

# SCIENTIFIC REPORTS



OPEN

## Global transcriptomic analysis suggests carbon dioxide as an environmental stressor in spaceflight: A systems biology GeneLab case study

Afshin Beheshti<sup>1</sup>, Egle Cekanaviciute<sup>2</sup>, David J. Smith<sup>3</sup> & Sylvain V. Costes<sup>3</sup>

Spaceflight introduces a combination of environmental stressors, including microgravity, ionizing radiation, changes in diet and altered atmospheric gas composition. In order to understand the impact of each environmental component on astronauts it is important to investigate potential influences in isolation. Rodent spaceflight experiments involve both standard vivarium cages and animal enclosure modules (AEMs), which are cages used to house rodents in spaceflight. Ground control AEMs are engineered to match the spaceflight environment. There are limited studies examining the biological response invariably due to the configuration of AEM and vivarium housing. To investigate the innate global transcriptomic patterns of rodents housed in spaceflight-matched AEM compared to standard vivarium cages we utilized publicly available data from the NASA GeneLab repository. Using a systems biology approach, we observed that AEM housing was associated with significant transcriptomic differences, including reduced metabolism, altered immune responses, and activation of possible tumorigenic pathways. Although we did not perform any functional studies, our findings revealed a mild hypoxic phenotype in AEM, possibly due to atmospheric carbon dioxide that was increased to match conditions in spaceflight. Our investigation illustrates the process of generating new hypotheses and informing future experimental research by repurposing multiple space-flown datasets.

Comprehensive analysis of molecular signatures, such as transcriptional profiling, has become a standard technique in space biosciences and typically generates more extensive data than is required for the specific topic of investigation. Making all spaceflight data publicly accessible ensures that biological experiments can be repurposed to answer novel research questions and generate hypotheses. Therefore, the GeneLab open science platform ([genelab.nasa.gov](http://genelab.nasa.gov)) was created to store raw molecular “omics” data from ground and spaceflight biology experiments supported by NASA. Here we present a case study using GeneLab datasets generated from ground controls associated with multiple rodent spaceflight datasets. Our overarching aim is to generate a hypothesis to drive future spaceflight rodent research and examine the potential impact of one known confounding factor in spaceflight, i.e. the environment in the animal habitat.

We proceeded by incorporating multiple, independent publicly available transcriptomic datasets from spaceflight experiments. Investigating spaceflight-induced changes in the transcriptome involves sending model organisms to orbit, such as rodents on the space shuttle (Space Transportation System program, STS), on satellites such as Bion-M1 (BF), or on the International Space Station (ISS). In these experiments multiple aspects of the environment are collectively altered. NASA space flown rodents are housed in a specific type of cage, called the Animal Enclosure Module (AEM). Within an AEM, animals will experience gravitational changes ranging from

<sup>1</sup>Wyle Labs, Space Biosciences Division, NASA Ames Research Center, Mountain View, CA, USA. <sup>2</sup>Universities Space Research Association, Space Biosciences Division, NASA Ames Research Center, Mountain View, CA, USA. <sup>3</sup>NASA, Space Biosciences Division, NASA Ames Research Center, Mountain View, CA, USA. Afshin Beheshti and Egle Cekanaviciute contributed equally to this work. Correspondence and requests for materials should be addressed to A.B. (email: [afshin.beheshti@nasa.gov](mailto:afshin.beheshti@nasa.gov)) or S.V.C. (email: [sylvain.v.costes@nasa.gov](mailto:sylvain.v.costes@nasa.gov))

## A) Cage Types



Animal Enclosure Module (AEM)

Sample vivarium cage

## B)

GeneLab Study	Mission	Species	CO <sub>2</sub> (ppm)	Duration (days)	Tissue Type
GLDS-21	STS-108	mouse	~3000	11.8	skeletal muscle (gastrocnemius)
GLDS-111	BF	mouse	~600	30	soleus muscle
GLDS-111	BF	mouse	~600	30	extensor digitorum
GLDS-25	STS-135	mouse	~3000	13	liver
GLDS-63	STS-70	rat	~3000 (est)	9	mammary gland

**Figure 1.** Illustration of the AEM and vivarium cages and GeneLab datasets used for analysis. **(A)** Images and dimensions of the both the AEM and vivarium cages used to house rodents. The upper two photos of the AEM cage was provided by NASA (Credits: NASA/Dominic Hart and <https://www.nasa.gov/ames/research/space-biosciences/rodent-research-1>). The vivarium cage photo was taken in our laboratory. **(B)** The GeneLab datasets which were used for analysis including the information on rodent species, the specific tissue type and its approximate location in a rodent, duration of experiment, and CO<sub>2</sub> levels in the AEM cage.

hypergravity during launch and landing to microgravity in orbit while simultaneously exposed to higher levels of ionizing radiation than found on Earth<sup>1</sup>.

Selecting appropriate controls for such multifactorial experiments is therefore complicated. The most frequently used experimental design is to keep all environmental conditions the same except for the flight by containing rodents in an AEM on spacecraft and using the same type of AEM hardware for ground controls<sup>2</sup>. An alternative and complementary approach is using regular vivarium cages for housing rodents as controls. AEM has been used as the standard rodent enclosure without major modifications in spaceflight experiments from STS<sup>2</sup> to Bion-M1 (BF)<sup>3</sup>, and its more modern version called the Habitat module of the Rodent Research Hardware System is currently used on the ISS. Both AEM and vivarium housing follow the standard guidelines for laboratory animal care, which require at least 15 square inches per >25 g adult rodent (NASA Johnson Space Center Animal Care and Use Handbook). AEM can either contain up to 10 mice in two compartments (5 mice per compartment) or 6 rats maximum (3 rats per compartment). This is in comparison to the vivarium cages which can house either 5 mice, or 2 rats maximum in a single compartment (depending on the rat's mass). The AEM has a combined larger surface area per rodent, because it includes climbable walls throughout the enclosure (Fig. 1A). Rodents housed in AEM and vivarium cages are typically kept on the same light/dark cycle and the air has the same oxygen concentration. However, CO<sub>2</sub> concentration in AEM ground controls replicates CO<sub>2</sub> concentration on spacecraft<sup>2,4</sup>, which tends to be up to an order of magnitude higher than on Earth. In general, CO<sub>2</sub> concentration on ground and therefore, in vivarium cages is approximately 300 ppm, while on the space shuttle and in matched ground AEM it reaches up to 3000 ppm<sup>5</sup>. Notably, it was lower in the AEM controls for the Bion-M1 satellite study: 682 ppm on average with 718 ppm standard deviation (range: 201–2096 ppm)<sup>4</sup>. Interestingly, previous studies on different CO<sub>2</sub> conditions on human health have revealed major impact on cognitive functions. Specifically, it was observed that a 400 ppm increase in CO<sub>2</sub> levels results in a 21% drop in cognitive scores<sup>6</sup> and specific research related to the increased CO<sub>2</sub> levels on the ISS have shown increased incidence of headaches with astronauts on the ISS<sup>7</sup>.

Here we analyze the impact of changing only the habitat, from the classic vivarium cage to the spaceflight-matched AEM, using ground control datasets located in the GeneLab depository (Fig. 1B). We combine multiple experiments and species (C57BL/6 mice (*Mus musculus*) and rats (*Rattus norvegicus*)) to increase the certainty that any common changes found across these datasets are consistent and reproducible. Additionally, by utilizing multiple tissues (muscle, breast and liver) in our analysis we are able to define a global systematic biological response to the different housing conditions. Finally, by using a previously-established systems biology approach to combine multiple analysis methods<sup>8</sup> we are able to uncover a statistical consensus on molecular pathways being specifically modified by the change in animal housing.

## Materials and Methods

**Data from GeneLab Platform.** Processing of animal tissues, RNA extraction, and microarray details were previously reported by NASA funded investigators associated with each dataset<sup>2,3,9,10</sup>. Our results are based on openly-available data housed in the NASA GeneLab platform ([genelab.nasa.gov](http://genelab.nasa.gov)). Although some of these datasets can also be found on NCBI's Gene Expression Omnibus (GEO), not all NASA funded transcriptomic data will be available on GEO. In general, the GeneLab platform will provide the most publicly available comprehensive NASA related “-omics” data. We analyzed the following datasets from GeneLab: GLDS-21, GLDS-111, GLDS-25, and GLDS-63 (Fig. 1B). These four datasets were chosen because they are the only datasets available for rodents on GeneLab that have both the AEM and vivarium controls. For each dataset we only performed analysis between the AEM and vivarium controls. The following replicates were available for each dataset for both conditions: for GLDS-21 there were 4 AEM controls and 5 vivarium controls, for GLDS-111 there were 3 AEM controls and 3 vivarium controls for both Soleus and Extensor digitorum muscles, for GLDS-25 there were 5 AEM controls and 6 vivarium controls, and for GLDS-63 there were 3 AEM controls and 4 vivarium controls. GLDS-25 and -21 datasets were obtained from 9 to 11-week old, female C57BL/6 mice. GLDS-63 dataset was derived from female pregnant Sprague-Dawley rats. GLDS-111 dataset was procured from 19–20 week-old male C57BL/N6 mice. It is assumed from the information available that all rodents were not in isolation during the duration of the experiments.

**GeneLab transcriptome analysis.** The microarray experiments from all datasets mentioned were previously performed on Affymetrix platforms. An in-depth experimental detail for each dataset is available on the GeneLab website. Raw data for all studies were first background adjusted and quantile normalized using RMAExpress<sup>11</sup>. Data were then imported into MultiExperiment Viewer<sup>12</sup> and statistically significant genes were determined by t-test with a P-value < 0.05 for all comparisons to take forward for pathway analysis. Independent and separate analysis was done for each dataset using these conditions due to platform incompatibility. Due to the small size of biological replicates available for the datasets on GeneLab we were only able to use P-value statistics. Although this might increase the chances of false positives it will also reduce the chances of false negatives. In addition, this analysis demonstrates that GeneLab datasets with less than optimal biological replicates will be able to generate a hypothesis and useful information to direct future experiments.

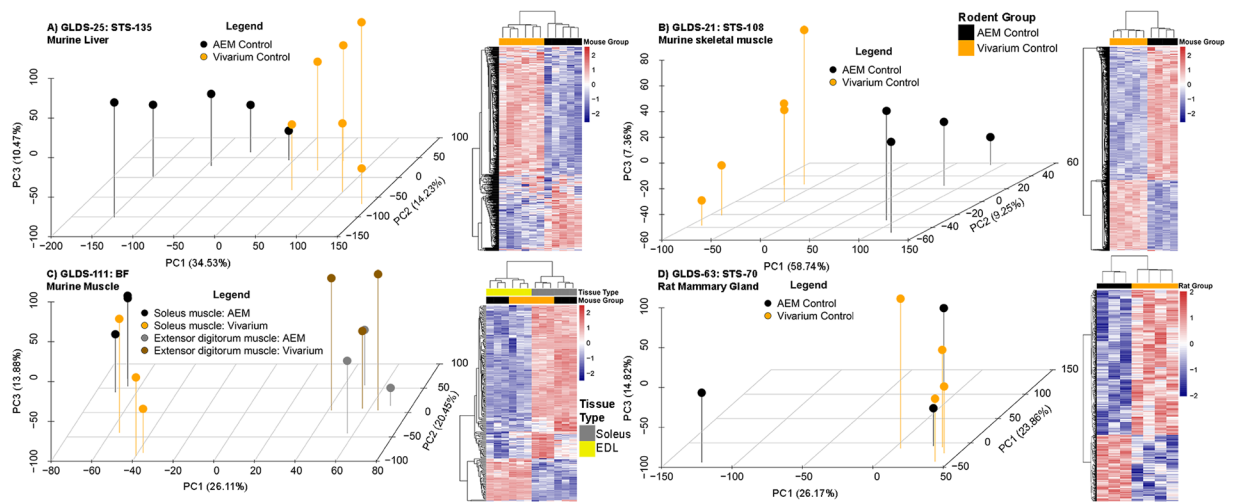
Next, a pathway analysis of the selected genes was performed by using a fold-change  $\geq 1.2$  (or  $\leq -1.2$ ) comparing AEM versus vivarium conditions. This fold-change is an arbitrary value which we believe will produce the maximum amount of the genes above the noise. One method for observing pathway relationships was done using Ingenuity Pathway Analysis (IPA) software (Ingenuity Systems). Upstream regulator analysis from IPA identified any molecule that affected the expression or function of the measured downstream target genes. The activation state of each upstream regulator from the experimental data set was determined by calculating the z-score ( $\geq 2$ , activated or  $\leq -2$ , inhibited). Similar analysis through IPA was done predicting biofunctional activity. Gene set enrichment analysis (GSEA) was performed using the entire list of genes and with leading edge analysis as previously described by Subramanian *et al.*<sup>13</sup>. Significant gene sets between age groups were considered with false discovery rate (FDR) of 0.05. Network representation of GSEA functions was done using Cytoscape.

An unbiased method to identify key genes/drivers were determined as previously reported in Beheshti *et al.*<sup>8</sup> by locating the overlapping genes involved in predicting significant upstream regulators, biofunctions, and GSEA gene sets (which includes the following gene sets: C2, C5, and hallmarks gene set). More specifically, for each set of genes under analysis, the association with statistically significant pathways and functions were determined through both IPA and GSEA. Common genes were determined to be involved in the analysis of both IPA upstream regulators and biofunctions. These sets of genes were further compared to the GSEA's leading edge genes (FDR < 0.05). The overlapping genes between these two analyses were considered to be the key genes involved and in control of the majority of predicted functions and activity with the system being analyzed. In previous publications, we have validated key genes determined with this method through experimental approaches involving Western blots, qPCR, and other functional methods<sup>8,14</sup>. These experiments proved that the key genes determined with this bioinformatics interrogation method were indeed involved in the system being studied.

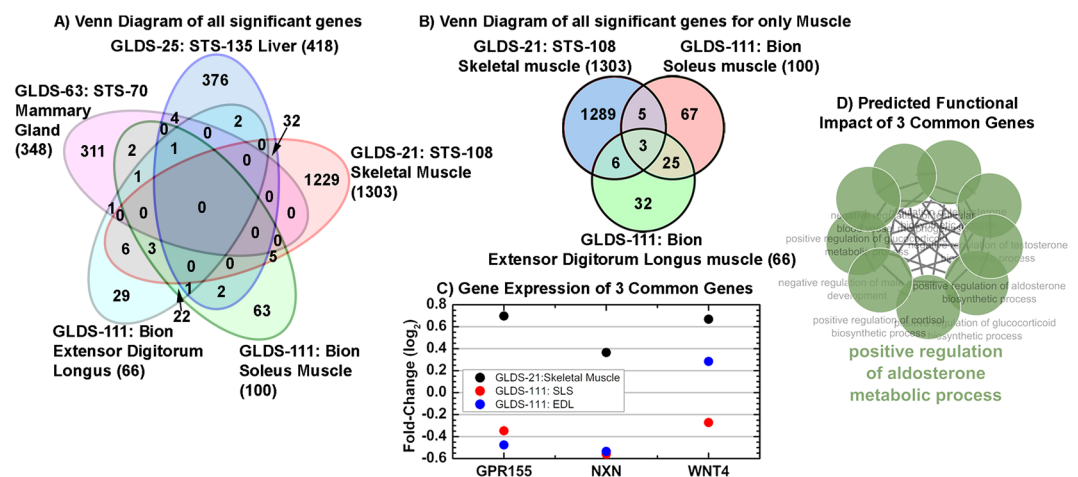
## Results

**Comparing transcriptome changes between AEM and vivarium ground controls of rodent spaceflight experiments.** We selected four rodent (3 mouse, 1 rat) spaceflight studies<sup>2,3,9,10</sup> that used both AEM and vivarium housing (Fig. 1B) as ground controls, and compared global transcriptomic changes associated with different housing conditions. We used publicly available GeneLab data collected from different tissues (Fig. 1) including liver<sup>10</sup>, muscle<sup>2,3</sup>, and mammary gland<sup>9</sup>. The liver (GLDS-25) and skeletal muscle (GLDS-21) datasets were generated from 11 week old C57BL/6 female mice, while the Bion-M1 (BF) muscles were from 19–20 week old C57BL/6 male mice. The mammary gland tissue was from female Sprague-Dawley rats.

Global differences were observed in tissue from rodents in AEM and vivarium housing (Fig. 2 and Supplemental Fig. 1). Principal component analysis (PCA) demonstrated obvious differences between individual AEM and vivarium samples for all conditions (Fig. 2A–C) except for mammary glands from rats (Fig. 2D). It is apparent from the percent variance in PC1 that clear separation occurs between the vivarium and AEM samples (especially for the GLDS-21 skeletal muscle tissue with 59% variance in PC1). Once we determined the significantly regulated genes (p-value < 0.05 or FDR < 0.05), clear differences were observed between AEM and vivarium conditions for all datasets as indicated by the clear division between the up (red) and downregulated (blue) genes represented in the heat maps (Fig. 2 heat maps). Lastly, we analyzed the significantly regulated genes for each dataset with Silhouette plots through K-means statistics (Supplemental Fig. 1). This revealed distinct population of genes which have some functional impact for each dataset. Interestingly, for both Extensor Digitorum Longus (EDL) and Soleus muscles the two functions related with one cluster of genes are shown in the literature



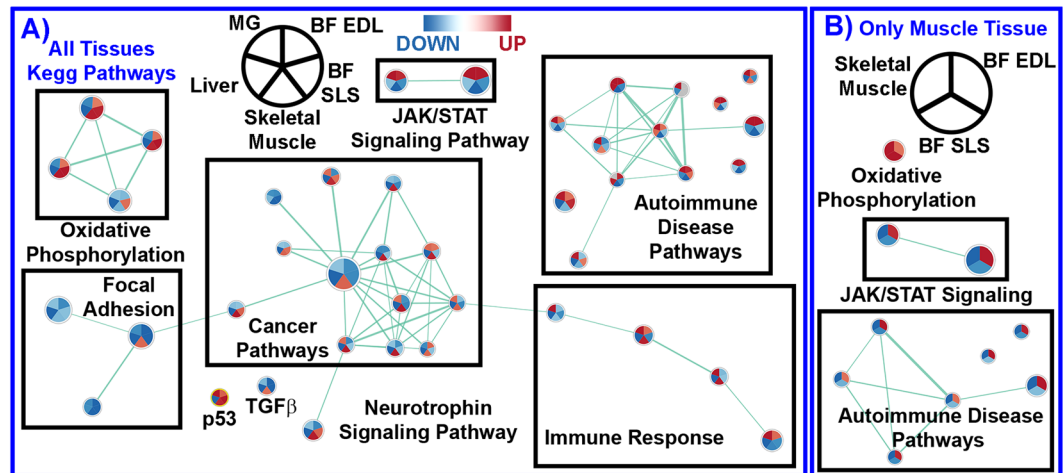
**Figure 2.** Global transcriptome analysis comparing tissue from rodents in AEM vs vivarium housing. Principal component analysis (PCA) and heatmaps representing hierarchical clustering of significantly regulated genes by complete linkage and Euclidean distance calculation (with either  $p$ -value  $< 0.05$  or FDR  $< 0.05$ ) for the following GeneLab datasets and tissue: (A) GLDS-25/STS-135 murine liver, (B) GLDS-21/STS-108 murine skeletal muscle, (C) GLDS-111/BF murine muscle (soleus and extensor digitorum longus), and (D) GLDS-63/STS-70 rat mammary gland. The percentage variance is shown for each PC in parenthesis next to PC axis.



**Figure 3.** Comparison of significantly regulated genes between AEM vs vivarium housed rodents. (A) Venn diagram of significantly regulated genes with 1.2 fold-change for AEM vs vivarium house rodents for all tissues. (B) Venn diagram of significantly regulated genes with 1.2 fold-change for AEM vs vivarium house rodents for only datasets containing muscle tissue. (C) The fold-change values for the 3 common genes between the 3 muscle tissues. (D) The predicted functions of the 3 common genes determined through ClueGO<sup>16</sup>. The overall common function is shown with the bold green text.

to be present with higher regulation of CO<sub>2</sub> in the environment<sup>15</sup>. This indicates that systematic biological differences occur in the tissue of rodents depending on the housing.

**Rodent housing conditions influence the transcriptomic pathways that regulate metabolism and immune responses.** By utilizing established systems biology techniques, we determined the specific biological pathways and key drivers responsible for the systematic differences observed due to housing. We analyzed each dataset separately to identify significantly regulated genes, since all results were generated by microarray assays performed on different Affymetrix platforms (the analysis methods are described further in Materials and Methods). The significantly regulated genes from each dataset were then compared to each other to find possible overlap of genes that might be in common when comparing AEM vs. vivarium conditions (Fig. 3). There was no overlap of significantly regulated genes between all the different tissues (Fig. 3A), which was not surprising given the differences in functions of sampled tissues. However, by focusing on muscle, which was the only tissue type used in multiple experiments, we identified three housing-dependent genes that were differentially



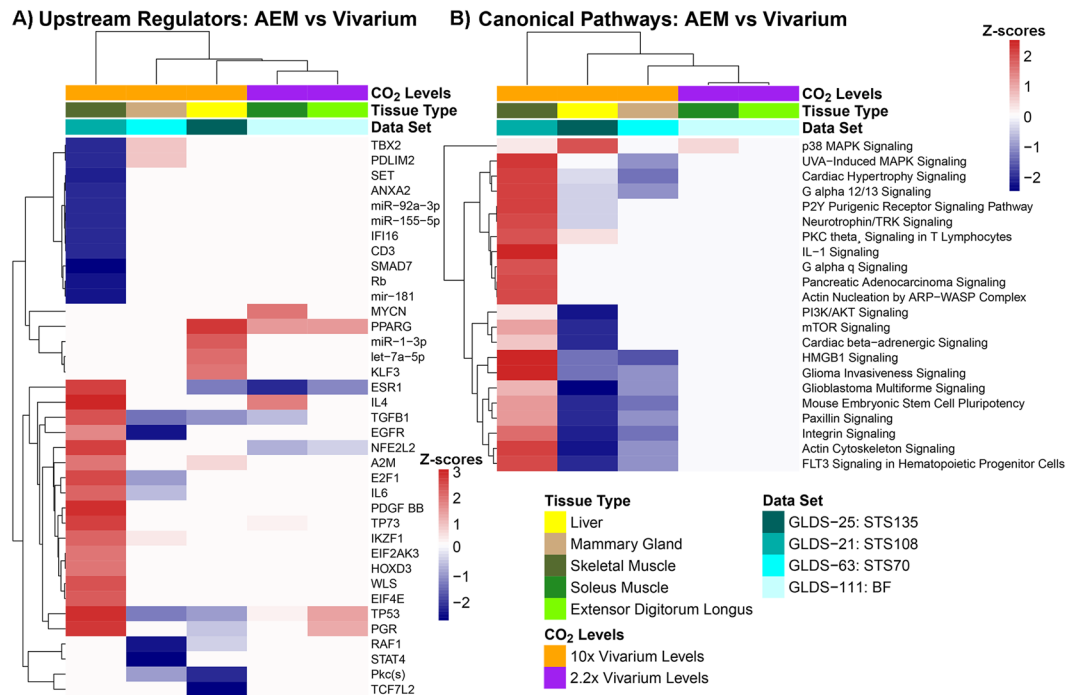
**Figure 4.** Partial network representation of gene sets listed in the KEGG pathway database from GSEA C2 gene set annotations. The GSEA network only shows the most relevant pathways. The complete network of KEGG pathways is shown in the Supplemental Fig. 2. (A) The statistically significant gene sets and the overlapping regulation of each gene set for all datasets and tissues. (B) The statistically significant gene sets and the overlapping regulation for only muscle related tissues. Leading edge analysis with an FDR < 0.05 determined significant gene sets enriched for each group. The size of each node reflects the amount of molecules involved in each gene set. The edge (green lines) represents the number of genes associated with the overlap of two gene sets (or nodes) that the edge connects. Clusters were named according to common function in each grouping. Upregulated gene sets were denoted with red color and downregulated gene sets were denoted by blue color. The grey color represents no change in regulation for that dataset. The legends on the top show which quadrant in the node is associated with a specific tissue with BF EDL = Bion, Extensor Digitorum Longus, BF SLS = Bion Soleus Muscle, and MG = Mammary Gland. The labels for the majority of single nodes without connections were not shown.

expressed in all three types of muscle (skeletal, soleus and extensor digitorum longus) in two studies, STS-108 and Bion-M1 (Fig. 3B): *GPR155* (G-protein coupled receptor 155), *NXN* (nucleoredoxin) and *WNT4* (wingless-type MMTV integration site family, member 4). Notably, gene expression in two muscles from the same GLDS-111/Bion-M1 study followed a similar pattern that was distinct from the muscle in GLDS-21/STS-108, which suggested a study-dependent effect (Fig. 3C).

The functional impact of these three genes in association of AEM environment was predicted by ClueGO<sup>16</sup> to be commonly regulated by the aldosterone metabolic process (Fig. 3D). Aldosterone is a mineralocorticoid hormone that regulates ion balance, kidney and liver function, and cardiovascular inflammation<sup>17</sup>. An increase in aldosterone is associated with metabolic syndrome, which is characterized by chronic inflammation, and aldosterone secretion can be triggered by hypoxia<sup>18</sup>. Further investigation is needed to validate aldosterone as a master regulator and to understand in which direction its metabolic process is regulated by AEM housing. However, based on our analysis we can generate a hypothesis that the association between AEM and inflammatory pathways is mediated by an increase in aldosterone due to a mild hypoxic environment in AEM compared to vivarium cages.

The global functional impact of AEM vs. vivarium cages on rodents was determined by utilizing Gene Set Enrichment Analysis (GSEA) with Kegg Pathway Gene Sets<sup>13</sup> (Fig. 4), upstream regulator predictions through Ingenuity Pathway Analysis (IPA) (Fig. 5A), and canonical pathways prediction also through IPA (Fig. 5B). IPA predictions were done by utilizing Z-score statistics with Z-scores  $\geq 2$  indicated activation while Z-scores  $\leq -2$  indicated inhibition. Using multiple pathway analysis tools allowed us to find central and key pathways that are being affected when comparing AEM with vivarium housing. It was revealed that tissue specific effects occur with AEM vs vivarium housing (Fig. 4A and Supplemental Fig. 2A). For example individual datasets revealed more functional similarities in the muscle transcriptome from mice housed in AEM vs. vivarium cages, including an upregulation of oxidative phosphorylation combined with a downregulation of cytokine-receptor interactions, complement cascades and JAK/STAT pathways, suggesting worse cellular damage and abnormal immune responses in AEM-housed mice (Fig. 4B and Supplemental Fig. 2B). Furthermore, AEM was associated with increased immune activation in the mammary gland as well as inhibited metabolism and cell cycle components in the liver (Fig. 4A and Supplemental Fig. 2A), emphasizing that the effects are tissue-specific.

When comparing multiple pathway tools we were able to observe that p53 and TGF $\beta$ 1 were altered in AEM conditions in the majority of the rodents (Figs 4A and 5A)<sup>19,20</sup>. Out of these factors, TGF $\beta$  is the only one strongly involved in inflammation<sup>21</sup> and oxidative stress responses<sup>22</sup>, and is known to be regulated by hypoxia. TGF $\beta$ 1 was downregulated for all tissues except for the skeletal muscle (and did not appear in the Extensor Digitorum Longus (EDL)), while p53 was up-regulated in the EDL and skeletal muscle and downregulated in liver. We also identified multiple canonical pathways that were strongly altered in high-CO<sub>2</sub> but not low CO<sub>2</sub> matched AEM environment (Fig. 5B). Interestingly, most of them were upregulated in skeletal muscle (Fig. 4B, Supplemental Figs 2B and 5B),



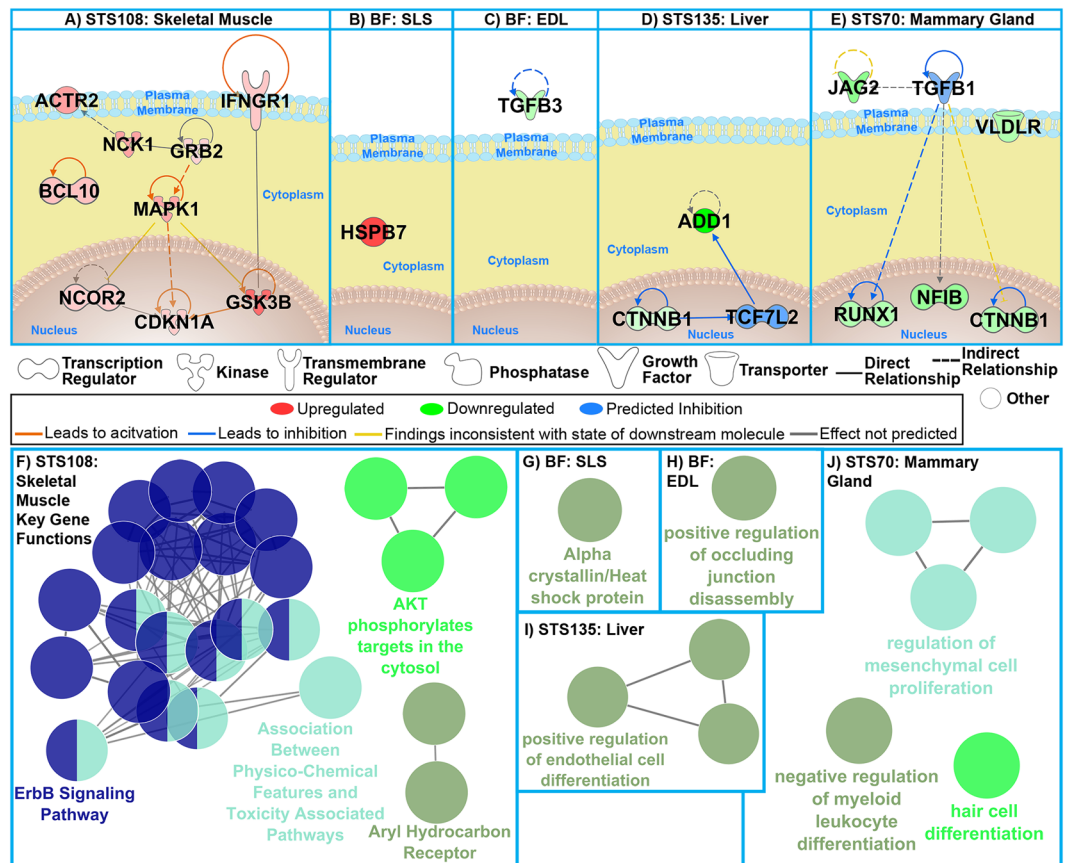
**Figure 5.** Upstream regulator and canonical pathways predictions for AEM vs vivarium housed rodents. Hierarchical clustering based on the Z-scores used to predict either activation (Z-score > 0) or inhibition (Z-score < 0) for predicted upstream regulators (A) or canonical pathways (B).

but downregulated in liver and mammary gland (Fig. 4A, Supplemental Figs 2A and 5b), emphasizing that environmental effects appear organ-specific.

**Identifying key cage-dependent transcriptional drivers.** An unbiased systems biology analysis was utilized to determine the key drivers of housing-dependent changes<sup>8,14</sup>. As previously reported<sup>8,14</sup>, the key drivers were determined by finding the overlapping genes which were involved predicting the GSEA functions, IPA Upstream regulators, Canonical pathways, and Biofunctions. These common genes can be thought of as the central drivers controlling many functions in the system that is being studied. As previously shown for other experiments, once these newly determined drivers are knocked-out, most of the functions will not be activated and the system will be deficient<sup>14</sup>.

Consistent with prior results, the driver genes in the muscle and the mammary gland (Fig. 6A–C,E) were associated with inflammation (e.g. NCK1, BCL10, IFNGR1, TGF $\beta$ 1, JAG2, VLDLR)<sup>21,23–26</sup>, and oxidative stress (NCK1, GSK3B, TGF $\beta$ 1)<sup>22,23,27</sup>. TGF $\beta$ 1 was added as the driver that provided the most connectivity between other genes in the network. Its expression was significantly different between AEM and vivarium controls across multiple tissues, and its expression is known to be associated with multiple spaceflight-relevant conditions and altered by hypoxia (Figs 4 and 5). For the liver (Fig. 6B) the two key drivers were ADD1 and CTNNB1 and a connection was made through TCF7L2. TCF7L2 was predicted to be the most statistically relevant gene that connects these two key genes. It was predicted that the influence of the two key liver drivers will inhibit TCF7L2. Interestingly decrease in TCF7L2 has been shown to have direct impact on hepatic glucose metabolism with the absence of TCF7L2 in the liver causing promotion of blood glucose levels and increased glucose intolerance<sup>28,29</sup>. Also, through predicted upstream regulator analysis both TGF $\beta$ 1 and TCF7L2 were also predicted to be significantly inhibited (Fig. 5A) for the specific tissue.

The key liver genes indicated AEM-dependent reduction in metabolism (ADD1)<sup>30</sup> and unexpectedly, reduced carcinogenesis via inhibition of CTNNB1, which is a major tumorigenic factor. On the other hand, we identified multiple pro-tumorigenic factors, such as NCK1, BCL10, IFNGR1, NCOR2, JAG2, TGF $\beta$ 1, VLDLR, NFIB, RUNX1<sup>19,31–33</sup> and CDKN1A/p21<sup>34</sup>, to be linked with AEM housing in skeletal muscle and in mammary gland (Fig. 6). Several driver genes were associated with a major carcinogenic p53 pathway (CDKN1A, HSPB7) and autophagy (BCL10), which is necessary for both immune regulation and tumor progression. Notably, a large proportion of the genes that were altered by cage type across multiple tissues (e.g. TGF $\beta$ 1, GSK3B, TGF $\beta$ 3, CTNNB1, CDKN1A, JAG2 and VLDLR) are regulated by hypoxia<sup>35–39</sup>. Furthermore, the main pathways associated with the key driver genes primarily regulate immune responses (AKT-dependent phosphorylation<sup>40</sup>, aryl hydrocarbon receptor<sup>41</sup>, alpha crystallin<sup>42</sup>, myeloid leukocyte differentiation<sup>43</sup>) and vascular morphology (occludin, endothelial cell differentiation<sup>44</sup>) (Fig. 6F–J), and interact with TGF $\beta$ <sup>40–44</sup>. The immune responses were also shown to be globally regulated from the GSEA analysis (Fig. 4 and Supplemental Fig. 2) which indicates that the overall immune response can be driven by these key genes. These results suggest that flight-matched AEM environment targets systemic immunity, via TGF $\beta$  signaling.



**Figure 6.** The impact of predicted key genes for tissue from rodents in AEM vs vivarium housing. (A–E) Network representation of key genes for each GeneLab dataset. The predicted relationships between all genes are also shown. Log<sub>2</sub> fold-changes (with a cutoff of 1.2 fold-change) to the gene expression were used to obtain different shades of green for fold-change in downregulated genes, while different shades of red depict fold-change in upregulated genes. The darker the shade of green or red, the greater the fold-change. TCF7L2 and TGFβ1 were added to the networks to provide connectivity and the predicted regulation based on the key genes is shown to inhibit both genes. (F–J) The predicted functions for the key genes for each dataset determined through ClueGO. Each node (circle) represents a specific predicted function with the colors representing the common predicted pathway shown by the text with the same color. Nodes which are associated with more than one common predicted function is shown as two colors.

## Discussion

Using global transcriptome analysis of publicly available GeneLab data from multiple spaceflight studies, we were able to determine key biological differences associated with rodent environmental habitat (spaceflight-matched AEM vs. vivarium cages) in liver, muscle and mammary gland tissues. An established unbiased systems biology approach<sup>8,14</sup> was used to determine the impact of rodent habitats on transcriptional pathways related to immune regulation, metabolic activity and tumorigenesis.

Although prior studies have analyzed behavioral and physiological responses in AEM housing, our study is the first to compare rodent adaptation to spaceflight-matched and standard ground environment at the transcriptional level<sup>45</sup>. A recent comparison of anatomical and physiological parameters of mice and rats housed in AEM and vivarium cages on ground (without a matching flight experiment)<sup>45</sup> indicated physiological changes associated with immune and metabolic functions. For example, AEM-housed rodents had enlarged spleens and increased cholesterol. The duration of this experiment, 35 days, was significantly longer than our studies (~13 days), which may have exacerbated AEM cage effects, but it is representative of current rodent research on the ISS.

Based on IACUC examination at NASA ARC, AEM-housed rodents showed no visible or physiological signs of stress compared to vivarium-housed controls. On the other hand, there have been no studies explicitly comparing stress responses at specific tissue levels for vivarium-housed controls. Chronic low-grade stress in AEM would be consistent both with increased systemic inflammation and reduced thymus<sup>45</sup>. However, contrary to the “stress hypothesis”, AEM allows housing more mice per cage, and higher housing density in mice tends to be associated with lower heart rate and other signs of reduced stress<sup>46</sup>.

Metabolic pathways were consistently shown to be altered in the tissues between AEM vs vivarium housed rodents (Figs 3–6). Cage-dependent differences in metabolism could be caused by different nutritional content of food or different metabolic responses to the same food. Flight food primarily differs from vivarium food in its

consistency: microgravity conditions require non-crumbing solid food bars, irradiated and treated with potassium sorbate to prevent mold and bacterial growth<sup>47</sup>. However, nutritionally flight and vivarium diets are almost identical except for vitamin content, which tends to be higher in the flight diet except for vitamin K<sup>47</sup> and a few differences in indispensable amino acids, including a reduction in methionine<sup>47</sup>. Although vitamin K has been inversely associated with inflammatory markers in an epidemiologic study<sup>48</sup>, and methionine and choline deficiency combined with high fat diet is used to cause liver inflammation in rodent models, these results were achieved by the complete absence of nutrients instead of a minor reduction, so the dietary effects are unlikely to play a major physiological role in AEM vs. vivarium environments.

Furthermore, any changes in diet or exercise can alter the microbiome, which has been shown to regulate both metabolism and immune responses<sup>49</sup>, and since mice are coprophagic, the microbiome is directly shared between cagemates. In the limited number of investigations of microbiome responses to spaceflight environment, the microbiome has been compared between flight and ground AEM-housed rodents, but not between AEM and vivarium. Thus, it may be advisable to include microbiome as a routine component for future studies.

We believe that the most plausible explanation accounting for housing related differences in rodents relates to carbon dioxide level differences between AEM versus vivarium cages. The concentration of carbon dioxide is slightly increased in AEM ground control cages to match the flight conditions. Although the CO<sub>2</sub> concentration in AEM cages on ground or in spaceflight does not exceed 0.6%, which is the maximal permitted level, its concentration in shuttle-matched AEM is approximately 10-fold higher and in Bion-M1-matched AEM is approximately 2-fold higher than the typical ground concentration. Increased CO<sub>2</sub> levels could cause either chronic hypoxia, or hypercapnia and mild acidosis<sup>28</sup>.

Mild chronic hypoxia or hypercapnia due to increased CO<sub>2</sub> levels could explain both the increase in immune responses and a reduction in metabolism, especially since the key driver genes associated with changing the housing conditions, including GSK3B, CTNNA1, JAG2, VLDLR and CDKN1A have also been reported to respond to hypoxia<sup>35–39</sup>. In general, a combination of reduced metabolism and systemic immune dysfunction that is predicted based on transcriptomic levels in AEM cages could contribute to abnormal responses to stressors, including infection, and play a role in systemic decline in organ functions in a manner that resembles the aging process. These deficits in metabolism and immunity can be caused by multiple factors specific to the cage environment, such as the changes in gas composition leading to a mild CO<sub>2</sub>-dependent hypoxia/hypercapnia<sup>50</sup>. Furthermore, the relatively few transcriptional changes associated with Bion-M1-matched AEM could be caused by a lower increase in CO<sub>2</sub> compared to shuttle-matched AEM. However, it is important to emphasize that the causal relationship between CO<sub>2</sub> levels and these physiological outcomes remains to be confirmed or rejected based on experimental evidence.

Chronic CO<sub>2</sub> exposure *in vitro* (10%, while standard cell culture media is kept at 5%) has been reported to induce inflammation via NFκB pathway upregulation<sup>51</sup>, although it is unclear whether this result would translate to the whole organism. There are conflicting reports on rodent responses to mild increases in CO<sub>2</sub> *in vivo*. Increasing CO<sub>2</sub> to 3000 ppm has previously been demonstrated not to affect rat physiological parameters with the exception of sodium levels;<sup>52</sup> however, increasing it to 2000–4000 ppm in mice induces systemic inflammatory responses and vascular leak in muscle<sup>53</sup>. Notably, neither of these studies performed a transcriptional-level analysis. Our results suggest that in order to fully examine the effect of CO<sub>2</sub> it would be advisable to study the responses at the transcriptional level to different CO<sub>2</sub> concentrations by keeping the rest of the environment constant (e.g. only using AEM cages).

In conclusion, through a systems biology approach we observed global transcriptomic changes in rodents induced by spaceflight-matched environment in AEM cages. We also identify spaceflight CO<sub>2</sub> levels as a potential environmental stressor that merits experimental investigation. More generally, systematically changing one environmental aspect at a time (gas concentration, radiation, microgravity, etc.) and analyzing and comparing transcriptional responses could be used to create a network that could predict the most relevant causes and countermeasures for spaceflight-associated conditions, as well as confounding factors for spaceflight experiments.

Finally, our work shows how one can generate new hypothesis utilizing a portion of raw experimental data available to the public in GeneLab platform. Our results indirectly suggest that before designing future experiments it is valuable to access publicly available datasets, such as GeneLab, in order to examine potential confounding factors from spaceflight conditions. This work also highlights that metadata completeness is critical in order to better identify these confounding factors, such as carbon dioxide levels. These specific results were limited to the data that was available on GeneLab which only included a small number of subjects. Although we were able to generate statistically significant results from these small numbers, there are limitations which occur when considering subject-to-subject heterogeneities, study-to-study variations, and tissue-to-tissue variations. For future studies investigators should consider increasing the number of replicates for each experimental condition to allow for better statistics and generation of additional hypothesis for future analysis on GeneLab.

Finally, similar studies could easily be conducted to compare the effects of spaceflight and housing/growing conditions on transcriptional changes in other model organisms such as *Drosophila* and *Arabidopsis*, which also require special hardware in flights. More generally, meta-studies using GeneLab data may inform experimental and engineering goals: from cage design to targeting a particular gene that was discovered to be significant across a variety of studies. Ultimately, integrated analyses are aimed at conserving previous resources in spaceflight, including time, effort, cost of course the crew members.

## References

1. Fernandez-Gonzalo, R., Baatout, S. & Moreels, M. Impact of Particle Irradiation on the Immune System: From the Clinic to Mars. *Front Immunol* **8**, 177, <https://doi.org/10.3389/fimmu.2017.00177> (2017).
2. Allen, D. L. *et al.* Effects of spaceflight on murine skeletal muscle gene expression. *J Appl Physiol* (1985) **106**, 582–595, <https://doi.org/10.1152/jappphysiol.90780.2008> (2009).



3. Gambara, G. *et al.* Gene Expression Profiling in Slow-Type Calf Soleus Muscle of 30 Days Space-Flown Mice. *PLoS One* **12**, e0169314, <https://doi.org/10.1371/journal.pone.0169314> (2017).
4. Andreev-Andrievskiy, A. *et al.* Mice in Bion-M 1 space mission: training and selection. *PLoS One* **9**, e104830, <https://doi.org/10.1371/journal.pone.0104830> (2014).
5. Pecaut, M. J. *et al.* Genetic models in applied physiology: selected contribution: effects of spaceflight on immunity in the C57BL/6 mouse. I. Immune population distributions. *J Appl Physiol* (1985) **94**, 2085–2094, <https://doi.org/10.1152/jappphysiol.01052.2002> (2003).
6. Allen, J. G. *et al.* Associations of Cognitive Function Scores with Carbon Dioxide, Ventilation, and Volatile Organic Compound Exposures in Office Workers: A Controlled Exposure Study of Green and Conventional Office Environments. *Environ Health Perspect* **124**, 805–812, <https://doi.org/10.1289/ehp.1510037> (2016).
7. Law, J. *et al.* Relationship between carbon dioxide levels and reported headaches on the international space station. *J Occup Environ Med* **56**, 477–483, <https://doi.org/10.1097/JOM.000000000000158> (2014).
8. Beheshti, A. *et al.* Host age is a systemic regulator of gene expression impacting cancer progression. *Cancer Res* **75**, 1134–1143, <https://doi.org/10.1158/0008-5472.CAN-14-1053> (2015).
9. Casey, T., Patel, O. V. & Plaut, K. Transcriptomes reveal alterations in gravity impact circadian clocks and activate mechanotransduction pathways with adaptation through epigenetic change. *Physiol Genomics* **47**, 113–128, <https://doi.org/10.1152/physiolgenomics.00117.2014> (2015).
10. Jonscher, K. R. *et al.* Spaceflight Activates Lipotoxic Pathways in Mouse Liver. *PLoS One* **11**, e0152877, <https://doi.org/10.1371/journal.pone.0152877> (2016).
11. Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185–193 (2003).
12. Saeed, A. I. *et al.* TM4 microarray software suite. *Methods Enzymol* **411**, 134–193, [https://doi.org/10.1016/S0076-6879\(06\)11009-5](https://doi.org/10.1016/S0076-6879(06)11009-5) (2006).
13. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102**, 15545–15550, <https://doi.org/10.1073/pnas.0506580102> (2005).
14. Ravi, D. *et al.* Proteasomal Inhibition by Ixazomib Induces CHK1 and MYC-Dependent Cell Death in T-cell and Hodgkin Lymphoma. *Cancer Res* **76**, 3319–3331, <https://doi.org/10.1158/0008-5472.CAN-15-2477> (2016).
15. Kimura, R. & Miller, W. M. Glycosylation of CHO-derived recombinant tPA produced under elevated pCO<sub>2</sub>. *Biotechnol Prog* **13**, 311–317, <https://doi.org/10.1021/bp9700162> (1997).
16. Bindea, G. *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093, <https://doi.org/10.1093/bioinformatics/btp101> (2009).
17. Jaissner, F. & Farman, N. Emerging Roles of the Mineralocorticoid Receptor in Pathology: Toward New Paradigms in Clinical Pharmacology. *Pharmacol Rev* **68**, 49–75, <https://doi.org/10.1124/pr.115.011106> (2016).
18. Maron, B. A. *et al.* Upregulation of steroidogenic acute regulatory protein by hypoxia stimulates aldosterone synthesis in pulmonary artery endothelial cells to promote pulmonary vascular fibrosis. *Circulation* **130**, 168–179, <https://doi.org/10.1161/CIRCULATIONAHA.113.007690> (2014).
19. Bellomo, C., Caja, L. & Moustakas, A. Transforming growth factor beta as regulator of cancer stemness and metastasis. *Br J Cancer* **115**, 761–769, <https://doi.org/10.1038/bjc.2016.255> (2016).
20. Fischer, M. Census and evaluation of p53 target genes. *Oncogene* **36**, 3943–3956, <https://doi.org/10.1038/onc.2016.502> (2017).
21. Cekanaviciute, E. *et al.* Astrocytic transforming growth factor-beta signaling reduces subacute neuroinflammation after stroke in mice. *Glia* **62**, 1227–1240, <https://doi.org/10.1002/glia.22675> (2014).
22. Richter, K., Konzack, A., Pihlajaniemi, T., Heljasvaara, R. & Kietzmann, T. Redox-fibrosis: Impact of TGFbeta1 on ROS generators, mediators and functional consequences. *Redox Biol* **6**, 344–352, <https://doi.org/10.1016/j.redox.2015.08.015> (2015).
23. Chen, J. *et al.* Recruitment of the adaptor protein Nck to PECAM-1 couples oxidative stress to canonical NF-kappaB signaling and inflammation. *Sci Signal* **8**, ra20, <https://doi.org/10.1126/scisignal.2005648> (2015).
24. Torres, J. M. *et al.* Inherited BCL10 deficiency impairs hematopoietic and nonhematopoietic immunity. *J Clin Invest* **124**, 5239–5248, <https://doi.org/10.1172/JCI77493> (2014).
25. Pollard, K. M. *et al.* Definition of IFN-gamma-related pathways critical for chemically-induced systemic autoimmunity. *J Autoimmun* **39**, 323–331, <https://doi.org/10.1016/j.jaut.2012.04.003> (2012).
26. Xishan, Z. *et al.* Jagged-2 enhances immunomodulatory activity in adipose derived mesenchymal stem cells. *Sci Rep* **5**, 14284, <https://doi.org/10.1038/srep14284> (2015).
27. Jellusova, J. *et al.* Gsk3 is a metabolic checkpoint regulator in B cells. *Nat Immunol* **18**, 303–312, <https://doi.org/10.1038/ni.3664> (2017).
28. Norton, L. *et al.* The mechanisms of genome-wide target gene regulation by TCF7L2 in liver cells. *Nucleic acids research* **42**, 13646–13661, <https://doi.org/10.1093/nar/gku1225> (2014).
29. Oh, K. J. *et al.* TCF7L2 modulates glucose homeostasis by regulating CREB- and FoxO1-dependent transcriptional pathway in the liver. *PLoS genetics* **8**, e1002986, <https://doi.org/10.1371/journal.pgen.1002986> (2012).
30. Wood, A. R. *et al.* A Genome-Wide Association Study of IVGTT-Based Measures of First-Phase Insulin Secretion Refines the Underlying Physiology of Type 2 Diabetes Variants. *Diabetes* **66**, 2296–2309, <https://doi.org/10.2337/db16-1452> (2017).
31. Martinez-Iglesias, O. A. *et al.* Autoregulatory loop of nuclear corepressor 1 expression controls invasion, tumor growth, and metastasis. *Proc Natl Acad Sci USA* **113**, E328–337, <https://doi.org/10.1073/pnas.1520469113> (2016).
32. Hu, Y. *et al.* The NOTCH ligand JAGGED2 promotes pancreatic cancer metastasis independent of NOTCH signaling activation. *Mol Cancer Ther* **14**, 289–297, <https://doi.org/10.1158/1535-7163.MCT-14-0501> (2015).
33. Riggio, A. I. & Blyth, K. The enigmatic role of RUNX1 in female-related cancers - current knowledge & future perspectives. *FEBS J* **284**, 2345–2362, <https://doi.org/10.1111/febs.14059> (2017).
34. Georgakilas, A. G., Martin, O. A. & Bonner, W. M. p21: A Two-Faced Genome Guardian. *Trends Mol Med* **23**, 310–319, <https://doi.org/10.1016/j.molmed.2017.02.001> (2017).
35. Huang, S. *et al.* GSK-3beta inhibitor TDZD-8 reduces neonatal hypoxic-ischemic brain injury in mice. *CNS Neurosci Ther* **23**, 405–415, <https://doi.org/10.1111/cns.12683> (2017).
36. Alig, S. K. *et al.* The tyrosine phosphatase SHP-1 regulates hypoxia inducible factor-1alpha (HIF-1alpha) protein levels in endothelial cells under hypoxia. *PLoS One* **10**, e0121113, <https://doi.org/10.1371/journal.pone.0121113> (2015).
37. Wang, J., Ni, Z., Duan, Z., Wang, G. & Li, F. Altered expression of hypoxia-inducible factor-1alpha (HIF-1alpha) and its regulatory genes in gastric cancer tissues. *PLoS One* **9**, e99835, <https://doi.org/10.1371/journal.pone.0099835> (2014).
38. Chen, H., Shen, Y., Gong, F., Jiang, Y. & Zhang, R. HIF-1alpha Promotes Chronic Myelogenous Leukemia Cell Proliferation by Upregulating p21 Expression. *Cell Biochem Biophys* **72**, 179–183, <https://doi.org/10.1007/s12013-014-0434-2> (2015).
39. Xing, F. *et al.* Hypoxia-induced Jagged2 promotes breast cancer metastasis and self-renewal of cancer stem-like cells. *Oncogene* **30**, 4075–4086, <https://doi.org/10.1038/nc.2011.122> (2011).
40. Kurebayashi, Y. *et al.* TGF-beta-induced phosphorylation of Akt and Foxo transcription factors negatively regulates induced regulatory T cell differentiation. *Biochem Biophys Res Commun* **480**, 114–119, <https://doi.org/10.1016/j.bbrc.2016.09.153> (2016).
41. Silgner, M. *et al.* The aryl hydrocarbon receptor links integrin signaling to the TGF-beta pathway. *Oncogene* **35**, 3260–3271, <https://doi.org/10.1038/nc.2015.387> (2016).

42. van Noort, J. M. *et al.* Alfab-crystallin is a target for adaptive immune responses and a trigger of innate responses in preactive multiple sclerosis lesions. *J Neuropathol Exp Neurol* **69**, 694–703, <https://doi.org/10.1097/NEN.0b013e3181e4939c> (2010).
43. Letterio, J. J. Murine models define the role of TGF-beta as a master regulator of immune cell function. *Cytokine Growth Factor Rev* **11**, 81–87 (2000).
44. Zavadil, J. & Bottinger, E. P. TGF-beta and epithelial-to-mesenchymal transitions. *Oncogene* **24**, 5764–5774, <https://doi.org/10.1038/sj.onc.1208927> (2005).
45. Moyer, E. L. *et al.* Evaluation of rodent spaceflight in the NASA animal enclosure module for an extended operational period (up to 35 days). *NPJ Microgravity* **2**, 16002, <https://doi.org/10.1038/npjmicrograv.2016.2> (2016).
46. Morgan, J. L. *et al.* Effects of housing density in five inbred strains of mice. *PLoS One* **9**, e90012, <https://doi.org/10.1371/journal.pone.0090012> (2014).
47. Tou, J. *et al.* Evaluation of NASA Foodbars as a standard diet for use in short-term rodent space flight studies. *Nutrition* **19**, 947–954 (2003).
48. Shea, M. K. *et al.* Vitamin K and vitamin D status: associations with inflammatory markers in the Framingham Offspring Study. *Am J Epidemiol* **167**, 313–320, <https://doi.org/10.1093/aje/kwm306> (2008).
49. Shi, Y. & Mu, L. An expanding stage for commensal microbes in host immune regulation. *Cell Mol Immunol* **14**, 339–348, <https://doi.org/10.1038/cmi.2016.64> (2017).
50. Dominguez-Alvarez, M., Gea, J. & Barreiro, E. Inflammatory Events and Oxidant Production in the Diaphragm, Gastrocnemius, and Blood of Rats Exposed to Chronic Intermittent Hypoxia: Therapeutic Strategies. *J Cell Physiol* **232**, 1165–1175, <https://doi.org/10.1002/jcp.25600> (2017).
51. Cummins, E. P. *et al.* NF-kappaB links CO<sub>2</sub> sensing to innate immunity and inflammation in mammalian cells. *J Immunol* **185**, 4439–4445, <https://doi.org/10.4049/jimmunol.1000701> (2010).
52. Wade, C. E., Wang, T. J., Lang, K. C., Corbin, B. J. & Steele, M. K. Rat growth, body composition, and renal function during 30 days increased ambient CO<sub>2</sub> exposure. *Aviat Space Environ Med* **71**, 599–609 (2000).
53. Thom, S. R., Bhopale, V. M., Hu, J. & Yang, M. Inflammatory responses to acute elevations of carbon dioxide in mice. *J Appl Physiol* (1985) **123**, 297–302, <https://doi.org/10.1152/jappphysiol.00343.2017> (2017).

## Acknowledgements

We would like to thank Dr. Ruth Globus for taking the time to provide insightful comments for this manuscript. Research funding was provided by the GeneLab Project at NASA Ames Research Center, through NASA's Space Biology Program in the Division of Space Life and Physical Sciences Research and Applications (SLPSRA). Any use of trade names is for descriptive purposes only and does not imply endorsement by the US Government.

## Author Contributions

A.B. provided experiment design, interpreted the data, analyzed all data and wrote manuscript. E.C. wrote the manuscript and interpreted the data. D.S. provided exchange of ideas and edited the manuscript. S.C. provided exchange of ideas and edited the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-22613-1>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018