Gene Expression Meta-Analysis Reveals mRNA Association with Spaceflight-induced Changes Across Muscle Types

Caris Kojima *caris.jordan.kojima@gmail.com*
Chapin School, NYC
Lumiere Education Inc.
New York, 10065, USA

Spencer Kojima *spencer.hayden.kojima@gmail.com*
Collegiate School, NYC
Lumiere Education Inc.
New York, 10065, USA

Laura K. Harris *oesterei@gmail.com*
Harris Interdisciplinary Education
Dewitt, 48820, USA
Lumiere Education Inc.
Wilmington, 48824, USA

Abstract

Background: Spaceflight-induced muscle changes result from exposure to microgravity conditions lacking oxygen and pressure. This is the initial study to implement a Gene Set Enrichment meta-analysis (GSEA) to highlight differentiations in gene expression, possibly detecting future targets for potential therapies.

Methods: This study defines 8 gene signatures between space and Earth-based samples of muscle atrophy in mice, comparing them through Gene Set Enrichment Analysis (GSEA). This method helped to define positive and negative panels by identifying the leading-edge genes in each query set. These panels were then utilized to reveal differentiations in gene expression of atrophy, and compare different gene signatures displaying alterations in muscle atrophy gene expression within different environments.

Results: Enrichment was shown among identification signatures where 149 genes from the positive panel and 69 genes from the negative panel were defined. When gene expression was compared across different muscle types, 15 significant up-regulated genes and 9 significant down-regulated genes were found across all muscle types. Ultimately, the data consisted of some genes which had similarities to previous literature and also showed new, novel associations.

Conclusion: The use of a gene expression meta-analysis based method highlighted alterations associated with spaceflight induced muscle atrophy. These results may have significant implications in the development of potential therapies.

Keywords: Muscle atrophy, Spaceflight, Microgravity, Gene Expression, Meta-analysis.

1. INTRODUCTION

Due to the evolution of humans originating on Earth, individuals have adapted to a specific environment. When astronauts are sent to travel into space, the novel environment can be severely impactful because of the lack of oxygen and pressure (Lee et al., 2022). There are many risks associated with spaceflight, including the exposure to microgravity, excess radiation, and confinement in a specific environment (Lee et al., 2022). These conditions have major

implications on the physiological systems of astronauts – particularly their muscle mass (Lee et al., 2022). These impacts not only reach humans, but also other organisms that are sent into space, including rodents such as mice. Muscle mass is particularly significant as the muscles of these organisms depend on their strength for both standing and moving; however, this weightless environment introduces the severe danger of muscle atrophy (Lee et al., 2022). Using models of microgravity conditions is crucial to more holistically understand the consequences of muscle atrophy (Bettis et al., 2018). Currently, no specific therapies exist to combat muscle atrophy, as the only currently identified treatment is the completion of daily exercise (He et al., 2020). Though regular exercise is a useful countermeasure in reducing bone and muscle loss, other solutions are still being examined for increased effectiveness (Bonanni et al., 2023).

It is imperative to comprehend gene expression and pathway activity alterations in accordance with spaceflight's effects on muscles because of presentation of knowledge on how other environmental factors impact the transcription and translation processes. This results in an alteration in the metabolic inclinations of muscles, and provides awareness for the extreme consequences of muscle atrophy (Qaisar et al., 2020). These effects contain further significance due to their relation to other models seen on Earth such as aging and immobility (Bonanni et al., 2023). These factors can be analyzed from their presentation in various studies. These impacts are examined further in the shoulder muscles of mice throughout three particular space missions: the short term missions (STS-131 and STS-135), and the long-term mission (Bion-M1) (Shen et al., 2017). This study fixated on the deltoid and rotator cuff muscles, concluding that the deltoid muscle had further sensitivity to microgravity than the rotator cuff (Shen et al., 2017). It further revealed a decrease in the adipogenic and myogenic gene expression for the short-term space flight missions. The results from this study are significant as they reveal comprehension of the molecular responses of shoulder muscles in microgravity conditions (Shen et al., 2017). These findings have significant implications into the broader conditions clinically for long-term spaceflight missions and examples of muscle disuse. The cumulative analysis of studies relating to the effects of microgravity contribute to the overall comprehension of how astronauts in space deal with muscle atrophy, and through individuals facing similar conditions on Earth (Bonanni et al., 2023).

Previous studies have used Fisher's exact test as a method of displaying the gene expression of mice in Space versus Earth, despite its inefficiency and ineffectiveness. The method utilized by this paper, Gene Set Enrichment Analysis (GSEA) has increased efficiency due to its inclusion of pathway dynamics, use of novel and differing techniques, and the incorporation of entire distributions (Maleki et al., 2020). The incorporation of entire gene sets and individual groups of pathways allows for GSEA to analyze alterations between various related genes (Subramanian et al., 2005). In contrast, Fisher's exact test typically only analyzes specific gene-level associations (Kim et al., 2017). Furthermore, GSEA utilizes a rank-based method, incorporates the gene set size in its analysis, and includes biological context with predefined gene sets, which aid in interpreting the data and generating hypotheses (Maleki et al., 2020). GSEA can analyze all genes in a dataset by examining gene distribution throughout differentially expressed groups (Maleki et al., 2020). These benefits reveal the importance of using GSEA for gene set analysis, and differentiate it further from Fisher's exact test's limitations to individual gene-level analysis.

The cumulative exploration of various muscle types aids in shaping a more holistic, comprehensive picture of muscle atrophy associated with spaceflight. Specifically, the dataset GSE94381 examines the longissimus dorsi (LD), a muscle holding a central position in the erector spinae group, and a fast twitch speed (Dietrich et al., 2021; Wu et al., 2023). This dataset also explores the tongue muscle (TN), which is attached to the hyoid bone, mandible, styloid process, soft palate, and pharynx, and has a similarly fast twitch speed (Dotiwala et al., 2022; Rasmussen et al., 2021). Additionally, the dataset GSE80223 examines the soleus (SOL) muscle, which is located in the posterior part of the ankle, whose muscle fibers have a slow twitch speed (Benito-de-Pedro et. al, 2021; Gong et al., 2022). This dataset also analyzes the extensor digitorum longus (EDL) muscle, located in the anterior compartment of the lower limb, and whose muscle fibers have a fast twitch speed (Lezak et al., 2022; Ağaşcıoğlu et al., 2022). Lastly, the

dataset GSE10533 highlights the gastrocnemius (G) muscle, which is located in the posterior part of the knee, and has a fast muscle twitch speed (Benito-de-Pedro et. al, 2021; Song et. al, 2018). The thorough analysis of these five prominent muscle types and their respective impacts help to form a more comprehensive understanding of the broader study of muscle atrophy in an environment of microgravity.

Slow and fast twitch fibers have many significant commonalities and differentiations. Slow twitch muscle fibers are more fatigue-resistant, though their muscle contraction response time is significantly slower than fast-twitch muscle fibers (Plotkin et al., 2021). In comparison, fast-twitch muscle fibers contain a high level of myosin ATPase activity, and faster contraction response (Bao et al., 2020). These prominent differences in the function of the different twitch speeds contribute to the contrast in gene expression and pathway activity. These differentiations were highlighted in two specific/previous studies comparing fast and slow twitch muscle fibers between two different animals. Through examining the transcriptome of the splenius and gluteus medius muscles of the Mongolian horse, researchers found contrasting expression levels for the fast and slow twitch fibers by analyzing the *Actn3, Sln,* and *Myoz2* genes (Bao et al., 2020). The *Actn3* and *Sln* gene displayed significant upregulation for fast-twitch fibers, while the *Myoz2* gene was upregulated in slow twitch fibers (Bao et al., 2020). Additionally, an alternate example examining the twitch speed of the Pseudocaranx dentex fish observed significant distinctions between the two (Wang et al., 2022). This study highlighted the increased mRNA expression levels in the *ACC, ACBP, Fabp3, LPL, PPARα, Acadm, Acad1, ACS, PDE3A* and *PKA PPARa* genes for the slow-twitch muscle fibers, in contrast to fast-twitch fibers (Wang et al., 2022). Conversely, the study emphasized the upregulation of the *AKT, Fbp1b, PGK, GPIA,* and *GYS* genes in fast-twitch muscle fibers (Wang et al., 2022). These differentiations supply molecular support for the distinctions between fast and slow twitch muscle types. The observations in these current literature examples suggest difficulty in comparing fast and slow twitch muscle types due to their stark contrast; however, this study's implementation of GSEA analysis can provide a bridge to this challenge.

Due to this knowledge gap, there are still many details unknown regarding this topic, and many more genes to be identified. Stronger conclusive evidence is required to support that the previously identified genes are truly responsible for spaceflight-induced muscle atrophy. This paper helps to fill in those gaps on the broader spectrum of scientific literature, as it is the primary study to use microarray technology for analysis. The utilization of a GSEA-based method of analysis in preceding studies enabled the identification of both previously discovered and new found genes associated with severe acute respiratory syndrome infection (SARS) by the comparison of differential gene expression with mRNA expression datasets (Park et al., 2021). Furthermore, GSEA was employed to detect gene expression changes associated with resistance to Diffuse Intrinsic Pontine Glioma (DIPG) treatments and in Human papillomavirus (HPV)-induced cervical cancer. (Harris, 2021; Ojo & Harris, 2023) In this study, GSEA is newly applied to specifically analyze effects of spaceflight-induced muscle atrophy using analysis from examining mRNA expression, comparing and defining various gene expression signatures. This novel implementation is useful in identifying new genes associated with spaceflight-induced muscle atrophy, and with it, possibly providing future insight into specific targets to ameliorate its impacts.

2. METHODS

2.1 Gene Expression Datasets Used

To identify gene expression changes associated with spaceflight in mice, the same research design, data collection, and data analysis was adopted from previously reported studies to inductively determine if a gene signature meta-analysis approach would include and expand upon previously reported findings that analyzed these datasets using other computational approaches (cite SARS Frontiers, DIPG from IJBB). The Gene Expression Omnibus (GEO) repository was searched (Edgar et al., 2002; Barrett et al., 2011; Barrett et al., 2013). Three published independent datasets examining gene expression in ground- and flown- mice (*i.e.*,

mousetronauts) that were profiled using similar probes - GSE80223, GSE94381, and GSE10533 - were selected for use in this study (Table 1). For GSE80223 and GSE94381, 19–20 weeks old C57BL/N6 mice were 1) flown aboard the Bion M1 biosatellite in a microgravity environment for 30 days, 2) housed in the same habitat of flown animals but exposed to Earth gravity (*i.e.*, ground), or 3) housed in a standard animal facility (*i.e.*, control). GSE80223 sampled soleus (SOL) and extensor digitorum longus (EDL) muscles while GSE94381 tested longissimus dorsi (LD) and tongue (TN). Both GSE80223 and GSE94381 were profiled on the [Mouse430A_2] Affymetrix Mouse Genome 430A 2.0 Array (GEO GPL8321). Finally, GSE10533 collected gastrocnemius samples from C57BL/N6 female mice that were 1) flown on the middeck of the space shuttle Endeavour (STS-108/UF-1) for 11 days and 19 h, or 2) housed at Cape Canaveral Air Force Station's Hangar L orbital environmental simulator for 11 days and 19 h with conditions mimicking the shuttle's middeck temperature, humidity, and CO2 levels (*i.e.*, ground). All GSE10533 animals were 77 days of age (11 weeks old) at the time of death. GSE10533 samples were profiled on [MOE430A] Affymetrix Mouse Expression 430A Array (GPL339). Probes were converted to gene symbols for each gene using the GEO provided platform data table when necessary.

Spaceflight Datasets Utilized for this Study			
Dataset	Description	Platform	Probes
GSE94381	Study analyzing gene expression in tongue and longissimus dorsi muscles in mice over 30 days in space, in a lab, or Earth.	GPL8321	22626
GSE80223	Study analyzing gene expression in soleus and extensor digitorum longus muscles in mice over 30 days in space, a lab, or Earth.	GPL8321	22690
GSE10533	Study analyzing gene expression in the gastrocnemius muscle of mice over an 11 day period in space, a lab, or Earth.	GPL339	22690

TABLE 1: Spaceflight Datasets Utilized for this Study.

2.2 Gene Signature Formation

Expression data provided by GEO for all datasets were unlogged intensities, so z-score normalization was employed across all samples within the dataset prior to use. Differential gene expression was measured by Welch's two-sample T-test score of normalized values and used to define 8 gene signatures (*i.e.*, gene lists ranked by differential gene expression between mousetronaut and Earth-based samples, Table 2), which included 4 fast-twitch muscles (EDL, TN, LD, and G) and 1 slow-twitch muscle (SOL).

Gene Signatures Defined in this Study							
Dataset	Group 1 (N)	Group 2 (N)	Signature	Use	High	Low	Cross
GSE94381	LD Space (5)	LD Earth (5)	LDSpacevsEarth	I	18.17	-12.97	11315

GSE94381	LD Space (5)	LD Lab (5)	LDSpacevsLab	I	21.65	-13.63	11315
GSE94381	TN Space (3)	TN Earth (3)	TNSpacevsEarth	C	44.38	-19.23	11315
GSE10533	G Space (4)	G Earth (4)	GSpacevsEarth	C	16.95	-23.42	11347
GSE80223	EDL Space (3)	EDL Earth (3)	EDLSpacevsEarth	C	24.67	-17.16	11347
GSE80223	EDL Space (3)	EDL Lab (3)	EDLSpacevsLab	C	36.41	-34.55	11347
GSE80223	SOL Space (3)	SOL Earth (3)	SOLSpacevsEarth	C	40.86	-43.89	11347
GSE80223	SOL Space (3)	SOL Lab (3)	SOLSpacevsLab	C	28.19	-34.32	11347

C, comparison; EDL, extensor digitorum longus; G, gastrocnemius; I, identification; LD, longissimus dorsi; N, number of samples; SOL, soleus; TN, tongue.

TABLE 2: Gene Signatures Defined in this Study.

2.3 Overview of Gene Set Enrichment Analysis

In this study, Gene Set Enrichment Analysis (GSEA) was utilized as a method of calculating the gene enrichment from a query gene set to a reference gene signature. Using T-score in this GSEA method found that “hits” in the enrichment score caused an increase, in corresponding to its T-score, while a “miss” caused the enrichment score to become smaller. As a result, GSEA was able to calculate a maximum enrichment score relating to this reference signature and individual query set. Leading-edge genes primarily aided in the attainment of this maximum enrichment score. GSEA additionally determined the normalized enrichment score (NES) from the 1000 permutations of the reference signature in order to establish the significance of enrichment while comparing the reference signature with the query gene set.

2.4 Identification of Spaceflight Induced Gene Expression Changes in Longissimus Dorsi

This study utilized two different gene signature tails (as shown in Figure 1), and chose 500 genes from the negative and positive tails from the GSE94381-derived LDSpacevsEarth gene signature. This was further implemented in the creation of two different query gene sets, using GSEA to contrast each of the sets with the GSE943810-derived LDSpacevsEarth gene reference signature. The resulting leading-edge genes were utilized to define the two panels — one panel for each tail, looking at genes in the panels as related to the gene expression of the mice in Space vs Earth. Both of the gene panels used pathway enrichment analysis using Database for Annotation, Visualization, and Integrated discovery (DAVID)— a comprehensive database containing various methods enabling researchers to comprehend the biological significance of many sets of genes (Huang et al., 2009; Sherman et al., 2022).

DAVID is used to formulate critical knowledge regarding gene set enrichment by processing the inputted data and implementing different algorithms of bioinformatics, emphasizing significant enrichment for various biological pathways, processes, and molecular operations (Huang et al., 2009). The analysis conducted on this study, the terms of proteins for probes using differential expression found statistically were implemented into DAVID as gene symbols in a Homo sapiens gene list. For the parameters that remained, default settings were applied.

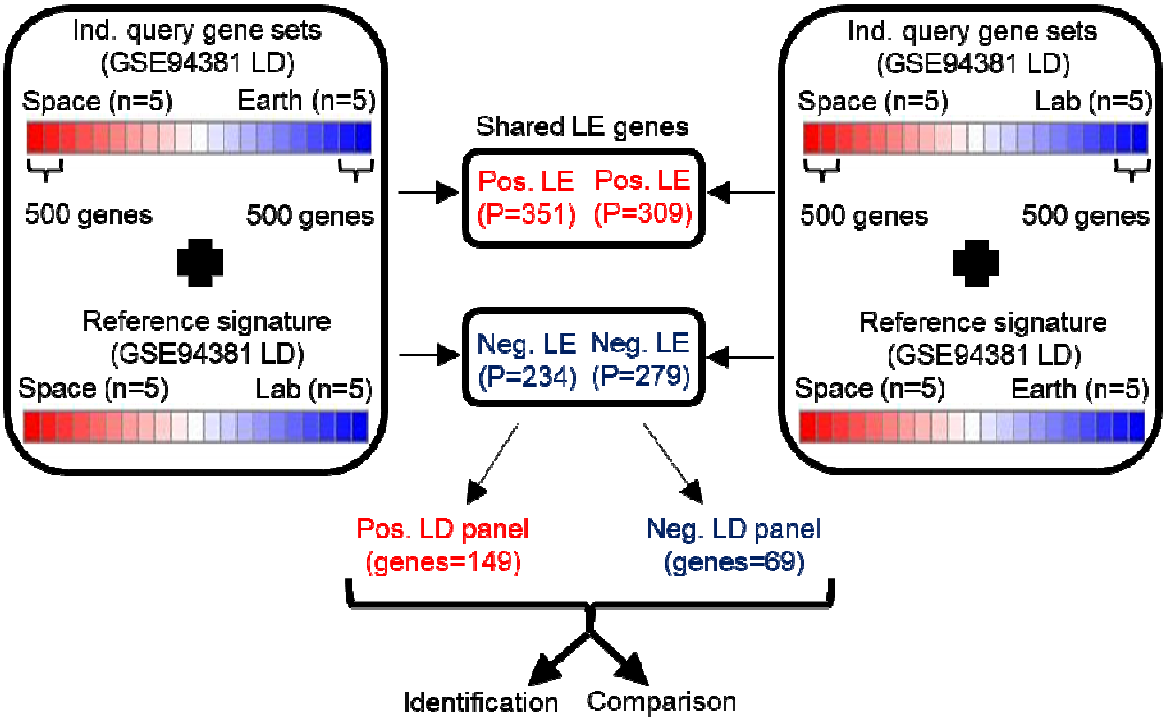

FIGURE 1: Gene Signature Definition and Generation Longissimus Dorsi (LD) Gene Panels. Illustrative representation of a gene signature. This figure depicts the differentiations in gene expression by comparing two groups: gene expression from Space and Earth as compared with gene expression from Space and Lab. The signatures are ranked lists of genes from the range of high (in red) and low (in blue) differential mRNA expression in comparing the separate groups. To identify differentially expressed genes associated with muscle atrophy in mice, query gene sets with the 500 most over- or under-expressed genes from positive or negative tails of the signature created from the Gene Expression Omnibus (GEO) accession number GSE94381 dataset. The positive and negative tail query sets were separately compared with the signature also from the GEO GSE94381 dataset, which was used as reference for Gene Set Enrichment Analysis (GSEA). GSEA identified leading-edge genes: genes that lead most to attaining the greatest possible enrichment. Two gene panels were defined from leading-edge genes identified in each query set. These gene panels were used in this study for two different purposes: 1) the identification of differences in gene expression associated with muscle atrophy in different environments, and 2) the comparison with different gene signatures showing changes in gene expression associated with other instances of muscle atrophy in different environments.

2.5 Comparison of Gene Expression Changes in LD to Changes in Other Muscles

To find and compare the gene expression of the various genes from the LD gene panels in mice, GSEA was used for the LD gene panels and GSE94381-derived SpacevsEarth reference signature. The leading edges determined in analyzing these signatures were used to contrast gene membership and find genes associated with atrophy and faster speed of muscle twitch fibers. Any gene panels having been excluded from this analysis were considered separately when examining shared leading-edge genes. The GSEA analysis method was reused to further the analysis by comparing gene expression in mice across various missions. The genes that correlated with muscle atrophy were determined by the gene membership in leading edges displayed throughout the analysis. The GSEA approach process was redone again for gene signatures coming from other datasets and comparing them to analyze their gene membership throughout leading-edges.

Throughout these comparisons, this study used random modeling to result in the production of heat maps using a Python script that is publicly available at https://github.com/oesterei/JNBs4astrobioIJBB. The methods utilized consisted of implementing the appropriate GEO datasets, where data cleaning was applied by removing data headers or

footers, separating data into subgroups, and removing rows with missing data. Next, data normalization was completed by performing the z-score of samples per probe, and signatures were generated by performing the t-score of samples per probe. Furthermore, panels were generated by selecting signature tails as query sets, where GSEA was run to find commonalities between leading-edges to define panels. Lastly, panels were compared by preparing necessary signatures, running the GSEA program, and finding commonalities between leading-edges for prominent genes (Park & Harris, 2021).

3. RESULTS

3.1 Gene Signature Based Approach Identified Gene Expression Changes Associated with Spaceflight in Mouse Longissimus Dorsi

To identify genes associated with muscle loss due to spaceflight, two LD gene signatures were defined from GSE94381 - LDSpacevsEarth and LDSpacevsLab (Table 2). The tails from each of these identification signatures was used to generate 4 gene sets containing the 500 most differentially expressed genes (T-score >3.91 and <-3.61 for positive and negative tails, respectively). 500 genes were selected - details in STable 1 - to maximize the signature's coverage that was allowable by GSEA (Subramanian et al., 2005). To assess similarity between these 2 gene signatures, GSEA was used to calculate enrichment between gene signatures individually and the query sets derived from them. Perfect alignment was observed and maximum enrichment achieved between the query set and signature it was derived from for both signatures (NES > 4.38 and <-5.03 for positive and negative query sets, respectively). Significant enrichment was seen between query sets and the opposite signature the sets were derived from (*e.g.*, LDSpacevsEarth query and LDSpacevsLab signature, Figure 2). Positive and negative gene panels were defined from the shared leading-edge genes identified by GSEA in this opposing signature analysis (Table 3, details in STable 2). These panels contained 149 up-regulated and 69 down-regulated genes associated with spaceflight induced muscle atrophy.

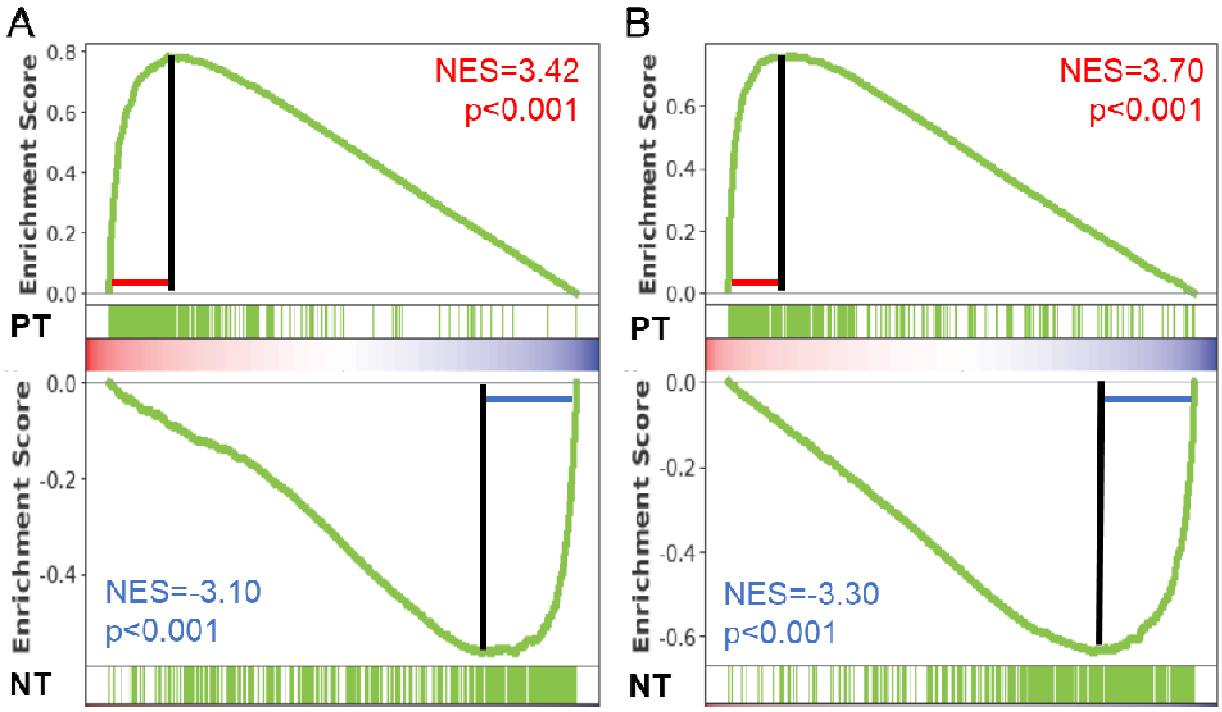

FIGURE 2: Enrichment Plots for Longissimus Dorsi (LD) Identification Signatures. A) Gene Set Enrichment Analysis (GSEA) calculated enrichment, as determined by normalized enrichment score (NES), between the positive tail (PT) and the GSE94381 LDspacevsEarth gene signature. B) GSEA between the negative tail (NT) and the GSE94381 LDspacevsEarth gene signature. C) GSEA between the PT and the GSE94381 LDspacevsLab. D) GSEA between the NT and the GSE94381 LDspacevsLab. These figures show significant enrichment scores which display strong similarities in the sequences which can be used for subsequent analysis.

Positive and Negative Longissimus Dorsi Gene Panels	
Panel	Gene symbols
Positive	44992, 2900097C17Rik, 6430548M08Rik*, Acot2, Ank3, Apbb3, Arl2bp, Atg13, Bcl10, Bcl2l11*, Cab39, Cbll1, Cdkal1, Cebpb, Chic2, Chp1, Cir1, Cnot7, Csnk2a1, Cyp2d22, Dapk3, Ddi2, Ddx10, Eif3a, Eif4ebp1, Eif4g1, Fam134b*, Fbxo32*, Fbxo9, Fbxw2, Fkbp5*, Fus*, Gabarapl1*, Gadd45a, Gm1821, Gm3579, Gna13, Gsk3b, Gsr, Gtpbp4, H1f0, Hectd1, Hmgb2, Ide, Il6ra, Jmy, Jph1, Kdm6a, Keap1, Kif1c, Kif2a, Klf9*, Klhl24, Lcn2, Luc7l2, Mef2a, Mettl7a1*, Mettl7a2, Mocs1, Nap1l1, Ncl, Nolc1, None, Npc1, Nrbf2, Nub1, Nufip1, Pcyt1a, Pdlim5, Pnmt, Pnpla2, Polr3c, Prkca, Prpf40a, Psmd11*, Psmd7, Pum1, Pus7, Rbbp4, Rbm25, Rnf138, Rora, Rpf1, Rrp1, Rrs1, Sec23ip, Sertad2, Sesn1, Sf3b1, Sf3b2, Sfxn2, Sin3a*, Slc39a14, Son, Srsf2, Synj2, Tacc2, Tapbpl, Tbc1d, Tbcel, Tcerg1, Tmem140, Tmem38a, Tns1, Top2b, Tspyl1, Txnip, Uba2, Ubc*, Ube2g2*, Ube4a*, Ubtf, Ubxn4*, Usp1, Usp3*, Wdr26, Wdr43, Yes1, Ythdf1, Ythdf2, Ywhag, Zfand5*, Zfp395, Zranb1, Zrsr1
Negative	1700021F05Rik, 1810037I17Rik, 3110009E18Rik, Ahrr, Ank1, Aurka, B9d2, BC031181, Bloc1s5, Bola2, Cab39l, Calm1, Cby1, Ccdc23, Ccdc43, Ccl27a, Ccrn4l, Chchd4, Cntnap4, Commd7, Ddit3, Dpm2, Dpy30, Dtymk, Dusp19, Elof1, Fam162a, Fnta, Gm17748, Gm19980, Gmppb, Gprc5c, Jkamp, Krt10, Krtcap2, Lsm3, Mafb, Mapre1, Mrps16, Naa35, Ndel1, Nme6, Nmi, Nprl2, P4ha1, Phxr4, Pigyl, Psmg1, Rala, Rdh13, Rnf14, Rps26, Runx2, Ruvbl2, Sec23b, Sec61b, Serpine2, Siva1, Slc25a19, Smarcal1, Smim11, Spcs1, Ssna1, St8sia5, Tpd52l2, Trappc2l, Uqcc2, Zfp426, Zmpste24

*More than one probe associated with the same gene symbol found in the panel.

TABLE 3: Positive and Negative Longissimus Dorsi Gene Panels.

3.2 Longissimus Dorsi Panel Genes Are Associated with RNA-related and Other Biological Processes

To expand on this analysis, DAVID was used to calculate enrichment between LD panels and biological processes/pathways; BioCarta, Gene Ontology Biological Processes (GO-BP), Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, and WikiPathways were used in this analysis (Huang et al., 2009). DAVID identified 1 significant (EASE score p-value<0.05) pathway for BioCarta, 31 for GO-BP, 9 for KEGG, 118 for Reactome, and 2 for WikiPathways from the positive LD panel (STable 3). DAVID also identified 4 significant pathways for GO-BP and 6 for Reactome from the negative LD panel (STable 3). To illustrate findings and focus this discussion without complicating result interpretation from overlapping pathways in different knowledge bases, only GO-BP results were illustrated in Figure 3.

Some significantly enriched pathways have experimentally established associations with muscle functions and atrophy, such as autophagy signaling pathway (GO:0006914, p−value=0.03297), and this shows the ability of this gene signature approach to detect pathways associated with muscle atrophy. (Xia et al., 2021) Other identified pathways were upregulated significantly but do not have nearly as strong associations to muscle atrophy compared to other pathways. For example, the up-regulation of in utero embryonic development (GO:0001701, p-value=0.01677) and the down-regulation of mitotic centrosome separation (GO:0007100, p-value=0.01522) did not have prior connections to spaceflight induced muscle atrophy. Thus, it can be speculated that down-regulated or up-regulated pathways without strong prior connections to muscle atrophy identified here also were involved in muscle atrophy, and that this paper reports a novel association of these pathway activity changes with atrophy.

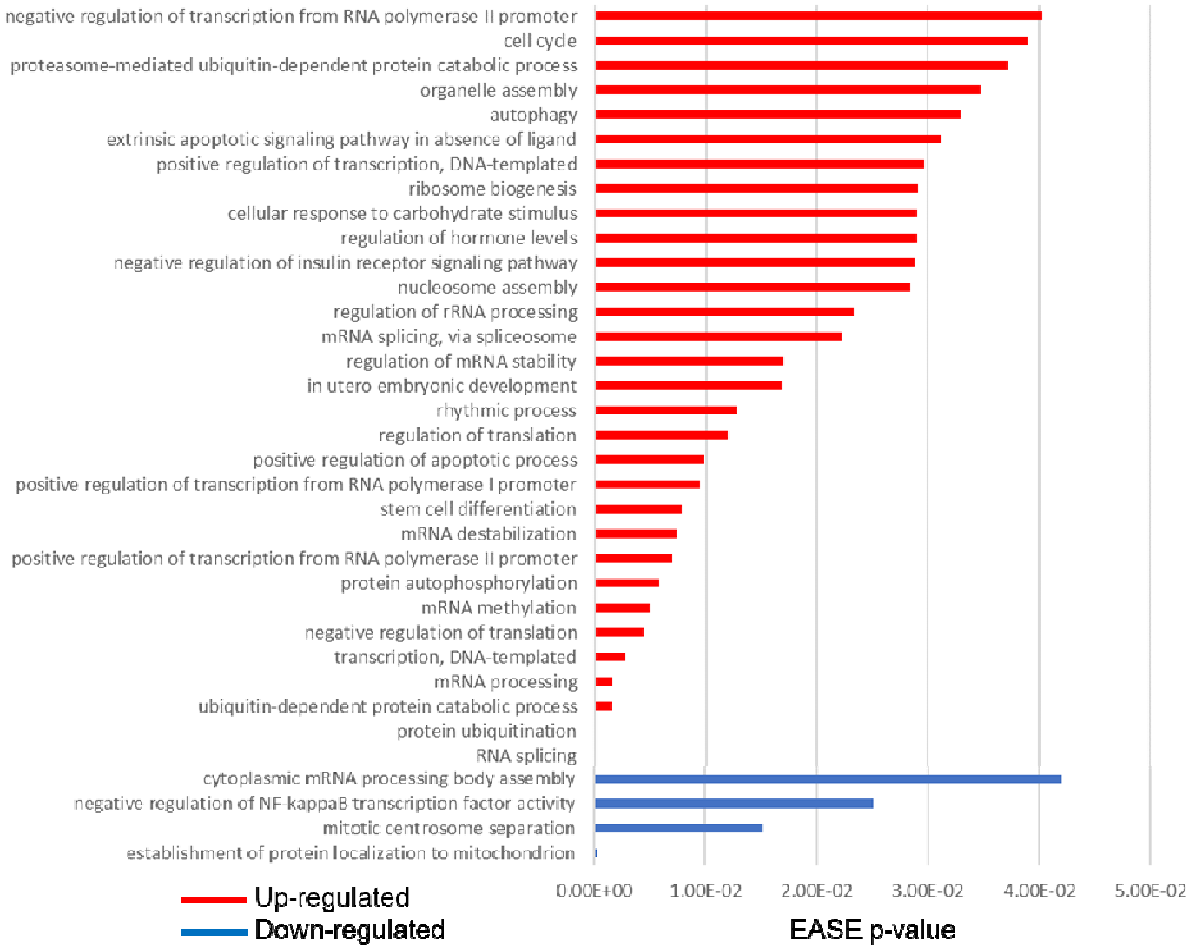

FIGURE 3: Bar Graph of EASE p-values for Significant Gene Ontology Biological Processes. Most of these pathways include negative regulation of RNA polymerase II promoter, upregulation of cell cycle, and down-regulation of cytoplasmic mRNA processing body assembly.

3.3 Significant, Non-random, Enrichment of Longissimus Dorsi Panels Observed in Different Muscle Types

With gene expression changes associated with spaceflight in LD established, the next aim was to determine if these changes were seen also in other muscles responding to spaceflight. To do this, GSEA was used to calculate enrichment between LD panels and 6 signatures derived from 4 different muscle types - soleus, tongue, gastrocnemius, and extensor digitorum longus (Table 2). A random model then was used to calculate significance of the calculated enrichment. As seen in Figure 4, the positive LD panel was significant for all signatures except GSE80223SOLspacevsEarth. The negative LD panel was significant for all signatures except GSE10533GspacevsEarth. Therefore, leading-edge genes identified by GSEA from these 2 analyses were not considered in subsequent analyses.

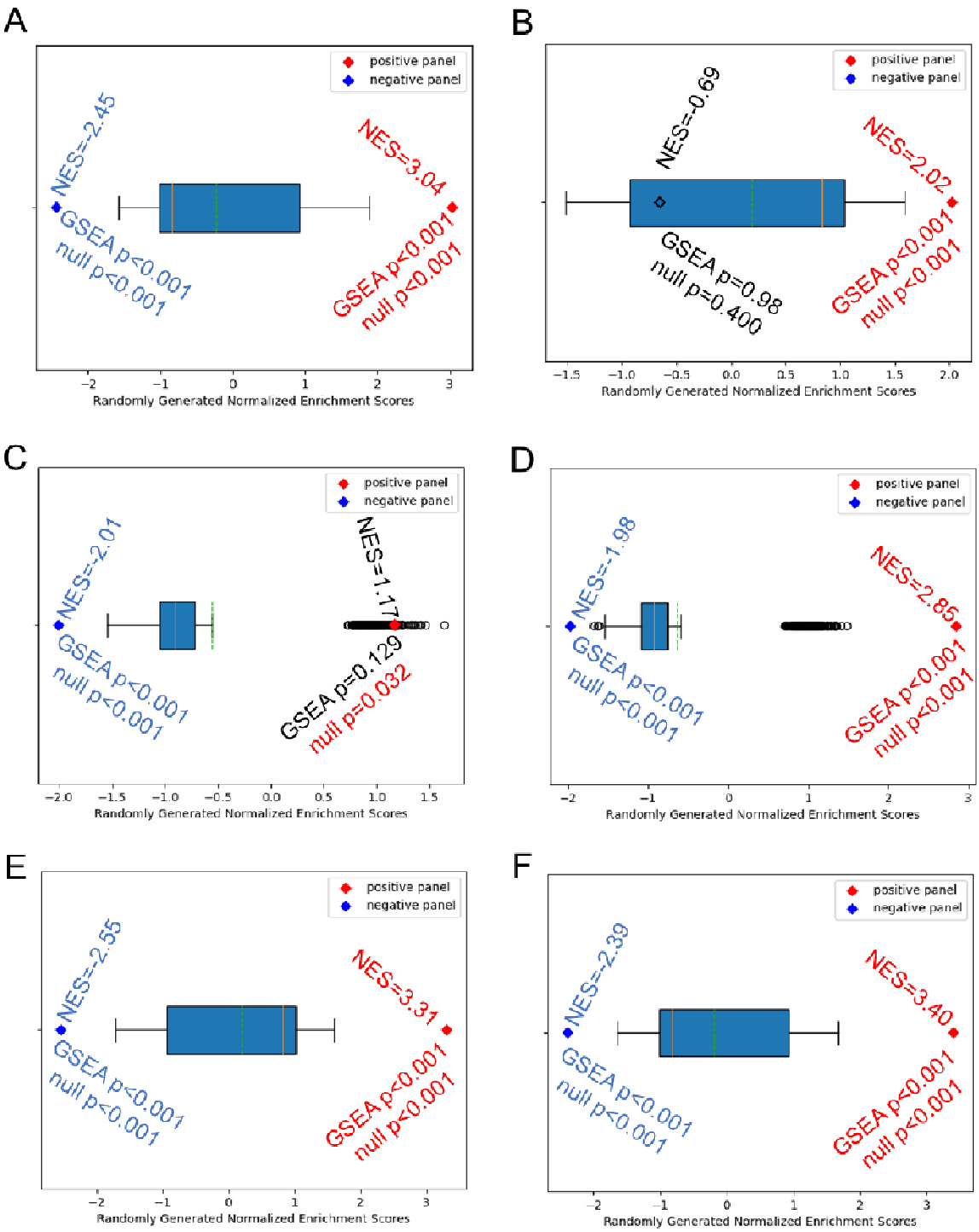

FIGURE 4: Randomly Generated Normalized Enrichment Scores per Comparison Signature. Random models were produced for A) GSE94381 TNspacevsEarth B) GSE80223SolspacevsEarth C) GSE80223EDLspacevsEarth D) GSE10533GspacevsEarth E) GSE80223SOLspacevsLab F) GSE80223EDLspacevsLab. They are shown here in box and whisker plots of NES from 1000 randomly generated gene panels containing 149 genes compared to gene signatures (individual references) and 69 genes compared to gene signatures (individual references).

3.4 Gene Expression Changes Associated with Spaceflight Found Across Different Muscle Types

When examining across leading-edge genes identified by GSEA from LD panels (STable 4 with details in STables 5 - 8), this meta-analysis approach found 15 of the 149 positive panel genes and 9 of the 69 negative panel genes were found in significantly enriched signatures despite muscle type (Figure 4). Of these, 12 positive panel and 3 negative panel genes were also found in the signatures that were insignificant for an LD panel. The identity of these leading-edge genes shared across all 5 muscles was considered (Tables 4 and 5). A heat map of differential gene expression (i.e., T-scores) for these 24 genes across all 8 gene signatures revealed none of these genes were statistically significant (Welch's two-sampled, two-sided T–test p–value<0.05) in all signatures (Figure 5), emphasizing the benefits of utilizing a GSEA meta-analysis method as compared to an individual gene approach. This could have potentially hindered identification of the other prominent genes due to borderline significance within signatures. Further, the differential gene expression heat map shows expression consistency across gene signatures despite lack of significance for most genes (Figure 5), further supporting the conclusion that these genes were associated with spaceflight regardless of muscle type.

Positive Panel Genes Shared Across Different Muscle Types		
ID	Symbol	Description
1416125_at	Fkbp5	FK506 binding protein 5
1417563_at	Eif4ebp1	eukaryotic translation initiation factor 4E binding protein 1
1421816_at	Gsr	glutathione reductase
1422560_at	Ddi2	DNA-damage inducible protein 2 /// regulatory solute carrier protein, family 1, member 1
1426446_at	6430548M08Rik	RIKEN cDNA 6430548M08 gene
1427747_a_at	Lcn2	lipocalin 2
1431697_at	Synj2	synaptojanin 2
1432827_x_at	Ubc	ubiquitin C
1437666_x_at	Ubc	ubiquitin C
1448231_at	Fkbp5	FK506 binding protein 5
1449354_at	Zrsr1	zinc finger (CCCH type), RNA binding motif and serine/arginine rich 1
1450505_a_at	Fam134b	family with sequence similarity 134, member B
1450606_at	Pnmt	phenylethanolamine-N-methyltransferase
1452082_at	6430548M08Rik	RIKEN cDNA 6430548M08 gene
1455892_x_at	None	None

TABLE 4: Positive Panel Genes Shared Across Different Muscle Types.

Negative Panel Genes Shared Across Different Muscle Types		
ID	Symbol	Description
1424316_at	Slc25a19	solute carrier family 25 (mitochondrial thiamine pyrophosphate carrier), member 19
1417465_at	Fnta	farnesyltransferase, CAAX box, alpha
1423820_at	Elof1	elongation factor 1 homolog (ELF1, S. cerevisiae)
1421019_at	1700021F05Rik	RIKEN cDNA 1700021F05 gene
1416666_at	Serpine2	serine (or cysteine) peptidase inhibitor, clade E, member 2
1417402_at	Smim11	small integral membrane protein 11
1426367_at	Cab39l	calcium binding protein 39-like
1448643_at	Ssna1	Sjogren's syndrome nuclear autoantigen 1
1451385_at	Fam162a	family with sequence similarity 162, member A

TABLE 5: Negative Panel Genes Shared Across Different Muscle Types.

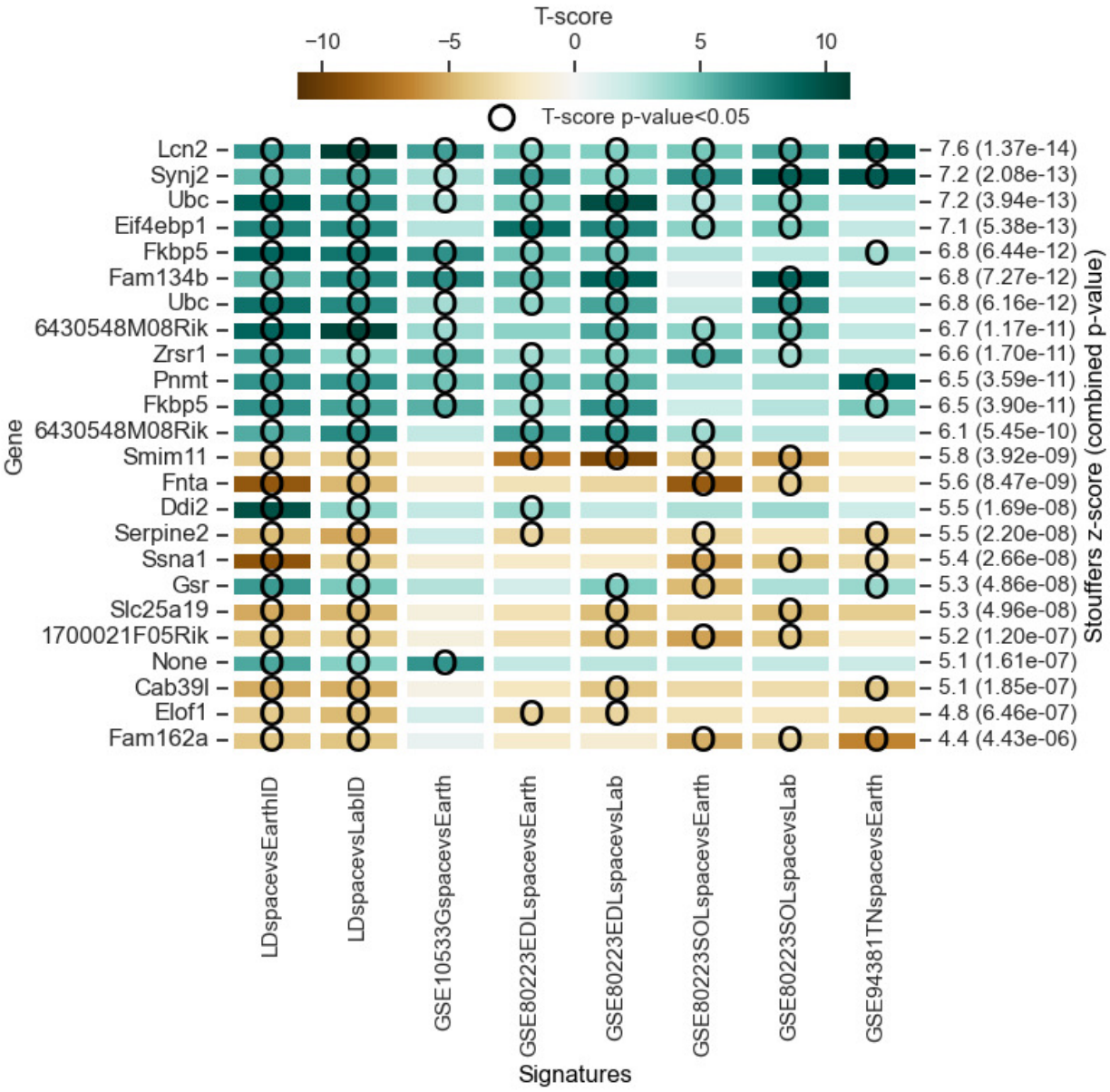

FIGURE 5: Heatmap of T-scores for the positive and negative LD panel leading-edge genes identified by GSEA. These are the top differential genes associated with spaceflight and are shared across all different signatures which were examined. This suggests that they may be potential therapeutic targets.

4. DISCUSSION

Spaceflight-induced muscle atrophy introduces many challenges for long-term space missions, enforcing the need for a detailed comprehension of the underlying molecular mechanisms. The gene signature approach was used to examine differential gene patterns in relation to muscle degeneration as a result of exposure to space. The results revealed distinct shifts in expression of multiple genes, pinpointed as the contributors to muscle atrophy in this environment. This analysis highlights the pathways and genes needed to be targeted to prevent the detrimental effects of microgravity conditions on muscle tissue.

There were 15 significant up-regulated genes and 9 significant down-regulated genes that this study found. There are few studies involving spaceflight-induced muscle atrophy; however, some

candidates had previously reported associations with muscle. LCN2, when floxed in mice, leads to increased intact muscle fibers and better grip strength. (Ponzetti et al., 2022) If the contrary is true, this implies that its overexpression could potentially block important muscle growth. Shown in this paper, its overexpression was related to muscle atrophy. CAB39L has an association with fast-fiber muscles. When downregulated—as it was in this study—it leads to altered contractile function in muscles. (Burton et al., 2023) In addition, Ubc is responsible for ubiquitin ligases, which are strongly associated with muscle atrophy. (Salvador et al., 2021) The examples of LCN2, UBC, and CAB39L reaffirm that the GSEA method is accurate in identifying genes related to muscle function or muscle atrophy. In addition, the GSEA method seems to have been effective in identifying genes with no prior connection to spaceflight-induced muscle atrophy. For example, Ssna1, 6430548M08Rik, Fnta, Elof1, and 1700021F05Rik did not have direct prior associations to spaceflight-induced muscular atrophy. While the GSEA analysis distributed some aforementioned genes responsible for other, less direct muscular processes—such as Ssna1, related to microtubule function which is involved in proper muscle function—there has not been conclusive prior evidence for these genes' direct relation to muscle atrophy itself. (Lévêque et al., 2016; Lucas et al., 2023) Interestingly, these results also conflict with some existing literature. FKBP5 is the most researched gene of the ones found connected to muscular changes due to gravity. However, a study showed that its expression might be responsible for increasing muscle mass, which would conflict with the data that suggested that its upregulation was associated with muscle atrophy. (Shimoide et al., 2016) Ultimately, the data of this paper reaffirmed previous perspectives, suggested new genes that could be associated with atrophy, and challenged some supposed functions.

It is important to note, however, that there were some detection limitations that could have biological implications. For instance, some commonly reported genes which interfere with protein synthesis—like atrogin-1, which codes for a ubiquitin ligase—were not found in this study. (Salvador et al., 2021) Some prominent genes with previous relations were not included in the original GSEA analysis. For other prominent genes a significant enrichment was not found—ie. their t-scores did not make the 500 gene cut-off—despite previously reported connections. This could be due to platform variations, which can greatly impact GSEA results. These are the drawbacks of GSEA. Significant t-scores are necessary for them to be involved in the enrichment process, and if not the genes can be neglected, even though they might still carry biological implications. However, GSEA in this study was perfectly sufficient, as the goal was to search for over- and under-expressed genes.

This study was the first to implement a GSEA-based method to detect gene expression changes in spaceflight-induced muscle atrophy. This study detected some genes that were not reported in previous literature. The up-regulated and down-regulated genes that did not have previous connections to spaceflight-induced muscle atrophy could potentially serve as new therapeutic targets. Thus, this study is important as it is the first to highlight these genes.

In order to verify these results further, laboratory testing would be needed. This could help facilitate experiments using qRT-PCR to confirm the overexpression and underexpression of the genes found in this paper. Living cells would be required in order to provide conclusive evidence for the importance of these aforementioned genes.

5. CONCLUSION

This is the primary study to analyze the expression of mRNA datasets to determine genes associated with spaceflight-induced muscle atrophy by implementing the GSEA method of analysis. The specific use of a GSEA-based method of analysis was successfully implemented in prior studies, such as the identification of genes associated with SARS. Through use of this method, this study outlined 15 significant overexpressed and 9 significant under-expressed genes that were most highly associated with spaceflight-induced muscle atrophy. These revelations are especially significant regarding the knowledge gap in scientific research highlighting the impacts of spaceflight-induced muscle atrophy of astronauts in the long-term. By obtaining knowledge on

the significant enrichment of these genes, researchers may be able to target specific genes in future treatments. Ultimately, this study displayed the effectiveness of the GSEA meta-analysis in identifying genes, and the application of novel information regarding mRNA expression data to ameliorate current treatment methods and introduce new targeted therapies.

6. SUPPLEMENTARY MATERIAL

https://docs.google.com/file/d/1Ox4ltDcQnVqbc0AOAAf-dlYzE6f65_Lp/edit?usp=docslist_api&filetype=msexcel

7. REFERENCES

Ağaşcıoğlu, E., Çolak, R., & Çakatay, U. (2022). Redox status biomarkers in the fast-twitch extensor digitorum longus resulting from the hypoxic exercise. *Nagoya journal of medical science*, *84*(2), 433–447. https://doi.org/10.18999/nagjms.84.2.433.

Bao, T., Han, H., Li, B., Zhao, Y., Bou, G., Zhang, X., Du, M., Zhao, R., Mongke, T., Laxima, Ding, W., Jia, Z., Dugarjaviin, M., & Bai, D. (2020). The distinct transcriptomes of fast-twitch and slow-twitch muscles in Mongolian horses. *Comparative biochemistry and physiology. Part D, Genomics & proteomics*, *33*, 100649. https://doi.org/10.1016/j.cbd.2019.100649.

Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, Kim IF, et al. NCBI GEO: Archive for Functional Genomics Data Sets–10 Years on. *Nucleic Acids Res* (2011) 39(Database issue):D1005–1010. doi: 10.1093/nar/gkq1184.

Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: Archive for Functional Genomics Data Sets–Update. *Nucleic Acids Res* (2013) 41(Database issue):D991–995. doi: 10.1093/nar/gks1193.

Benito-de-Pedro, M., Calvo-Lobo, C., López-López, D., Benito-de-Pedro, A. I., Romero-Morales, C., San-Antolín, M., Vicente-Campos, D., & Rodríguez-Sanz, D. (2021). Electromyographic Assessment of the Efficacy of Deep Dry Needling versus the Ischemic Compression Technique in Gastrocnemius of Medium-Distance Triathletes. *Sensors (Basel, Switzerland)*, *21*(9), 2906. https://doi.org/10.3390/s21092906.

Bettis, T., Kim, B. J., & Hamrick, M. W. (2018). Impact of muscle atrophy on bone metabolism and bone strength: implications for muscle-bone crosstalk with aging and disuse. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA*, *29*(8), 1713–1720. https://doi.org/10.1007/s00198-018-4570-1.

Bilodeau, P. A., Coyne, E. S., & Wing, S. S. (2016). The ubiquitin proteasome system in atrophying skeletal muscle: roles and regulation. *American journal of physiology. Cell physiology*, *311*(3), C392–C403. https://doi.org/10.1152/ajpcell.00125.2016.

Bonanni, R., Cariati, I., Marini, M., Tarantino, U., & Tancredi, V. (2023). Microgravity and Musculoskeletal Health: What Strategies Should Be Used for a Great Challenge?. *Life (Basel, Switzerland)*, *13*(7), 1423. https://doi.org/10.3390/life13071423.

Burton, M. A., Antoun, E., Garratt, E. S., Westbury, L., Baczynska, A., Dennison, E. M., Harvey, N. C., Cooper, C., Patel, H. P., Godfrey, K. M., Lillycrop, K. A., & EpiGen Global Research Consortium (2023). Adiposity is associated with widespread transcriptional changes and downregulation of longevity pathways in aged skeletal muscle. *Journal of cachexia, sarcopenia and muscle*, *14*(4), 1762–1774. https://doi.org/10.1002/jcsm.13255.

Dietrich, J., Handschuh, S., Steidl, R., Böhler, A., Forstenpointner, G., Egerbacher, M., Peham, C., & Schöpper, H. (2021). Muscle Fibre Architecture of Thoracic and Lumbar Longissimus Dorsi Muscle in the Horse. *Animals : an open access journal from MDPI*, *11*(3), 915. https://doi.org/10.3390/ani11030915.

Dotiwala AK, Samra NS. Anatomy, Head and Neck, Tongue. [Updated 2022 Aug 27]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK507782/.

Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI Gene Expression and Hybridization Array Data Repository. *Nucleic Acids Res* (2002) 30(1):207–10. doi: 10.1093/nar/30.1.207.

Gong, H. M., Ma, W., Regnier, M., & Irving, T. C. (2022). Thick filament activation is different in fast- and slow-twitch skeletal muscle. *The Journal of physiology*, *600*(24), 5247–5266. https://doi.org/10.1113/JP283574.

He, N., & Ye, H. (2020). Exercise and Muscle Atrophy. *Advances in experimental medicine and biology*, *1228*, 255–267. https://doi.org/10.1007/978-981-15-1792-1_17.

Huang, daW., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*, *4*(1), 44–57. https://doi.org/10.1038/nprot.2008.211.

Harris, L.K. (2021). Gene Expression Changes Associated with Developing Resistance to Diffuse Intrinsic Pontine Glioma Treatments. *International Journal of Biometrics and Bioinformatics*, 14(3), 40-57.

Kim H. Y. (2017). Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. *Restorative dentistry & endodontics*, *42*(2), 152–155. https://doi.org/10.5395/rde.2017.42.2.152.

Lee, P. H. U., Chung, M., Ren, Z., Mair, D. B., & Kim, D. H. (2022). Factors mediating spaceflight-induced skeletal muscle atrophy. *American journal of physiology. Cell physiology*, *322*(3), C567–C580. https://doi.org/10.1152/ajpcell.00203.2021.

Lévêque, M. F., Berry, L., & Besteiro, S. (2016). An evolutionarily conserved SSNA1/DIP13 homologue is a component of both basal and apical complexes of Toxoplasma gondii. *Scientific reports*, *6*, 27809. https://doi.org/10.1038/srep27809.

Lezak B, Summers S. Anatomy, Bony Pelvis and Lower Limb: Extensor Hallucis Longus Muscle. [Updated 2022 Sep 26]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK539875/.

Lucas, L., & Cooper, T. A. (2023). Insights into Cell-Specific Functions of Microtubules in Skeletal Muscle Development and Homeostasis. *International journal of molecular sciences*, *24*(3), 2903. https://doi.org/10.3390/ijms24032903.

Maleki, F., Ovens, K., Hogan, D. J., & Kusalik, A. J. (2020). Gene Set Analysis: Challenges, Opportunities, and Future Research. *Frontiers in genetics*, *11*, 654. https://doi.org/10.3389/fgene.2020.00654.

Ojo, D.O. & Harris, L. K. (2023). Gene Expression Insights into HPV-Induced Cervical Cancer: A GSEA Meta-analysis Approach. *International Journal of Biometrics and Bioinformatics*, 16(1).

Park, A., Nwobi, M., Harris L.K. (2021). Meta-Analysis Identifies Type I Interferon Response as Top Pathway Associated with SARS Infection. *International Journal of Biometrics and Bioinformatics*, 14(2), 14-39.

Plotkin, D. L., Roberts, M. D., Haun, C. T., & Schoenfeld, B. J. (2021). Muscle Fiber Type Transitions with Exercise Training: Shifting Perspectives. *Sports (Basel, Switzerland)*, *9*(9), 127. https://doi.org/10.3390/sports9090127.

Ponzetti, M., Ucci, A., Maurizi, A., Giacchi, L., Teti, A., & Rucci, N. (2022). Lipocalin 2 Influences Bone and Muscle Phenotype in the *MDX* Mouse Model of Duchenne Muscular Dystrophy. *International journal of molecular sciences*, *23*(2), 958. https://doi.org/10.3390/ijms23020958.

Qaisar, R., Karim, A., & Elmoselhi, A. B. (2020). Muscle unloading: A comparison between spaceflight and ground-based models. *Acta physiologica (Oxford, England)*, *228*(3), e13431. https://doi.org/10.1111/apha.13431.

Rasmussen, M., & Jin, J. P. (2021). Troponin Variants as Markers of Skeletal Muscle Health and Diseases. *Frontiers in physiology*, *12*, 747214. https://doi.org/10.3389/fphys.2021.747214.

Shen, H., Lim, C., Schwartz, A. G., Andreev-Andrievskiy, A., Deymier, A. C., & Thomopoulos, S. (2017). Effects of spaceflight on the muscles of the murine shoulder. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, *31*(12), 5466–5477. https://doi.org/10.1096/fj.201700320R.

Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., Imamichi, T., & Chang, W. (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic acids research*, *50*(W1), W216–W221. https://doi.org/10.1093/nar/gkac194.

Shimoide, T., Kawao, N., Tamura, Y., Morita, H., & Kaji, H. (2016). Novel roles of FKBP5 in muscle alteration induced by gravity change in mice. *Biochemical and biophysical research communications*, *479*(3), 602–606. https://doi.org/10.1016/j.bbrc.2016.09.126.

Song, H., Cho, S., Lee, H. Y., Lee, H., & Song, W. (2018). The Effects of Progressive Resistance Exercise on Recovery Rate of Bone and Muscle in a Rodent Model of Hindlimb Suspension. *Frontiers in physiology*, *9*, 1085. https://doi.org/10.3389/fphys.2018.01085.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(43), 15545–15550. https://doi.org/10.1073/pnas.0506580102.

Wang, H., Li, B., Yang, L., Jiang, C., Zhang, T., Liu, S., & Zhuang, Z. (2022). Expression profiles and transcript properties of fast-twitch and slow-twitch muscles in a deep-sea highly migratory fish, *Pseudocaranx dentex*. *PeerJ*, *10*, e12720. https://doi.org/10.7717/peerj.12720].

Wu, L. L., Zhang, X. Y., Li, X., Jin, J. J., Yang, G. S., & Shi, X. (2023). miR-196b-5p promotes myoblast proliferation and differentiation. *Yi chuan = Hereditas*, *45*(5), 435–446. https://doi.org/10.16288/j.yczz.23-025.

Xia, Q., Huang, X., Huang, J., Zheng, Y., March, M. E., Li, J., & Wei, Y. (2021). The Role of Autophagy in Skeletal Muscle Diseases. *Frontiers in physiology*, *12*, 638983. https://doi.org/10.3389/fphys.2021.638983.