and *FLG,* are of importance.

At present, there is a lack of widely accepted countermeasures to effectively mitigate the damage caused to the skin induced by exposure to the space environment. However, this analysis has identified potential candidates capable of altering similar molecular changes that manifest in the skin during spaceflight (**Figure 6.12**). Several drugs, including calcitriol, tretinoin, and tamoxifen showed significant positive correlation with regulatory changes in all data subsets from the MHU-2 and RR-7 missions except for the C3H/HeJ mice samples collected at the 25-day time point (**Figure 6.12**). Tretinoin, a vitamin A derivative is a retinoid known for its applications in treating acne and leukaemia, exerts its effects through the activation of retinoic acid receptors, influencing various skin processes such as increased proliferation and inhibited differentiation of epidermal cells, increased keratinocyte turnover, increased collagen synthesis, inhibited UV radiation induced metalloproteinase activity, and melanin reduction (Baldwin et al., 2013; Lee et al., 2009; Voorhees, 1990). Tretinoin treatment on reconstructed human epidermis was shown to induce loss of keratohyalin granules and reduce filaggrin and Keratin 10 expression (Hsia et al., 2008). Retinoids have also been observed to downregulate caspase-14 expression for *in vitro* reconstructed skin models (Rendl et al., 2002). Dry skin and exfoliation are common side effects during its use (Baldwin et al., 2013). Tamoxifen, a selective oestrogen receptor modulator, has been shown to exhibit both

agonistic and antagonistic effects depending on the tissue (Gallo and Kaufman, 1997). Its use in breast cancer patients is linked to cutaneous adverse events, including dry skin, itching, and occasionally radiation recall dermatitis and baboon syndrome (a Type IV hypersensitivity reaction) (Andrew et al., 2014; Mofarrah et al., 2021; Moseson et al., 1978). Furthermore, tamoxifen induces androgen receptor expression in hair follicles, leading to tamoxifen-induced alopecia in some cases (Nonomura et al., 2012), and abnormal hair follicles, epidermal atrophy and dermal fibrosis have been observed in rat models (Inalöz et al., 2002). On the other hand, only olaparib, oestrogen, and diethylstilbestrol (a synthetic form of oestrogen) showed an opposing pattern, with significant negative correlation in all data subsets except for those from the RR-5 mission and the C3H/HeJ mice samples collected at the 25-day time point of the RR-7 mission (**Figure 6.12**). Oestrogen plays a crucial role in skin health, influencing keratinocyte function, collagen production, moisture retention and telogen-anagen follicle transition (Movérare et al., 2002; Murakami et al., 2022; Oh and Smart, 1996; Stevenson and Thornton, 2007). In an explant model of foetal rat skin development, oestrogen was shown to increase filaggrin expression (Kömüves et al., 1998). Conversely, oestrogen deficiency during menopause is associated with adverse skin changes such as loss of collagen, skin dryness and atrophy (Kamp et al., 2022; Lephart and Naftolin, 2021). Overall, molecular alterations observed in the space-flown mice and astronaut skin data show similarities with molecular effects associated with drugs like tretinoin and tamoxifen (Baldwin et al., 2013), such as *FLG* and *CASP14* downregulation, while symptoms reported in astronauts including dry skin and dermatitis also resemble side effects associated with drugs like tretinoin and tamoxifen (Baldwin et al., 2013) (**Figure 6.12**). Topical moisturisers and emollients, as commonly used by astronauts (Braun et al., 2019; Nguyen and Urquieta, 2023), may also help to counteract the physiological consequences of disrupted barrier function genes, and may help to explain why results on impact of spaceflight on skin physiological parameters in

astronauts have been mixed (Braun et al., 2019; Heinrich, 2008), despite physiological evidence of skin damage in space-flown rodents (Neutelings et al., 2015).

In conclusion, a comprehensive molecular study on the impact of spaceflight on skin health has been outlined within this chapter. Currently, there is a gap in knowledge for how space radiation and microgravity affects skin biology, yet many of the common themes of spaceflight dysfunction emerged in this analysis, suggesting that skin could be an easily-accessible candidate for studying the biological impact of spaceflight. This systems biology analysis revealed some genes of interest and potential mechanisms to investigate in the context of countermeasures that can be utilised in future studies. The findings of this study begin to address the current gaps and provide some clues on how to potentially mitigate the adverse effects of the space environment to the skin.

6.5. Limitations of Study

The rodent datasets used come from three different spaceflight experiments with a variety of confounding variables associated with differences in study design. This means that direct comparison between the experiments is challenging, but indeed this also presents an opportunity to hypothesise how these factors may influence biology, as done in this manuscript. A lack of physiological data, such as dermal thickness, and lack of individual skin layer data from the rodent datasets means that generated hypotheses relating to mouse physiology cannot be confirmed without follow-up investigations, but the comparison to astronaut data helps translate the results to human relevance. Lastly, gene expression levels differ between skin compartments so an identical fold-change value in a layer where the specific gene expression is generally low may not represent the same magnitude in a layer where the same gene's expression is relatively greater.

As such, for the i4 skin data, the focus was on changes in the compartments where the gene is established to be predominantly expressed.

6.6. Materials and Methods

The following supporting information for Chapter VI can be downloaded via the associated publication (Henry Cope et al., 2024): Supplemental Data 6.1. Differential Gene Expression Results, Supplemental Data 6.2. Gene Set Enrichment Analysis Results, Supplemental Data 6.3. Supplemental IPA Drug Results, Table S6.1. Curated pathways for directed pathway analysis. Pathway lists used in this analysis for Figures 6.7, 6.9, and 6.10.

6.6.1. GeneLab spaceflight murine datasets

Five RNA-Seq datasets (OSD-238, OSD-239, OSD-240, OSD-241, OSD-254) were downloaded from the NASA OSDR (https://osdr.nasa.gov/bio/repo) via the API, and full dataset descriptions can be found on the dataset pages. These datasets are derived from whole murine skin samples from the MHU-2, RR-5, and RR-7 spaceflight experiments. For MHU-2, singly-housed male C57BL/6J mice were 9 weeks of age when flown on the ISS for 30 days; they were euthanized less than one day after return to Earth. Dorsal and femoral skin samples were extracted from MHU-2 mice, and six replicates in spaceflight microgravity and the matching six replicates in the 1G ground control were split into two sets of three, with half fed a JAXA chow diet and the other half fed JAXA chow with supplemental FOS (Matsuda et al., 2019). For RR-5, 30 week old BALB/c mice in shared housing were flown on the ISS for 30 days; following return to Earth, mice were given 30 days to recover before euthanasia and extraction of dorsal and femoral skin tissue. Finally, for RR-7, 11 week old C57BL/6J mice and C3H/HeJ mice were flown on the ISS for either 25 or 75 days before being euthanized on-orbit.

6.6.2. RNA-Seq analysis of rodent datasets

Raw counts data, derived via a previously reported pipeline (Overbey et al., 2021), were downloaded from the NASA OSDR and loaded with tximport (v1.28.0). ERCC genes were filtered out, as were genes with zero counts in the minimum number of samples corresponding to a single condition. Differential Gene Expression (DGE) analysis using the Wald significance test was performed on spaceflight microgravity and ground control samples using the R package DESeq2 (Love et al., 2014) (v1.40.2). For each data subset, microgravity spaceflight samples were contrasted with ground control samples and cook's cut off and independent filtering were not used. Mapping of mouse ENSEMBL gene IDs to gene symbols and human orthologs was done using the R Package biomaRt (v2.56.1) and ENSEMBL 105. For human orthologs, one-to-many mapping was supported. All heatmaps were generated using the R package ComplexHeatmap (v2.16.0). All DGE statistics for the rodent data, including the list of cross-mission genes, can be found in **Supplemental Data 6.1.**

6.6.3. Pathway analysis of rodent datasets

For Over Representation Analysis (ORA), the R package clusterProfiler (v4.8.3) was used on vectors of genes and bar plots were made via ggplot2 (v3.4.4). For Gene Set Enrichment Analysis (GSEA), the R package FGSEA (v1.26.0) was used. GSEA was performed on all DEGs for all data subsets, with rank vectors consisting of HGNC symbols and corresponding sorted t-scores in the data subset. MSigDB human collections H, C2, and C5 were used (v7.4) and set sizes >15 were permitted. For the mitochondrial pathway analysis, human pathways from MitoPathways (v3.0), with some additional custom curated pathways (**Table S6.1**), were used and set sizes >1 were permitted. All heatmaps were generated using the R package ComplexHeatmap (v2.16.0). All GSEA results, including mitochondrial results, can be found in **Supplemental Data 6.2.**

6.6.4. RNA-Seq data analysis on Twin Study samples

Longitudinal samples were collected from a male astronaut aboard the ISS and his identical twin on Earth during a 340 day mission including 6 months pre-flight and 6 months post-flight follow-up, for a total of 19 time points for the flight subject and 13 time points for the ground subject. Blood was collected using CPT vacutainers (BD Biosciences Cat # 362760) per manufacturer's recommendations. For full details of sample separation and processing see (Garrett-Bakelman et al., 2019). Briefly, samples collected on ISS were either frozen in -80°C after separation of mononuclear cells by centrifugation (referred to as CPT), or returned to Earth in 4°C in a Soyuz capsule and sorted into CD4, CD8, CD19 populations and a lymphocyte depleted (LD) fraction. Samples collected on Earth were either frozen for mononuclear cells or processed when fresh for sorted cell populations. To correct for the effects of ambient temperature exposure on RNA (approximately 36 hours including landing and repatriation) control samples were created by simulating similar conditions to those that may occur during the ambient return and were compared to fresh blood collections from the same individual. RNA extraction, library prep and sequencing were completed per (Garrett-Bakelman et al., 2019) using both ribodepletion or polyA selection kits.

As previously reported (Garrett-Bakelman et al., 2019), generated sequences were trimmed using Trim Galore! (v0.4.1) and quantified to genes using kallisto(Bray et al., 2016) on ENSEMBL transcripts. Differentially expressed genes were called using DESeq2 (Love et al., 2014) on each cell type separately by comparing pre-flight, in-flight and post-flight groups, controlling for the normal biological variance within the 24 months using the longitudinal data of the ground twin and using the simulated ambient control samples as another covariate for sorted cells (Love et al., 2014). Reads were

mapped onto an updated reference genome (GRCh38) (Rosenfeld et al., 2012).

6.6.5. JAXA Cell-Free Epigenome (CFE) Study RNA quantification data

Aggregated RNA differential expression data and study protocols were shared through the NASA OSDR with accession number: OSD-530 (Husna et al., 2024). Plasma cell-free RNA samples for RNA-seq analysis were derived from blood samples collected from six astronauts before, during, and after the spaceflight on the ISS. Mean expression values were obtained from normalised read counts of six astronauts for each time point. In total two to three pre-flight samples (L-168 ± 14 days, L-112 ± 14 days, L-56 ± 14 days), four in-flight samples (L+5 ± 1 day, L+30 ± 7 days, L+60 ± 14 days, L+120 ± 14d or R-8 (-14d/+0d day)), and three to four post-flight samples (R+3 ± 1 day, R+30 ± 7 days, R+60 ± 14 days, R+120 ± 14 days) were collected per astronaut. As provided by OSDR, data points for all subjects were pooled into one group per mission phase.

6.6.6. JAXA astronaut hair follicle data

Gene expression data from 10 JAXA astronauts' hair follicles(Terada et al., 2016) was downloaded from the NASA OSDR (OSD-174). Raw data for 60 total samples was processed using LIMMA with R/bioconductor (Ritchie et al., 2015). Briefly, duplicate sample single-colour Agilent microarray data was background corrected, filtered for low expression probes, and quantile normalised. Differential gene expression was measured between pre-flight, in-flight, and post-flight data points using p-values adjusted for False Discovery Rates (FDR) with the Benjamini–Hochberg method (Benjamini and Hochberg, 1995). As previously described (Terada et al., 2016) hair follicle samples were collected from each astronaut on 6-month ISS missions at two pre-flight time points (firstly between L-180 and L-90, and secondly between L-60 and L-14), two in-flight time points (firstly between L+20 and

L+37, and secondly between R-20 and R-7), and two post-flight time points (firstly between R+2 and R+7, and secondly between R+30 and R+90).

6.6.7. Inspiration4 (i4) astronaut sample collection

Four civilians, two males and two females, spent three days in Low-Earth Orbit (LEO) at 585 km above Earth. The mission launched from NASA Kennedy Space Center on September 15th, 2021 and splashed down in the Atlantic Ocean near Cape Canaveral on September 18th, 2021. Several human health and performance related experiments were carried out in collaboration with SpaceX, the Translational Research Institute for Space Health (TRISH) at Baylor College of Medicine (BCM), and Weill Cornell Medicine. Experiments were performed in accordance with the relevant guidelines at the principal investigators' institutions. Moreover, the different study designs and the corresponding methods to collect and analyse the biological samples were approved by BCM IRB (Protocol 21-05023569) and Weill Cornell Medicine (IRB #21-05023569). All biological data derived from the Inspiration4 mission were collected pre- and post-flight, and will be accessible via the Space Omics and Medical Atlas (SOMA) (Overbey et al., 2024). For this study, only data from blood samples and skin biopsies were used. Pre-flight samples were collected at L-92, L-44, and L-3 days prior to launch to space. Upon return, post-flight samples were collected at R+1, R+45, and R+82 days.

Blood samples were collected before (Pre-launch: L-92, L-44, and L-3) and after (Return; R+1, R+45, and R+82) the spaceflight. Chromium Next GEM Single Cell 5' v2, 10x Genomics was used to generate single cell data from isolated PBMCs. Subpopulations were annotated based on Azimuth human PBMC reference (Stuart et al., 2019). Reads were mapped onto a recent reference genome (GRCh38) (Rosenfeld et al., 2012).

For skin spatial transcriptomics data, 4mm diameter skin biopsies were obtained from all Inspiration4 crew members, once before flight and as soon as possible after return (L-44 and R+1). As previously described in detail (Dupras and Bunnik, 2021), these biopsies were flash frozen and processed with the NanoString GeoMx platform. Based on staining images using fluorescent antibodies, a total of 95 freeform regions of interest (ROIs) were profiled across outer epidermal (OE), inner epidermal (IE), outer dermal (OD) and vascular (VA) regions. The IE and OE regions represented stratum granulosum, stratum spinosum and stratum basale, respectively. GeoMx WTA sequencing reads from NovaSeq6000 were compiled into FASTQ files corresponding to each ROI and converted to digital count conversion files using the NanoString GeoMx NGS DnD Pipeline. From the Q3 normalised count matrix that accounts for factors such as capture area, cellularity, and read quality, the DESeq2 method was used to perform DGE analysis.

Records of vitamin D supplement consumption were not available for the i4 crew.

6.6.8. Astronaut Physiological data

Data are reported from three human subject experiments conducted on the International Space Station: Nutritional Status Assessment (SMO 016E; 2006-2012), Dietary Intake Can Predict and Protect Against Changes in Bone Metabolism During Space Flight and Recovery (Pro K) (2010-2015), and Biochemical Profile (2013-2018). All protocols were reviewed and approved by the NASA IRB and all subjects provided written informed consent.

These missions were 4-6 months in length, and these studies included blood and urine collections before, during, and after flight, with analysis of an array of nutritional and biochemical markers. Blood and urine samples were collected two or three times before flight: approximately Launch minus (L-)

180 days and L-45 days. In some cases, a third blood sample was collected (typically along with the L-45 collection), and these tubes were centrifuged and frozen for aliquoting after flight batched with the samples collected in-flight. Blood samples were collected in-flight, at approximately Flight Day (FD) 15, 30, 60, 120, and 180. post-flight samples were collected in the first 24-h after landing (designated return+0, R+0) and again 30-d later (R+30). The R+0 samples were not necessarily fasting, given the time of day and nature of return from flight. Of the 59 crewmembers reported herein: eight returned on the Space Shuttle, with blood collection 2-4 hours after landing; 51 landed in Kazakhstan, with seven of them returning to Star City, Russia, with blood collection 8-10 hours after landing; 44 were transported directly back to the Johnson Space Center in Houston, with blood collection approximately 24-h after landing. Pre- and post-flight collections included two 24-h urine collections, and in-flight collections included one 24-h urine collection. These collection techniques have been previously described (Smith, M Heer, et al., 2012).

Here, data on vitamins and metabolites, oxidative stress and damage markers, inflammatory markers and cytokines, liver enzymes and endocrine indices are reported. These were analysed using standard techniques as previously reported (Zwart et al., 2016).

Data were available for 59 crewmembers (47 male, 12 female). Age at launch was 47.0 ± 5.6 y, body mass at launch was 79.2 ± 11.8 kg (M: 83.3 ± 9.3 kg; F: 63.0 ± 4.5 kg). Body mass index was 25.5 ± 2.9 kg/m^2 (M: 26.4 ± 2.6 kg/m^2 ; F: 22.3 ± 1.5 kg/m^2).

All available data are reported here, although the reported n for any given test or session varies for a number of reasons, including: not all experiments had all analytes included, mission length differences for some crewmembers, schedule or other issues occasionally precluded sample

collection, and methods changes over time. Repeated measures analysis of variance was conducted to test for differences during and after flight compared to pre-flight, and comparisons among time points were made using a Bonferroni t-test. Multiple comparisons were accounted for, and only those tests with $p<0.001$ are reported as statistically significant results. The data was plotted using R package ggplot2 (v3.3.5).

6.6.9. QLattice symbolic regression modelling

Symbolic regression (QLattice v3.0.1) was used to construct models involving combinations of synergistic genes which map from the gene expressions to spaceflight status. These models only distinguish between mice that went to space and mice that did not. Conventional statistical methods typically allow for calculating the effect of a single gene at a time through metrics such as p-values and false discovery rates. In contrast, symbolic regression models can reveal combinations of genes and modules that best predict spaceflight status. These could be linear combinations involving two or more genes that have previously also been shown to be statistically significant, or it could be non-linear combinations that reveal features that on their own were non-significant but in synergy with a second feature become highly predictive. In addition, known biological functions of genes included in models, as well as the resultant mathematical relationship between them, can potentially be interpreted to reveal regulations or interactions that are affected by spaceflight.

In biological pathway analysis, it is well-known that up- or down-regulation of one gene can have cascading effects, such that the function of one gene becomes sensitive to that of another (Itadani et al., 2008). It has previously been demonstrated that parsimonious machine learning models are able to provide accurate outcome prediction in omics data, while preserving interpretability (Christensen et al., 2022; Elsborg and Salvatore, 2023). The

interpretability often results from the fact that models might demonstrate otherwise opaque relations, which become clear when combined effects are considered. An example could be an interaction where the regulation of a single gene is itself unimportant for functional changes, *unless* another gene is simultaneously regulated. This would indicate a synergistic compound effect, which naturally can expand to a 3-, 4- or 5-gene synergy and beyond. The combined effect of two or more individually insignificant gene expression levels may thus theoretically be more powerful than the effect of a single gene with low p-value. Thus, traditional statistical analysis masks such effects by the use of metrics related to the individual gene.

6.6.10. UMAP dimensional reduction and gene clustering

For the clustering of genes shown in **Figure 6.1B**, a uniqueness filtering was performed by first dropping all genes where more than a quarter of the samples had identical expression levels. Subsequently, all genes with a variance of less than 1.7 were filtered out, resulting in 2184 filtered genes with high variability.

A dimensional reduction was performed subsequent to filtering by uniformly distributing the filtered data on a Riemannian manifold, using the Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP). UMAP is a general purpose manifold learning and dimension reduction algorithm which is similar to t-SNE in that it predominantly preserves local structure. Yet, UMAP preserves more global structure, which makes it a more suited algorithm when the objective of the dimensional reduction is more than simple visualisation (in this case the objective is clustering).

Genes were then clustered using the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) (McInnes et al., 2017). A primary advantage of HDBSCAN is that it is always deterministic for the

same hyperparameters, and will thus always return the same clustering, all else being equal. In addition, comparable clustering algorithms such as k-means do not perform well unless clusters are of equal size and density with few outliers. With biological data such as gene expressions, one expects large variation in cluster size and density, making HDBSCAN the ideal choice.

Once genes were clustered, the gene sets belonging to each cluster were extracted and analysed using three ontology databases in the Python implementation, GSEAPY, of the Gene Set Enrichment Analysis (Subramanian et al., 2005) tool Enrichr (Chen et al., 2013). The Elsevier Pathway Collection, the 2021 WikiPathway Collection, and the 2021 MGI Mammalian Phenotype Level 4 Collection were used. These enrichment analyses were used to provide context to each cluster by appending an annotation if a notable amount of hits showed up for a particular association.

6.6.11. Regulatory Network Analysis

Differential gene expression data from all cross-mission genes across each comparison was used to perform an upstream regulator analysis using QIAGEN Ingenuity Pathway Analysis (IPA) (version 01-22-01 (Krämer et al., 2014)). Biological or chemical drugs in the QIAGEN IPA library of regulators were identified, defined as having a Benjamini-Hochberg adjusted P value < 0.05 and an absolute Z-score > 1. Drugs from the dataset that met these criteria and were associated with similar $\log_2$ fold-change patterns in the QIAGEN Knowledge Base were visualised by hierarchical clustering of the unstandardized Z-scores using Euclidean distance as the similarity metric, and complete-linkage method for agglomeration. Statistics on the full list of drug targets can be found in **Supplemental Data 6.3.**

Chapter VII: General Discussion

As documented in this thesis, the thesis author investigated strategies to improve data generation capabilities and data science capabilities within the field of space biology. Particular focus is given to the opportunities and challenges of employing omics data to bridge gaps between biological and phenotypic alterations underlying health risks associated with the hostile spaceflight environment. The overarching goal of this work is to accelerate the rate of scientific discovery within space biology as a means to inform space medicine strategies. These new strategies are necessary for delivering the health and safety requirements of ambitious missions beyond Low Earth Orbit (BLEO) and an increasingly diverse spacefaring population, predominantly arising from commercial space endeavours.

7.1. Contributions in Human Space Omics

One clear research gap identified was the lack of omics studies performed in space using humans as test subjects. Although the feasibility and research value of such studies has now been demonstrated with NASA astronauts (Garrett-Bakelman et al., 2019), commercial astronauts (Jones et al., 2024), and JAXA astronauts (Husna et al., 2024) - there remains a lack of clarity on how the European Space Agency (ESA) and its constituent national space agencies plan to rally to perform similar studies, or better still, routine programmes, in their astronauts.

In this thesis, efforts were made to address this gap by engaging with space biology expert scientific communities, including the International Standards for Space Omics Processing (ISSOP) consortium and the ESA-funded Space Omics Topical Team. The desired outcome of these engagements was to distil expert recommendations from within these communities into reports and publications which could then be disseminated to relevant stakeholders, including space agencies and the wider space life sciences community. To

this end, the thesis author led an investigation into the European opportunity for performing these types of studies, and then a subsequent investigation into the UK national opportunity for these studies.

7.1.1 European Human Space Omics

The investigation into the European opportunity for performing omics studies in human spaceflight was disseminated to the space life sciences community through a publication in the Cell Press “Space omics in Europe” special issue. This publication was incorporated into the thesis as Chapter III “Astronaut Omics - The European Opportunity” (Cope et al., 2022). Briefly, this work discusses different types of omics and sample collection methods, considerations for sample processing, data governance practices that are preparatory for AI-analyses, and the importance of incorporating phenotypic measures alongside omics collection. The work also discusses the regulation of astronaut omics data in the context of the GDPR. In the conclusion of the paper, a set of potential questions for follow-up discussion were put forward, as shown in **Table 3.1**.

Secondly, through in-person discussion sessions led by the thesis author, these questions were individually considered by experts within the ESA Space Omics Topical Team. These expert agreed considerations and recommendations on the topic were written up by the thesis author and included within the final report of the ESA space omics topical team, which was disseminated to ESA as an internal report and to the space life sciences community via a publication (Manzano et al., 2023):
Manzano A, Weging S, Bezdan D, Borg J, Cahill T, Carnero-Diaz E, ***Cope H****, Deane CS, Etheridge T, Giacomello S, et al. Enhancing European capabilities for application of multi-omics studies in biology and biomedicine space research. iScience. 2023 Sep 15;26(9).*

The following text is incorporated into this thesis in accordance with the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). Modifications have been made for its inclusion within this thesis, which include adjustments to the formatting as well as conversion of the text to British English.

The first subset of questions aimed to initiate discussion of the possible collection routines from astronauts and spaceflight participants, decide on the omics data to collect, from which samples, and at which time points. These questions and the initial considerations and recommendations are listed below:

1. **Which omic types and technologies have the highest potential for scientific return and clinical actionability (with additional consideration to multi-omic combinations)? -** Genomic sequencing is likely to be the most immediately clinically actionable. Other omic types, including transcriptomics, microbiomics, and proteomics have potential clinical applications and may be useful for monitoring the effect of spaceflight throughout the mission.

2. **Which sample types are the most practical to collect (i.e., cost, sample processing procedures, and invasiveness)? -** For frequent sampling, particularly in-flight, non-invasive measures including urine, saliva, skin swabs, and stool will be useful. Blood samples are also a great option for omics, due to being minimally invasive and well established for spaceflight collection and medicine on Earth.

3. **What metadata would be the most useful and practical to standardise alongside omics collection (e.g., physiological, environmental, and lifestyle)? -** Spacecraft environmental data will be essential as will regular performance data, such as those collected via

NASA's Spaceflight Standard Measures Program (e.g., cognition, exercise, diet).

4. **At which time points should omics be collected? -** Pre-flight screening is recommended, and multiple pre-flight time points for dynamic omics are important for building a normalised baseline. Multiple in-flight omic time points are useful for monitoring change throughout the flight, including within the first few days where significant physiological changes already occur (Trudel et al., 2022). Post-flight data will be important for monitoring recovery and lingering effects.

5. **How can the artificial intelligence-readiness of the generated datasets be optimised? -** Data and metadata should be standardised between human space omic studies and stored in machine-readable formats. Due to personal data challenges, it seems logical to ensure support for federated learning.

The second subset of questions pertained to the ethical, legal, and regulatory considerations of designing policy for European spaceflight omics studies in humans. The discussion questions and some initial considerations and recommendations are shown below:

1. **What is the potential identifiably of different omic types, in the specific context of astronauts? -** Astronauts are particularly easy to identify since they are a small cohort of public figures, with phenotypic information on individuals readily available online. This easily accessible phenotypic information could be combined with omic information through linkage attacks, to identify individuals and breach their privacy. Genomic data are likely to hold higher identification risk compared to other omic types, yet other data types and combinations cannot be considered to be low risk without thorough analysis (Dupras and Bunnik, 2021).

2. **How should ESA go about obtaining meaningful informed consent for omics research? -** Broad consent is often used for omics research (e.g., UK Biobank), since analysis of omics data is not always hypothesis driven and future data uses are sometimes unknown. However, efforts should be taken to educate subjects as much as possible on the potential risks of having their omics data collected, and standard procedures such as oversight via ethical review boards should be followed.

3. **How should ESA handle potential ethical issues, such as incidental findings and discrimination? -** Similar to omic initiatives on Earth, clear policy which includes genetic counselling will be required for handling incidental findings, and regarding discrimination, focus should be given to using findings from omics to tailor countermeasures for individual crew members during spaceflight, especially if they are at a higher risk for a medical incident. An ESA Topical Team has recently discussed the opportunities for personalised medicine in space (Pavez et al., 2021).

4. **How should data storage and data sharing between ESA's European and international network be handled? -** Data sharing is desirable to enable open-science efforts, and NASA GeneLab is well established for data sharing of non-sensitive space omics data. However, most of the omics data are likely to be highly sensitive, so uploading sensitive data to a controlled access biobank, such as the UK Biobank and Federated European Genome-phenome Archive, to provide access only to approved researchers may make sense. Federated approaches, with the human omics data stored in local secure European biobanks, may also serve as an appropriate option for European-wide and even international data sharing.

Ultimately, these considerations and recommendations present a launching point for ESA to begin to integrate omics as a routine programme for astronauts. ESA has extensive expertise in conducting human analogue

studies for space life sciences, and has developed a set of core measurements to be taken across these analogue studies. The ISM measures and ESA core measures include phenotypic data, such as DEXA scan body composition, bone mineral density, and exercise performance including vertical jump and treadmill tests (Clément, Crucian, et al., 2022; Clément, Rittweger, et al., 2022; Sundblad et al., 2016). Where possible, it makes sense for ESA to collect the same set of measurements throughout human spaceflight studies, to enable comparability across analogue and spacelight studies. Furthermore, ESA has an opportunity to incorporate a set of omics within these core measurements for both analogue studies and spaceflight studies. Indeed, ESA has recently increased the generation of omics from human analogue studies (Archer et al., 2024; Stratis et al., 2023) and spaceflight studies (Blottner et al., 2023), but it is currently unclear whether a standard set of omics will be integrated into a routine programme such as via extension of the existing core measurements programme. Since ESA astronauts can opt into the NASA standard measures programme on the ISS, and there have been statements that the programme will begin to include omics, there could be a salient opportunity to combine efforts. However, with ESA being subject to the GDPR, any collaborations should be carefully designed to ensure appropriate handling and use of data from ESA astronauts.

7.1.2 UK Human Space Omics

Following a solicitation sent to the UK space science and exploration community via the UK Space Agency, the thesis author led a project to consider the opportunity of generating space omics data from the UK national perspective.

This project led to a presentation and associated publication at the British Interplanetary Society (BIS) “Future of UK Human Spaceflight Policy

Symposium". The publication was disseminated to the scientific community via open-access publication and disseminated to a representative of the UK Space Agency directly via the BIS. This publication was incorporated into this thesis as Chapter IV "Astronaut Omics - The UK Opportunity" (Cope et al., 2023). Briefly, this work discusses the opportunity for the UK as a world-leader in genomics to capitalise on experience from programmes like the UK Biobank and NHS Genomic Medicine service by collecting omics and phenotypic data from UK participants in spaceflight programmes. In particular, focus was given to routine whole genome sequencing (WGS) of astronauts prior to flight to uncover the role of individual genomic differences in spaceflight response, and to enable personalised medicine strategies (**Figure 4.1**). The paper concluded with a suggestion that the UK launch a pilot study with ESA astronauts or commercial astronauts that are also UK citizens as the study participants, potentially following an initial early-stage pilot study with UK citizens within a remote environment analogue (**Figure 4.2**). The goal of this pilot study or set of pilot studies would be to establish an agreed set of data for collection, validate collection protocols and data governance policies, and analyse the data to retrospectively assess opportunities of clinical actionability that could have been made from the data, such as individualised pharmaceutical supplies (i.e. pharmacogenomic) or increased monitoring for health risks with elevated susceptibility.

Overall, this investigation into the UK opportunity for collecting omics data from UK citizens venturing into space via ESA or commercial missions presents a path forward for UK space science researchers and the UK Space Agency to engage with human space omics. While human spaceflight studies are expensive, the suggested early-stage pilot study in an analogue environment presents a relatively low-cost path for the UK to explore this opportunity further and to build UK expertise in an approach that is gaining traction internationally. As an ESA member state, these genomic-focussed

studies in the UK could lay the groundwork for, or be part of a wider study towards, the implementation of a routine omics programme for ESA astronauts.

7.2. Contributions to Data Science Approaches

The second research gap highlighted a need to explore approaches for maximising the utility of rare space biology datasets in order to advance the understanding of how spaceflight impacts biology and how these biological changes link to the undesirable phenotypes associated with spaceflight health issues. A few specific challenges were considered in the context of space omics analysis, including optimal strategies for integrating associated phenotypic information, performing analyses across multiple organisms, and managing the confounding factors in experiment design introduced when pooling data from multiple missions

7.2.1 Transcriptomic analysis of worms in space

The first case study included within this thesis was a phenotypic and transcriptomic analysis of *C. elegans* nematodes, a popular model organism, flown to space during the micro-16 mission. The study aimed to elucidate the biological changes underpinning loss of muscle performance in space, which is a well established spaceflight risk in astronauts. This work resulted in a publication, which is also incorporated within this thesis as Chapter V “Data Analysis Case Study - Transcriptomic Analysis of Worms” (Soni et al., 2023). Briefly, this work included quantification of muscle strength from pillar displacement within a microfluidics device (Soni et al., 2022), and other phenotypic measures like worm size, performed by another author. This phenotypic analysis was followed by transcriptomic analysis of the data to reveal biological effects, which was performed in-part by the thesis author. The phenotypic analysis concluded that worm strength was reduced in space compared to on ground, as also observed in astronauts (English et al.,

2015) while the transcriptomics analysis of the worms, which were frozen on the ISS, revealed that spaceflight induced a stress and immune response, cytoskeletal and mitochondrial dysregulation, and altered neuromuscular signalling.

This work has contributed to fundamental space biology knowledge by validating *C. elegans* as a model organism for studying muscular strength decline in space, and by generating new hypotheses for spaceflight-induced strength decline which are potentially translational to other organisms, including humans. For example, while loss of muscle performance in astronauts is known to be primarily caused by mechanical unloading and is significantly mitigated via nutrition and resistive exercise, there are likely other factors that play a role. Mitochondrial dysregulation is a consistent molecular signature of spaceflight exposure across organisms and tissues (da Silveira et al., 2020), and is known to contribute to muscle strength decline on Earth (Ellwood et al., 2021; Vintila et al., 2023). Thus, the observation of mitochondrial dysregulation in this study where strength decline was also observed, suggests mitochondrial dysregulation as a candidate mechanism for driving loss of muscle performance in space; a hypothesis which can now be explored and validated in follow-up studies through the use of mitochondrial interventions, such as mitochondrial hydrogen sulphide (H_2S) (Ellwood et al., 2021; Vintila et al., 2023). Potentially, these follow-up studies could provide sufficient evidence for testing and incorporating mitochondrial countermeasures for mitigating muscle performance decline into occupational health programmes for human spaceflight.

7.2.2 Transcriptomic analysis of skin in space

The second case study included within this thesis was a broader analysis project, focussed primarily on transcriptomics changes in skin samples from mice flown to space, but also including transcriptomic and blood

biochemistry changes in astronauts. The study aimed to elucidate the biological changes underpinning dermatological issues, such as atopic dermatitis and skin rashes. This work resulted in a publication, which is also incorporated within this thesis as Chapter VI “Data Analysis Case Study - Transcriptomic Analysis of Mice” (H. Cope et al., 2024). Briefly, the study involved the analysis of space-flown mouse skin transcriptomic datasets from multiple missions, biochemical profiles of 50 NASA astronauts and human transcriptomic datasets generated from blood and hair samples of Japan Aerospace Exploration Agency (JAXA) astronauts, as well as blood samples obtained from the NASA Twins Study, and skin and blood samples from the first civilian commercial mission, Inspiration4. Key biological changes related to skin health, DNA damage & repair, and mitochondrial dysregulation were identified, including signatures of dysregulated collagen formation and skin barrier function.

This work has contributed to fundamental space biology knowledge by pooling together existing data from space biology studies to corroborate existing hypotheses of molecular hallmarks of spaceflight, and to generate hypotheses pertaining to molecular alterations potentially underlying dermatological issues in spaceflight. Notably, skin issues are one of the most common reported issues in astronauts and the pathology is not well understood. Thus, this work is an example of extracting additional scientific value from archived space biology data. For example, one hypothesis posed by the work implicates the downregulation of *FLG* and *CASP14* genes as potential drivers of some of the reported skin degradation and atopic dermatitis symptoms in space. *FLG* and *CASP14* are both genes responsible for the formation and maintenance of the skin barrier, so dysregulation of these genes on Earth has been linked to impaired skin barrier function with associated symptoms like dryness and atopic dermatitis. Changes to skin barrier integrity may also interact with other spaceflight-induced biological alterations to cause these skin issues, like immune dysregulation and

microbiome shifts in space, including colonisation of *Staphylococcus aureus*. The value of this study is that it shows that *FLG* and *CASP14* downregulation appears to persist across mice and humans, and appears to persist regardless of experimental differences like age, strain and sex; giving credence to the idea that this could be a common molecular signature underlying some spaceflight skin issues. Thus, the analysis of archived data has generated novel hypotheses which can be tested in follow-up targeted studies. For example, an investigation of skin barrier gene regulation alongside skin barrier phenotypic measures in analogue studies could help to dissect the role that individual spaceflight stressors play in skin barrier health, such as microgravity and air quality. If spaceflight-induced skin barrier dysfunction is eventually linked to dermatological issues in space via follow-up studies, this could provide evidence for potential countermeasures, such as programmatic use of topical moisturisers and emollients.

Another potentially important finding from this study was the downregulation/suppression of collagen/extracellular matrix (ECM) genes and pathways in the MHU-2 dataset, where samples were collected 1-day post-flight. From this data it can be hypothesised that the genes were heavily dysregulated by the hypergravity event of return to Earth, as the RR-5 (30 days post-flight recovery) and RR-7 (samples collected in space) datasets showed an opposing pattern. This hypothesis of collagen/ECM function suppression due to hypergravity and activation due to microgravity (as with the RR-7 dataset) is consistent with a previous study which showed a reduction in collagen synthesis occurred under hypergravity conditions in cultured human fibroblasts a 143% increase in collagen synthesis during microgravity (Seitzer et al., 1995). Further investigation is needed to explore whether such changes in collagen and ECM pathways are linked to phenotypic effects, such as rashes and skin sensitivity issues that occur within a day post-landing (Law et al., 2020).

7.3. Reflections on the Future of Data Generation and Data Science in Space Omics

7.3.1 Space biology requires international harmonisation of data governance

Within the last decade, the NASA GeneLab has emerged as the *de facto* data repository for housing space omics data (Scott et al., 2020). NASA GeneLab is specialised for the curation of space biology studies, which often have unique metadata parameters that are not easily supported by general repositories housing omics from terrestrial studies. In spite of the repository being NASA funded, it houses a significant quantity of datasets generated from international partners, including ESA and JAXA. NASA GeneLab serves as a cornerstone of the international space omics community, with a significant portion of the space omics community involved in NASA GeneLab analysis working groups (AWGs). Although international groups like the ESA Space Omics Topical Team and International Standards for Space Omics Processing (ISSOP) consortium have formed with no direct affiliation to NASA GeneLab, the majority of members from these groups are also integrated within the NASA GeneLab AWGs, and alignment with NASA GeneLab is embedded within the groups' manifestos (Madrigal et al., 2020; Rutter et al., 2020). Recent integration of NASA GeneLab into the wider NASA Open Science Data Repository - alongside biospecimen repositories and the NASA Ames Life Science Data Archive for phenotypic and environmental data - aims to further unite the international space biology community, opening the doors to researchers who have been primarily focussed on phenotypic space biology studies. Ultimately, the monumental efforts of NASA GeneLab and the new NASA Open Science Data Repository have served to unite the international space biology community, which has led to significant research output and improved impact through initiatives like the Cell Press "The biology of spaceflight" paper package, and an upcoming Nature Portfolio package.

This united open-science approach has clear benefits and presents a logical path forwards for the field.

Recent efforts in performing omics studies with human subjects in space have occurred primarily at the level of individual space agencies or commercial companies. Indeed, within this thesis, considerations and recommendations for performing similar studies in European and UK astronauts have been put forward. Currently, the data generated from these existing studies is housed in a variety of locations (Gupta et al., 2024). Data from JAXA human space omics studies is available via the OSDR, data from the NASA Twins study is not currently publicly accessible, and data from commercial initiatives has been housed in a new commercial database (Urquieta et al., 2022). Furthermore, it has recently been announced that the commercial astronaut data will be available via NASA OSDR and the Space Omics and Medical Atlas (SOMA), a new database for commercial studies which will also house data from JAXA astronauts, and from the NASA Twins study due to a re-consent process. While there has been some alignment in measures collected - due to an overlap in the research team for the NASA twins study and recent commercial studies like Inspiration4, with new data being generated and new databases being created - it is crucial that the community seeks to adopt a united international approach. As such, any new human space omics initiatives should strive to collect methods that are comparable across all human omics studies and strive for aligned data governance policies, with data stored in the same repositories, curated similarly, and comparable in terms of policies regarding access and usage. Achieving this alignment will require standardisation efforts that overcome differences in laws and regulations internationally. This ambition is embedded within the 2020 Artemis accords, signed by several European countries, the UK, the US, and Japan, amongst others; the accords include a commitment to “the open sharing of scientific data”.

The following text is adapted from a ipublication (Rutter et al., 2024); the text is the sole work of the thesis author.

Rutter LA, **Cope H***, MacKay MJ*, Herranz R, Das S, Ponomarev SA, Costes SV, Paul AM, Barker R, Taylor DM, Bezdan D, et al. Astronaut omics and the impact of space on the human body at scale. Nature Communications. 2024 Jun 11;15(1):4952.*

*These authors contributed equally to this work.

In consequence to the increasing adoption of human omics research in space, it has become crucial to increase standardisation of policies for regulating the collection, storage, access, and usage of astronauts' (sometimes called spaceflight participants, or SFPs) omics data. Based on recommendations from a 2014 report by the National Academy of Medicine (formerly the Institute of Medicine) (Committee on Ethics Principles and Guidelines for Health Standards for Long Duration and Exploration Spaceflights et al., 2014), NASA instituted a policy (NPD 7170.1) regarding the collection and usage of genomic data for human research. However, the thesis author is not currently aware of other prominent space agencies - or commercial spaceflight companies - publishing public policies regarding astronaut or crew omics data. Notably, the United Nations recently formed a working group on "Space and Global Health". In their draft resolution (A/AC.105/C.1/L.402), they encourage Member States to "establish a policy-enabled environment and governance mechanisms, with due consideration of legal and ethical issues, for removing barriers to the effective use of space-based technologies, including telemedicine solutions". The current rarity of human omics collection and curation can be viewed as a barrier to the development and deployment of emerging space-based health technologies, including precision healthcare. Thus, in the following section, legal and ethical challenges pertaining to policy making in this context are considered, concluding with some thoughts on

how best to achieve balanced policies which support space omics research while protecting the rights of the participants.

Dangers of restrictive policies: Scientific potential under utilised

As mentioned throughout this manuscript, standardised collection and curation of human space omics data has great potential to accelerate scientific research into human health in space and enable the deployment of precision medicine technologies to improve occupational healthcare during spaceflight missions. Therefore, underutilisation of the scientific and healthcare potential for human space omics data is the primary danger of overly restrictive policies.

Firstly, if policies restrict the operational usage of human space omics data in occupational health processes, such as countermeasure design and long-term health monitoring, the full potential for reducing health risks might not be realised. Consequently, this could endanger crew members, reduce likelihood of mission success, and potentially infringe upon the duty of care space agencies have for their astronaut employees. This duty of care includes not just the immediate health and safety of astronauts during missions, but also their long-term health due to occupational exposures such as cosmic radiation (Committee on Ethics Principles and Guidelines for Health Standards for Long Duration and Exploration Spaceflights et al., 2014).

Secondly, if policies do not allow for sufficient data to be collected and processed from humans in space for research purposes, the pace of space life sciences research will be slower, with increased dependency on ground-based analogs. Similarly, restrictive policies for data accessibility/sharing would also slow the pace of research. When investigators contribute data to the scientific community through open-science initiatives, it can be analysed via varied approaches, including comparatively to other datasets to yield novel insights. In the context of

spaceflight omics data, the value of open-science has recently been demonstrated through a large-scale multi-omics analysis comparing datasets from the NASA GeneLab biorepository, to uncover mitochondrial dysregulation as a key hub of the biological response to spaceflight (da Silveira et al., 2020). Biorepositories can be leveraged with “Big data” approaches, where patterns such as radiation sensitivity, can be learned via application of machine learning methods to collections of multi-omics datasets (Lewis and Kemp, 2021). With increased collection of omics data from humans in space, these kinds of approaches could be applied to human spaceflight, such as to predict susceptibility to radiation-based health issues in individual crewmembers (Belli et al., 2002). Limiting accessibility to human space omics data creates a barrier to analysis efforts, thus every decrement from complete open access reduces the scientific utility of the data.

A final consequence of restrictive policies is reductions in the scope and precision of the data provided to the wider scientific community. For the NASA Twins study, authors of one follow-up paper that accessed and reanalyzed the original study data indicated that single nucleotide polymorphisms (SNPs), indels, copy number variations (CNVs), and structural genomic elements could not be included in the publication for confidentiality purposes (Schmidt et al., 2020), thus limiting the scope of their public facing analysis. Regarding precision, efforts to anonymize and sanitise the data prior to data sharing - including data aggregation approaches and withholding metadata - could reduce the scientific utility, particularly as individual differences, such as gene variants associated with susceptibility to health issues, are essential information for unlocking precision healthcare approaches (Jones et al., 2024).

Dangers of permissible policies: Exposure to ethical harms

On the other hand, while scientific utilisation of human space omics data could reduce health risks for crews, if policies are overly permissible, omics research could also lead to harm by exposing human subjects to ethical issues. In the case of space agencies, this would also infringe on their duty of care to their astronaut employees. Due to the familial nature of genomics, these issues may also affect family members. Notably, compared to typical terrestrial cohorts, many of these ethical challenges are amplified by the unique nature of the spacefaring population; astronauts are easily identifiable due to their current rarity and status as public figures (Reed and Antonsen, 2018).

One ethical issue that is particularly amplified by the standing of astronauts is that of privacy. The term “genetic privacy” is used ubiquitously, yet privacy breaches could apply to any identifiable omic-based personal data. Privacy breaches would involve disclosure of this data against the will of the participants, which could lead to psychological harm. Specifically, policies may be considered overly permissible if they fail to account for the risks of identifiability from the full range of omic data types and data formats. It is well understood that individuals can be identified using genomics data, such as SNPs (Lippert et al., 2017). However, identification via other omic data types, such as transcriptomics (Gürsoy et al., 2020), proteomics (Bandeira et al., 2021; Geyer et al., 2021), microbiomics (Elhaik et al., 2021), and combinations of omic types (multi-omics) (Dupras and Bunnik, 2021), is a developing area of the literature. Identifying specific participants becomes possible through “linkage attacks” when phenotypic information is available in addition to the omics data (Gürsoy et al., 2020). Due to the nature of astronauts as public figures, phenotypic information such as ethnicity, age, and biological sex is readily available to the public. The small population size also means that the number of possible matches is limited, increasing the likelihood of successful linkage attacks. Aside from linkage attacks, privacy may also be violated via attacks on the data storage system. Thus, policies

may be considered overly permissible if they do not ensure appropriate security and safeguarding mechanisms against data breaches.

Relating to the issue of privacy is the ethical issue of genetic discrimination, which refers to individuals receiving differential or unfair treatment based on their genetic data. Policies permitting space agencies to use omics data predictive of health issues, such as genetic variants, in astronaut selection or during processes such as flight assignment, could be seen as discriminatory. Indeed, many countries worldwide have enacted laws to prevent the use of genetic information in employment decisions (Joly et al., 2020). It is worth noting that this is somewhat of a controversial and complex topic; for example, in the United States, the military can use routine genetic screening results to inform assignment decisions, such as to withhold deploying troops with G6PD deficiency to locations which would require them to take antimalarial drugs, since doing so can cause life-threatening hemolytic reactions (Mehlman and Parasidis, 2021). While the United States Space Force could present an interesting exception, NASA is not a branch of the military, so it is subject to the Genetic Information Nondiscrimination Act 2008 (GINA), and therefore it would seem that it cannot use omic information in assignment and employment decisions (Reed and Antonsen, 2018), which is also aligned with the current NASA policy (NPD 7170.1). However, even with precision medicine approaches, sending an individual with a genetic predisposition to spaceflight-associated risks, such as radiation susceptibility, on a long-duration mission to Mars may still increase the likelihood of a serious medical incident. Thus, it has been suggested that omic information would be useful for spaceflight selection and assignment processes (Cortese et al., 2018). This presents an ethical dilemma, where on the one hand, using omic information for flight assignment without the individual's consent could be seen as discriminatory, and on the other hand, not doing so could potentially endanger the crew and reduce likelihood of mission safety and success. An additional consideration

here is that some spaceflight-relevant genetic mutations may be linked to certain populations on Earth, which raises further issues of equity (Jones et al., 2024). Using omics for precision healthcare does appear to be supported under the current NASA policy and GINA (Antonsen and Reed, 2019), and this could hopefully decrease health risks associated with individual differences during spaceflight missions to an acceptable level, while refraining from using omics information for employment selection and flight assignment. Altogether, this balance may align with what is perhaps the current consensus for an appropriate compromise, based on regulations in many regions worldwide.

One further ethical issue that will be faced by space agencies and commercial spaceflight companies collecting astronaut omics is the risk of incidental findings (Antonsen and Reed, 2019). For example, collection and analysis of astronaut omics data could reveal unexpected findings, such as predisposition to late onset Alzheimer's disease. In some cases, these findings may not be medically actionable. Disclosing these findings to the astronaut could cause them and their family members psychological harm. Policies would be seen as overly permissible if they do not factor in the risk of incidental findings and fail to implement appropriate and clear procedures for disclosure and genetic counselling.

Keys to developing balanced policies

Ultimately, policymaking for human space omics should be considered as a balancing act between developing policies that are not so overly restrictive that they limit the scientific potential of the data, and not so overly permissible that they invoke risk and expose participants and their families to harm. Standardisation of policies could ensure that an appropriate balance is struck, safeguarding against ethical risks, while enabling appropriate accessibility for scientific and healthcare utilisation of the data. This need for balance is embedded into the mixed role of space agencies,

as research institutions and clinical care providers for astronaut employees (Antonsen and Reed, 2019). While each dataset may still need to be considered on a case-by-case basis, such as for assessing identifiability of the data, clear policy frameworks could reduce bias and help to prevent unwarranted delays associated with unstandardized policy. As the number of humans and enterprises in space increases, and the understanding of omics increases in the context of ethical issues including privacy and discrimination, standards will need to be continually reassessed and updated.

It is worth noting that while ethical challenges such as privacy are certainly amplified at present by the relatively unique standing of astronauts as a small cohort of figures in the public spotlight, to some extent these issues are also shared by rare disease cohorts and elite athlete cohorts on Earth, and so existing policies in areas including handling incidental findings and data sharing appear to present an apt starting point for discussion (Nguyen et al., 2019).

It is essential that standardisation of policymaking for human space omics be considered with an international perspective, as space agencies and commercial companies are subject to different laws based on geographical region, and may also be influenced by cultural differences (Arnould, 2011) (Boyd et al., 2009). For example, laws surrounding genetic discrimination and handling of incidental findings vary internationally (Joly et al., 2020) (Thorogood et al., 2019). Where appropriate, policies should aim to find harmonious solutions within the framework of internationally established legislation. For example, while notably not signed by the US, the Declaration of Helsinki (World Medical Association, 2013) respecting the rights of the human data subjects. Careful consideration must also be given to the specific challenges of the astronaut population and the societal role of human spaceflight. On the ISS, the IGA (intergovernmental agreement) extends the jurisdiction of countries to their registered elements and

personnel; for example, European law applies to European astronauts and the Columbus laboratory module on the ISS, with the existence of member state national laws adding further complexity (Cope et al., 2022). Conversely, the Human Research Multilateral Review Board (HRMRB) is a review board between international partners to ensure that all human subject research onboard the ISS is conducted ethically according to internationally agreed principles. It is likely that future missions, including commercial missions involving multi-national participants, will require a similar navigation of international laws to create policies. This need to design harmonious international policy raises concerns over fairness; for example, if a country has more restrictive laws in areas such as data protection, it may become challenging for citizens of that country to participate in international space missions, omic studies, and omic-based technologies such as precision healthcare.

Similarly, given the role of international collaboration in human spaceflight research, establishing clear and harmonious policies for lawful sharing and processing of human space omics data will be crucial (Saulnier et al., 2019). Where human space omics data is deemed to be identifiable, databases should adhere to high standards of security, and access levels or other technical solutions for reducing ethical harm risks should be considered (Kaye et al., 2016). For example, data sanitisation methods can manipulate the data to reduce the risk of identification (Gürsoy et al., 2020), but should be balanced against the potential loss of scientific utility. Additionally, federated approaches (Thorogood et al., 2021), such as federating learning for training AI models without moving the data across jurisdictions, could prove useful (Casaletto et al., 2022; Rieke et al., 2020). In adherence to FAIR (Findable, Accessible, Interoperable and Reusable) principles (Wilkinson et al., 2016), human space omics data should be as “as open as possible, as closed as necessary”, in order to maximise scientific utility while protecting the rights of the participants.

Furthermore, an important aspect of developing appropriate policies will be striving for meaningful informed consent from participants, both in terms of consenting to venture into space, and consenting to the collection, usage, and sharing of their omics data. To the former point, omics technologies, such as predicted risk based on individual differences, could help to provide participants with additional data, enabling greater decision autonomy to improve the informed consent process for space travel. Broad consent has been proposed as a potential solution to gain consent in the context of omics research, where it may be hard to define future uses of the data at a granular level, and the identifiability risks of different omics data types may be challenging to accurately quantify (Steinsbekk et al., 2013). When carefully implemented, it has been argued that broad consent may be an appropriate choice for obtaining consent in the context of space omics studies and has been used in the recent commercial spaceflight health data repository established by the Translational Research Institute for Space Health (TRISH) (Urquieta et al., 2022).

Discussions regarding the development of appropriate policies need to commence early, so that anticipatory policies are ready as the science is ready (Antonsen and Reed, 2019). With deep molecular profiling of astronauts now at the forefront of space biology research, and a key part of the climate of commercial spaceflight and ambitious missions outside of LEO, now would seem the opportune time to kick-start discussions into appropriate policy design. ISSOP, with its international membership and links to NASA GeneLab and the ESA Space Omics Topical Team, is well positioned to help guide standardisation of policy making regarding astronaut or commercial crews' omics data, particularly from a scientific perspective. ISSOP can work closely with groups such as the Global Alliance for Genomic Health (GA4GH) to merge expertise in spaceflight omics research and terrestrial omics policymaking, which could help to

ensure that policies maximise scientific utility while protecting the rights of the astronauts and crews.

7.3.2 Space biology data should be generated and curated with AI and other Data Science approaches in mind

Efforts by the NASA OSDR in metadata curation have significantly improved the ease of comparing across multiple archived datasets. In the skin analysis project, reported in this thesis, curated metadata on mission variables including the spaceflight exposure and post-spaceflight recovery time was particularly useful for studying the potential role of confounding variables on the biological changes observed. For example, the hypergravity event of returning to Earth was identified as a potential explanation for opposing collagen biosynthesis signatures in one of the missions compared to the other two missions. This also highlights the dangers of pooling together multiple space biology datasets without consideration for these variables which may have significant confounding effects. As such, it is recommended that future analysis projects explore the potential biological impacts of these metadata differences prior to pooling multiple space biology datasets.

Another key takeaway from working with archived space biology data, via these two case studies and other analysis projects performed within the context of the thesis, is the need to consider data science approaches during the design stages of the space biology experiments themselves. For example, when analysing the data from the micro-16 mission, due to the culturing protocols it was unfortunately impossible to perform any omic-phenotypic analysis at an individual level, such as linking the strength loss observed in an individual worm and then linking it to the gene expression changes within that worm specifically. Similarly, this issue was encountered when working with data from the CARA plant study; it was challenging to link the gene expression of an individual growth plate directly

to the picture of the individual growth plate, likely because the transcriptomic data and image data had until recently been housed in two separate repositories with different identifiers. Eventually, it was only possible to resolve the linkage issue for the CARA analysis through contacting the study's original Principal Investigator, an opportunity which may not always be available. Further to this point, developing computer vision models to extract strength and movement measures from worms in the micro-16 experiment was within the original scope of this thesis. However, this proved to be impractical compared to manual analysis because changes to the engineering and science design meant that multiple worms were included in the spaceflight data but not the training data for the model; meaning that complicated interactions like overlapping worms and worms touching the same pillars were introduced into the data and were beyond the capabilities of the model. Had the data science requirements been included within the planning stage, this issue could have been spotted earlier or potentially avoided. Similarly, the involvement of a data scientist in planning of the experimental hardware could have led to improved analysis capabilities (e.g. putting marks in the centre of pillars for easier detection of pillar displacement). As such, it makes sense to include a data science requirements plan within the project planning stage for spaceflight experiments, to ensure that the data from the experiment is aligned with downstream analysis capabilities.

Similarly, machine learning (ML) and artificial intelligence (AI) approaches are seeing increased adoption within space life sciences (Sanders et al., 2023; RT Scott et al., 2023), but there are considerable challenges to applying these compared to terrestrial life sciences; namely the limited sample sizes (and therefore training data) and differences between space biology data and terrestrial data limiting the generalisability of models trained on terrestrial data. Trivially, when space biology data has the same properties as terrestrial data, ML and AI approaches pre-trained on terrestrial data can

be used for inference on the space biology data. For example DNA sequencing data produced by the same machine on samples processed in the same workflow has the same properties for space biology experiments as it does for terrestrial experiments, so AI/ML tools like DeepVariant for calling germline variants (Poplin et al., 2018) will generalise to the space biology data. Thus, when designing experiments for space biology it makes sense to ensure that data generated has as much alignment as possible with terrestrial data. Striving for relative similarity between terrestrial data and space biology data also increases the potential for transfer learning approaches, which require a significantly reduced quantity of data for training a specialised model. For example, fine-tuning takes a neural network pre-trained on lots of data (e.g. biological data) and uses a small amount of new data (e.g. space biology data) to tune a subset of the existing weights or all of the weights over a small number of iterations, improving performance on data similar (e.g. other space biology data) to the new data used for the fine-tuning process. Data augmentation also presents a strategy for overcoming low sample sizes, when carefully designed (Vermeulen et al., 2023).

7.4. Closing thought

Space agencies and commercial entities have set forth unprecedented plans for expansion beyond Earth, yet the mission planning and risk mitigation requirements for travel and settlement of a diverse human population beyond low earth orbit (BLEO) exposes significant gaps in our understanding of biology and medicine both on Earth and in space. As discussed within this thesis, accelerated data generation and advanced integrated data analyses present a clear path forward for elucidating these gaps and for informing the logistics and engineering to ensure the safe and successful expansion of humanity into space.

References

Achouri Y, Noël G, Vertommen D, et al. (2004) Identification of a dehydrogenase acting on D-2-hydroxyglutarate. *Biochemical Journal* 381(Pt 1): 35–42.

Adav SS, Subbaiaih RS, Kerk SK, et al. (2018) Studies on the Proteome of Human Hair - Identification of Histones and Deamidated Keratins. *Scientific reports* 8(1): 1599.

Afshinnekoo E, Scott RT, MacKay MJ, et al. (2020) Fundamental Biological Features of Spaceflight: Advancing the Field to Enable Deep-Space Exploration. *Cell* 183(5): 1162–1184.

Ahsan MM, Luna SA and Siddique Z (2022) Machine-Learning-Based Disease Diagnosis: A Comprehensive Review. *Healthcare (Basel, Switzerland)* 10(3).

Akiyama T, Horie K, Hinoi E, et al. (2020) How does spaceflight affect the acquired immune system? *NPJ microgravity* 6: 14.

Alavi A, Bogu GK, Wang M, et al. (2022) Real-time alerting system for COVID-19 and other stress events using wearable data. *Nature medicine* 28(1): 175–184.

Alia E and Kerr PE (2021) Vitamin D: Skin, sunshine, and beyond. *Clinics in dermatology* 39(5): 840–846.

Andersson A, Bergenstråhle J, Asp M, et al. (2020) Single-cell and spatial transcriptomics enables probabilistic inference of cell type topography. *Communications biology* 3(1): 565.

Andreev-Andrievskiy A, Popova A, Boyle R, et al. (2014) Mice in Bion-M 1 space mission: training and selection. *PloS one* 9(8): e104830.

Andrew P, Valiani S, MacIsaac J, et al. (2014) Tamoxifen-associated skin reactions in breast cancer patients: from case report to literature review. *Breast cancer research and treatment* 148(1): 1–5.

Angeloni D, Baatout S, Benchoua A, et al. (2021) ROADMAP #9: Biology in space and analogue environments. In: *ESA SciSpacE white papers – final version*, 15 December 2021. Available at: https://esamultimedia.esa.int/docs/HRE/09_Biology_Biology_in_Space_Analogue_Environments.pdf (accessed 7 September 2023).

Antonsen EL and Reed RD (2019) Policy Considerations for Precision Medicine in Human Spaceflight. *Hous. J. Health L. & Pol'y* 19. HeinOnline: 1.

Antonsen EL, Myers JG, Boley L, et al. (2022) Estimating medical risk in human spaceflight. *NPJ microgravity* 8(1): 8.

Arc-Chagnaud C, Py G, Fovet T, et al. (2020) Evaluation of an Antioxidant and Anti-inflammatory Cocktail Against Human Hypoactivity-Induced Skeletal Muscle Deconditioning. *Frontiers in physiology* 11: 71.

Archer SN, Möller-Levet C, Bonmatí-Carrión M-Á, et al. (2024) Extensive dynamic changes in the human transcriptome and its circadian organization during prolonged bed rest. *iScience* 27(3): 109331.

Arnould J (2011) *Icarus' Second Chance: The Basis and Perspectives of Space Ethics*. Springer Science & Business Media. Available at: https://play.google.com/store/books/details?id=lLbfQfW4zzkC.

Arora S, Pattwell SS, Holland EC, et al. (2020) Variability in estimated gene expression among commonly used RNA-seq pipelines. *Scientific reports* 10(1): 2734.

Auñón-Chancellor SM, Pattarini JM, Moll S, et al. (2020) Venous Thrombosis during Spaceflight. *The New England journal of medicine* 382(1): 89–90.

Avila-Herrera A, Thissen J, Urbaniak C, et al. (2020) Crewmember microbiome may influence microbial composition of ISS habitable surfaces. *PloS one* 15(4): e0231838.

Aziz S, Raza MA, Noreen M, et al. (2022) Astropharmacy: Roles of Pharmacist in Space. *Innovations in pharmacy* 13(3).

Baldwin HE, Nighland M, Kendall C, et al. (2013) 40 years of topical tretinoin use in review. *Journal of drugs in dermatology: JDD* 12(6): 638–642.

Bandeira N, Deutsch EW, Kohlbacher O, et al. (2021) Data Management of Sensitive Human Proteomics Data: Current Practices, Recommendations, and Perspectives for the Future. *Molecular & cellular proteomics: MCP* 20: 100071.

Barker R, Kruse CPS, Johnson C, et al. (2023) Meta-analysis of the space flight and microgravity response of the Arabidopsis plant transcriptome. *NPJ microgravity* 9(1): 21.

Barrila J, Sarker SF, Hansmeier N, et al. (2021) Evaluating the effect of spaceflight on the host-pathogen interaction between human intestinal epithelial cells and Salmonella Typhimurium. *NPJ microgravity* 7(1): 9.

Bass JJ, Hardy EJO, Inns TB, et al. (2021) Atrophy Resistant vs. Atrophy Susceptible Skeletal Muscles: 'aRaS' as a Novel Experimental Paradigm to Study the Mechanisms of Human Disuse Atrophy. *Frontiers in physiology* 12: 653060.

Beheshti A, Cekanaviciute E, Smith DJ, et al. (2018) Global transcriptomic analysis suggests carbon dioxide as an environmental stressor in spaceflight: A systems biology GeneLab case study. *Scientific reports* 8(1): 4191.

Beheshti A, McDonald JT, Hada M, et al. (2021) Genomic Changes Driven by Radiation-Induced DNA Damage and Microgravity in Human Cells. *International journal of molecular sciences* 22(19).

Behjati S, Gundem G, Wedge DC, et al. (2016) Mutational signatures of ionizing radiation in second malignancies. *Nature communications* 7: 12605.

Belli M, Sapora O and Tabocchini MA (2002) Molecular targets in cellular response to ionizing radiation and implications in space radiation protection. *Journal of radiation research* 43 Suppl: S13–9.

Be NA, Avila-Herrera A, Allen JE, et al. (2017) Whole metagenome profiles of particulates collected from the International Space Station. *Microbiome* 5(1): 81.

Benjamini Y and Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society* 57(1). Wiley: 289–300.

Berekméri A, Tiganescu A, Alase AA, et al. (2019) Non-invasive Approaches for the Diagnosis of Autoimmune/Autoinflammatory Skin Diseases-A Focus on Psoriasis and Lupus erythematosus. *Frontiers in immunology* 10: 1931.

Bergboer JGM, Tjabringa GS, Kamsteeg M, et al. (2011) Psoriasis risk genes of the late

cornified envelope-3 group are distinctly expressed compared with genes of other LCE groups. *The American journal of pathology* 178(4): 1470–1477.

Bergström A, Stringer C, Hajdinjak M, et al. (2021) Origins of modern human ancestry. *Nature* 590(7845): 229–237.

Berrios DC, Galazka J, Grigorev K, et al. (2021) NASA GeneLab: interfaces for the exploration of space omics data. *Nucleic acids research* 49(D1): D1515–D1522.

Bezdan D, Grigorev K, Meydan C, et al. (2020) Cell-free DNA (cfDNA) and Exosome Profiling from a Year-Long Human Spaceflight Reveals Circulating Biomarkers. *iScience* 23(12): 101844.

Björnsson B, Borrebaeck C, Elander N, et al. (2019) Digital twins to personalize medicine. *Genome medicine* 12(1): 4.

Blottner D, Moriggi M, Trautmann G, et al. (2023) Space Omics and Tissue Response in Astronaut Skeletal Muscle after Short and Long Duration Missions. *International journal of molecular sciences* 24(4).

Blum HE (2017) The human microbiome. *Advances in medical sciences* 62(2): 414–420.

Bohr A and Memarzadeh K (2020) Chapter 2 - The rise of artificial intelligence in healthcare applications. In: Bohr A and Memarzadeh K (eds) *Artificial Intelligence in Healthcare*. Academic Press, pp. 25–60. Available at: http://dx.doi.org/10.1016/B978-0-12-818438-7.00002-2.

Boscolo D and Durante M (2022) Dose Limits and Countermeasures for Mitigating Radiation Risk in Moon and Mars Exploration. *Physics* 4(1). Multidisciplinary Digital Publishing Institute: 172–184.

Boyd AS and King LE Jr (1999) Tamoxifen-induced remission of psoriasis. *Journal of the American Academy of Dermatology* 41(5 Pt 2): 887–889.

Boyd JE, Kanas NA, Salnitskiy VP, et al. (2009) Cultural differences in crewmembers and mission control personnel during two space station programs. *Aviation, space, and environmental medicine* 80(6): 532–540.

Brady AF, Demirdas S, Fournel-Gigleux S, et al. (2017) The Ehlers-Danlos syndromes, rare types. *American journal of medical genetics. Part C, Seminars in medical genetics* 175(1): 70–115.

Braun N, Binder S, Grosch H, et al. (2019) Current Data on Effects of Long-Term Missions on the International Space Station on Skin Physiological Parameters. *Skin pharmacology and physiology* 32(1): 43–51.

Bray NL, Pimentel H, Melsted P, et al. (2016) Near-optimal probabilistic RNA-seq quantification. *Nature biotechnology* 34(5): 525–527.

Brereton NJB, Pitre FE and Gonzalez E (2021) Reanalysis of the Mars500 experiment reveals common gut microbiome alterations in astronauts induced by long-duration confinement. *Computational and structural biotechnology journal* 19: 2223–2235.

Broome JD (1981) L-Asparaginase: discovery and development as a tumor-inhibitory agent. *Cancer treatment reports* 65 Suppl 4: 111–114.

Brown SJ and Irvine AD (2008) Atopic eczema and the filaggrin story. *Seminars in cutaneous medicine and surgery* 27(2): 128–137.

Bruhns S (2012) *Yuri Gagarin: The Spaceman*. Hyperink. Available at: https://books.google.co.uk/books?hl=en&lr=&id=2Ix-BwAAQBAJ&oi=fnd&pg=PT2&dq=yuri+gagarin+8g+reentry&ots=-SBRahysTN&sig=mFx4CYqA2ZUrhfwyD5eKGjjc2DI&redir_esc=y#v=onepage&q&f=false.

Brzhozovskiy A, Kononikhin A, Indeykina M, et al. (2017) Label-free study of cosmonaut's urinary proteome changes after long-duration spaceflights. *European journal of mass spectrometry* 23(4): 225–229.

Brzhozovskiy AG, Kononikhin AS, Pastushkova LC, et al. (2019) The Effects of Spaceflight Factors on the Human Plasma Proteome, Including Both Real Space Missions and Ground-Based Experiments. *International journal of molecular sciences* 20(13).

Buchheim J-I, Ghislin S, Ouzren N, et al. (2020) Plasticity of the human IgM repertoire in response to long-term spaceflight. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology* 34(12): 16144–16162.

Bult CJ and Sternberg PW (2023) The alliance of genome resources: transforming comparative genomics. *Mammalian genome: official journal of the International Mammalian Genome Society* 34(4): 531–544.

Burgess C and Dubbs C (2007) *Animals in Space: From Research Rockets to the Space Shuttle*. Springer-Praxis Books in Space Exploration. New York, NY: Springer.

Burke MJ and Zalewska-Szewczyk B (2022) Hypersensitivity reactions to asparaginase therapy in acute lymphoblastic leukemia: immunology and clinical consequences. *Future oncology* 18(10): 1285–1299.

Burton AS, Stahl SE, John KK, et al. (2020) Off Earth Identification of Bacterial Populations Using 16S rDNA Nanopore Sequencing. *Genes* 11(1).

Cadena SM, Zhang Y, Fang J, et al. (2019) Skeletal muscle in MuRF1 null mice is not spared in low-gravity conditions, indicating atrophy proceeds by unique mechanisms in space. *Scientific reports* 9(1): 9397.

Cahill T, Cope H, Bass JJ, et al. (2021) Mammalian and Invertebrate Models as Complementary Tools for Gaining Mechanistic Insight on Muscle Responses to Spaceflight. *International journal of molecular sciences* 22(17).

Campanella A, De Summa S and Tommasi S (2019) Exhaled breath condensate biomarkers for lung cancer. *Journal of breath research* 13(4): 044002.

Cancedda R, Liu Y, Ruggiu A, et al. (2012) The Mice Drawer System (MDS) experiment and the space endurance record-breaking mice. *PloS one* 7(5): e32243.

Cary M, Podshivalova K and Kenyon C (2020) Application of Transcriptional Gene Modules to Analysis of Caenorhabditis elegans' Gene Expression Data. *G3* 10(10): 3623–3638.

Casaletto J, Parsons M, Markello C, et al. (2022) Federated analysis of BRCA1 and BRCA2 variation in a Japanese cohort. *Cell genomics* 2(3).

Castro-Wallace SL, Chiu CY, John KK, et al. (2017) Nanopore DNA Sequencing and Genome Assembly on the International Space Station. *Scientific reports* 7(1): 18022.

Cekanaviciute E, Rosi S and Costes SV (2018) Central Nervous System Responses to Simulated Galactic Cosmic Rays. *International journal of molecular sciences* 19(11).

Cekanaviciute E, Tran D, Nguyen H, et al. (2023) Mouse genomic associations with in vitro

sensitivity to simulated space radiation. *Life sciences and space research* 36: 47–58.

Čepelak I, Dodig S and Pavić I (2019) Filaggrin and atopic march. *Biochemia medica: casopis Hrvatskoga drustva medicinskih biokemicara / HDMB* 29(2): 020501.

Chai Z, Brereton P, Suzuki T, et al. (2003) 17 beta-hydroxysteroid dehydrogenase type XI localizes to human steroidogenic cells. *Endocrinology* 144(5): 2084–2091.

Checinska Sielaff A, Urbaniak C, Mohan GBM, et al. (2019) Characterization of the total and viable bacterial and fungal communities associated with the International Space Station surfaces. *Microbiome* 7(1): 50.

Chen EY, Tan CM, Kou Y, et al. (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC bioinformatics* 14: 128.

Cheng J, Novati G, Pan J, et al. (2023) Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science* 381(6664): eadg7492.

Chen Rui, Mias GI, Li-Pook-Than J, et al. (2012) Personal omics profiling reveals dynamic molecular and medical phenotypes. *Cell* 148(6): 1293–1307.

Chen Y, Guo J, Xing S, et al. (2021) Global-Scale Metabolomic Profiling of Human Hair for Simultaneous Monitoring of Endogenous Metabolome, Short- and Long-Term Exposome. *Frontiers in chemistry* 9: 674265.

Chen Yuxin, Chen Yongsheng, Shi C, et al. (2018) SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *GigaScience* 7(1): 1–6.

Choi SY, Saravia-Butler A, Shirazi-Fard Y, et al. (2020) Validation of a New Rodent Experimental System to Investigate Consequences of Long Duration Space Habitation. *Scientific reports* 10(1): 2336.

Christensen NJ, Demharter S, Machado M, et al. (2022) Identifying interactions in omics data for clinical biomarker discovery using symbolic regression. *Bioinformatics* 38(15): 3749–3758.

Chyba CF and Hand KP (2005) Astrobiology: The Study of the Living Universe. *Annual review of astronomy and astrophysics* 43(Volume 43, 2005). Annual Reviews: 31–74.

Cikach FS Jr and Dweik RA (2012) Cardiovascular biomarkers in exhaled breath. *Progress in cardiovascular diseases* 55(1): 34–43.

Clément G, Rittweger J, Nitsche A, et al. (2022) Assessing the effects of artificial gravity in an analog of long-duration spaceflight: The protocol and implementation of the AGBRESA bed rest study. *Frontiers in physiology* 13: 976926.

Clément G, Crucian BE, Downs M, et al. (2022) International standard measures during the VaPER bed rest study. *Acta astronautica* 190: 208–217.

Clough E and Barrett T (2016) The Gene Expression Omnibus Database. *Methods in molecular biology* 1418: 93–110.

Cohain A, Divaraniya AA, Zhu K, et al. (2017) Exploring the Reproducibility of Probabilistic Causal Molecular Network Models. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing* 22: 120–131.

Committee on Ethics Principles and Guidelines for Health Standards for Long Duration and Exploration Spaceflights, Board on Health Sciences Policy and Institute of Medicine

(2014) *Health Standards for Long Duration and Exploration Spaceflight: Ethics Principles, Responsibilities, and Decision Framework* (J Kahn et al.eds). Washington (DC): National Academies Press (US). Available at: http://dx.doi.org/10.17226/18576.

Cook SJ, Jarrell TA, Brittin CA, et al. (2019) Whole-animal connectomes of both Caenorhabditis elegans sexes. *Nature* 571(7763): 63–71.

Cope H, Willis CRG, MacKay MJ, et al. (2022) Routine omics collection is a golden opportunity for European human research in space and analog environments. *Patterns of prejudice*: 100550.

Cope H, Deane C, Szewczyk NJ, et al. (2023) A data collection programme for improving healthcare in UK human spaceflight ventures. *Journal of the British Interplanetary Society* 76(6). British Interplanetary Society: 213–220.

Cope H., Elsborg J, Demharter S, et al. (2024) More than a Feeling: Transcriptomics Analysis Reveals Molecular Alterations Underpinning Spaceflight Dermatology. *Communications Medicine* In press.

Cope Henry, Elsborg J, Demharter S, et al. (2024) Transcriptomics analysis reveals molecular alterations underpinning spaceflight dermatology. *Communication & medicine* 4(1). Springer Science and Business Media LLC: 106.

Corbett AH (2018) Post-transcriptional regulation of gene expression and human disease. *Current opinion in cell biology* 52: 96–104.

Cortese F, Klokov D, Osipov A, et al. (2018) Vive la radiorésistance!: converging research in radiobiology and biogerontology to enhance human radioresistance for deep space exploration and colonization. *Oncotarget* 9(18): 14692–14722.

Corvelyn M, De Beukelaer N, Duelen R, et al. (2020) Muscle Microbiopsy to Delineate Stem Cell Involvement in Young Patients: A Novel Approach for Children With Cerebral Palsy. *Frontiers in physiology* 11: 945.

Cronin NJ, Finni T and Seynnes O (2020) Fully automated analysis of muscle architecture from B-mode ultrasound images with deep learning. *arXiv [eess.IV]*. Available at: http://arxiv.org/abs/2009.04790.

Crucian B, Stowe R, Mehta S, et al. (2013) Immune system dysregulation occurs during short duration spaceflight on board the space shuttle. *Journal of clinical immunology* 33(2): 456–465.

Crucian B, Stowe RP, Mehta S, et al. (2015) Alterations in adaptive immunity persist during long-duration spaceflight. *NPJ microgravity* 1: 15013.

Crucian B, Johnston S, Mehta S, et al. (2016) A case of persistent skin rash and rhinitis with immune system dysregulation onboard the International Space Station. *The journal of allergy and clinical immunology in practice* 4(4): 759–762.e8.

Crucian B, Babiak-Vazquez A, Johnston S, et al. (2016) Incidence of clinical symptoms during long-duration orbital spaceflight. *International journal of general medicine* 9: 383–391.

Crucian BE, Zwart SR, Mehta S, et al. (2014) Plasma cytokine concentrations indicate that in vivo hormonal regulation of immunity is altered during long-duration spaceflight. *Journal of interferon & cytokine research: the official journal of the International Society for Interferon and Cytokine Research* 34(10): 778–786.

Dalmaijer ES, Nord CL and Astle DE (2022) Statistical power for cluster analysis. *BMC bioinformatics* 23(1): 205.

Danko DC, Singh N, Butler DJ, et al. (2020) Genetic and Immunological Evidence for Microbial Transfer Between the International Space Station and an Astronaut. *bioRxiv*. Available at: http://dx.doi.org/10.1101/2020.11.10.376954 (accessed 26 February 2022).

Däpp C, Schmutz S, Hoppeler H, et al. (2004) Transcriptional reprogramming and ultrastructure during atrophy and recovery of mouse soleus muscle. *Physiological genomics* 20(1): 97–107.

da Silveira WA, Fazelinia H, Rosenthal SB, et al. (2020) Comprehensive Multi-omics Analysis Reveals Mitochondrial Stress as a Central Biological Hub for Spaceflight Impact. *Cell* 183(5): 1185–1201.e20.

Deane CS, Borg J, Cahill T, et al. (2022) Space omics research in Europe: contributions, geographical distribution and ESA member state funding schemes. *iScience* 0(0). Elsevier.

Deane CS, Phillips BE, Willis CRG, et al. (2023) Proteomic features of skeletal muscle adaptation to resistance exercise training as a function of age. *GeroScience* 45(3): 1271–1287.

Deing V, Roggenkamp D, Kühnl J, et al. (2013) Oxytocin modulates proliferation and stress responses of human skin cells: implications for atopic dermatitis. *Experimental dermatology* 22(6): 399–405.

Denecker G, Ovaere P, Vandenabeele P, et al. (2008) Caspase-14 reveals its secrets. *The Journal of cell biology* 180(3): 451–458.

de Souza GO, Wasinski F and Donato J Jr (2022) Characterization of the metabolic differences between male and female C57BL/6 mice. *Life sciences* 301: 120636.

D'hondt S, Guillemyn B, Syx D, et al. (2018) Type III collagen affects dermal and vascular collagen fibrillogenesis and tissue integrity in a mutant Col3a1 transgenic mouse model. *Matrix biology: journal of the International Society for Matrix Biology* 70: 72–83.

Donovan BM, Bastarache L, Turi KN, et al. (2019) The current state of omics technologies in the clinical management of asthma and allergic diseases. *Annals of allergy, asthma & immunology: official publication of the American College of Allergy, Asthma, & Immunology* 123(6): 550–557.

Douglas GL, Zwart SR and Smith SM (2020) Space food for thought: Challenges and considerations for food and nutrition on exploration missions. *The Journal of nutrition* 150(9). Elsevier BV: 2242–2244.

Dunn J, Kidzinski L, Runge R, et al. (2021) Wearable sensors enable personalized predictions of clinical laboratory measurements. *Nature medicine* 27(6): 1105–1112.

Dupras C and Bunnik EM (2021) Toward a Framework for Assessing Privacy Risks in Multi-Omic Research and Databases. *The American journal of bioethics: AJOB* 21(12): 46–64.

Edqvist P-HD, Fagerberg L, Hallström BM, et al. (2015) Expression of human skin-specific genes defined by transcriptomics and antibody-based profiling. *The journal of histochemistry and cytochemistry: official journal of the Histochemistry Society* 63(2): 129–141.

Elhaik E, Ahsanuddin S, Robinson JM, et al. (2021) The impact of cross-kingdom molecular forensics on genetic privacy. *Microbiome* 9(1): 114.

Ellwood RA, Hewitt JE, Torregrossa R, et al. (2021) Mitochondrial hydrogen sulfide supplementation improves health in the C. elegans Duchenne muscular dystrophy model. *Proceedings of the National Academy of Sciences of the United States of America* 118(9).

Ellwood RA, Slade L, Lewis J, et al. (2022) Sulfur amino acid supplementation displays therapeutic potential in a C. elegans model of Duchenne muscular dystrophy. *Communications biology* 5(1): 1255.

Elsborg J and Salvatore M (2023) Using LLMs and Explainable ML to Analyze Biomarkers at Single-Cell Level for Improved Understanding of Diseases. *Biomolecules* 13(10). Multidisciplinary Digital Publishing Institute: 1516.

Emms A, Castleman J, Allen S, et al. (2022) Next Generation Sequencing after Invasive Prenatal Testing in Fetuses with Congenital Malformations: Prenatal or Neonatal Investigation. *Genes* 13(9).

English KL, Lee SMC, Loehr JA, et al. (2015) Isokinetic Strength Changes Following Long-Duration Spaceflight on the ISS. *Aerospace medicine and human performance* 86(12 Suppl): A68–A77.

English KL, Downs M, Goetchius E, et al. (2020) High intensity training during spaceflight: results from the NASA Sprint Study. *NPJ microgravity* 6: 21.

Erwin DH, Laflamme M, Tweedt SM, et al. (2011) The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science* 334(6059): 1091–1097.

Etheridge T, Rahman M, Gaffney CJ, et al. (2015) The integrin-adhesome is required to maintain muscle structure, mitochondrial ATP production, and movement forces in Caenorhabditis elegans. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology* 29(4): 1235–1246.

Everroad C, Foster J, Galazka JM, et al. (2021) Space Biology Beyond LEO Instrumentation & Science Series - Science Working Group 2021 Annual Report. Epub ahead of print 8 November 2021.

Fan Y and Pedersen O (2021) Gut microbiota in human metabolic health and disease. *Nature reviews. Microbiology* 19(1): 55–71.

Ferranti F, Del Bianco M and Pacelli C (2020) Advantages and Limitations of Current Microgravity Platforms for Space Biology Research. *NATO Advanced Science Institutes series E: Applied sciences* 11(1). Multidisciplinary Digital Publishing Institute: 68.

Fitts RH, Trappe SW, Costill DL, et al. (2010) Prolonged space flight-induced alterations in the structure and function of human skeletal muscle fibres. *The Journal of physiology* 588(Pt 18): 3567–3592.

Fitts RH, Colloton PA, Trappe SW, et al. (2013) Effects of prolonged space flight on human skeletal muscle enzyme and substrate profiles. *Journal of applied physiology* 115(5): 667–679.

Fuat A, Adlen E, Monane M, et al. (2024) A polygenic risk score added to a QRISK®2 cardiovascular disease risk calculator demonstrated robust clinical acceptance and clinical utility in the primary care setting. *European journal of preventive cardiology*

31(6). Oxford Academic: 716–722.

Gabel L, Liphardt AM, Hulme PA, et al. (2022) Pre-flight exercise and bone metabolism predict unloading-induced bone loss due to spaceflight. *British journal of sports medicine* 56(4). Br J Sports Med.

Gaffney CJ, Etheridge T and Szewczyk NJ (2020) Minimizing Muscle Atrophy. In: *Handbook of Life Support Systems for Spacecraft and Extraterrestrial Habitats*. Cham: Springer International Publishing, pp. 1–27. Available at: https://link.springer.com/referenceworkentry/10.1007/978-3-319-09575-2_28-1.

Gallego Romero I, Pai AA, Tung J, et al. (2014) RNA-seq: impact of RNA degradation on transcript quantification. *BMC biology* 12(1). Springer Science and Business Media LLC: 42.

Gallo MA and Kaufman D (1997) Antagonistic and agonistic effects of tamoxifen: significance in human cancer. *Seminars in oncology* 24(1 Suppl 1): S1–71–S1–80.

Gao H, Hamp T, Ede J, et al. (2023) The landscape of tolerated genetic variation in humans and primates. *bioRxiv : the preprint server for biology*. Epub ahead of print 2 May 2023. DOI: 10.1101/2023.05.01.538953.

Gao P, Shen X, Zhang X, et al. (2022) Precision environmental health monitoring by longitudinal exposome and multi-omics profiling. *Genome research* 32(6): 1199–1214.

Gao Y (2015) Urine is a better biomarker source than blood especially for kidney diseases. *Advances in experimental medicine and biology* 845: 3–12.

Garrett-Bakelman FE, Darshi M, Green SJ, et al. (2019) The NASA Twins Study: A multidimensional analysis of a year-long human spaceflight. *Science* 364(6436).

Geyer PE, Mann SP, Treit PV, et al. (2021) Plasma Proteomes Can Be Reidentifiable and Potentially Contain Personally Sensitive and Incidental Findings. *Molecular & cellular proteomics: MCP* 20. American Society for Biochemistry and Molecular Biology.

Giesen M, Gruedl S, Holtkoetter O, et al. (2011) Ageing processes influence keratin and KAP expression in human hair follicles. *Experimental dermatology* 20(9): 759–761.

Gilbert JA, Blaser MJ, Caporaso JG, et al. (2018) Current understanding of the human microbiome. *Nature medicine* 24(4): 392–400.

Ginsburg GS and Phillips KA (2018) Precision Medicine: From Science To Value. *Health Affairs*. Available at: http://dx.doi.org/10.1377/hlthaff.2017.1624.

Goetz LH and Schork NJ (2018) Personalized medicine: motivation, challenges, and progress. *Fertility and sterility* 109(6): 952–963.

Granger E, Davies G and Keogh RH (2022) Treatment patterns in people with cystic fibrosis: have they changed since the introduction of ivacaftor? *Journal of cystic fibrosis: official journal of the European Cystic Fibrosis Society* 21(2). Elsevier BV: 316–322.

Green MJ, Shivakumar T, Robson DJ, et al. (2022) Astropharmacy - Healthcare for Earth, the Moon, Mars and beyond. In: *27th European Low Gravity Research Association Biennial Symposium and General Assembly*, 2022.

Griko YV, Loftus DJ, Stolc V, et al. (2022) Private Spaceflight: A New Landscape for Dealing with Medical Risk. *Life Sciences in Space Research* 33: 41–47.

Guarnieri JW, Dybas JM, Fazelinia H, et al. (2023) Core mitochondrial genes are

down-regulated during SARS-CoV-2 infection of rodent and human hosts. *Science translational medicine* 15(708): eabq1533.

Gupta A, Barker R, Madrigal P, et al. (2024) Emerging Trends in Space Omics. *In review*. Epub ahead of print 2024.

Gürsoy G, Emani P, Brannon CM, et al. (2020) Data Sanitization to Reduce Private Information Leakage from Functional Genomics. *Cell* 183(4): 905–917.e16.

Gürsoy G, Li T, Liu S, et al. (2021) Functional genomics data: privacy risk assessment and technological mitigation. *Nature reviews. Genetics*. Epub ahead of print 10 November 2021. DOI: 10.1038/s41576-021-00428-7.

Haendel M, Vasilevsky N, Unni D, et al. (2020) How many rare diseases are there? *Nature reviews. Drug discovery* 19(2): 77–78.

Halldorsson BV, Eggertsson HP, Moore KHS, et al. (2022) The sequences of 150,119 genomes in the UK Biobank. *Nature* 607(7920): 732–740.

Hallinan D (2020) Broad consent under the GDPR: an optimistic perspective on a bright future. *Life sciences, society and policy* 16(1): 1.

Hallinan D, Bernier A, Cambon-Thomsen A, et al. (2021) International transfers of personal data for health research following Schrems II: a problem in need of a solution. *European journal of human genetics: EJHG* 29(10): 1502–1509.

Harada S, Hashizume T, Nemoto K, et al. (2016) Fluid dynamics alter Caenorhabditis elegans body length via TGF-β/DBL-1 neuromuscular signaling. *NPJ microgravity* 2: 16006.

Harmanci A and Gerstein M (2016) Quantification of private information leakage from phenotype-genotype data: linking attacks. *Nature methods* 13(3): 251–256.

Harris KM, Weber T, Greaves D, et al. (2022) Going against the flow: are venous thromboembolism and impaired cerebral drainage critical risks for spaceflight? *Journal of applied physiology* 132(1): 270–273.

Hart KL, Kimura SL, Mushailov V, et al. (2013) Improved eye- and skin-color prediction based on 8 SNPs. *Croatian medical journal* 54(3): 248–256.

Hasin Y, Seldin M and Lusis A (2017) Multi-omics approaches to disease. *Genome biology* 18(1): 83.

Hauschild A-C, Frisch T, Baumbach JI, et al. (2015) Carotta: Revealing Hidden Confounder Markers in Metabolic Breath Profiles. *Metabolites* 5(2): 344–363.

Heer M and Paloski WH (2006) Space motion sickness: incidence, etiology, and countermeasures. *Autonomic neuroscience: basic & clinical* 129(1-2): 77–79.

Heinrich HTMW (2008) Change in skin physiological parameters in space - report on and results of the first study on man. *Skin pharmacology and physiology* 21(5): 283–292.

Herranz R, da Silveira W, Bezdan D, et al. (2022) Building the Space Omics Topical Team to boost European space researchers' role in the international consortia redefining spaceflight-generated datasets. *iScience* 25(9): 104868.

Hewitt JE, Pollard AK, Lesanpezeshki L, et al. (2018) Muscle strength deficiency and mitochondrial dysfunction in a muscular dystrophy model of Caenorhabditis elegans and its functional response to drugs. *Disease models & mechanisms* 11(12).

Higashibata A, Szewczyk NJ, Conley CA, et al. (2006) Decreased expression of myogenic transcription factors and myosin heavy chains in Caenorhabditis elegans muscles developed during spaceflight. *The Journal of experimental biology* 209(Pt 16): 3209–3218.

Higashibata A, Hashizume T, Nemoto K, et al. (2016) Microgravity elicits reproducible alterations in cytoskeletal and metabolic gene and protein expression in space-flown Caenorhabditis elegans. *NPJ microgravity* 2: 15022.

Higashitani A, Hashizume T, Sugimoto T, et al. (2009) C. elegans RNAi space experiment (CERISE) in Japanese Experiment Module KIBO. *Uchu Seibutsu Kagaku* 23(4): 183–187.

Higashitani A, Teranishi M, Nakagawa Y, et al. (2023) Increased mitochondrial Ca2+ contributes to health decline with age and Duchene muscular dystrophy in C. elegans. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology* 37(4): e22851.

High-Level Advisory Group on Human and Robotic Space Exploration for Europe (2023) *Revolution Space: Europe's Mission for Space Exploration*. March. Available at: https://esamultimedia.esa.int/docs/corporate/h-lag_brochure.pdf.

Hirahata T, Ul Quraish R, Quraish AU, et al. (2022) Liquid Biopsy: A Distinctive Approach to the Diagnosis and Prognosis of Cancer. *Cancer informatics* 21: 11769351221076062.

Hodkinson PD, Anderton RA, Posselt BN, et al. (2017) An overview of space medicine. *British journal of anaesthesia* 119(suppl_1): i143–i153.

Honda Y, Higashibata A, Matsunaga Y, et al. (2012) Genes down-regulated in spaceflight are involved in the control of longevity in Caenorhabditis elegans. *Scientific reports* 2: 487.

Hong SP, Oh Y, Jung M, et al. (2010) Topical calcitriol restores the impairment of epidermal permeability and antimicrobial barriers induced by corticosteroids. *The British journal of dermatology* 162(6): 1251–1260.

Horie K, Kato T, Kudo T, et al. (2019) Impact of spaceflight on the murine thymus and mitigation by exposure to artificial gravity during spaceflight. *Scientific reports* 9(1): 19866.

Horii A, Mitani K, Masumura C, et al. (2017) Hippocampal gene expression, serum cortisol level, and spatial memory in rats exposed to hypergravity. *Journal of vestibular research: equilibrium & orientation* 27(4): 209–215.

Hoste E, Kemperman P, Devos M, et al. (2011) Caspase-14 is required for filaggrin degradation to natural moisturizing factors in the skin. *The Journal of investigative dermatology* 131(11). Elsevier BV: 2233–2241.

Howell GR, Soto I, Zhu X, et al. (2012) Radiation treatment inhibits monocyte entry into the optic nerve head and prevents neuronal damage in a mouse model of glaucoma. *The Journal of clinical investigation* 122(4): 1246–1261.

Hrach HC, O'Brien S, Steber HS, et al. (2020) Transcriptome changes during the initiation and progression of Duchenne muscular dystrophy in Caenorhabditis elegans. *Human molecular genetics* 29(10): 1607–1623.

Hsia E, Johnston MJ, Houlden RJ, et al. (2008) Effects of topically applied acitretin in reconstructed human epidermis and the rhino mouse. *The Journal of investigative*

dermatology 128(1): 125–130.

Huang Y-T, Vanderweele TJ and Lin X (2014) JOINT ANALYSIS OF SNP AND GENE EXPRESSION DATA IN GENETIC ASSOCIATION STUDIES OF COMPLEX DISEASES. *The annals of applied statistics* 8(1): 352–376.

Hu J, Li C, Wang S, et al. (2021) Genetic variants are identified to increase risk of COVID-19 related mortality from UK Biobank data. *Human genomics* 15(1): 10.

Husna N, Aiba T, Fujita S-I, et al. (2024) Release of CD36-associated cell-free mitochondrial DNA and RNA as a hallmark of space environment response. *Nature communications* 15(1). Springer Science and Business Media LLC: 4814.

Hvid M, Johansen C, Deleuran B, et al. (2011) Regulation of caspase 14 expression in keratinocytes by inflammatory cytokines--a possible link between reduced skin barrier function and inflammation? *Experimental dermatology* 20(8): 633–636.

Ibrahim W, Carr L, Cordell R, et al. (2021) Breathomics for the clinician: the use of volatile organic compounds in respiratory diseases. *Thorax* 76(5): 514–521.

Igawa S, Choi JE, Wang Z, et al. (2019) Human Keratinocytes Use Sphingosine 1-Phosphate and its Receptors to Communicate Staphylococcus aureus Invasion and Activate Host Defense. *The Journal of investigative dermatology* 139(8): 1743–1752.e5.

Inalöz HS, Deveci E, Inalöz SS, et al. (2002) The effects of tamoxifen on rat skin. *European journal of gynaecological oncology* 23(1): 50–52.

Ishitsuka Y and Roop DR (2022) Loricrin at the Boundary between Inside and Outside. *Biomolecules* 12(5).

Itadani H, Mizuarai S and Kotani H (2008) Can systems biology understand pathway activation? Gene expression signatures as surrogate markers for understanding the complexity of pathway activation. *Current genomics* 9(5): 349–360.

Jackson B, Tilli CMLJ, Hardman MJ, et al. (2005) Late cornified envelope family in differentiating epithelia--response to calcium and ultraviolet irradiation. *The Journal of investigative dermatology* 124(5): 1062–1070.

Jacobsen A, de Miranda Azevedo R, Juty N, et al. (2020) FAIR principles: interpretations and implementation considerations. *Data Intelligence* 2(1-2). MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info …: 10–29.

Jing J and Gao Y (2018) Urine biomarkers in the early stages of diseases: current status and perspective. *Discovery medicine* 25(136): 57–65.

Jochems A, Deist TM, van Soest J, et al. (2016) Distributed learning: Developing a predictive model based on data from multiple hospitals without data leaving the hospital - A real life proof of concept. *Radiotherapy and oncology: journal of the European Society for Therapeutic Radiology and Oncology* 121(3): 459–467.

Joly Y, Dupras C, Pinkesz M, et al. (2020) Looking Beyond GINA: Policy Approaches to Address Genetic Discrimination. *Annual review of genomics and human genetics* 21: 491–507.

Jones C, Overbey E, Author T, et al. (2024) The SpaceX Inspiration4 mission reveals inflight molecular and physiological metrics from an all-civilian crew. *Nature (in review)*. Epub ahead of print 2024.

Jönsson KI, Rabbow E, Schill RO, et al. (2008) Tardigrades survive exposure to space in low Earth orbit. *Current biology: CB* 18(17): R729–R731.

Jost PD (2008) Simulating human space physiology with bed rest. *Hippokratia* 12 Suppl 1(Suppl 1): 37–40.

Jumper J, Evans R, Pritzel A, et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596(7873): 583–589.

Jung M, Choi J, Lee S-A, et al. (2014) Pyrrolidone carboxylic acid levels or caspase-14 expression in the corneocytes of lesional skin correlates with clinical severity, skin barrier function and lesional inflammation in atopic dermatitis. *Journal of dermatological science* 76(3): 231–239.

Kahan V, Andersen ML, Tomimori J, et al. (2009) Stress, immunity and skin collagen integrity: evidence from animal models and clinical conditions. *Brain, behavior, and immunity* 23(8): 1089–1095.

Kamp E, Ashraf M, Musbahi E, et al. (2022) Menopause, skin and common dermatoses. Part 2: skin disorders. *Clinical and experimental dermatology* 47(12): 2117–2122.

Katahira K (2001) [Global harmonization in the care and use of laboratory animals for the space life science research]. *Uchu Seibutsu Kagaku* 15(1): 30–34.

Kaye J, Briceño Moraia L, Mitchell C, et al. (2016) Access Governance for Biobanks: The Case of the BioSHaRE-EU Cohorts. *Biopreservation and biobanking* 14(3): 201–206.

Kim B-S, Alcantara AV Jr, Moon J-H, et al. (2023) Comparative Analysis of Muscle Atrophy During Spaceflight, Nutritional Deficiency and Disuse in the Nematode Caenorhabditis elegans. *International journal of molecular sciences* 24(16).

Kim J, Author A, Author A, et al. (2024) Single-cell multi-ome and immune profiles of the Inspiration4 crew reveal cell-type and sex-specific responses to spaceflight. *Nature (in review)*. Epub ahead of print 2024.

Kim SK, Lund J, Kiraly M, et al. (2001) A gene expression map for Caenorhabditis elegans. *Science* 293(5537): 2087–2092.

Kim W, Underwood RS, Greenwald I, et al. (2018) OrthoList 2: A New Comparative Genomic Analysis of Human and Caenorhabditis elegans Genes. *Genetics* 210(2): 445–461.

Kinross JM, Darzi AW and Nicholson JK (2011) Gut microbiome-host interactions in health and disease. *Genome medicine* 3(3): 14.

Kirchmeier P, Zimmer A, Bouadjar B, et al. (2017) Whole-Exome-Sequencing Reveals Small Deletions in CASP14 in Patients with Autosomal Recessive Inherited Ichthyosis. *Acta dermato-venereologica* 97(1): 102–104.

Kitano H (2002) Systems biology: a brief overview. *Science* 295(5560): 1662–1664.

Klaus R, Niyazi M and Lange-Sperandio B (2021) Radiation-induced kidney toxicity: molecular and cellular pathogenesis. *Radiation oncology* 16(1): 43.

Kömüves LG, Hanley K, Jiang Y, et al. (1998) Ligands and activators of nuclear hormone receptors regulate epidermal differentiation during fetal rat skin development. *The Journal of investigative dermatology* 111(3): 429–433.

Kononikhin AS, Brzhozovskiy AG, Ryabokon AM, et al. (2019) Proteome Profiling of the Exhaled Breath Condensate after Long-Term Spaceflights. *International journal of*

molecular sciences 20(18).

Kousathanas A, Pairo-Castineira E, Rawlik K, et al. (2022) Whole genome sequencing reveals host factors underlying critical Covid-19. *Nature*. Epub ahead of print 7 March 2022. DOI: 10.1038/s41586-022-04576-6.

Krämer A, Green J, Pollard J Jr, et al. (2014) Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* 30(4): 523–530.

Kretz R, Bozorgmehr B, Kariminejad MH, et al. (2011) Defect in proline synthesis: pyrroline-5-carboxylate reductase 1 deficiency leads to a complex clinical phenotype with collagen and elastin abnormalities. *Journal of inherited metabolic disease* 34(3): 731–739.

Kuivaniemi H and Tromp G (2019) Type III collagen (COL3A1): Gene and protein structure, tissue distribution, and associated diseases. *Gene* 707: 151–171.

Kuo T-R and Chen C-H (2017) Bone biomarker for the clinical assessment of osteoporosis: recent developments and future perspectives. *Biomarker research* 5: 18.

Lagrelius M, Wahlgren C-F, Bradley M, et al. (2020) Filaggrin gene mutations in relation to contact allergy and hand eczema in adolescence. *Contact dermatitis* 82(3): 147–152.

Lai Polo S-H, Saravia-Butler AM, Boyko V, et al. (2020) RNAseq Analysis of Rodent Spaceflight Experiments Is Confounded by Sample Collection Techniques. *iScience* 23(12): 101733.

Lander ES, Linton LM, Birren B, et al. (2001) Initial sequencing and analysis of the human genome. *Nature* 409(6822): 860–921.

Landrum MJ, Lee JM, Benson M, et al. (2016) ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic acids research* 44(D1): D862–8.

Lane HW, Bourland C, Barrett A, et al. (2013) The role of nutritional research in the success of human space flight. *Advances in nutrition* 4(5): 521–523.

Laranjeiro R, Harinath G, Pollard AK, et al. (2021) Spaceflight affects neuronal morphology and alters transcellular degradation of neuronal debris in adult Caenorhabditis elegans. *iScience* 24(2): 102105.

Law J, Gilmore S and Kelly S (2020) Postflight Rash and Skin Sensitivity Following a Year-Long Spaceflight Mission. *Aerospace medicine and human performance* 91(7): 604–607.

Laws JM, Caplan N, Bruce C, et al. (2020) Systematic review of the technical and physiological constraints of the Orion Multi-Purpose Crew Vehicle that affect the capability of astronauts to exercise effectively during spaceflight. *Acta astronautica* 170: 665–677.

Lee AG, Mader TH, Gibson CR, et al. (2017) Space Flight-Associated Neuro-ocular Syndrome. *JAMA ophthalmology* 135(9): 992–994.

Lee AG, Mader TH, Gibson CR, et al. (2020) Spaceflight associated neuro-ocular syndrome (SANS) and the neuro-ophthalmologic effects of microgravity: a review and an update. *NPJ microgravity* 6: 7.

Lee D-D, Stojadinovic O, Krzyzanowska A, et al. (2009) Retinoid-responsive transcriptional changes in epidermal keratinocytes. *Journal of cellular physiology* 220(2): 427–439.

Lee I-M, Shiroma EJ, Lobelo F, et al. (2012) Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The Lancet* 380(9838): 219–229.

Lee MD, O'Rourke A, Lorenzi H, et al. (2021) Reference-guided metagenomics reveals genome-level evidence of potential microbial transmission from the ISS environment to an astronaut's microbiome. *iScience* 24(2): 102114.

Lee O, Page K, Ivancic D, et al. (2014) A randomized phase II presurgical trial of transdermal 4-hydroxytamoxifen gel versus oral tamoxifen in women with ductal carcinoma in situ of the breast. *Clinical cancer research: an official journal of the American Association for Cancer Research* 20(14): 3672–3682.

Lee PHU, Chung M, Ren Z, et al. (2022) Factors mediating spaceflight-induced skeletal muscle atrophy. *American journal of physiology. Cell physiology* 322(3): C567–C580.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in ecology & evolution* 17(4). Elsevier: 183–189.

Lephart ED and Naftolin F (2021) Menopause and the Skin: Old Favorites and New Innovations in Cosmeceuticals for Estrogen-Deficient Skin. *Dermatology and therapy* 11(1): 53–69.

Lesanpezeshki L, Qadota H, Darabad MN, et al. (2021) Investigating the correlation of muscle function tests and sarcomere organization in C. elegans. *Skeletal muscle* 11(1): 20.

Lewis JE and Kemp ML (2021) Integration of machine learning and genome-scale metabolic modeling identifies multi-omics biomarkers for radiation resistance. *Nature communications* 12(1): 2700.

Li M, Hener P, Zhang Z, et al. (2006) Topical vitamin D3 and low-calcemic analogs induce thymic stromal lymphopoietin in mouse keratinocytes and trigger an atopic dermatitis. *Proceedings of the National Academy of Sciences of the United States of America* 103(31): 11736–11741.

Limper U, Tank J, Ahnert T, et al. (2021) The thrombotic risk of spaceflight: has a serious problem been overlooked for more than half of a century? *European heart journal* 42(1): 97–100.

Linnarsson D, Carpenter J, Fubini B, et al. (2012) Toxicity of lunar dust. *Planetary and space science* 74(1): 57–71.

Lin Z, Akin H, Rao R, et al. (2023) Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 379(6637): 1123–1130.

Lippert C, Sabatini R, Maher MC, et al. (2017) Identification of individuals by trait prediction using whole-genome sequencing data. *Proceedings of the National Academy of Sciences of the United States of America* 114(38): 10166–10171.

Liu M, Zhang J, Wang Y, et al. (2020) Non-invasive proteome-wide quantification of skin barrier-related proteins using label-free LC-MS/MS analysis. *Molecular medicine reports* 21(5): 2227–2235.

Liu Z, Luo G, Du R, et al. (2020) Effects of spaceflight on the composition and function of the human gut microbiota. *Gut microbes* 11(4): 807–819.

Li X, Dunn J, Salins D, et al. (2017) Digital Health: Tracking Physiomes and Activity Using

Wearable Biosensors Reveals Useful Health-Related Information. *PLoS biology* 15(1): e2001402.

Loehr JA, Lee SMC, English KL, et al. (2011) Musculoskeletal adaptations to training with the advanced resistive exercise device. *Medicine and science in sports and exercise* 43(1): 146–156.

Love MI, Huber W and Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 15(12): 550.

Lulli M, Cialdai F, Vignali L, et al. (2018) The Coenzyme Q10 as an antiapoptotic countermeasure for retinal lesions onboard the International Space Station. *Frontiers in physiology* 9. Frontiers Media SA.

Lynch M, Ackerman MS, Gout J-F, et al. (2016) Genetic drift, selection and the evolution of the mutation rate. *Nature reviews. Genetics* 17(11): 704–714.

Mackay LK, Rahimpour A, Ma JZ, et al. (2013) The developmental pathway for CD103(+)CD8+ tissue-resident memory T cells of skin. *Nature immunology* 14(12): 1294–1301.

Mader TH, Gibson CR, Pass AF, et al. (2011) Optic disc edema, globe flattening, choroidal folds, and hyperopic shifts observed in astronauts after long-duration space flight. *Ophthalmology* 118(10): 2058–2069.

Madrigal P, Gabel A, Villacampa A, et al. (2020) Revamping Space-omics in Europe. *Cell systems* 11(6): 555–556.

Ma J, Fong SH, Luo Y, et al. (2021) Few-shot learning creates predictive models of drug response that translate from high-throughput screens to individual patients. *Nature cancer* 2(2): 233–244.

Malkani S, Chin CR, Cekanaviciute E, et al. (2020) Circulating miRNA Spaceflight Signature Reveals Targets for Countermeasure Development. *Cell reports* 33(10): 108448.

Manzano A, Weging S, Bezdan D, et al. (2023) Enhancing European capabilities for application of multi-omics studies in biology and biomedicine space research. *iScience* 26(9): 107289.

Mao XW, Pecaut MJ, Stodieck LS, et al. (2014) Biological and metabolic response in STS-135 space-flown mouse skin. *Free radical research* 48(8): 890–897.

Markaki M and Tavernarakis N (2010) Modeling human diseases in Caenorhabditis elegans. *Biotechnology journal* 5(12): 1261–1276.

Markiewicz A, Sigorski D, Markiewicz M, et al. (2021) Caspase-14-From Biomolecular Basics to Clinical Approach. A Review of Available Data. *International journal of molecular sciences* 22(11).

Marshall-Goebel K, Laurie SS, Alferova IV, et al. (2019) Assessment of Jugular Venous Blood Flow Stasis and Thrombosis During Spaceflight. *JAMA network open* 2(11): e1915011.

Martin MA and Kroetz DL (2013) Abacavir pharmacogenetics--from initial reports to standard of care. *Pharmacotherapy* 33(7): 765–775.

Massaro Tieze S, Liddell LC, Santa Maria SR, et al. (2023) BioSentinel: A Biological CubeSat for Deep Space Exploration. *Astrobiology* 23(6): 631–636.

Mastrandrea C, Greaves D and Hughson R (2021) Tracking astronaut physical activity and cardiorespiratory responses with the Bio-Monitor sensor shirt. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology* 35(S1). Wiley.

Masuda-Kuroki K and Di Nardo A (2022) Sphingosine 1-Phosphate Signaling at the Skin Barrier Interface. *Biology* 11(6).

Matey ET, Ragan AK, Oyen LJ, et al. (2022) Nine-gene pharmacogenomics profile service: The Mayo Clinic experience. *The pharmacogenomics journal* 22(1): 69–74.

Mathyk BA, Tabetah M, Karim R, et al. (2024) Spaceflight induces changes in gene expression profiles linked to insulin and estrogen. *Communications Biology* In press.

Matsuda C, Kato T, Inoue-Suzuki S, et al. (2019) Dietary intervention of mice using an improved Multiple Artificial-gravity Research System (MARS) under artificial 1 g. *NPJ microgravity* 5: 16.

Mazzolai L, Aboyans V, Ageno W, et al. (2018) Diagnosis and management of acute deep vein thrombosis: a joint consensus document from the European Society of Cardiology working groups of aorta and peripheral vascular diseases and pulmonary circulation and right ventricular function. *European heart journal* 39(47): 4208–4218.

McGregor HR, Lee JK, Mulder ER, et al. (2021) Ophthalmic changes in a spaceflight analog are associated with brain functional reorganization. *Human brain mapping* 42(13): 4281–4297.

McInnes L, Healy J and Astels S (2017) hdbscan: Hierarchical density based clustering. *Journal of open source software* 2(11). The Open Journal: 205.

McKnight G, Shah J and Hargest R (2022) Physiology of the skin. *Surgery* 40(1). Elsevier BV: 8–12.

Mehlman MJ and Parasidis E (2021) Predictive Genetic Testing by the U.S. Military: Legal and Ethical Issues. *Military medicine* 186(7-8): 726–732.

Mehta SK, Szpara ML, Rooney BV, et al. (2022) Dermatitis during Spaceflight Associated with HSV-1 Reactivation. *Viruses* 14(4).

Mencia-Trinchant N, MacKay MJ, Chin C, et al. (2021) Clonal hematopoiesis before, during, and after human spaceflight. *Cell reports* 34(6): 108740.

Metwally AA, Zhang T, Wu S, et al. (2022) Robust identification of temporal biomarkers in longitudinal omics studies. *Bioinformatics* 38(15): 3802–3811.

Mias GI, Singh VV, Rogers LRK, et al. (2021) Longitudinal saliva omics responses to immune perturbation: a case study. *Scientific reports* 11(1): 710.

Min M, Chen X-B, Wang P, et al. (2017) Role of keratin 24 in human epidermal keratinocytes. *PloS one* 12(3): e0174626.

Mirastschijski U, Lupše B, Maedler K, et al. (2019) Matrix Metalloproteinase-3 is Key Effector of TNF-α-Induced Collagen Degradation in Skin. *International journal of molecular sciences* 20(20).

Mitchell C, Ordish J, Johnson E, et al. (2020) *The GDPR and genomic data—the impact of the GDPR and DPA 2018 on genomic healthcare and research*.

Mofarrah Ramin, Mofarrah Ramina, Kränke B, et al. (2021) First report of tamoxifen-induced

baboon syndrome. *Journal of cosmetic dermatology* 20(8): 2574–2578.

Molnár-Gábor F and Korbel JO (2020) Genomic data sharing in Europe is stumbling-Could a code of conduct prevent its fall? *EMBO molecular medicine* 12(3): e11421.

Molnár-Gábor F, Sellner J, Pagil S, et al. (2021) Harmonization after the GDPR? Divergences in the rules for genetic and health data sharing in four member states and ways to overcome them by EU measures: Insights from Germany, Greece, Latvia and Sweden. *Seminars in cancer biology*. Epub ahead of print 9 December 2021. DOI: 10.1016/j.semcancer.2021.12.001.

Morrison MD, Thissen JB, Karouia F, et al. (2021) Investigation of Spaceflight Induced Changes to Astronaut Microbiomes. *Frontiers in microbiology* 12: 659179.

Moseson DL, Sasaki GH, Kraybill WG, et al. (1978) The use of antiestrogens tamoxifen and nafoxidine in the treatment of human breast cancer in correlation with estrogen receptor values. A phase II study. *Cancer* 41(3): 797–802.

Mostafa WZ and Hegazy RA (2015) Vitamin D and the skin: Focus on a complex relationship: A review. *Journal of advertising research* 6(6): 793–804.

Mostavi M, Chiu Y-C, Chen Y, et al. (2021) CancerSiamese: one-shot learning for predicting primary and metastatic tumor types unseen during model training. *BMC bioinformatics* 22(1): 244.

Movérare S, Lindberg MK, Faergemann J, et al. (2002) Estrogen receptor alpha, but not estrogen receptor beta, is involved in the regulation of the hair follicle cycling as well as the thickness of epidermis in male mice. *The Journal of investigative dermatology* 119(5): 1053–1058.

Mukherjee S, Sainis KB and Deobagkar DD (2014) F1 hybrids of BALB/c and C57BL/6 mouse strains respond differently to low-dose ionizing radiation exposure. *Journal of genetics* 93(3): 667–682.

Muller DC, Johansson M and Brennan P (2017) Lung Cancer Risk Prediction Model Incorporating Lung Function: Development and Validation in the UK Biobank Prospective Cohort Study. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology* 35(8): 861–869.

Murakami K, Sawada A, Mori T, et al. (2022) Effect of estrogen/progesterone ratio on the differentiation and the barrier function of epidermal keratinocyte and three-dimensional cultured human epidermis. *Life sciences* 293: 120356.

Murgia M, Ciciliot S, Nagaraj N, et al. (2022) Signatures of muscle disuse in spaceflight and bed rest revealed by single muscle fiber proteomics. *PNAS nexus* 1(3): gac086.

Murray JC (1994) Radiation-induced fibrosis: the structure/function relationship. *Scanning microscopy* 8(1): 79–85; discussion 85–7.

Neutelings T, Nusgens BV, Liu Y, et al. (2015) Skin physiology in microgravity: a 3-month stay aboard ISS induces dermal atrophy and affects cutaneous muscle and hair follicles cycling in mice. *NPJ microgravity* 1: 15002.

Neve B, Fernandez-Zapico ME, Ashkenazi-Katalan V, et al. (2005) Role of transcription factor KLF11 and its diabetes-associated gene variants in pancreatic beta cell function. *Proceedings of the National Academy of Sciences of the United States of America* 102(13): 4807–4812.

Nguengang Wakap S, Lambert DM, Olry A, et al. (2020) Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *European journal of human genetics: EJHG* 28(2): 165–173.

Nguyen CN and Urquieta E (2023) Contemporary review of dermatologic conditions in space flight and future implications for long-duration exploration missions. *Life sciences and space research* 36: 147–156.

Nguyen MT, Goldblatt J, Isasi R, et al. (2019) Model consent clauses for rare disease research. *BMC medical ethics* 20(1): 55.

Nithya S, Radhika T and Jeddy N (2015) Loricrin - an overview. *Journal of oral and maxillofacial pathology: JOMFP* 19(1): 64–68.

Nonomura Y, Otsuka A, Inui S, et al. (2012) Female pattern hair loss possibly caused by tamoxifen: Androgen receptor expression in the outer root sheath in the affected area. *The Journal of dermatology* 39(12): 1060–1061.

Oczypok EA, Etheridge T, Freeman J, et al. (2012) Remote automated multi-generational growth and observation of an animal in low Earth orbit. *Journal of the Royal Society, Interface / the Royal Society* 9(68): 596–599.

Ogonowska P, Gilaberte Y, Barańska-Rybak W, et al. (2020) Colonization With Staphylococcus aureus in Atopic Dermatitis Patients: Attempts to Reveal the Unknown. *Frontiers in microbiology* 11: 567090.

Oh HS and Smart RC (1996) An estrogen receptor pathway regulates the telogen-anagen hair follicle transition and influences epidermal cell proliferation. *Proceedings of the National Academy of Sciences of the United States of America* 93(22): 12525–12530.

Oka S, Leon J, Sakumi K, et al. (2016) Human mitochondrial transcriptional factor A breaks the mitochondria-mediated vicious cycle in Alzheimer's disease. *Scientific Reports*. Available at: http://dx.doi.org/10.1038/srep37889.

Ong J, Mader TH, Gibson CR, et al. (2023) Spaceflight associated neuro-ocular syndrome (SANS): an update on potential microgravity-based pathophysiology and mitigation development. *Eye* 37(12): 2409–2415.

Ornstein DL and Cushman M (2003) Cardiology patient page. Factor V Leiden. *Circulation* 107(15): e94–7.

Overbey E, Mason C, Mason C, et al. (2024) The Space Omics and Medical Atlas (SOMA): A comprehensive data resource and biobank for astronauts. *Nature (in review)*. Epub ahead of print 2024.

Overbey EG, Saravia-Butler AM, Zhang Z, et al. (2021) NASA GeneLab RNA-seq consensus pipeline: standardized processing of short-read RNA-seq data. *iScience* 24(4): 102361.

Overbey EG, Das S, Cope H, et al. (2022) Challenges and considerations for single-cell and spatially resolved transcriptomics sample collection during spaceflight. *Cell reports methods* 2(11): 100325.

Pagel O, Loroch S, Sickmann A, et al. (2015) Current strategies and findings in clinically relevant post-translational modification-specific proteomics. *Expert review of proteomics* 12(3): 235–253.

Palmer CNA, Irvine AD, Terron-Kwiatkowski A, et al. (2006) Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for

atopic dermatitis. *Nature genetics* 38(4): 441–446.

Parafati M, Giza S, Shenoy TS, et al. (2023) Human skeletal muscle tissue chip autonomous payload reveals changes in fiber type and metabolic gene expression due to spaceflight. *NPJ microgravity* 9(1): 77.

Parikh H, Carlsson E, Chutkow WA, et al. (2007) TXNIP regulates peripheral glucose metabolism in humans. *PLoS medicine* 4(5): e158.

Pastushkova LK, Kashirina DN, Brzhozovskiy AG, et al. (2019) Evaluation of cardiovascular system state by urine proteome after manned space flight. *Acta astronautica* 160: 594–600.

Pastushkova LK, Rusanov VB, Goncharova AG, et al. (2019) Urine proteome changes associated with autonomic regulation of heart rate in cosmonauts. *BMC systems biology* 13(Suppl 1): 17.

Pastushkova LK, Rusanov VB, Goncharova AG, et al. (2021) Blood Plasma Proteins Associated With Heart Rate Variability in Cosmonauts Who Have Completed Long-Duration Space Missions. *Frontiers in physiology* 12: 760875.

Patel ZS, Brunstetter TJ, Tarver WJ, et al. (2020) Red risks for a journey to the red planet: The highest priority human health risks for a mission to Mars. *NPJ microgravity* 6(1): 33.

Paul A-L, Sng NJ, Zupanska AK, et al. (2017) Genetic dissection of the Arabidopsis spaceflight transcriptome: Are some responses dispensable for the physiological adaptation of plants to spaceflight? *PloS one* 12(6): e0180186.

Paul A-L, Elardo SM and Ferl R (2022) Plants grown in Apollo lunar regolith present stress-associated transcriptomes that inform prospects for lunar exploration. *Communications biology* 5(1): 382.

Paul AM, Overbey EG, da Silveira WA, et al. (2021) Immunological and hematological outcomes following protracted low dose/low dose rate ionizing radiation and simulated microgravity. *Scientific reports* 11(1): 11452.

Pavan WJ and Sturm RA (2019) The Genetics of Human Skin and Hair Pigmentation. *Annual review of genomics and human genetics* 20: 41–72.

Pavela J, Sargsyan A, Bedi D, et al. (2022) Surveillance for jugular venous thrombosis in astronauts. *Vascular medicine* 27(4): 365–372.

Pavez LE, Baatout S, Choukér A, et al. (2021) The Future of Personalized Medicine in Space: From Observations to Countermeasures. *Frontiers in bioengineering and biotechnology* 9. Front Bioeng Biotechnol.

Perakakis N, Yazdani A, Karniadakis GE, et al. (2018) Omics, big data and machine learning as tools to propel understanding of biological mechanisms and to discover novel diagnostics and therapeutics. *Metabolism: clinical and experimental* 87: A1–A9.

Perlman RL (2016) Mouse models of human disease: An evolutionary perspective. *Evolution, medicine, and public health* 2016(1): 170–176.

Perry JRB, Voight BF, Yengo L, et al. (2012) Stratifying type 2 diabetes cases by BMI identifies genetic risk variants in LAMA1 and enrichment for risk variants in lean compared to obese cases. *PLoS genetics* 8(5): e1002741.

Petersen N, Jaekel P, Rosenberger A, et al. (2016) Exercise in space: the European Space

Agency approach to in-flight exercise countermeasures for long-duration missions on ISS. *Extreme physiology & medicine* 5: 9.

Petersen N, Lambrecht G, Scott J, et al. (2017) Postflight reconditioning for European Astronauts - A case report of recovery after six months in space. *Musculoskeletal science & practice* 27 Suppl 1. Musculoskelet Sci Pract.

Peterson JM, Wei Z, Seldin MM, et al. (2013) CTRP9 transgenic mice are protected from diet-induced obesity and metabolic dysfunction. *American journal of physiology. Regulatory, integrative and comparative physiology* 305(5): R522–33.

Poplin R, Chang P-C, Alexander D, et al. (2018) A universal SNP and small-indel variant caller using deep neural networks. *Nature biotechnology* 36(10): 983–987.

Pulley JM, Rhoads JP, Jerome RN, et al. (2020) Using What We Already Have: Uncovering New Drug Repurposing Strategies in Existing Omics Data. *Annual review of pharmacology and toxicology* 60: 333–352.

Radstake WE, Baselet B, Baatout S, et al. (2022) Spaceflight Stressors and Skin Health. *Biomedicines* 10(2).

Radugina EA, Almeida EAC, Blaber E, et al. (2018) Exposure to microgravity for 30 days onboard Bion M1 caused muscle atrophy and impaired regeneration in murine femoral Quadriceps. *Life sciences and space research* 16: 18–25.

Rahimzadeh V, Fogarty J, Caulfield T, et al. (2023) Ethically cleared to launch? *Science* 381(6665): 1408–1411.

Rahman M, Hewitt JE, Van-Bussel F, et al. (2018) NemaFlex: a microfluidics-based technology for standardized measurement of muscular strength of C. elegans. *Lab on a chip* 18(15): 2187–2201.

Ray S, Gebre S, Fogle H, et al. (2019) GeneLab: Omics database for spaceflight experiments. *Bioinformatics* 35(10): 1753–1759.

Reed RD and Antonsen EL (2018) Should NASA Collect Astronauts' Genetic Information for Occupational Surveillance and Research? *AMA journal of ethics* 20(9): E849–856.

Rehm HL, Bale SJ, Bayrak-Toydemir P, et al. (2013) ACMG clinical laboratory standards for next-generation sequencing. *Genetics in medicine: official journal of the American College of Medical Genetics* 15(9): 733–747.

Rehm HL, Page AJH, Smith L, et al. (2021) GA4GH: International policies and standards for data sharing across genomic research and healthcare. *Cell genomics* 1(2).

Rendl M, Ban J, Mrass P, et al. (2002) Caspase-14 expression by epidermal keratinocytes is regulated by retinoids in a differentiation-associated manner. *The Journal of investigative dermatology* 119(5): 1150–1155.

Resing KA, Walsh KA, Haugen-Scofield J, et al. (1989) Identification of proteolytic cleavage sites in the conversion of profilaggrin to filaggrin in mammalian epidermis. *The Journal of biological chemistry* 264(3): 1837–1845.

Restier-Verlet J, El-Nachef L, Ferlazzo ML, et al. (2021) Radiation on Earth or in Space: What Does It Change? *International journal of molecular sciences* 22(7). Multidisciplinary Digital Publishing Institute (MDPI).

Reynolds R, Little MP, Day S, et al. (2021) Cancer incidence and mortality in the USA

Astronaut Corps, 1959-2017. *Occupational and environmental medicine* 78(12): 869–875.

Richerson PJ and Boyd R (2008) *Not By Genes Alone: How Culture Transformed Human Evolution*. University of Chicago Press. Available at: https://play.google.com/store/books/details?id=dU-KtEVgK6sC.

Rid R, Wagner M, Maier CJ, et al. (2013) Deciphering the calcitriol-induced transcriptomic response in keratinocytes: presentation of novel target genes. *Journal of molecular endocrinology* 50(2): 131–149.

Rieke N, Hancox J, Li W, et al. (2020) The future of digital health with federated learning. *NPJ digital medicine* 3: 119.

Rinn JL, Bondre C, Gladstone HB, et al. (2006) Anatomic demarcation by positional variation in fibroblast gene expression programs. *PLoS genetics* 2(7): e119.

Ritchie ME, Phipson B, Wu D, et al. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research* 43(7): e47.

Rittweger J, Albracht K, Flück M, et al. (2018) Sarcolab pilot study into skeletal muscle's adaptation to long-term spaceflight. *npj Microgravity*. Available at: http://dx.doi.org/10.1038/s41526-018-0052-1.

Robin A, Van Ombergen A, Laurens C, et al. (2023) Comprehensive assessment of physiological responses in women during the ESA dry immersion VIVALDI microgravity simulation. *Nature communications* 14(1): 6311.

Robson DJ and Cappelletti C (2022) Biomedical payloads: A maturing application for CubeSats. *Acta astronautica* 191: 394–403.

Robson DJ, Ferreira Y, Cope H, et al. (2021) The WormSail CubeSat - An International Educational Project To Elevate Space Science and Education. In: *72nd International Astronautical Congress*, 2021.

Roda A, Mirasoli M, Guardigli M, et al. (2018) Advanced biosensors for monitoring astronauts' health during long-duration space missions. *Biosensors & bioelectronics* 111: 18–26.

Roderick TH (1963) THE RESPONSE OF TWENTY-SEVEN INBRED STRAINS OF MICE TO DAILY DOSES OF WHOLE-BODY X-IRRADIATION. *Radiation research* 20: 631–639.

Ronneberger O, Fischer P and Brox T (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv [cs.CV]*. Available at: http://arxiv.org/abs/1505.04597.

Rooney BV, Crucian BE, Pierson DL, et al. (2019) Herpes Virus Reactivation in Astronauts During Spaceflight and Its Application on Earth. *Frontiers in microbiology* 10: 16.

Rosenfeld JA, Mason CE and Smith TM (2012) Limitations of the human reference genome for personalized genomics. *PloS one* 7(7): e40294.

Rosso JD, Zeichner J, Alexis A, et al. (2016) Understanding the Epidermal Barrier in Healthy and Compromised Skin: Clinically Relevant Information for the Dermatology Practitioner: Proceedings of an Expert Panel Roundtable Meeting. *The Journal of clinical and aesthetic dermatology* 9(4 Suppl 1): S2–S8.

Roy-O'Reilly M, Mulavara A and Williams T (2021) A review of alterations to the brain during spaceflight and the potential relevance to crew in long-duration space exploration. *NPJ*

microgravity 7(1): 5.

Rutter L, Barker R, Bezdan D, et al. (2020) A New Era for Space Life Science: International Standards for Space Omics Processing. *Patterns (New York, N.Y.)* 1(9). Patterns (N Y).

Rutter LA, Cope H, MacKay MJ, et al. (2024) Astronaut omics and the impact of space on the human body at scale. *Nature Communications* In press.

Salama SA, Serrana M and Au WW (1999) Biomonitoring using accessible human cells for exposure and health risk assessment. *Mutation research* 436(1): 99–112.

Salvucci E (2019) The human-microbiome superorganism and its modulation to restore health. *International journal of food sciences and nutrition* 70(7): 781–795.

Sanders LM, Scott RT, Yang JH, et al. (2023) Biological research and self-driving labs in deep space supported by artificial intelligence. *Nature Machine Intelligence* 5(3). Nature Publishing Group: 208–219.

Satam H, Joshi K, Mangrolia U, et al. (2023) Next-Generation Sequencing Technology: Current Trends and Advancements. *Biology* 12(7).

Saulnier KM, Bujold D, Dyke SOM, et al. (2019) Benefits and barriers in the design of harmonized access agreements for international data sharing. *Scientific data* 6(1): 297.

Saunders G, Baudis M, Becker R, et al. (2019) Leveraging European infrastructures to access 1 million human genomes by 2022. *Nature reviews. Genetics* 20(11): 693–701.

Sauro F, De Waele J, Payler SJ, et al. (2021) Speleology as an analogue to space exploration: The ESA CAVES training programme. *Acta astronautica* 184: 150–166.

Sawyers L, Anderson C, Boyd MJ, et al. (2022) Astropharmacy: Pushing the boundaries of the pharmacists' role for sustainable space exploration. *Research in social & administrative pharmacy: RSAP*. Epub ahead of print 12 February 2022. DOI: 10.1016/j.sapharm.2022.02.002.

Schmidt MA and Goodwin TJ (2013) Personalized medicine in human space flight: using Omics based analyses to develop individualized countermeasures that enhance astronaut safety and performance. *Metabolomics: Official journal of the Metabolomic Society* 9(6): 1134–1156.

Schmidt MA, Meydan C, Schmidt CM, et al. (2020) The NASA Twins Study: The Effect of One Year in Space on Long-Chain Fatty Acid Desaturases and Elongases. *Lifestyle genomics* 13(3): 107–121.

Schopf JW, Kudryavtsev AB, Agresti DG, et al. (2002) Laser--Raman imagery of Earth's earliest fossils. *Nature* 416(6876): 73–76.

Scott A, Willis CRG, Muratani M, et al. (2023) Caenorhabditis elegans in microgravity: an omics perspective. *iScience* 7(26). Elsevier.

Scott RT, Grigorev K, Mackintosh G, et al. (2020) Advancing the Integration of Biosciences Data Sharing to Further Enable Space Exploration. *Cell reports* 33(10): 108441.

Scott RT, Antonsen EL, Sanders LM, et al. (2021) Beyond Low Earth Orbit: Biomonitoring, Artificial Intelligence, and Precision Space Health. *arXiv [q-bio.OT]*. Available at: http://arxiv.org/abs/2112.12554.

Scott RT, Sanders LM, Antonsen EL, et al. (2023) Biomonitoring and precision health in deep space supported by artificial intelligence. *Nature Machine Intelligence* 5(3). Nature

Publishing Group: 196–207.

Seitzer U, Bodo M, Müller PK, et al. (1995) Microgravity and hypergravity effects on collagen biosynthesis of human dermal fibroblasts. *Cell and tissue research* 282(3): 513–517.

Selch F, Higashibata A, Imamizo-Sato M, et al. (2008) Genomic response of the nematode Caenorhabditis elegans to spaceflight. *Advances in space research: the official journal of the Committee on Space Research* 41(5): 807–815.

Shabani M and Borry P (2018) Rules for processing genetic data for research purposes in view of the new EU General Data Protection Regulation. *European journal of human genetics: EJHG* 26(2): 149–156.

Shabani M and Marelli L (2019) Re-identifiability of genomic data and the GDPR: Assessing the re-identifiability of genomic data in light of the EU General Data Protection Regulation. *EMBO reports* 20(6).

Shackelford LC (2019) Musculoskeletal Response to Space Flight. In: Barratt MR, Baker ES, and Pool SL (eds) *Principles of Clinical Medicine for Space Flight*. New York, NY: Springer New York, pp. 581–607. Available at: https://doi.org/10.1007/978-1-4939-9889-0_19.

Shavers M, Semones E, Tomi L, et al. (2024) Space agency-specific standards for crew dose and risk assessment of ionising radiation exposures for the International Space Station. *Zeitschrift fur medizinische Physik* 34(1). Elsevier BV: 14–30.

Shelhamer M, Bloomberg J, LeBlanc A, et al. (2020) Selected discoveries from human research in space that are relevant to human health on Earth. *NPJ microgravity* 6: 5.

Shiba D, Mizuno H, Yumoto A, et al. (2017) Development of new experimental platform 'MARS'-Multiple Artificial-gravity Research System-to elucidate the impacts of micro/partial gravity on mice. *Scientific reports* 7(1): 10837.

Shirah B, Bukhari H, Pandya S, et al. (2023) Benefits of Space Medicine Research for Healthcare on Earth. *Cureus* 15(5): e39174.

Simonsen LC, Slaba TC, Guida P, et al. (2020) NASA's first ground-based Galactic Cosmic Ray Simulator: Enabling a new era in space radiobiology research. *PLoS biology* 18(5): e3000669.

Singh NK, Wood JM, Karouia F, et al. (2018) Succession and persistence of microbial communities and antimicrobial resistance genes associated with International Space Station environmental surfaces. *Microbiome* 6(1): 204.

Skon CN, Lee J-Y, Anderson KG, et al. (2013) Transcriptional downregulation of S1pr1 is required for the establishment of resident memory CD8+ T cells. *Nature immunology* 14(12): 1285–1293.

Smith FJD, Irvine AD, Terron-Kwiatkowski A, et al. (2006) Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nature genetics* 38(3): 337–342.

Smith MG, Kelley M and Basner M (2020) A brief history of spaceflight from 1961 to 2020: An analysis of missions and astronaut demographics. *Acta astronautica* 175. Elsevier BV: 290–299.

Smith SM, Heer MA, Shackelford LC, et al. (2012) Benefits for bone from resistance exercise and nutrition in long-duration spaceflight: Evidence from biochemistry and densitometry. *Journal of bone and mineral research: the official journal of the American*

Society for Bone and Mineral Research 27(9). Wiley: 1896–1906.

Smith SM, Heer M, Wang Z, et al. (2012) Long-duration space flight and bed rest effects on testosterone and other steroids. *The Journal of clinical endocrinology and metabolism* 97(1): 270–278.

Smith SM, Heer M, Shackelford LC, et al. (2015) Bone metabolism and renal stone risk during International Space Station missions. *Bone* 81: 712–720.

Soneson C, Love MI and Robinson MD (2015) Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* 4: 1521.

Soni P, Anupom T, Lesanpezeshki L, et al. (2022) Microfluidics-integrated spaceflight hardware for measuring muscle strength of Caenorhabditis elegans on the International Space Station. *NPJ microgravity* 8(1): 50.

Soni P, Edwards H, Anupom T, et al. (2023) Spaceflight Induces Strength Decline in Caenorhabditis elegans. *Cells* 12(20).

Space Research Today (2022) Extended Abstracts. 215: 72–74.

Sreedhar A, Aguilera-Aguirre L and Singh KK (2020) Mitochondria in skin health, aging, and disease. *Cell death & disease* 11(6): 444.

Stabler RA, Rosado H, Doyle R, et al. (2017) Impact of the Mk VI SkinSuit on skin microbiota of terrestrial volunteers and an International Space Station-bound astronaut. *NPJ microgravity* 3: 23.

Stahl-Rommel S, Jain M, Nguyen HN, et al. (2021) Real-Time Culture-Independent Microbial Profiling Onboard the International Space Station Using Nanopore Sequencing. *Genes* 12(1).

Stark Z, Dolman L, Manolio TA, et al. (2019) Integrating Genomics into Healthcare: A Global Responsibility. *American journal of human genetics* 104(1): 13–20.

Steinsbekk KS, Kåre Myskja B and Solberg B (2013) Broad consent versus dynamic consent in biobank research: is passive participation an ethical problem? *European journal of human genetics: EJHG* 21(9): 897–902.

Stepanek J, Blue RS and Parazynski S (2019) Space Medicine in the Era of Civilian Spaceflight. *The New England journal of medicine* 380(11): 1053–1060.

Stephens M (2017) False discovery rates: a new deal. *Biostatistics* 18(2): 275–294.

Stevenson S and Thornton J (2007) Effect of estrogens on skin aging and the potential role of SERMs. *Clinical interventions in aging* 2(3): 283–297.

Stingl JC, Welker S, Hartmann G, et al. (2015) Where Failure Is Not an Option -Personalized Medicine in Astronauts. *PloS one* 10(10). PLoS One.

Stokes M, Evetts S, Rittweger J, et al. (2016) *Recommendations for Future Post-mission Neuro-musculoskeletal Reconditioning Research and Practice Post-mission Exercise (Reconditioning) Topical Team Report*. Topical Team Report. European Space Agency. Available at: https://elib.dlr.de/105291.

Stowe RP, Pierson DL and Barrett AD (2001) Elevated stress hormone levels relate to Epstein-Barr virus reactivation in astronauts. *Psychosomatic medicine* 63(6): 891–895.

Strangman GE, Sawyer A, Fabre KM, et al. (2019) Deep-space applications for point-of-care

technologies. *Current Opinion in Biomedical Engineering* 11. Elsevier: 45–50.

Stratis D, Trudel G, Rocheleau L, et al. (2023) The Characteristic Response of the Human Leukocyte Transcriptome to 60 Days of Bed Rest and to Reambulation. *Medicine and science in sports and exercise* 55(3): 365–375.

Stroud JE, Gale MS, Zwart SR, et al. (2022) Longitudinal metabolomic profiles reveal sex-specific adjustments to long-duration spaceflight and return to Earth. *Cellular and molecular life sciences: CMLS* 79(11): 578.

Stuart T, Butler A, Hoffman P, et al. (2019) Comprehensive Integration of Single-Cell Data. *Cell* 177(7): 1888–1902.e21.

Subramanian A, Tamayo P, Mootha VK, et al. (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*. Available at: http://dx.doi.org/10.1073/pnas.0506580102.

Sudevan S, Muto K, Higashitani N, et al. (2022) Loss of physical contact in space alters the dopamine system in C. elegans. *iScience* 25(2): 103762.

Sudlow C, Gallacher J, Allen N, et al. (2015) UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine* 12(3): e1001779.

Sugita T, Yamazaki T, Makimura K, et al. (2016) Comprehensive analysis of the skin fungal microbiota of astronauts during a half-year stay at the International Space Station. *Medical mycology: official publication of the International Society for Human and Animal Mycology* 54(3): 232–239.

Sundblad P, Orlov O, Angerer O, et al. (2016) Standardization of bed rest studies in the spaceflight context. *Journal of applied physiology* 121(1): 348–349.

Swensen AC, He J, Fang AC, et al. (2021) A Comprehensive Urine Proteome Database Generated From Patients With Various Renal Conditions and Prostate Cancer. *Frontiers of medicine* 8: 548212.

Szewczyk NJ, Kozak E and Conley CA (2003) Chemically defined medium and Caenorhabditis elegans. *BMC biotechnology* 3: 19.

Szewczyk NJ, Udranszky IA, Kozak E, et al. (2006) Delayed development and lifespan extension as features of metabolic lifestyle alteration in C. elegans under dietary restriction. *The Journal of experimental biology* 209(Pt 20): 4129–4139.

Tang H, Rising HH, Majji M, et al. (2021) Long-Term Space Nutrition: A Scoping Review. *Nutrients* 14(1).

Tang Q, Jin G, Wang G, et al. (2020) Current Sampling Methods for Gut Microbiota: A Call for More Precise Devices. *Frontiers in cellular and infection microbiology* 10: 151.

Terada M, Seki M, Higashibata A, et al. (2013) Genetic analysis of the human hair roots as a tool for spaceflight experiments. *Advances in bioscience and biotechnology* 04(10). Scientific Research Publishing, Inc,: 75–88.

Terada M, Seki M, Takahashi R, et al. (2016) Effects of a Closed Space Environment on Gene Expression in Hair Follicles of Astronauts in the International Space Station. *PloS one* 11(3): e0150801.

Thornton MJ (2013) Estrogens and aging skin. *Dermato-endocrinology* 5(2): 264–270.

Thorogood A, Dalpé G and Knoppers BM (2019) Return of individual genomic research results: are laws and policies keeping step? *European journal of human genetics: EJHG* 27(4): 535–546.

Thorogood A, Rehm HL, Goodhand P, et al. (2021) International federation of genomic medicine databases using GA4GH standards. *Cell genomics* 1(2).

Tian J, Pecaut MJ, Slater JM, et al. (2010) Spaceflight modulates expression of extracellular matrix, adhesion, and profibrotic molecules in mouse lung. *Journal of applied physiology* 108(1): 162–171.

Tiffon C (2018) The Impact of Nutrition and Environmental Epigenetics on Human Health and Disease. *International journal of molecular sciences* 19(11).

Tikkanen E, Gustafsson S, Amar D, et al. (2018) Biological Insights Into Muscular Strength: Genetic Findings in the UK Biobank. *Scientific Reports* 8(6451).

Tozzi L, Staveland B, Holt-Gosselin B, et al. (2020) The human connectome project for disordered emotional states: Protocol and rationale for a research domain criteria study of brain connectivity in young adult anxiety and depression. *NeuroImage* 214: 116715.

Trappe S, Costill D, Gallagher P, et al. (2009) Exercise in space: human skeletal muscle after 6 months aboard the International Space Station. *Journal of applied physiology* 106(4): 1159–1168.

Trudel G, Shahin N, Ramsay T, et al. (2022) Hemolysis contributes to anemia during long-duration space flight. *Nature medicine* 28(1): 59–62.

Truong N, Sun K, Wang S, et al. (2021) Privacy preservation in federated learning: An insightful survey from the GDPR perspective. *Computers & Security* 110: 102402.

Turnbull C, Scott RH, Thomas E, et al. (2018) The 100 000 Genomes Project: bringing whole genome sequencing to the NHS. *BMJ* 361: k1687.

Turroni S, Magnani M, Kc P, et al. (2020) Gut Microbiome and Space Travelers' Health: State of the Art and Possible Pro/Prebiotic Strategies for Long-Term Space Missions. *Frontiers in physiology* 11: 553929.

Urbaniak C, Lorenzi H, Thissen J, et al. (2020) The influence of spaceflight on the astronaut salivary microbiome and the search for a microbiome biomarker for viral reactivation. *Microbiome* 8(1): 56.

Urquieta E, Wu J, Hury J, et al. (2022) Establishment of an open biomedical database for commercial spaceflight. *Nature medicine* 28(4): 611–612.

Ursell LK, Metcalf JL, Parfrey LW, et al. (2012) Defining the human microbiome. *Nutrition reviews* 70 Suppl 1: S38–44.

Uruno A, Saigusa D, Suzuki T, et al. (2021) Nrf2 plays a critical role in the metabolic response during and after spaceflight. *Communications biology* 4(1): 1381.

Valdes AM, Walter J, Segal E, et al. (2018) Role of the gut microbiota in nutrition and health. *BMJ* 361: k2179.

van Campen JC, Sollars ESA, Thomas RC, et al. (2019) Next generation sequencing in newborn screening in the United Kingdom National Health Service. *Screening: journal of the International Society of Neonatal Screening* 5(4). MDPI AG: 40.

Van Ombergen A, Rossiter A and Ngo-Anh TJ (2021) 'White Mars' - nearly two decades of biomedical research at the Antarctic Concordia station. *Experimental physiology* 106(1): 6–17.

Varga EA and Moll S (2004) Cardiology patient pages. Prothrombin 20210 mutation (factor II mutation). *Circulation* 110(3): e15–8.

Vermeulen C, Pagès-Gallego M, Kester L, et al. (2023) Ultra-fast deep-learned CNS tumour classification during surgery. *Nature* 622(7984): 842–849.

Vilà-Brau A, De Sousa-Coelho AL, Mayordomo C, et al. (2011) Human HMGCS2 regulates mitochondrial fatty acid oxidation and FGF21 expression in HepG2 cell line. *The Journal of biological chemistry* 286(23): 20423–20430.

Vintila AR, Slade L, Cooke M, et al. (2023) Mitochondrial sulfide promotes life span and health span through distinct mechanisms in developing versus adult treated Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America* 120(32): e2216141120.

Vitry G, Finch R, Mcstay G, et al. (2022) Muscle atrophy phenotype gene expression during spaceflight is linked to a metabolic crosstalk in both the liver and the muscle in mice. *iScience* 25(10): 105213.

Voorhees JJ (1990) Clinical effects of long-term therapy with topical tretinoin and cellular mode of action. *The Journal of international medical research* 18 Suppl 3: 26C–28C.

Voorhies AA, Mark Ott C, Mehta S, et al. (2019) Study of the impact of long-duration space missions at the International Space Station on the astronaut microbiome. *Scientific reports* 9(1): 9911.

Waisberg E, Ong J, Masalkhi M, et al. (2024) Mitochondrial dysfunction in Spaceflight Associated Neuro-Ocular Syndrome (SANS): a molecular hypothesis in pathogenesis. *Eye* . Epub ahead of print 7 February 2024. DOI: 10.1038/s41433-024-02951-3.

Wallmeyer L, Dietert K, Sochorová M, et al. (2017) TSLP is a direct trigger for T cell migration in filaggrin-deficient skin equivalents. *Scientific reports* 7(1): 774.

Walsh L, Hafner L, Straube U, et al. (2021) A bespoke health risk assessment methodology for the radiation protection of astronauts. *Radiation and environmental biophysics* 60(2): 213–231.

Ward NE, Pellis NR, Risin SA, et al. (2006) Gene expression alterations in activated human T-cells induced by modeled microgravity. *Journal of cellular biochemistry* 99(4): 1187–1202.

Wilkinson MD, Dumontier M, Aalbersberg IJJ, et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data* 3: 160018.

Willis CRG, Szewczyk NJ, Costes SV, et al. (2020) Comparative Transcriptomics Identifies Neuronal and Metabolic Adaptations to Hypergravity and Microgravity in Caenorhabditis elegans. *iScience* 23(12): 101734.

Willis CRG, Gallagher IJ, Wilkinson DJ, et al. (2021) Transcriptomic links to muscle mass loss and declines in cumulative muscle protein synthesis during short-term disuse in healthy younger humans. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology* 35(9): e21830.

Wilson JW, Ott CM, Höner zu Bentrup K, et al. (2007) Space flight alters bacterial gene

expression and virulence and reveals a role for global regulator Hfq. *Proceedings of the National Academy of Sciences of the United States of America* 104(41): 16299–16304.

Winge MCG, Hoppe T, Berne B, et al. (2011) Filaggrin genotype determines functional and molecular alterations in skin of patients with atopic dermatitis and ichthyosis vulgaris. *PloS one* 6(12): e28254.

Wnorowski A, Sharma A, Chen H, et al. (2019) Effects of Spaceflight on Human Induced Pluripotent Stem Cell-Derived Cardiomyocyte Structure and Function. *Stem cell reports* 13(6): 960–969.

World Medical Association (2013) World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA: the journal of the American Medical Association* 310(20): 2191–2194.

Wu T, Hu E, Xu S, et al. (2021) clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Cambridge (Mass.))* 2(3): 100141.

Xiao S, Venkateswaran KJ and Jiang SC (2019) The risk of Staphylococcus skin infection during space travel and mitigation strategies. *Microbial Risk Analysis* 11: 23–30.

Xu D, Gao Y, Guo L, et al. (2018) Effect of dys-1 mutation on gene expression profile in space-flown Caenorhabditis elegans. *Muscle & nerve*. Epub ahead of print 18 January 2018. DOI: 10.1002/mus.26076.

Xu W, Yang H, Liu Y, et al. (2011) Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. *Cancer cell* 19(1): 17–30.

Yamamoto S, Kanca O, Wangler MF, et al. (2023) Integrating non-mammalian model organisms in the diagnosis of rare genetic diseases in humans. *Nature reviews. Genetics*. Epub ahead of print 25 July 2023. DOI: 10.1038/s41576-023-00633-6.

Yan Y, Tian J, Wang Y, et al. (2023) Transcriptomic Heterogeneity of Skin Across Different Anatomic Sites. *The Journal of investigative dermatology* 143(3): 398–407.e5.

Ye YM, Kim BE, Shin YS, et al. (2013) Overexpression of Epidermal Filaggrin in Patients with Chronic Idiopathic Urticaria Correlates with Urticaria Severity. *The Journal of allergy and clinical immunology* 131(2). Elsevier: AB56.

Ye Y-M, Kim BE, Shin Y-S, et al. (2014) Increased epidermal filaggrin in chronic idiopathic urticaria is associated with severity of urticaria. *Annals of allergy, asthma & immunology: official publication of the American College of Allergy, Asthma, & Immunology* 112(6): 533–538.

Yoham AL and Casadesus D (2023) Tretinoin. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing. Available at: https://www.ncbi.nlm.nih.gov/pubmed/32491410.

Yu CT, Chao BN, Barajas R, et al. (2022) An evaluation of the National Institutes of Health grants portfolio: identifying opportunities and challenges for multi-omics research that leverage metabolomics data. *Metabolomics: Official journal of the Metabolomic Society* 18(5): 29.

Zahedani AD, McLaughlin T, Veluvali A, et al. (2023) Digital health application integrating wearable data and behavioral patterns improves metabolic health. *NPJ digital medicine* 6(1): 216.

Zeitlin C, Hassler DM, Cucinotta FA, et al. (2013) Measurements of energetic particle radiation in transit to Mars on the Mars Science Laboratory. *Science* 340(6136).

Science.

Zhang S, Wimmer-Schweingruber RF, Yu J, et al. (2020) First measurements of the radiation dose on the lunar surface. *Science advances* 6(39). Sci Adv.

Zhang X, Johnson AD, Hendricks AE, et al. (2014) Genetic associations with expression for genes implicated in GWAS studies for atherosclerotic cardiovascular disease and blood phenotypes. *Human molecular genetics* 23(3): 782–795.

Zheng M, Charvat J, Zwart SR, et al. (2023) Time-resolved molecular measurements reveal changes in astronauts during spaceflight. *Frontiers in physiology* 14: 1219221.

Zierer J, Jackson MA, Kastenmüller G, et al. (2018) The fecal metabolome as a functional readout of the gut microbiome. *Nature genetics* 50(6): 790–795.

Zschokke S, Countryman S and Cushing PE (2020) Spiders in space-orb-web-related behaviour in zero gravity. *Die Naturwissenschaften* 108(1): 1.

Zwart SR and Smith SM (2022) Chapter 5 - Genetics, vitamins, and Spaceflight Associated Neuro-Ocular Syndrome. In: Lee AG and Ong J (eds) *Spaceflight Associated Neuro-Ocular Syndrome*. Academic Press, pp. 55–65. Available at: https://www.sciencedirect.com/science/article/pii/B978032391524300017X.

Zwart SR, Gregory JF, Zeisel SH, et al. (2016) Genotype, B-vitamin status, and androgens affect spaceflight-induced ophthalmic changes. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology* 30(1). FASEB J.

Zwart SR, Auñón-Chancellor SM, Heer M, et al. (2022) Albumin, oral contraceptives, and venous thromboembolism risk in astronauts. *Journal of applied physiology* 132(5): 1232–1239.