Supplementary Note

Supplementary note 1 – includes additional notes on *Trichinella* found here in all murine faecal samples, a technical comparison of amplicons and WMS species data, notes on low abundance species and expression of intestinal cadherin genes
Supplementary figures 1-21 – including mouse carcass weights, extended amplicon processing statistics, metagenomics processing statistics and host differential expression pathway maps.
Supplementary table 1 - WMS functional pathways of interest
Supplementary data 1 – RR6 amplicon-based metagenomics
Supplementary data 2 – RR6 WMS processing summary
Supplementary data 3 – RR6 WMS-based taxonomy
Supplementary data 4 – RR6 WMS-based function
Supplementary data 5 – RR6 host transcriptomics

Identification of *Trichinella* in all murine faecal samples

The identification of parasitic worms (*T. nativa*) across all mice samples could indicate widespread trichinosis potentially linked to the built-environment of rodent laboratory settings[51-53]. The life cycle of *Trichinella* includes an enteral phase where larval worms within the small intestine penetrate epithelial cells, activating Th1 type immune responses and inflammation before maturing, mating and producing larva. These larva can then induce a dominant Th2 type (helminth) immune responses in hosts aimed towards expulsing the parasite[54]. During the parental phase, larva circulate through the lymphatic and blood systems to enter and damage skeletal muscle. Within the context of spaceflight, exploration of changes in *Trichinella* relative abundance or gene expression are important due to potential interactions with mammalian muscle development, which is compromised in crew during to long-term confinement despite exercise intervention in ground-based experiments, such as MARS500[55,56], and as a consequence of microgravity in space.

16S rRNA gene amplification and whole metagenome sequencing comparison of DA species

To validate the putative species annotation from amplicon analysis, ESVs annotated as species were aligned to WMS contigs. These ESVs could be aligned to WMS contigs at 99.3% ANI and 28/35 species identified using amplicon analysis were independently identified within the WMS co-assembly (including improved resolution of 10 ambiguous calls using WMS). Three of the seven species not identified in WMS were reclassified to a different species within the same genus based on WMS findings (database errors), two were present in WMS and likely represents a strain <97% identity to a known genome (one resolved during analyses as *Acinetobacter courvalinii*, genome published 2023), while three remaining species could represent limited WMS depth or regent contaminants unique to amplification (Supplementary file 3).

Observing low abundance bacteria is important for statistical microbiome analysis

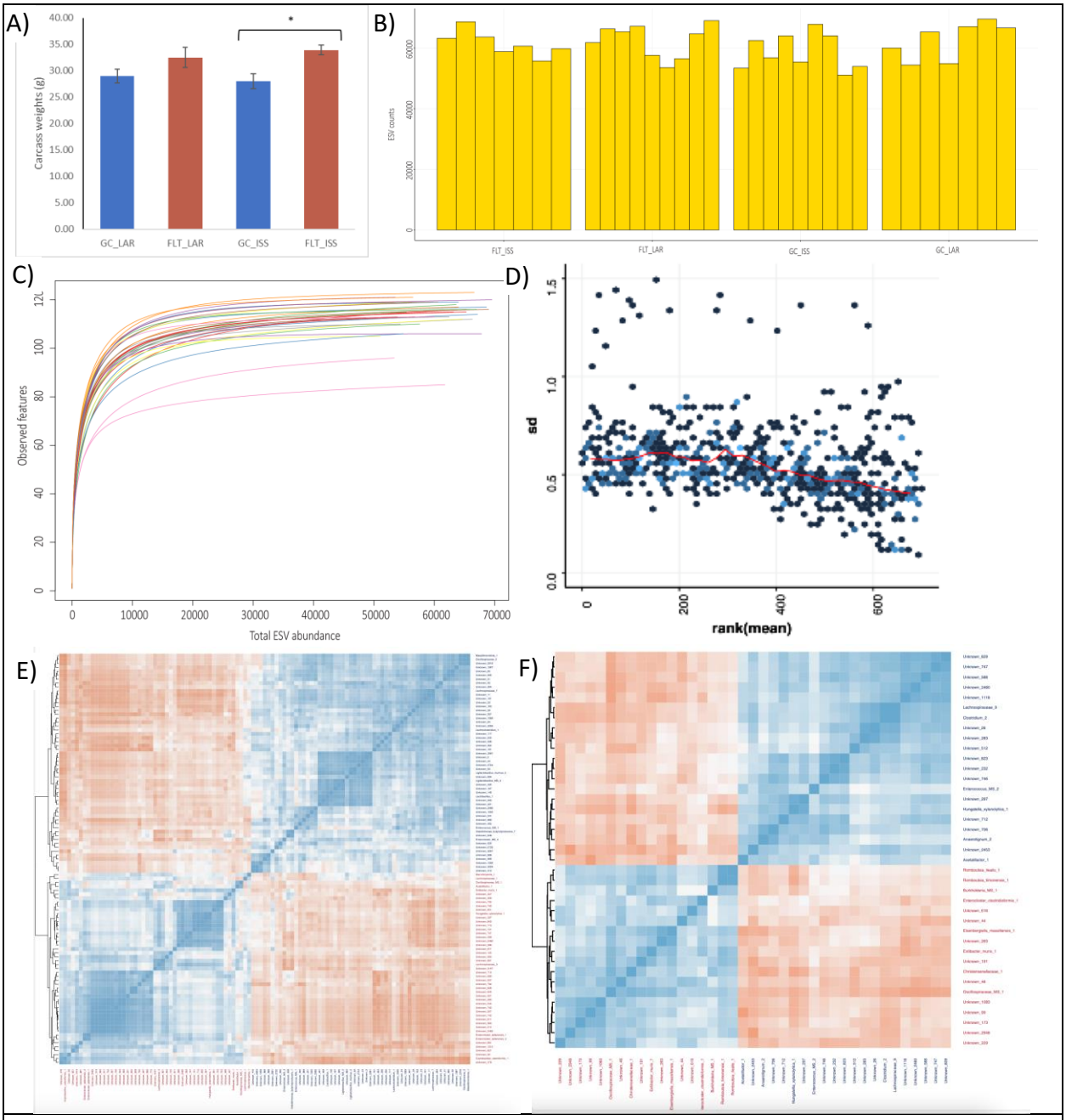
The majority of contigs were relatively short in length (Supplementary file 3) and associated to a small number of species with low relative abundance. Interestingly, contigs from some of these low abundance species were strictly associated to spaceflight, being present in mice after 29 and 56 days of spaceflight, and absent or below detection (structural zeros[51]) in their respective ground controls, including *D. welbionis*, *E. muris* and, in FLT_ISS samples, *G. tenuis* (Supplementary file 3). A limitation to this study was the moderate sequencing depth, with an average 8.7 M reads per sample. Minimum WMS depth requirements to capture most metagenomic gene content has been estimated as high as 80M reads in complex samples, and 200M reads may be insufficient to capture all genetic diversity[52]. Consequently, while more shallow WMS can be informative[53], the assembly and quantification across biological replicates of metagenome assembled genomes (MAGs) from low abundance species is challenging here due to insufficient sequencing coverage (Supplementary file 2). Recent high depth sequencing and culturing efforts are providing insight into the presence[54,55] and importance[56-58] of these lower abundance microbiota. As an example, *G. tenuis* is the type species of the recently described genus characterised within the human Gut Microbial Biobank constructed by Liu et al.[59] exploring "taxonomic dark matter" within the human gut microbiome. The authors predicted association of this previously uncultured species to weight-loss and highlighted it as worthy of further study in being low abundance but extremely widespread in humans (found in all datasets they investigated). The findings generated here, with common significant microbiome changes observed using 16 rRNA amplicon sequencing and WMS across a replicated biological question, reaffirm this importance of low abundance bacteria by suggesting these species could have roles in currently opaque pathologies, such as that induced by spaceflight.

Intestinal cadherin gene expression during spaceflight

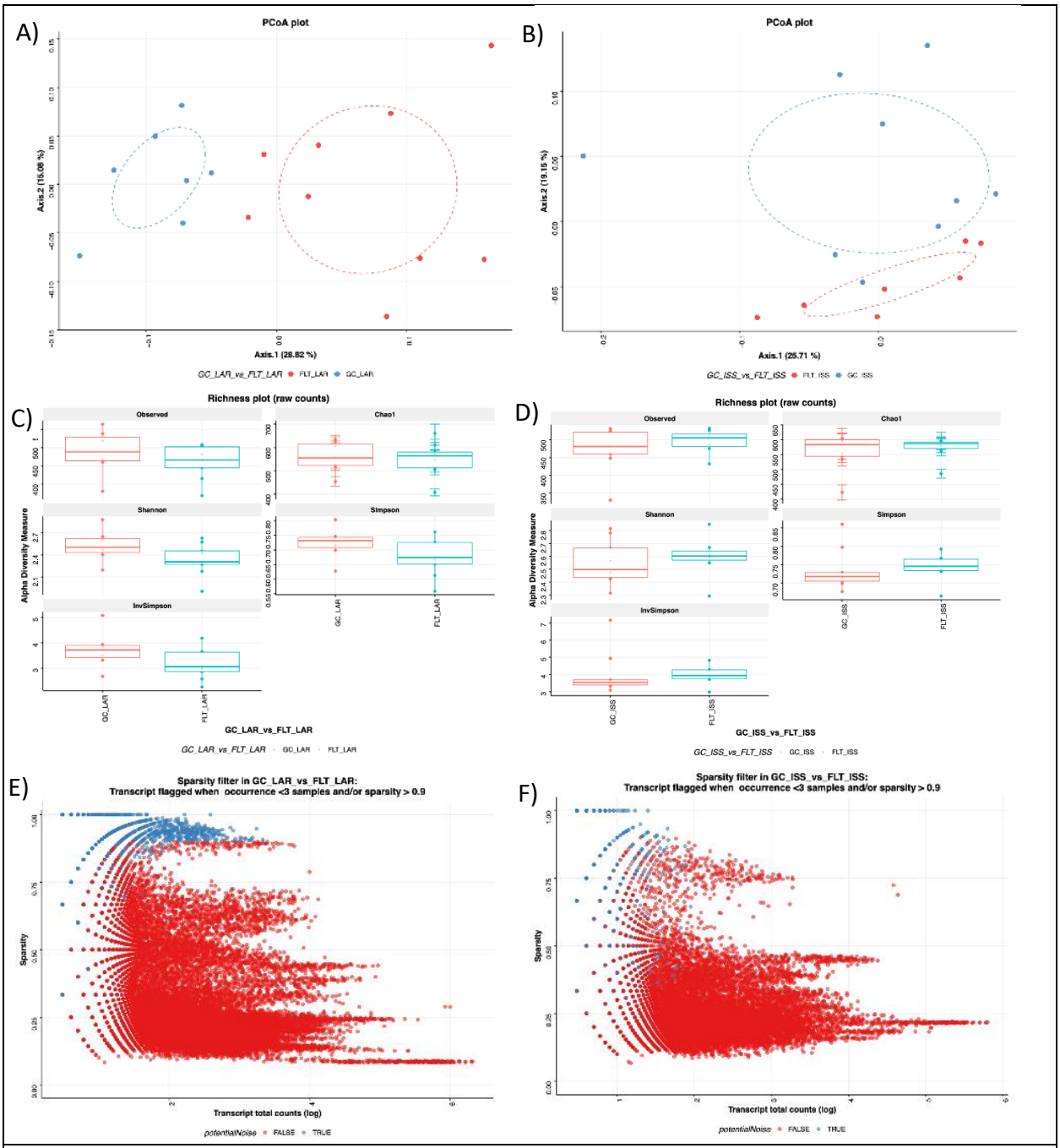
An exception to downregulation of cell adhesion molecules during spaceflight were cadherins, which included significant downregulation of *CLDN2* but up-regulation of *CLDN8*, *CLDN17* and *CLDN23*. Cadherins are transmembrane proteins concentrated in intestinal epithelia at tight junctions where they can form paracellular channels to mediate transport through intracellular spaces. Dysregulation of cadherins is associated with intestinal diseases such as IBD and disruption of epithelial integrity[110], and increases of the steroid hormone-regulated Claudin-8 (*CLDN8*) and Claudin-17 (*CLDN17*) have both been associated with tumorigenesis and cancer proliferation [111,112]. Interestingly, over-expression of *CLDN23* has recently been shown to improve intestinal barrier function but can decrease expression of *CLDN2*, notably involved in epithelial tight junction calcium absorption and bone mineralisation[113,114].

Supplementary figures

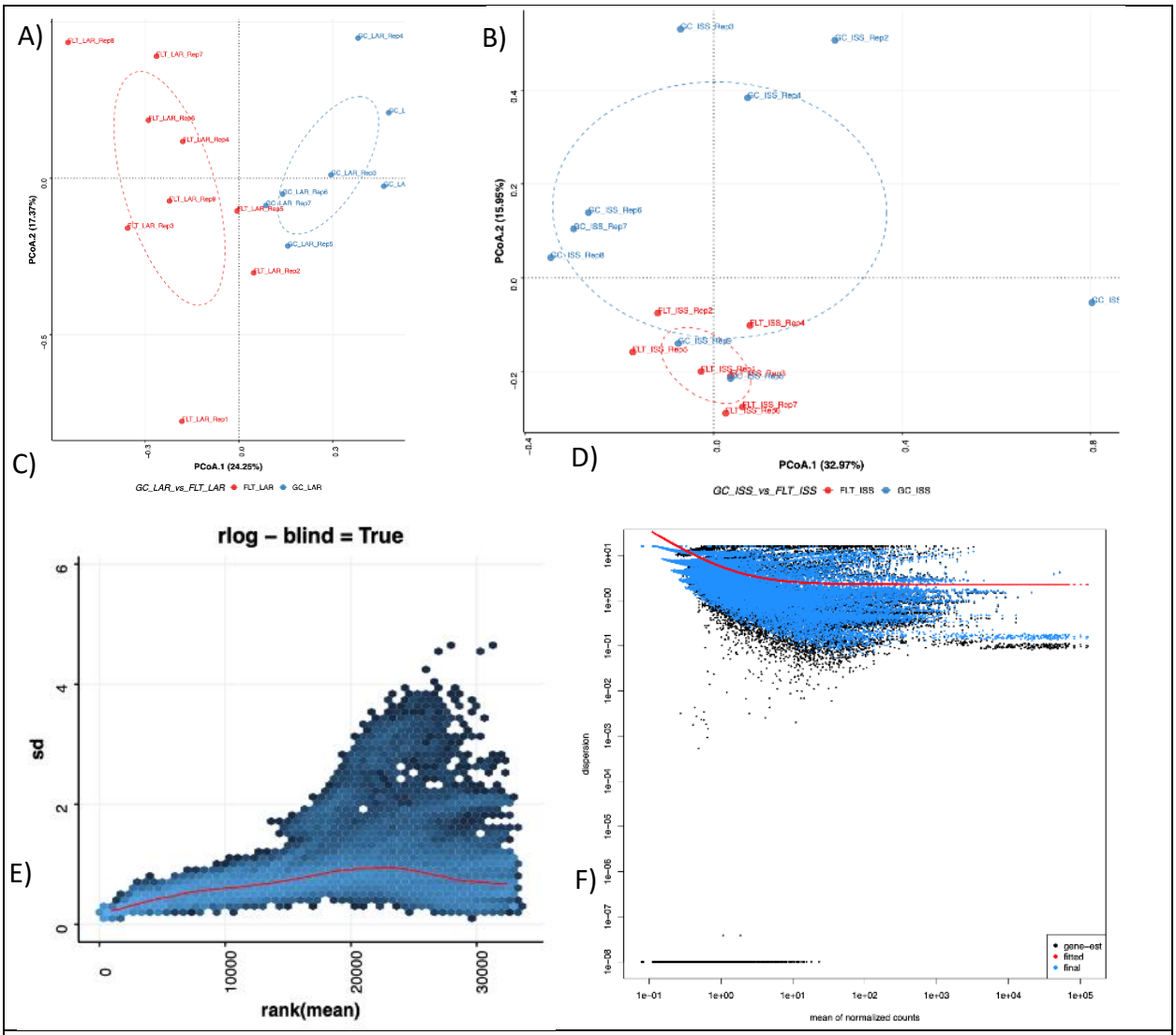
1.1 Spaceflight alters murine gut microbiota



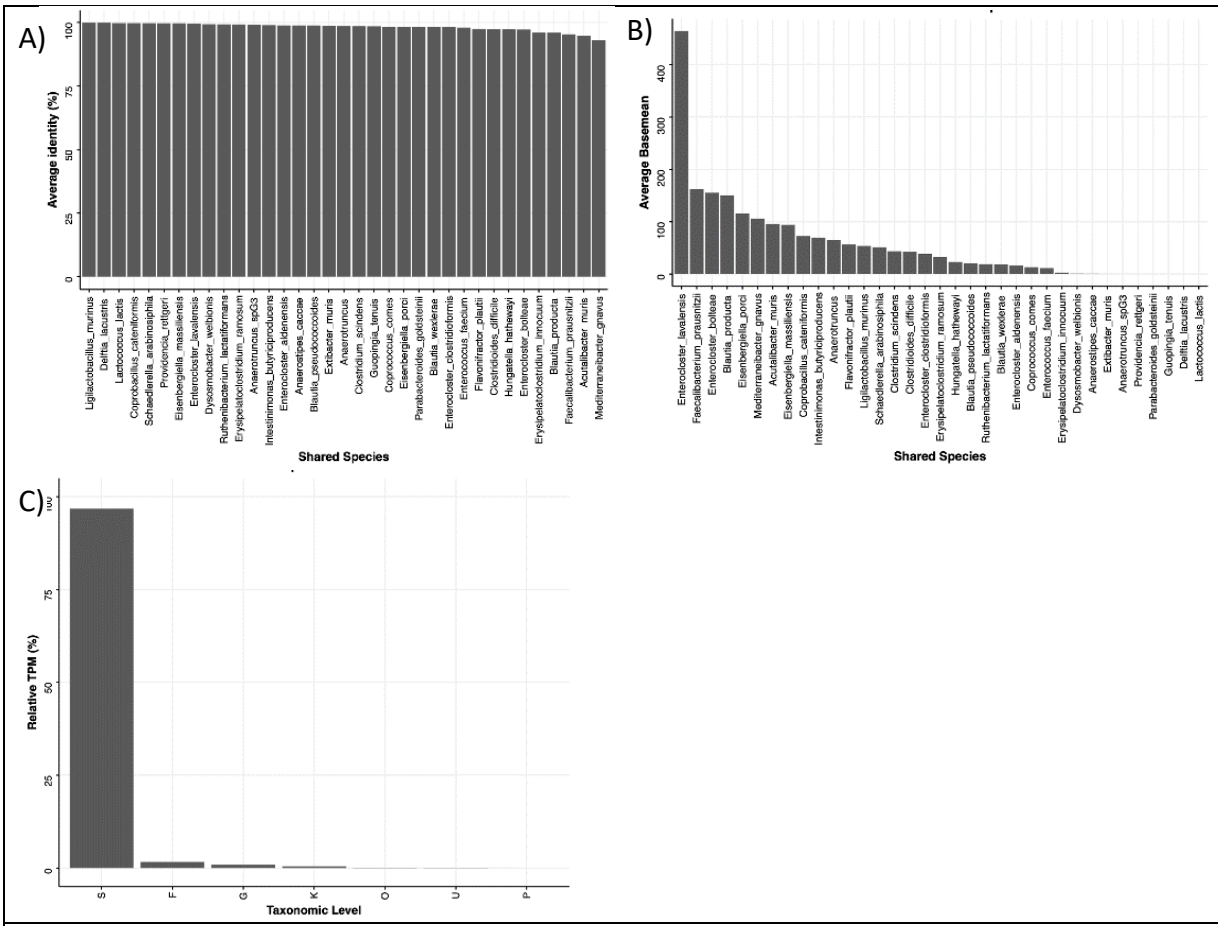
Supplementary figure 1 Carcass weights and 16S rRNA gene amplification statistics A) Carcass weights measured at dissection. B) 16S sample depth (i.e. ESV total abundance). Samples are grouped by replication. C) 16S rRNA rarefaction curves for each sample. D) Standard deviation of 16S normalized counts, i.e. the effect of rlog transformation on the variance. E) Pearson correlation of rlog transformed ESV abundance. Only differentially abundant ESVs are represented. The 2 colors represent the two different factors (GC LAR in blue and FLT LAR in red) F) Pearson correlation of rlog transformed ESV abundance. Only differentially abundant ESVs are represented. The 2 colors represent the two different factors (GC ISS in blue and FLT ISS in red).



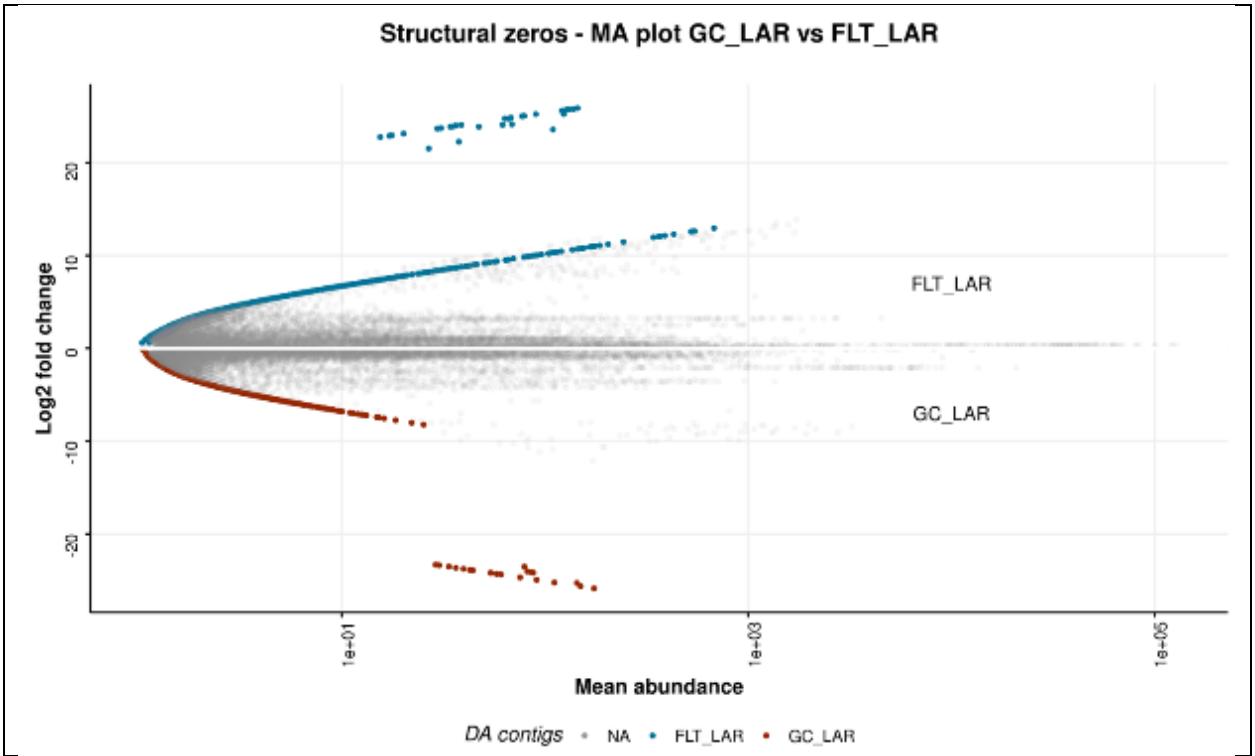
Supplementary figure 2 16S rRNA gene amplification statistics A) Unsupervised ordination using rlog transformed ESV counts in LAR (16S rRNA). B) Unsupervised ordination using rlog transformed ESV counts in ISS (16S rRNA). C) 16S rRNA GC-LAR vs FLT-LAR alpha-diversity of ESVs using multiple indexes. D) 16S rRNA GC-ISS vs FLT-ISS alpha-diversity using multiple indexes. E) Contig sparsity analysis in WMS LAR. Contig in blue were rejected from WMS statistical analysis. F) Contig sparsity analysis in WMS ISS. Contig in blue were rejected from WMS statistical analysis.



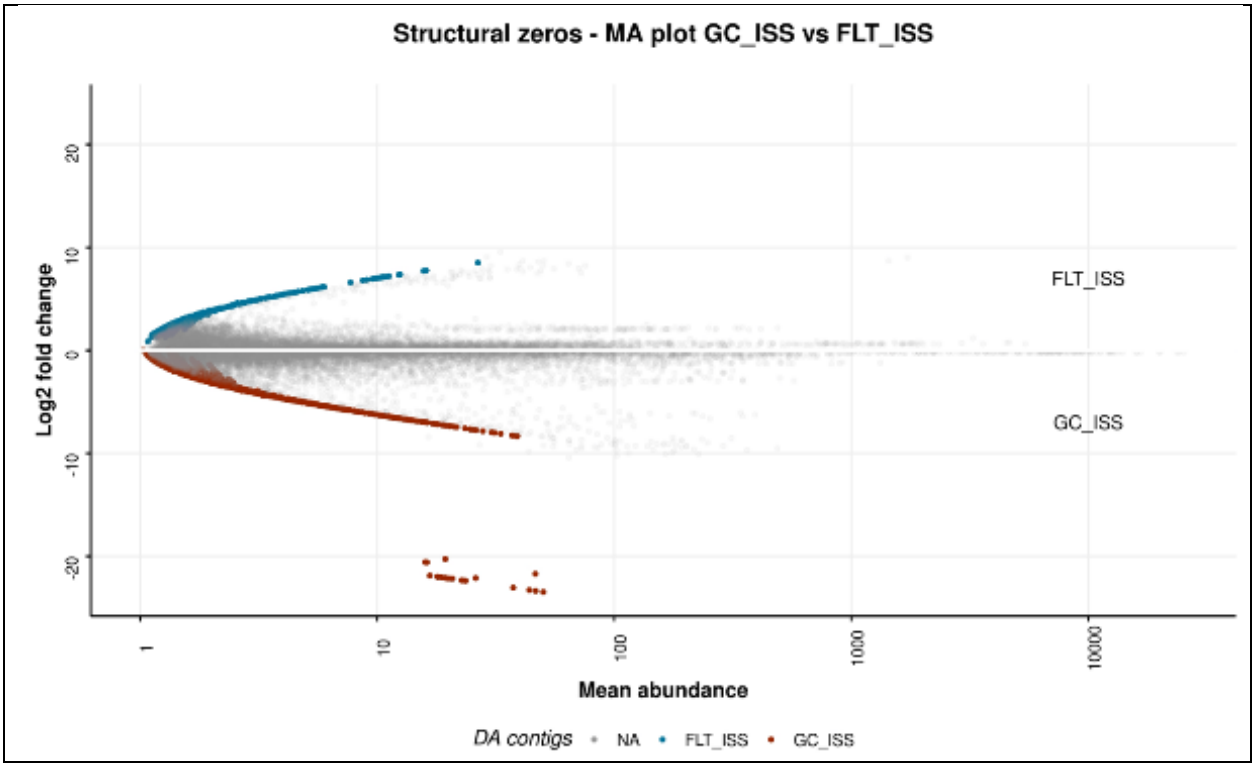
Supplementary figure 3 16S rRNA gene amplification statistics (cont.) A) Unsupervised ordination using cpm transformed contig counts in LARs (WMS). B) Unsupervised ordination using cpm transformed contig counts in ISS (WMS). C) Standard deviation of WMS normalized counts, i.e. the effect of cpm transformation on the variance. D) DESeq2 dispersion plot with the final estimates shrunk from the gene-wise estimates towards the fitted estimates.



Supplementary figure 4 WMS annotation statistics A) WMS annotation. Average ANI per WMS species. B) Average normalized abundance (DESeq2 basemean) across WMS species. C) Relative CPM per taxonomic level

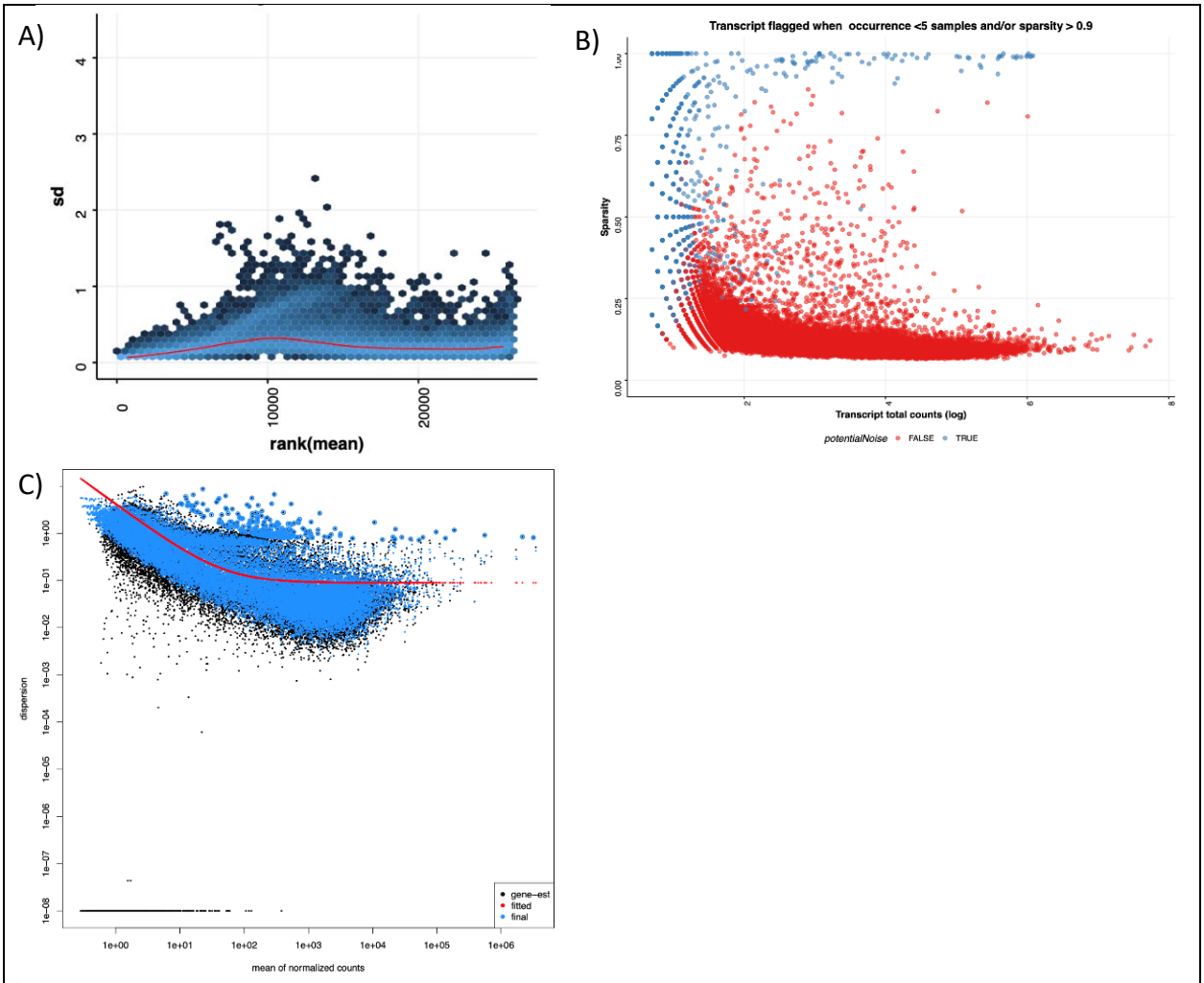


Supplementary figure 5 MA plot structural zeros in GC_LAR vs FLT_LAR. Differential abundance analysis (WMS) MA plots per species are explorable: https://github.com/gonzalezem/Spaceflight_host_microbiome_interactions

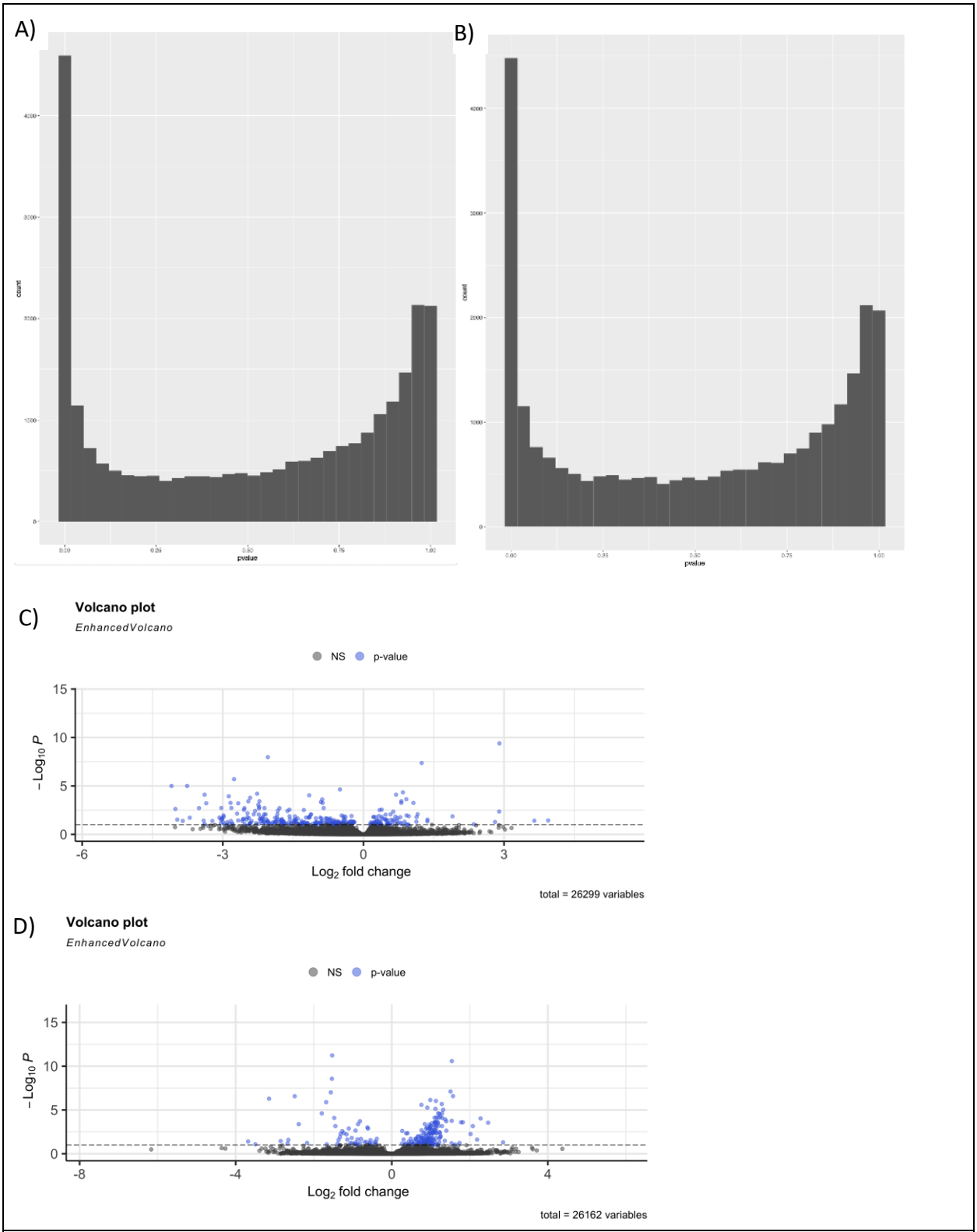


Supplementary figure 6 MA plot structural zeros in GC_ISS vs FLT_ISS. Differential abundance analysis (WMS) MA plots per species are explorable: https://github.com/gonzalezem/Spaceflight_host_microbiome_interactions

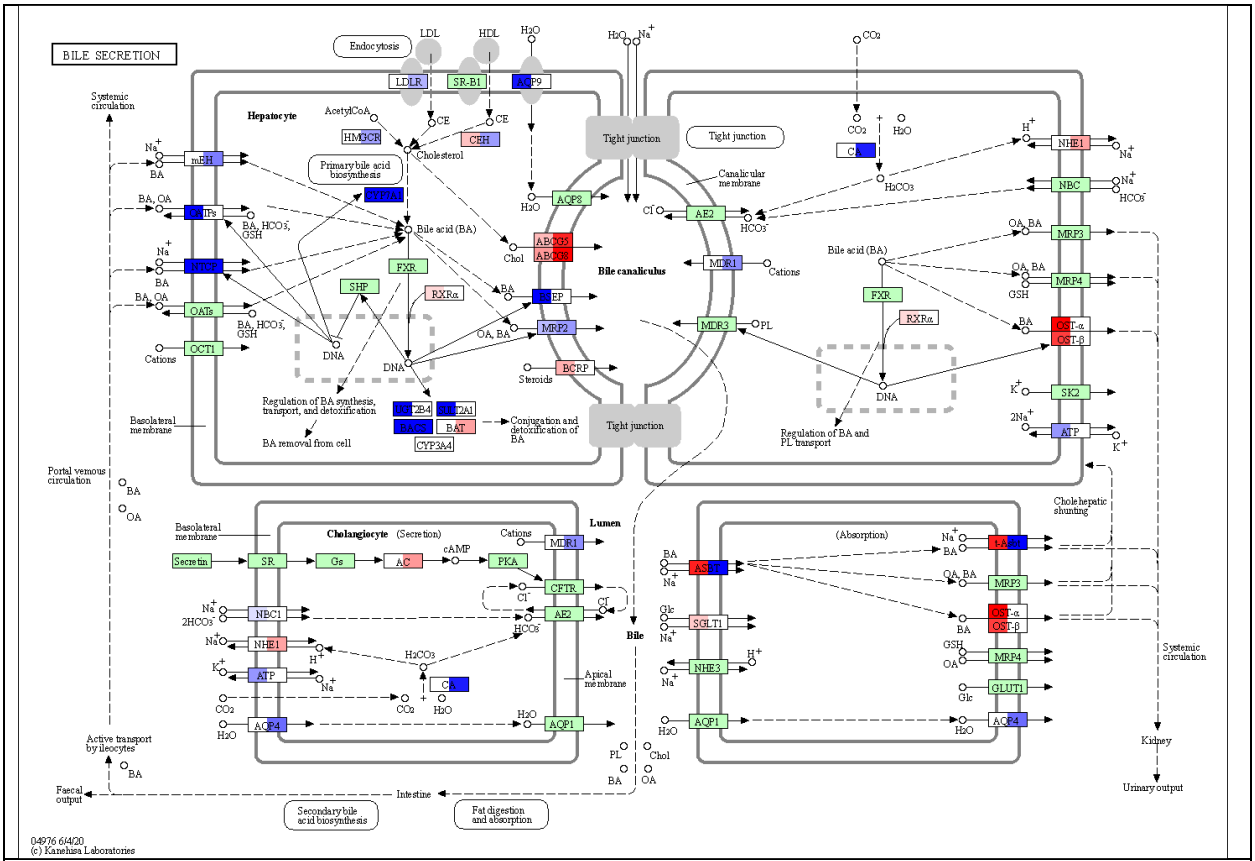
1.2 Spaceflight alters host intestinal gene expression



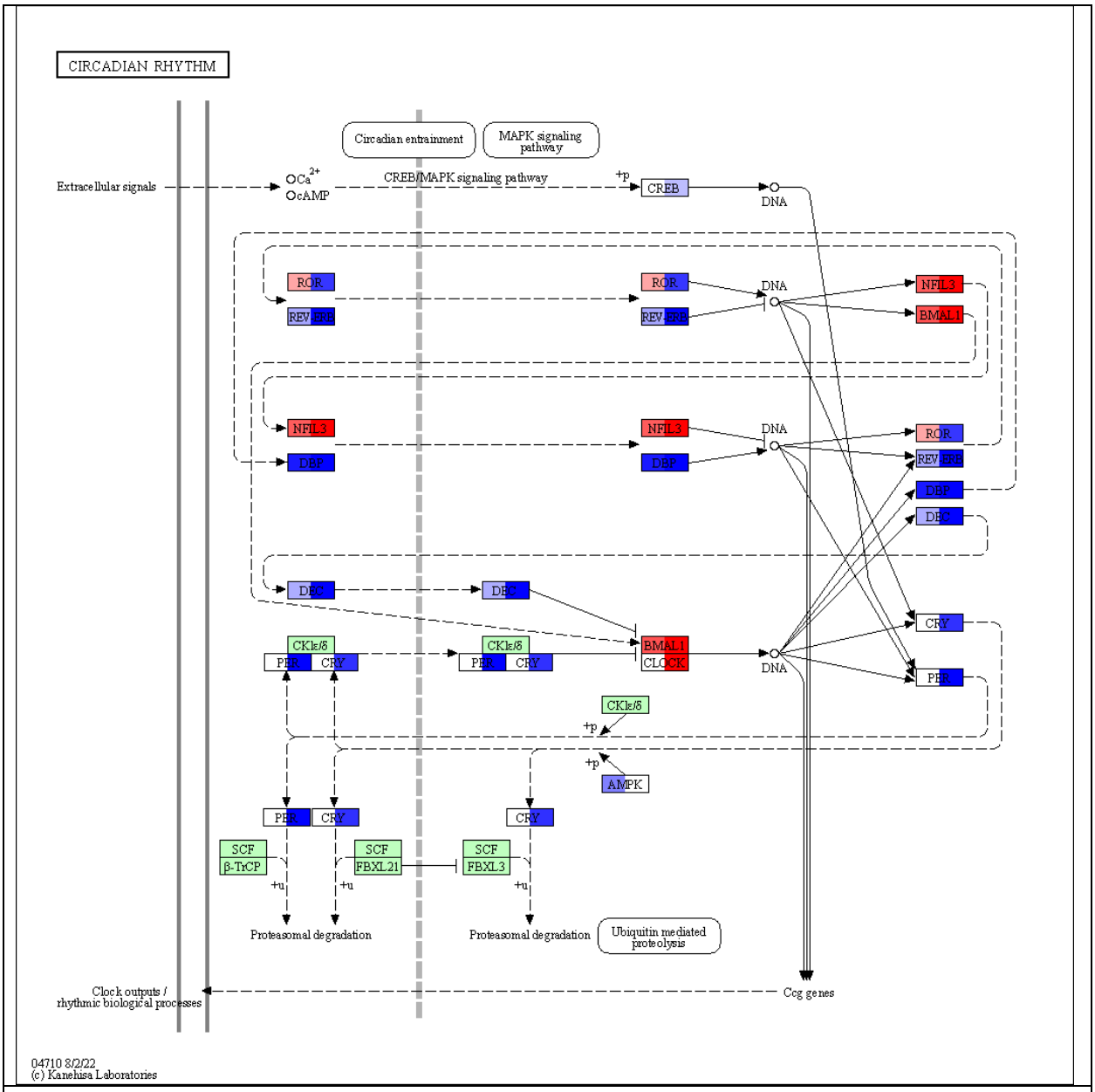
Supplementary figure 7 Colon RNASeq statistics A) Standard deviation of RNA-Seq normalized transcript counts (colon), i.e. the effect of rlog transformation on the variance. B) Sparsity analysis RNA-seq colon. Transcript in blue were rejected from statistical analysis. C) RNA-Seq DESeq2 dispersion plot (colon).



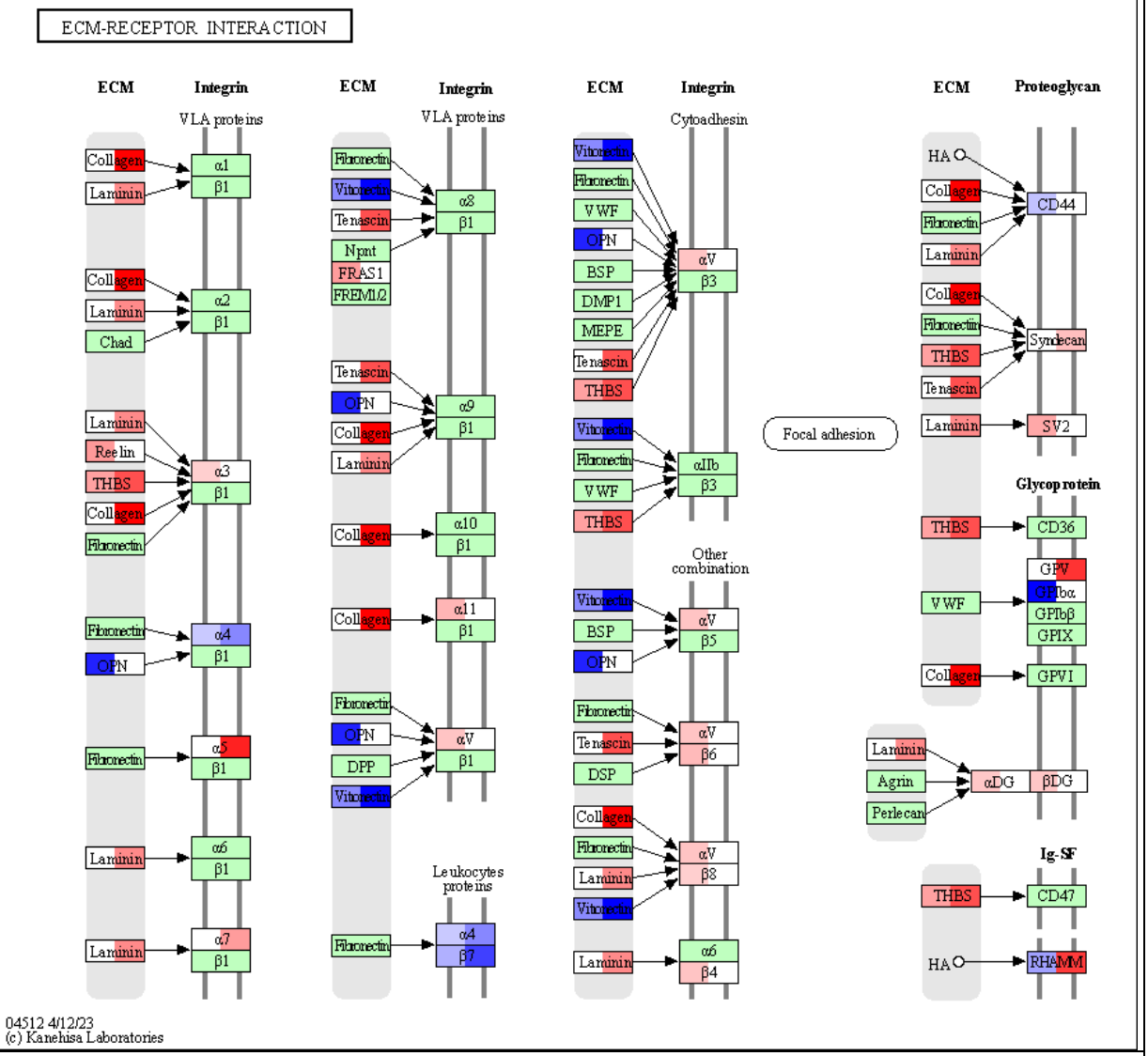
Supplementary figure 8 Colon RNASeq statistics (cont.) A) DESeq2 P-value histogram in GC-LAR vs FLT-LAR (colon). B) DESeq2 P-value histogram in GC-ISS vs FLT-ISS (colon). C) Volcano plot of differential expression analysis of GC-LAR vs FLT-LAR. D) Volcano plot of differential expression analysis of GC-ISS vs FLT-ISS.



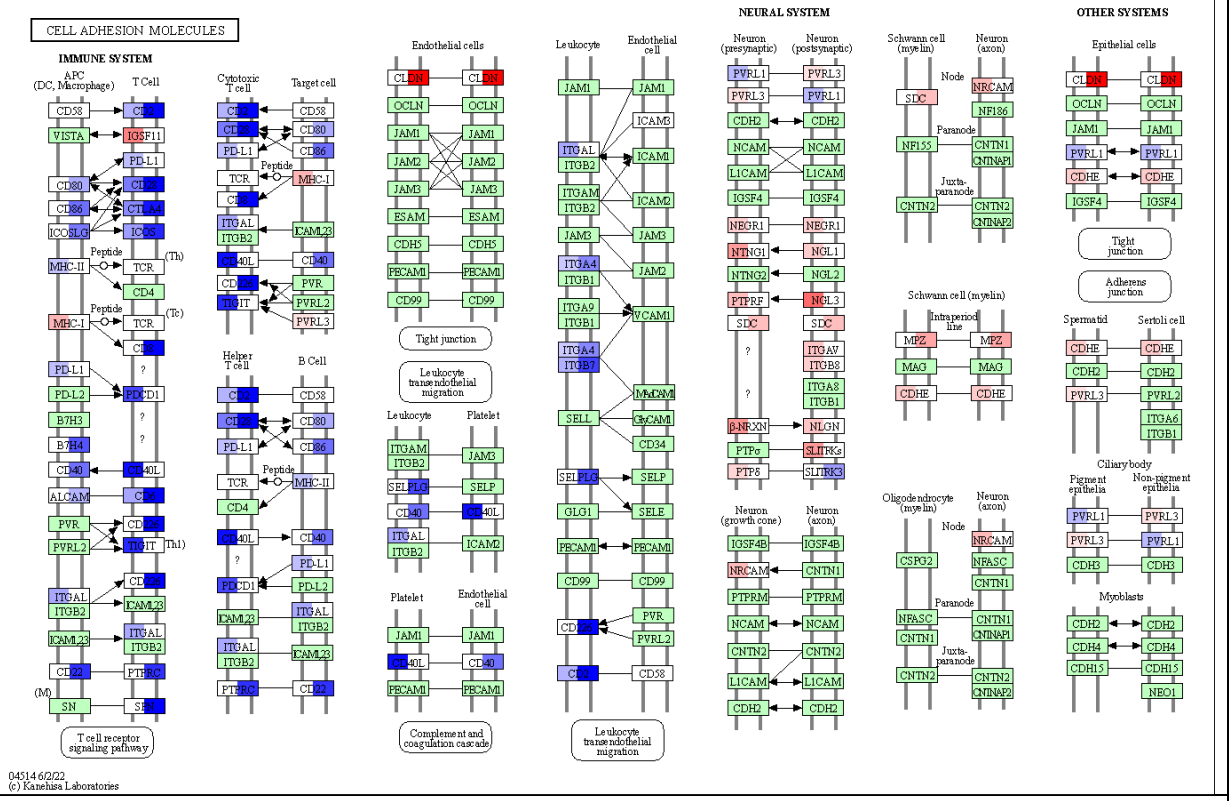
Supplementary figure 9 Selected KEGG pathways with colon differentially expressed genes - Bile secretion. Blue is downregulated and red is upregulated (padj<0.1); left half of gene is GC_LAR vs FLT_LAR (29 days of spaceflight) and right half of gene is GC_ISS vs FLT_ISS (56 days of spaceflight). See *RR6 host transcriptomics* file for full DE gene list.



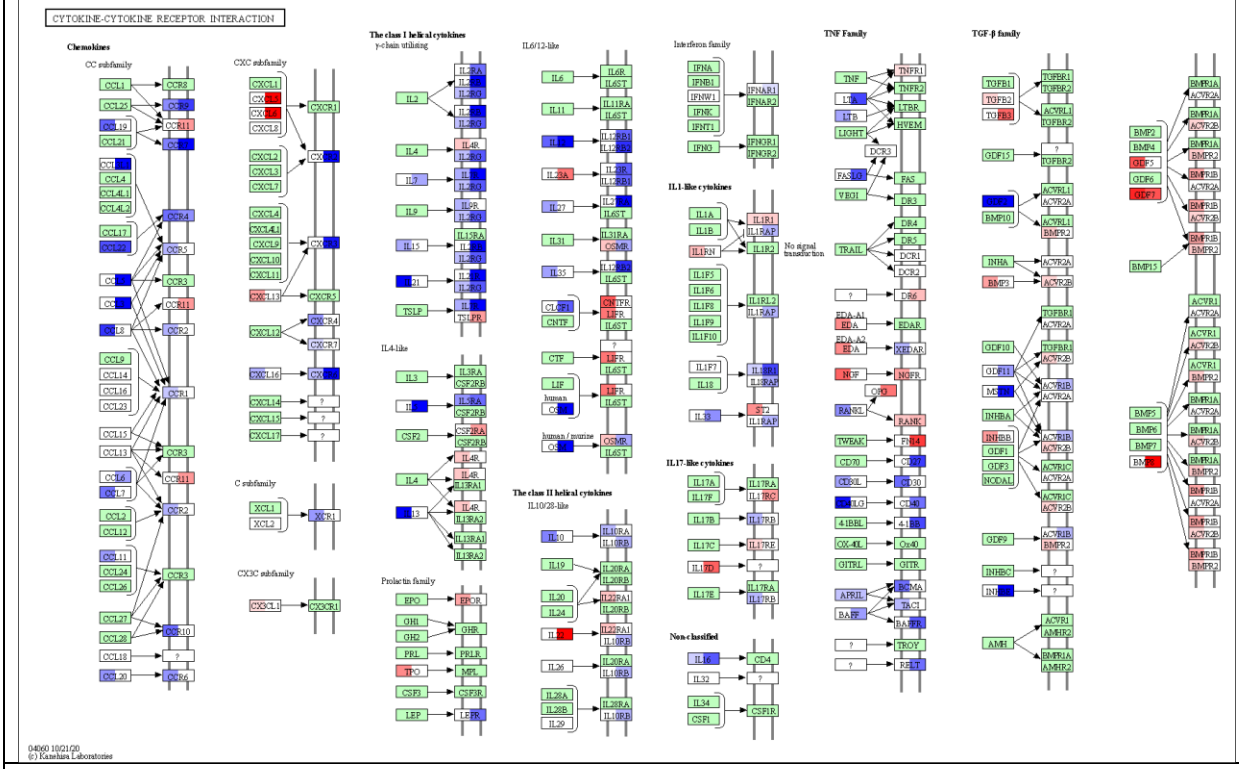
Supplementary figure 10 Selected KEGG pathways with colon differentially expressed genes – Circadian rhythm. Blue is downregulated and red is upregulated (padj<0.1); left half of gene is GC_LAR vs FLT_LAR (29 days of spaceflight) and right half of gene is GC_ISS vs FLT_ISS (56 days of spaceflight). See *RR6 host transcriptomics* file for full DE gene list.



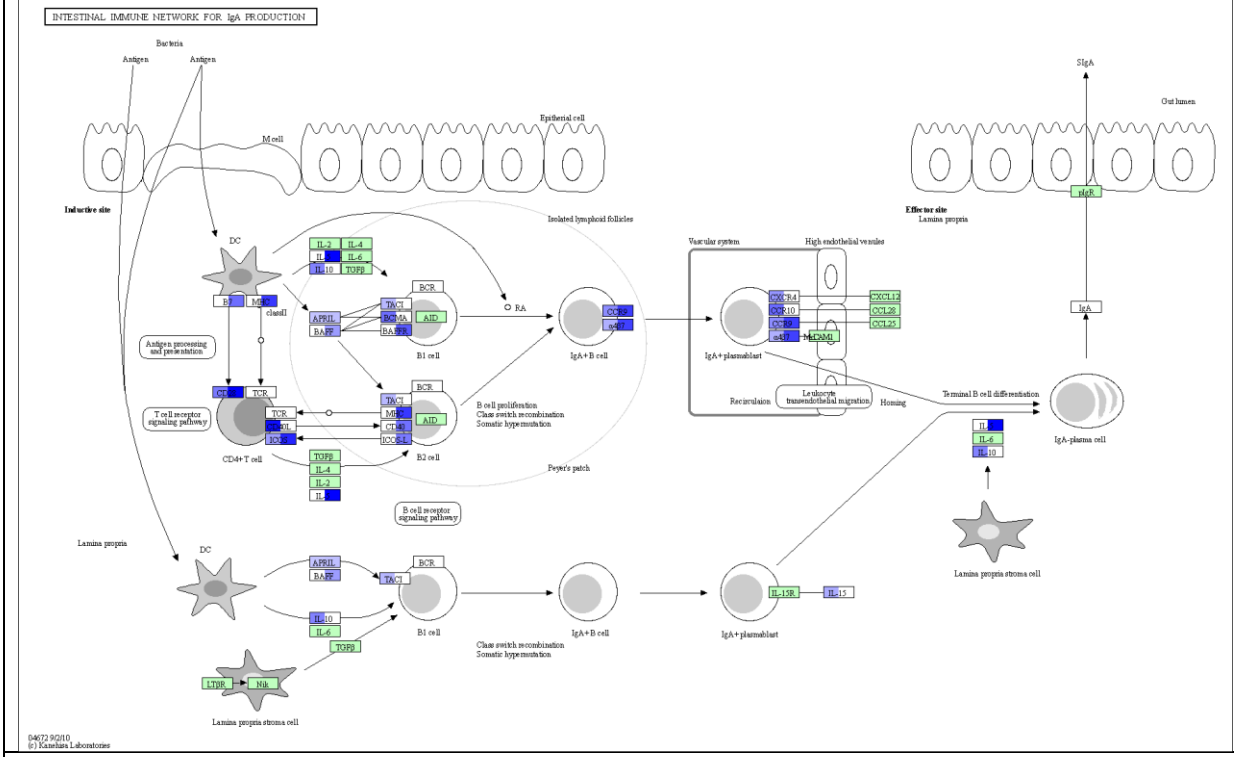
Supplementary figure 11 Selected KEGG pathways with colon differentially expressed genes – ECM-receptor interaction. Blue is downregulated and red is upregulated (padj<0.1); left half of gene is GC_LAR vs FLT_LAR (29 days of spaceflight) and right half of gene is GC_ISS vs FLT_ISS (56 days of spaceflight). See *RR6 host transcriptomics* file for full DE gene list.



Supplementary figure 12 Selected KEGG pathways with colon differentially expressed genes – Cell adhesion. Blue is downregulated and red is upregulated (padj<0.1); left half of gene is GC_LAR vs FLT_LAR (29 days of spaceflight) and right half of gene is GC_ISS vs FLT_ISS (56 days of spaceflight). See *RR6 host transcriptomics* file for full DE gene list.

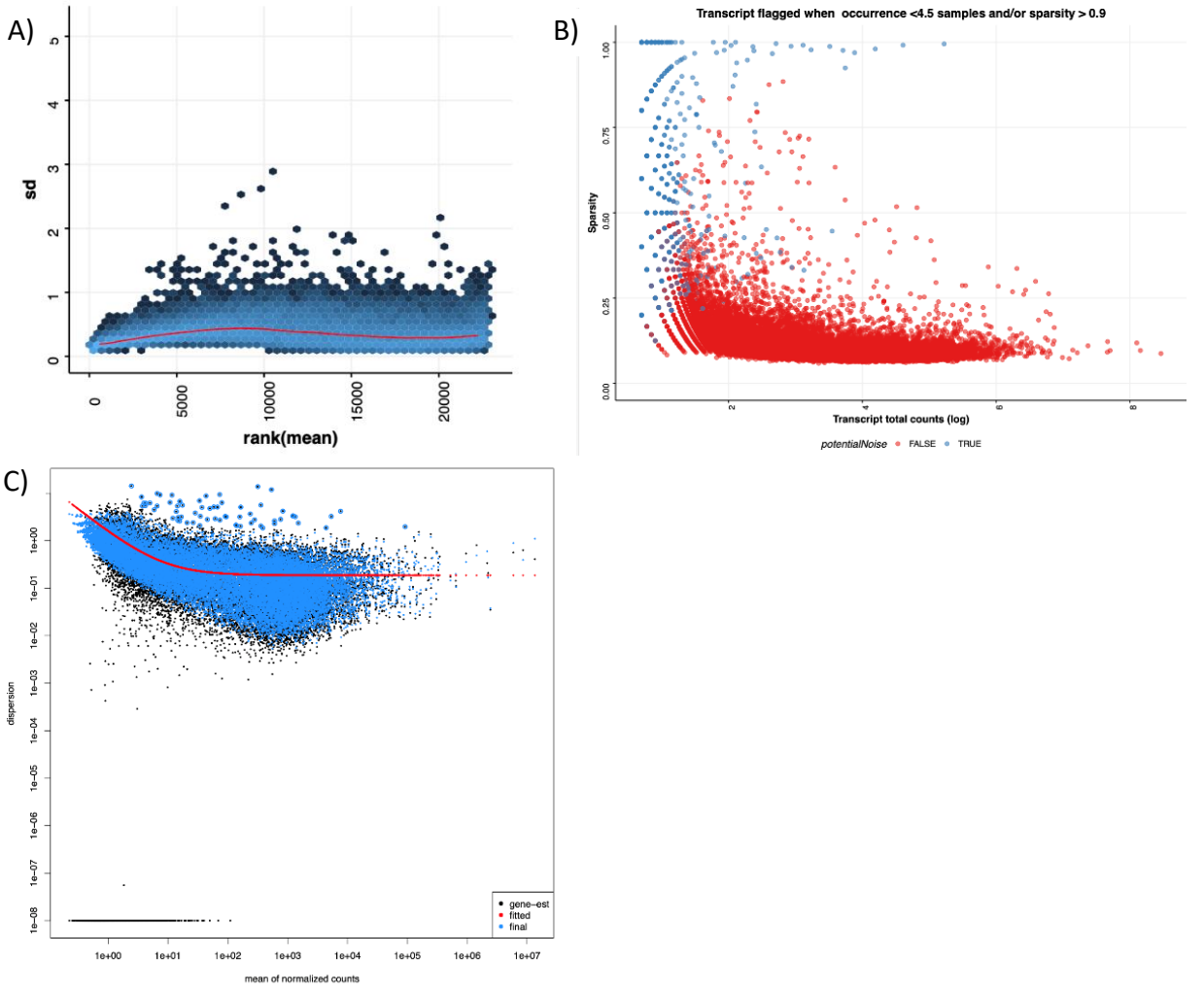


Supplementary figure 13 Selected KEGG pathways with colon differentially expressed genes – Cytokine-cytokine receptor interaction. Blue is downregulated and red is upregulated (padj<0.1); left half of gene is GC_LAR vs FLT_LAR (29 days of spaceflight) and right half of gene is GC_ISS vs FLT_ISS (56 days of spaceflight). See *RR6 host transcriptomics* file for full DE gene list.

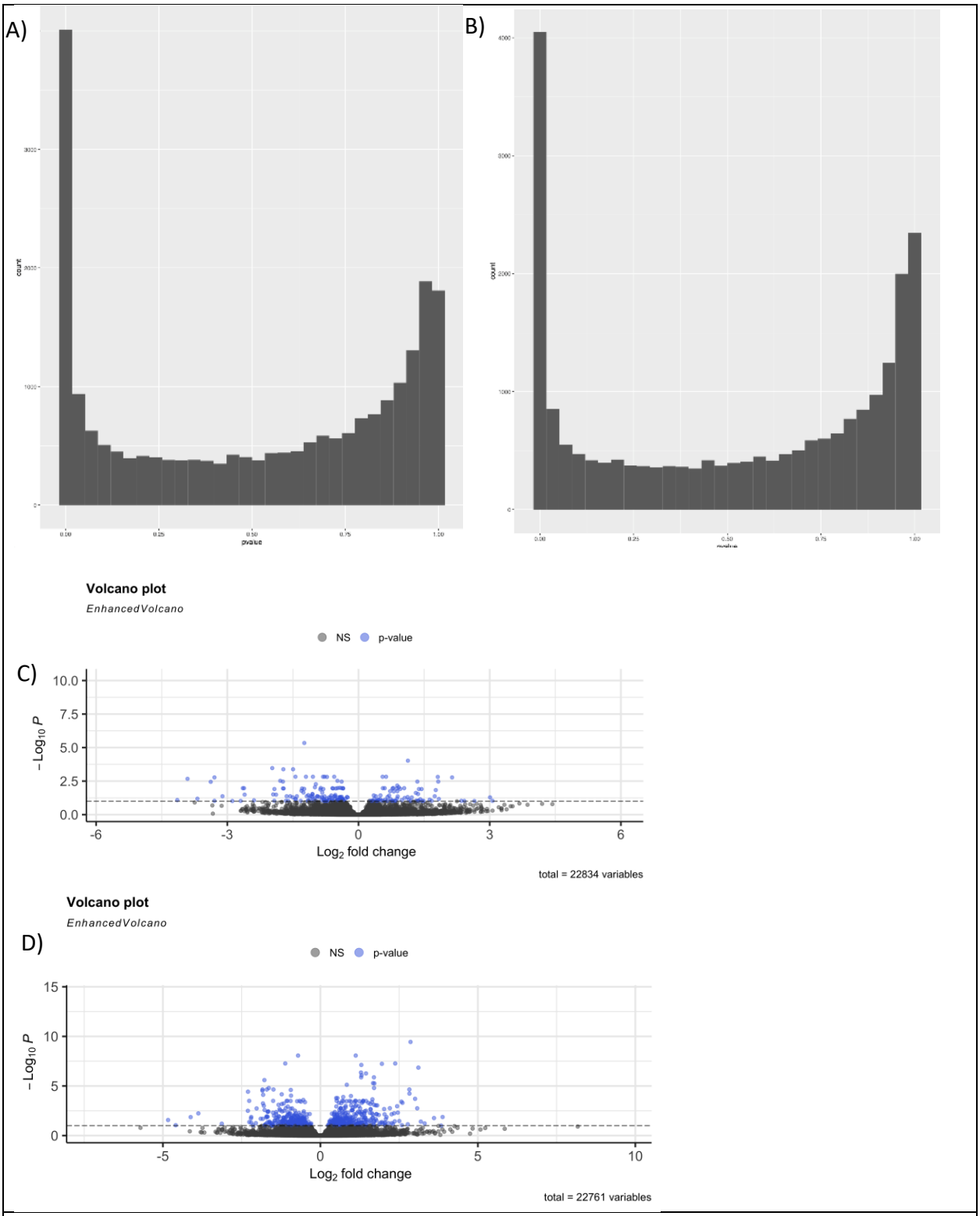


Supplementary figure 14 Selected KEGG pathways with colon differentially expressed genes – IgA production. Blue is downregulated and red is upregulated (padj<0.1); left half of gene is GC_LAR vs FLT_LAR (29 days of spaceflight) and right half of gene is GC_ISS vs FLT_ISS (56 days of spaceflight). See *RR6 host transcriptomics* file for full DE gene list.

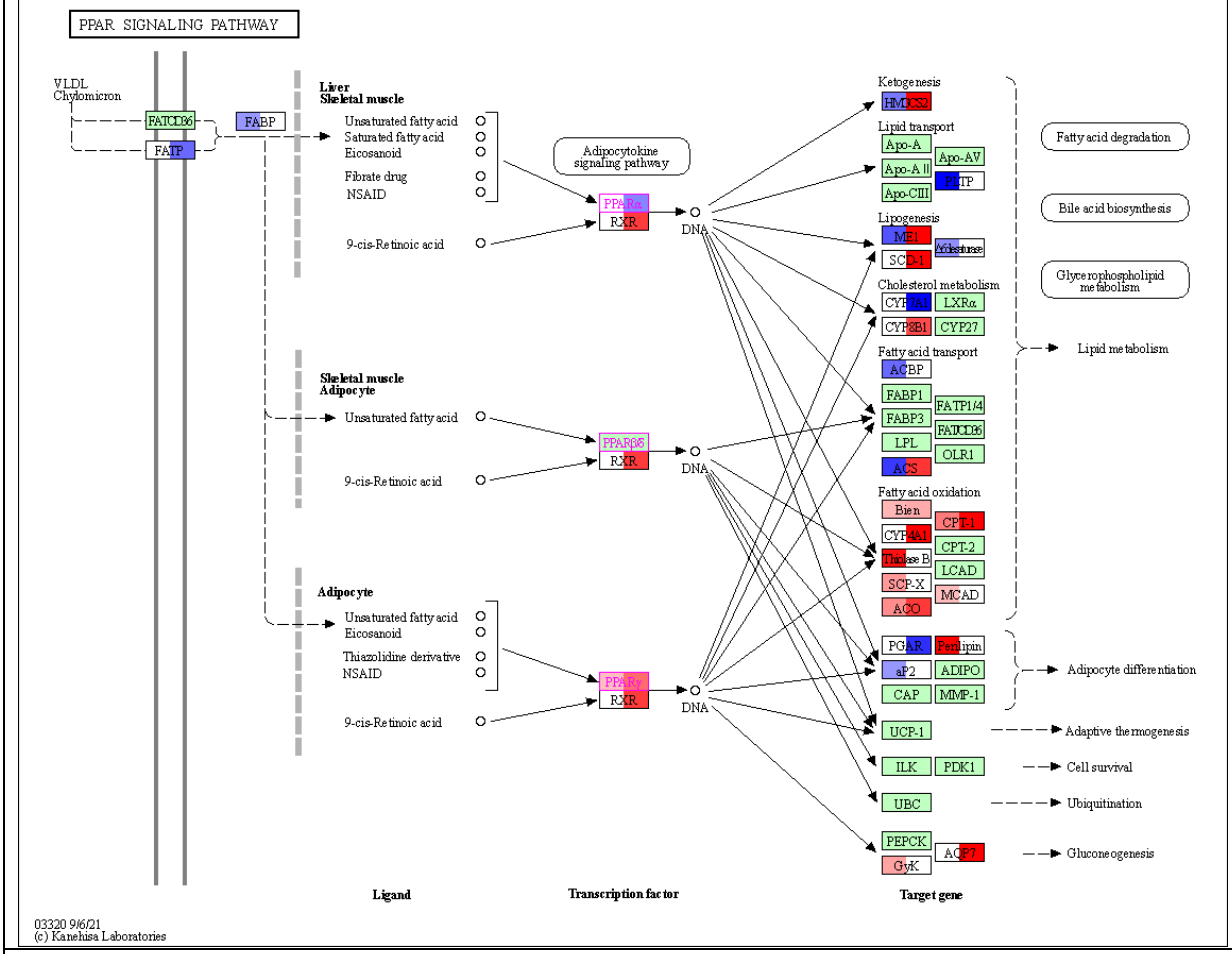
1.3 Spaceflight alters gene expression in the liver



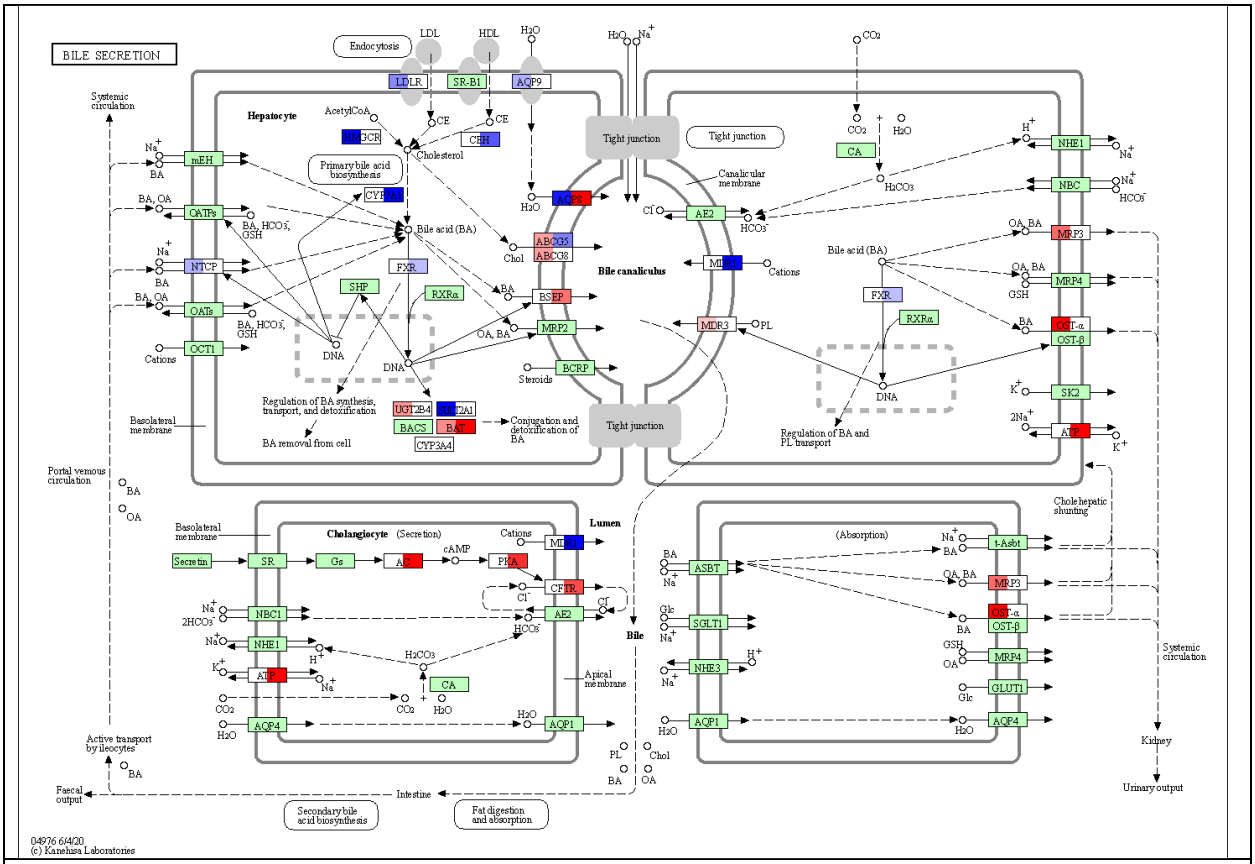
Supplementary figure 15 Liver RNASeq statistics A) Standard deviation of RNA-Seq normalized transcript counts (liver), i.e. the effect of rlog transformation on the variance. B) Sparsity analysis RNA-seq liver. Transcript in blue were rejected from statistical analysis. C) RNA-Seq DESeq2 dispersion plot (liver).



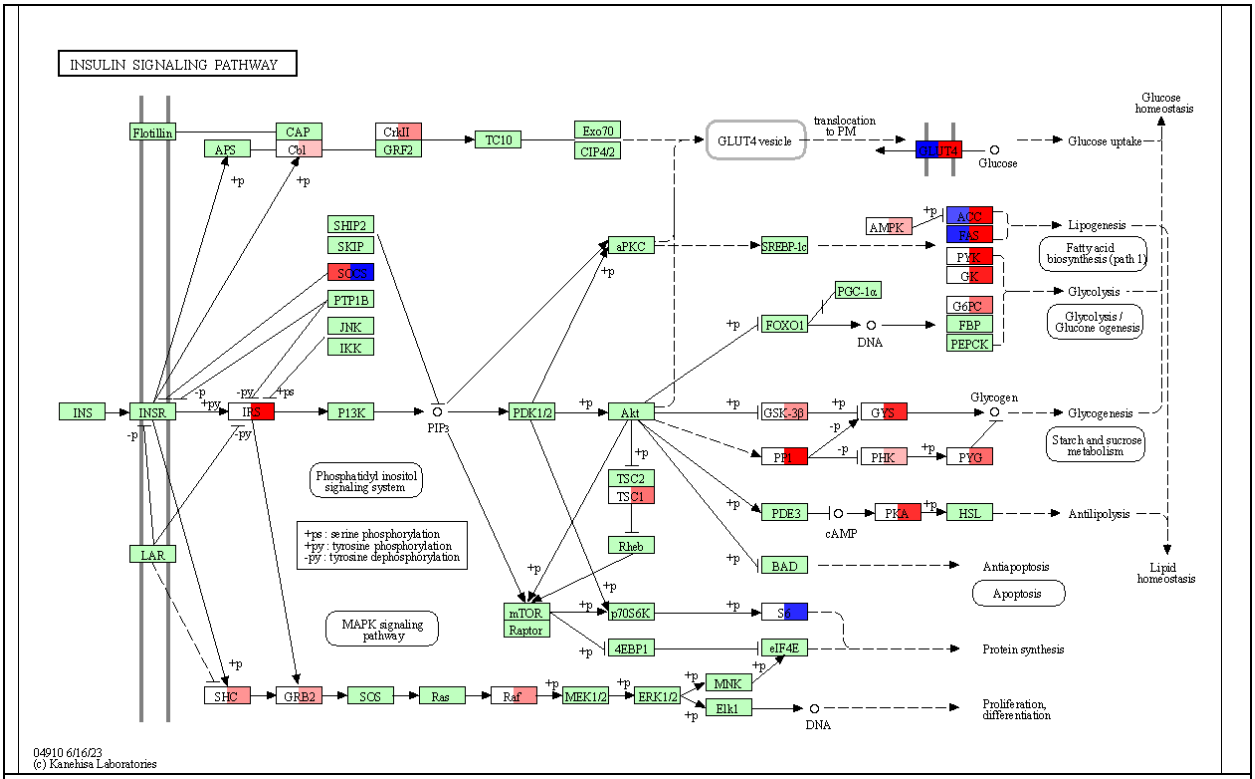
Supplementary figure 16 Liver RNASeq statistics (cont.) A) DESeq2 P-value histogram in GC-LAR vs FLT-LAR (liver). B) DESeq2 P-value histogram in GC-ISS vs FLT-ISS (liver). C) Volcano plot of differential expression analysis of GC-LAR vs FLT-LAR. D) Volcano plot of differential expression analysis of GC-ISS vs FLT-ISS.



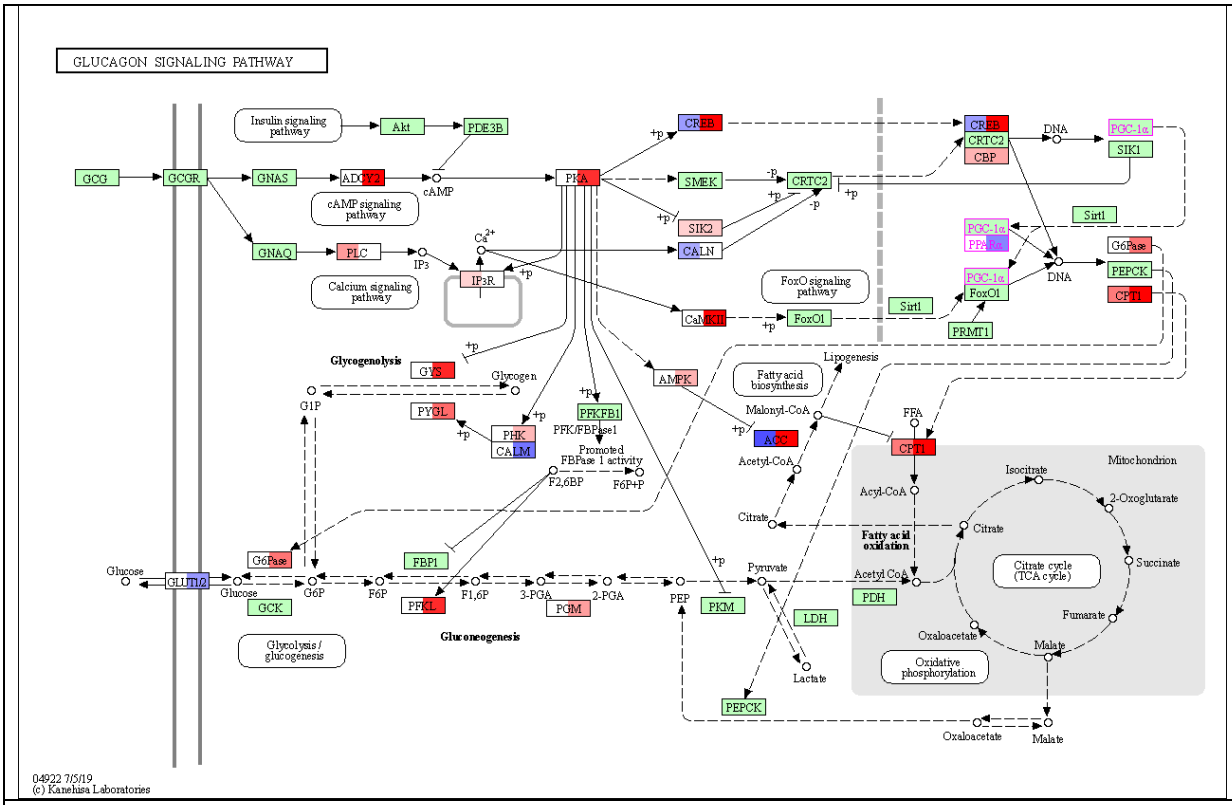
Supplementary figure 17 Selected KEGG pathways with liver differentially expressed genes – PPAR signaling. Blue is downregulated and red is upregulated (padj<0.1); left half of gene is GC_LAR vs FLT_LAR (29 days of spaceflight) and right half of gene is GC_ISS vs FLT_ISS (56 days of spaceflight). See *RR6 host transcriptomics* file for full DE gene list.



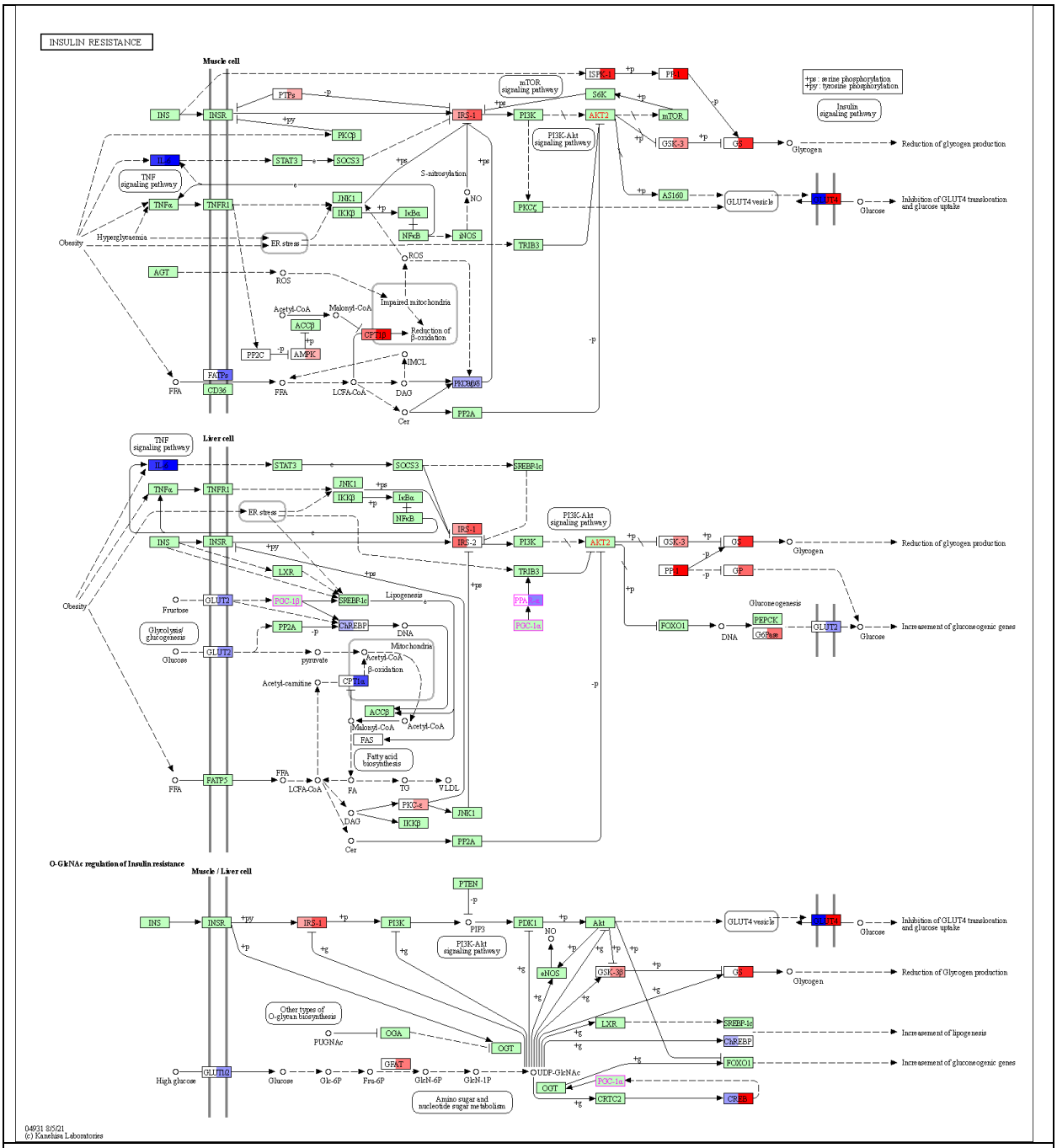
Supplementary figure 18 Selected KEGG pathways with liver differentially expressed genes – Bile secretion. Blue is downregulated and red is upregulated (padj<0.1); left half of gene is GC_LAR vs FLT_LAR (29 days of spaceflight) and right half of gene is GC_ISS vs FLT_ISS (56 days of spaceflight). See *RR6 host transcriptomics* file for full DE gene list.



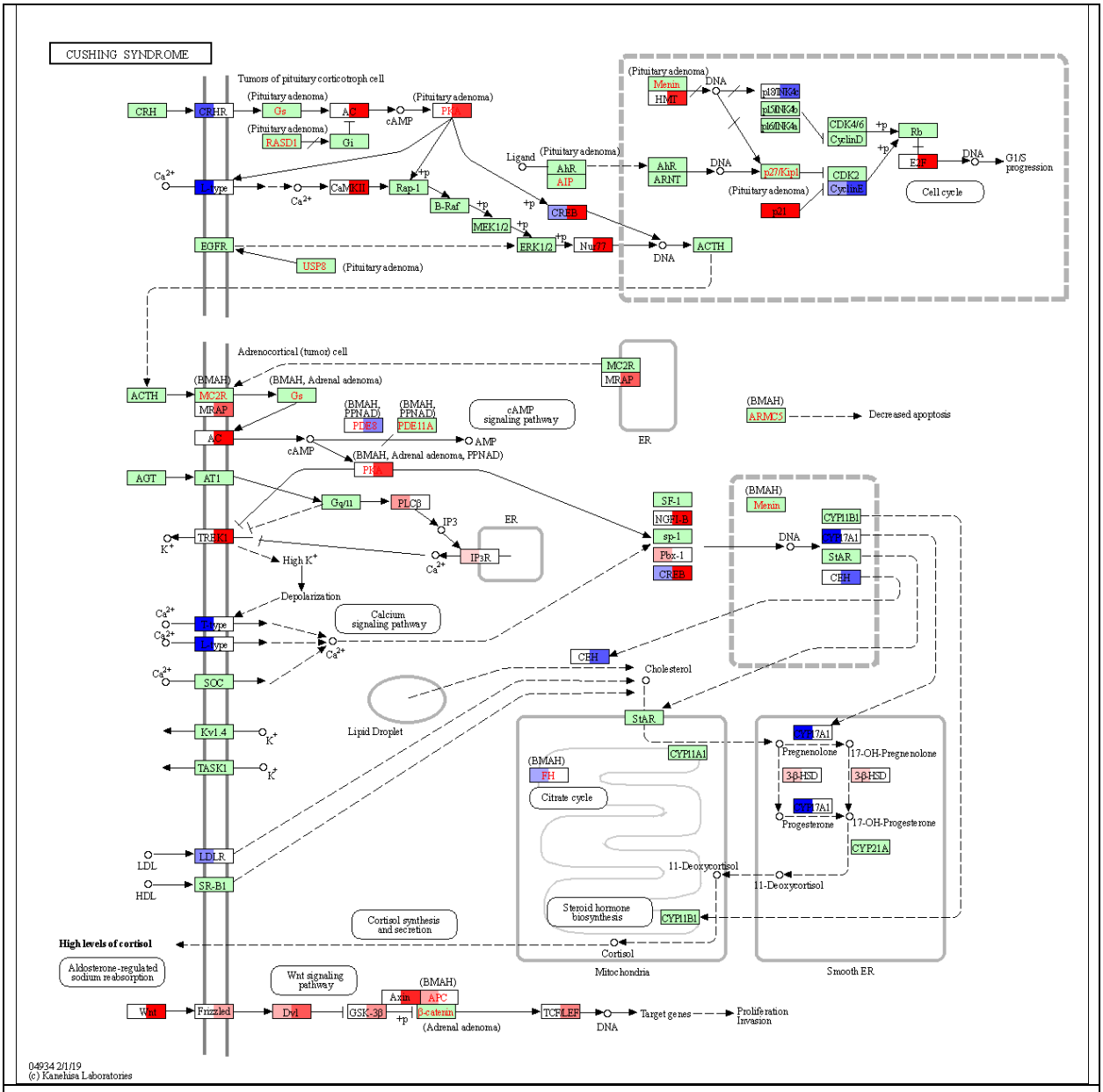
Supplementary figure 19 Selected KEGG pathways with liver differentially expressed genes – Insulin signaling. Blue is downregulated and red is upregulated (padj<0.1); left half of gene is GC_LAR vs FLT_LAR (29 days of spaceflight) and right half of gene is GC_ISS vs FLT_ISS (56 days of spaceflight). See *RR6 host transcriptomics* file for full DE gene list.



Supplementary figure 20 Selected KEGG pathways with liver differentially expressed genes – Glucagon signaling. Blue is downregulated and red is upregulated (padj<0.1); left half of gene is GC_LAR vs FLT_LAR (29 days of spaceflight) and right half of gene is GC_ISS vs FLT_ISS (56 days of spaceflight). See *RR6 host transcriptomics* file for full DE gene list.



Supplementary figure 21 Selected KEGG pathways with liver differentially expressed genes – Insulin resistance. Blue is downregulated and red is upregulated (padj<0.1); left half of gene is GC_LAR vs FLT_LAR (29 days of spaceflight) and right half of gene is GC_ISS vs FLT_ISS (56 days of spaceflight). See *RR6 host transcriptomics* file for full DE gene list.



Supplementary figure 22 Selected KEGG pathways with liver differentially expressed genes – Cushing syndrome. Blue is downregulated and red is upregulated (padj<0.1); left half of gene is GC_LAR vs FLT_LAR (29 days of spaceflight) and right half of gene is GC_ISS vs FLT_ISS (56 days of spaceflight). See *RR6 host transcriptomics* file for full DE gene list.

Supplementary Tables

Supplementary table 1 WMS functional pathways of interest	
KEGG Pathway	Description
52	Galactose metabolism
511	Other glycan degradation
531	Glycosaminoglycan degradation
520	Amino sugar and nucleotide sugar metabolism
500	Starch and sucrose metabolism
4512	ECM-receptor interaction
4350	TGF-beta signaling pathway
4976	Bile secretion
1212	Fatty acid metabolism
61	Fatty acid biosynthesis
51	Fructose and mannose metabolism
512	Mucin type O-glycan biosynthesis
511	Other glycan degradation
4973	Carbohydrate digestion and absorption
650	Butanoate metabolism
531	Glycosaminoglycan degradation
71	Fatty acid degradation
2025	Biofilm formation - Pseudomonas aeruginosa
121	Secondary bile acid biosynthesis
5100	Bacterial invasion of epithelial cells
1501	beta-Lactam resistance
1055	Biosynthesis of vancomycin group antibiotics
998	Biosynthesis of various antibiotics
4979	Cholesterol metabolism
515	Mannose type O-glycan biosynthesis
5130	Pathogenic Escherichia coli infection
ko02042	Bacterial toxins
ko01003	Glycosyltransferases
ko01504	Antimicrobial resistance genes
ko01011	Peptidoglycan biosynthesis and degradation proteins
ko00536	Glycosaminoglycan binding proteins
M00374	Dicarboxylate-hydroxybutyrate cycle
M00375	Hydroxypropionate-hydroxybutylate cycle
M00106	Conjugated bile acid biosynthesis; cholate => taurocholate_glycocholate
M00082	Fatty acid biosynthesis; initiation
M00083	Fatty acid biosynthesis; elongation
M00061	D-Glucuronate degradation; D-glucuronate => pyruvate + D-glyceraldehyde 3P
M00631	D-Galacturonate degradation
M00552	D-galactonate degradation; De Ley-Doudoroff pathway; D-galactonate => glycerate-3P
M00104	Bile acid biosynthesis; cholesterol => cholate_chenodeoxycholate
M00056	O-glycan biosynthesis; mucin type core
M00872	O-glycan biosynthesis; mannose type
M00019	Valine_isoleucine biosynthesis; pyruvate => valine _ 2-oxobutanoate => isoleucine
M00537	Xylene degradation; xylene => methylbenzoate