Microbial Community Structure and Composition is Associated with Host Species and Sex in Sigmodon Cotton Rats

Britton Strickland
Vanderbilt University
Mira Patel
Sigmovir Biosystems Inc
Meghan H. Shilts
Vanderbilt University Medical Center
Helen H. Boone
Vanderbilt University Medical Center
Arash Kamali
Sigmovir Biosystems Inc.
Wei Zhang
Sigmovir Biosystems Inc
Daniel Stylos
Sigmovir Biosystems Inc.
Marina S. Boukhvalova
Sigmovir Biosystems Inc.
Christian Rosas-Salazar
Vanderbilt University Medical Center
Shibu Yooseph
University of Central Florida
Seesandra V. Rajagopala
Vanderbilt University Medical Center
Jorge C. G. Blanco
Sigmovir Biosystems Inc.
Suman Ranjan Das (suman.r.das@vanderbilt.edu)
Vanderbilt University Medical Center https://orcid.org/0000-0003-2496-9724

Research Article

Keywords: microbiome, metagenomics, cotton rat, Sigmodon, 16S rRNA gene, gut, skin, respiratory

Posted Date: November 19th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-108250/v1

Version of Record: A version of this preprint was published at Animal Microbiome on April 16th, 2021. See the published version at https://doi.org/10.1186/s42523-021-00090-8.

Abstract

Background: The cotton rat (genus *Sigmodon*) is an essential small animal model for the study of human infectious disease and viral therapeutic development. However, the impact of the host microbiome on infection outcomes has not been explored in this model, partly due to the lack of a comprehensive characterization of microbial communities across different cotton rat species. Understanding the dynamics of their microbiome could significantly help to better understand its role during when modeling viral infections in this small animal model.

Results: We examined the bacterial communities of the gut and three external sites (skin, ear, and nose) of two inbred species of cotton rats commonly used in research (*S. hispidus* and *S. fulviventer*) by using 16S rRNA gene sequencing, constituting the first comprehensive catalog of the cotton rat microbiome. We showed that *S. fulviventer* maintained higher alpha diversity and richness than *S. hispidus* at external sites (skin, ear, nose), but there were no differentially abundant genera. However, *S. fulviventer* and *S. hispidus* had distinct fecal microbiomes composed of several significantly differentially abundant genera. Whole metagenomic shotgun sequencing of fecal samples identified species-level differences between *S. hispidus* and *S. fulviventer*, as well as different metabolic pathway functions as a result of differential host microbiome contributions. Furthermore, the microbiome composition of the external sites showed significant sex-based differences while fecal communities were not largely different.

Conclusions: Our study shows that host genetic background potentially exerts homeostatic pressures, resulting in distinct microbiomes for two different inbred cotton rat species. Because of the numerous studies that have uncovered strong relationships between host microbiome, viral infection outcomes, and immune responses, our findings represent a strong contribution for understanding the impact of different microbial communities on viral pathogenesis. Furthermore, we provide novel cotton rat microbiome data as a springboard to uncover the full therapeutic potential of the microbiome against viral infections.

Background

The commensal microbiome can dramatically influence many aspects of host health and disease, such as homeostatic signaling, nutrient acquisition, and protection from or exacerbation of infections (1–3). The majority of early studies established that environmental factors play the major role in shaping and modulating the host microbiome (4–7). In addition, recent studies have emphasized that there is a significant role of host genetics on co-evolution of the host and its associated microbiome (8–10). For example, the murine genetic background is a stronger determinant of microbiome composition and structure than environmental stimuli (11). Similarly, genetic polymorphisms, heritability, and overall host genetics in humans can also shape how commensal bacteria evolve alongside the host (12–14). The microbiome has also been instrumental in predicting and protecting against severe viral disease outcomes (15, 16). However, this burgeoning field of bacteria-host-virus interactions has been limited by a lack of translational models to study mechanisms of virus-microbiome interaction.

Cotton rats (genus *Sigmodon*) are an important small animal model to study various respiratory diseases, including Respiratory Syncytial Virus (RSV) (17), influenza A Virus (IAV) (18, 19), parainfluenza virus (20, 21), measles (22), human metapneumovirus (23), enterovirus (24), and rhinovirus (25) due to comparable human disease outcomes (26). Cotton rats have also provided a useful model for nasal colonization studies (especially with *Staphylococcus aureus*) due to their human-like nasal histology (27). Furthermore, cotton rats are a useful tool for research since they harbor zoonotic viruses in the wild like Alphavirus (Equine Encephalitis virus), Hantavirus (Black Creek Canal virus, Bayou virus), Cardiovirus, Arenavirus (Tamiami virus), and Flavivirus (West Nile virus) (28–33). However, viral-microbiota interactions in this model have not been explored due to the lack of a comprehensive analysis of the normal microbiota in healthy cotton rats and microbiome comparison between different species used in laboratory settings. To date, only one study has examined the nasal microflora of healthy *S. hispidus* but was limited by the sample number and lack of longitudinal timepoints (34).

To comprehensively characterize and establish the structure and composition of the cotton rat microbiome, we collected longitudinal samples from four different body sites of two commonly used inbred cotton rat species, *Sigmodon hispidus* and *S. fulviventer*, maintained under the same environment and dietary conditions. Our microbiome characterization using both 16S rRNA gene and whole metagenomic sequencing (WMS) comprehensively establishes the microbiome community structure and composition of different body sites in cotton rats and showed distinct structure based on the cotton rat species and sex. WMS also showed differential metabolic potential of the community between species. Overall, this study not only adds to the small but rapidly expanding literature of the influence of host genetics on the microbiome, but also describes an appropriate animal model for studying microbiome influences on viral and bacterial diseases.

Results

Characterization of cotton rat microbiome from multiple body sites.

Two groups of 10 young male cotton rats of *S. fulviventer and S. hispidus* were observed longitudinally for 111 days to characterize the healthy cotton rat microbiome structure and composition. A total of 140 samples were collected and processed for 16S rRNA gene sequencing: ear swabs (20 swabs/day 95), nasal brushes (20 swabs/day 95), skin swabs (20 swabs/day 95), and fecal samples (80 swabs/days 0, 4, 34, and 111) (Fig. 1A). DNA was extracted, and the V4 region of the 16S rRNA gene was amplified then sequenced on the Illumina MiSeq platform with 2 × 250 base pair reads, generating an average of 35,194 reads per sample with 96.4% of samples passing quality control (i.e., > 1000 reads/sample) for analysis. Microbiome data was processed by following the mothur SOP, and operational taxonomic units (OTUs) were clustered at 97% identity. We then examined the association of the alpha diversity, beta diversity, and abundance of taxa at each site using R version 3.5.1.

Differences in the microbiome community structure and composition between cotton rat species.

Phyla within *S. fulviventer* and *S. hispidus* external sites (i.e., ear, nose, skin) were similarly dominated by *Proteobacteria*, *Actinobacteria, Tenericutes*, and *Firmicutes*. Fecal communities consisted mostly of two dominant phyla with opposite abundances between cotton rat species: higher *Bacteroidetes* abundance in *S. fulviventer* compared to *S. hispidus* (50.0% vs. 42.4% respectively) and higher *Firmicutes* abundance in *S. hispidus* compared to S. *fulviventer* (36.2% vs. 31.8% respectively) (Table 1). The distribution of top 20 genera in each cotton rat species is shown in Figure S1.

We found that the cotton rat gut microbiome was stable over time in both cotton rat species. Richness and alpha diversity did not significantly change over time, no taxa were significantly differentially abundant when experimental day was set as the outcome variable, and beta diversity testing revealed no significant shifts in microbiome composition over time (data not shown). Subsequently, we analyzed groups by computing mean counts for individual cotton rats across time points. Comparison of individual body site microbiomes of both *S. fulviventer* and *S. hispidus* across all time points showed community distinctions between species. All sites from *S. fulviventer* consistently had higher richness (Observed OTUs, S.chao1 index) and alpha diversity (Shannon Index, Simpson's Index) when compared to *S. hispidus* (Fig. 1B and C, Figure S2; all values $p < 0.05$). Beta diversity, computed by calculating Bray-Curtis dissimilarity between samples at the OTU level, showed unique composition between the fecal communities of *S. fulviventer* and *S. hispidus* (Fig. 1D; PERMANOVA $p = 0.00025$, permutations = 4000, $R^2 = 0.26284$; beta dispersion, $p = 0.00025$, $R^2 = 0.34564$, Figure S3, S4D). However, comparison of beta diversity metrics of individual external sites from *S. fulviventer* and *S. hispidus* did not show significant differences (Fig. 1D; Figure S4A-C).

Analysis using the DESeq2(35) package identified several bacterial genera that were differentially abundant in the two cotton rat species. These differences were most apparent in the gut (Fig. 2A). *S. hispidus* had a higher abundance of 18 unique genera in the gut ($q < 0.05$), including *Lactobacillus* ($\log_2$ fold change = 3.12, q = 3.27E-13), *Helicobacter* ($\log_2$ fold change = 2.45, q = 2.33E-33), *Anaerostipes* ($\log_2$ fold change = 2.35, q = 0.029), and *Bifidobacterium* ($\log_2$ fold change = 1.99, q = 1.03E-06). *Escherichia/Shigella* was more abundant in the *S. hispidus* gut ($\log_2$ fold change = 7.30, q = 3.79E-11) but had very low relative abundance compared to other genera. *S. fulviventer* had a higher abundance of 18 unique genera in the gut ($q < 0.05$), including *Clostridium sensu stricto* ($\log_2$ fold change=-7.19, q = 7.77E-12), *Elusimicrobium* ($\log_2$ fold change=-6.22, q = 8.04E-18), and *Hespellia* ($\log_2$ fold change=-5.11, q = 2.21E-04) (Fig. 2A). Full data of differentially abundant taxa at both the genus and family levels are shown in Supplemental File 1. A total of 32 of the 36 DESeq2-calculated differentially abundant genera were also confirmed using the GeneSelector (36) R package (Figure S5, Supplemental File 2). A stability selection model showed *Lactobacillus* as one of the top genera (as well as those unclassified within the phyla *Bacteroidetes*) with a high probability of predicting whether a fecal sample was from *S. hispidus* or *S. fulviventer* (Fig. 2B). While *Lactobacillus* was one of the top 20 most abundant bacteria of the skin, ear,

and nose microbiomes of both *S. fulviventer* and *S. hispidus* (Figure S1), it was not significantly differentially abundant between the two species at any other sites except feces (Fig. 2C).

There were few specific taxa at both the genus and family levels that were significantly differentially abundant ($p < 0.05$, $q < 0.10$) between cotton rat species when examining external communities (ear, nose, skin). *Enterobacteriaceae* was more abundant in the *S. hispidus* nose (log_2 fold change = 1.21, q = 1.05E-06), ear (log_2 fold change = 0.48, q = 0.022), and skin (log_2 fold change = 0.47, q = 0.0031) compared to the equivalent sites in *S. fulviventer*. *Corynebacteriaceae* was more abundant in the *S. hispidus* nose (log_2 fold change = 0.532, q = 0.0305) and ear (log_2 fold change = 0.571, q = 0.0470) compared to the equivalent sites in *S. fulviventer*. *Leptotrichiaceae* was more abundant in the *S. fulviventer* nose (log_2 fold change=-1.41, q = 0.0456) and ear (log_2 fold change=-1.41, q = 0.047) compared to the equivalent sites in *S. hispidus*. *Pseudomonas* was more abundant in the *S. hispidus* ear (log_2 fold change = 1.33, q = 0.070) than in *S. fulviventer*. *Barnesiella* and *Porphyromonadaceae* were more abundant in the *S. fulviventer* ear (log_2 fold change=-4.58, q = 0.066; log_2 fold change=-2.11, q = 0.063). *Leuconostocaceae* was more abundant in the *S. fulviventer* nose (log_2 fold change=-1.56, q = 0.064). Full data at both genus and family levels are shown in Supplementary File 1.

Confirmation of 16S rRNA gene sequencing data using traditional culture methods.

For quantitative comparison of bacterial load between cotton rat species, we used qPCR analysis of total bacterial DNA extracted from homogenized stool (equal weight/volume) and found that the bacterial load was significantly higher in *S. hispidus* than *S. fulviventer* (Fig. 2C). We also plated an aliquot of normalized, homogenized stool on *Lactobacillus*-specific agar (De Man, Rogosa and Sharpe agar) and observed that the number of colony-forming units (CFU) per gram in *S. hispidus* stool was significantly higher than in *S. fulviventer* stool (Fig. 2D). We found that 86% of colonies grown from *S. hispidus* stool were *Lactobacillus*-positive, compared to zero *Lactobacillus*-positive colonies grown from *S. fulviventer* stool (Fig. 2E). Sequencing of colonies showed *Lactobacillus gasseri* and *Lactobacillus. reuteri* to be the two prominent bacterial species found in *S. hispidus* stool (Fig. 2E). This significant trend supports the relative abundance of *Lactobacillus* as determined by 16S rRNA gene sequencing, where *Lactobacillus* was significantly more abundant in *S. hispidus* compared to *S. fulviventer* (Fig. 2F). Additionally, *Corynebacterium* and *Bacteroides* species were also identified in *S. hispidus* stool samples. Sanger sequencing of isolates from *S. fulviventer* stool identified the presence of *Enterococcus gallinarum* and *E. casselifavus*.

Differences in the microbiome community structure and composition based on sex.

We conducted a follow-up experiment to assess if there were any associations between host sex and microbiome community structure and composition. This cohort included 13 *S. fulviventer* cotton rats (10 males, 3 females) and 16 *S. hispidus* cotton rats (5 males, 9 females). Animals in both groups were 4–6 weeks old and weighed approximately 100 g and were observed for 28 days, with fecal samples collected on days 0, 7, 13, 21, and 28 and nose, ear, and skin swabs collected on days 7 and 28. We performed 16S rRNA gene sequencing to examine any effect of host sex. Alpha diversity metrics indicated significant differences in richness (Observed OTUs, Chao1) and diversity (Shannon Index, Simpson Index) between male and female *S. fulviventer* at both the ear and fecal microbiomes (Figure S6A-D), but there were no significant differences in richness and diversity of the microbiomes of male and female cotton rats in the nose and skin for both *S. fulviventer* and *S. hispidus* (with the exception of *S. hispidus* skin diversity, Figure S6D). Overall, differences between host sex were most pronounced in the gut compared to external sites (Figure S6) but only in *S. fulviventer*. Beta-diversity measurements of each species revealed that microbial composition of the gut was significantly dissimilar between male and female cotton rats for both *S. fulviventer* and *S. hispidus* (Fig. 3A, B; *S. hispidus* PERMANOVA p = 0.00025, beta-dispersion p = 0.1116; *S. fulviventer* PERMANOVA p = 0.00025, beta-dispersion p = 7E-04). There were also notable differences between sex at *S. hispidus* skin and nose (Fig. 3E, G). Differential abundance analysis using DESeq2 was conducted between males and females at each site (Table S1). While the fecal community structure differed, there were only 3 differentially abundant genera due to sex in the *S. hispidus* gut and no different genera in *S. fulviventer*. There were differential taxa between sexes at external sites of both *S. hispidus* (21 genera) and *S. fulviventer* (13 genera). Full results at genus and family levels are listed in Supplementary File 3.

Differences in the microbiome between cotton rat species assessed by whole metagenomic sequencing.

Due to the dramatic differences between the *S. fulviventer* and *S. hispidus* gut microbiomes detected by 16S rRNA gene sequencing, we pursued further analysis in order to understand community differences at the species and strain level as well as differences in microbiome functional potential. DNA extracted from 10 male cotton rat stool samples (*S. hispidus* = 5, *S. fulviventer* = 5) at both days 34 and 111 (20 samples total) from the first experimental group were processed for shotgun metagenomic sequencing, which generated 1.11×10^9 reads (5.57×10^7 average reads per sample), comprising of 334,037 megabases (16,701 average megabases per sample) and 17.2% duplicate reads.

Whole metagenomic sequencing data showed differences at the species level that validated the 16S rRNA gene sequencing data. Abundances of several bacterial species were found to be statistically different ($q < 0.05$) between cotton rat species based on taxonomic classification as performed by MetaPhlAn2 followed by differential abundance analysis by both hierarchical clustering (based on Bray-Curtis dissimilarity) of the top 25 most abundant species (Fig. 4A) and DESeq2 (Supplemental File 4). *Lactobacillus reuteri*, *L. gasseri*, and the novel *L.* sp. ASF360 predominated the gut of *S. hipsidus* (Fig. 4A), and many other *Lactobacillus* species were significantly more abundant in *S. hispidus* samples

compared to *S. fulviventer* (Figure S7). Total *Lactobacillus* within the *S. fulviventer* gut was significantly less abundant but included species unique to *S. fulviventer*, including *L. murinus*, *L.* BHWM-4, and *L. animalis*. *Akkermansia muciniphilia* was significantly more abundant in *S. fulviventer* compared to *S. hispidus*. *Ruminococcus torques, Helicobacter cinaedi*, and *Oscillibacter* spp. were of higher abundance in the *S. hispidus* gut. *Parabacteroides* spp. (including *P. johnsonii*) and *Odoribacter* were more abundant in the *S. fulviventer* gut (Supplementary File 4). Proportional counts and raw counts can be found in Supplementary Files 5 and 6 respectively.

Differential functional potential between cotton rat species microbiome.

To understand the biological implications of these differences, Humann2 (37) was used to map any functional differences (MetaCyc pathway database (38)) defined by identified gene families and bacterial profiles. We identified 415 pathways (with nearly all associated with bacteria present in the sample) in the two cotton rat species based on the MetaCyc database (Fig. 4B). The majority of these pathways included biosynthesis (39.60%) and degradation/utilization/assimilation (18.79%) pathways, as well as several overarching superpathways (27.18%) and energy/metabolite production pathways (10.57%). More specifically, these pathways were part of several instrumental superclass ontologies that metabolize (including *de novo* pathways) electron carriers, vitamins, fatty acids, lipids, amino acids, carbohydrates, secondary metabolites, and fermentation-derived energy (Fig. 4C). Interestingly, several pathways were differentially abundant between cotton rat species. Each cotton rat species had unique pathways contributed to by their microbiomes (*S. fulviventer* = 14, and *S. hispidus* = 27), and most of these involved biosynthesis (Supplemental File 7).

In relation to differentially abundant bacteria species, we found that 44 pathways were solely driven by *Lactobacillus gasseri, L. reuteri*, and L. *ASF360* by matching reads from MetaPhlAn2 bacterial identifications with Humann2 predicted pathways. Several of these pathways were more highly expressed in *S. hispidus* (Fig. 5). These included L-proline biosynthesis from arginine (catalyzed by bacterial enzymes), inosine-5'-phosphate biosynthesis (for *de novo* synthesis of purines), pyruvate fermentation to acetate/lactate (for anaerobic energy production), adenosine deoxyribonuclease *de novo* biosynthesis (to promote ADP production), and D-galactose degradation (breakdown of D-galactose to a useable form in glycolysis). *Akkermansia municiphilia* was the driver of 25 other pathways, many of which were highly expressed in *S. fulviventer* compared to *S. hispidus* (Figure S8). These included L-isoleucine biosynthesis (for production of leucine and isoleucine), phosphopantothenate biosynthesis (to produce vitamin B5 *de novo*, of which animals cannot produce, and to feed production of coenzyme A and acyl carrier protein), glycolysis (particularly the degradation of starches for reductants and energy for anabolic pathways), and L-valine biosynthesis. Statistical comparison of all pathways can be found in Supplemental File 8.

Discussion

Here we have comprehensively characterized the cotton rat microbiome and compared bacterial communities of two different species (*S. hispidus* and *S. fulviventer*) that were housed in the same facility with identical diets. From these analyses, we were able to uncover species-specific differences of gut bacteria, even though the two species had the same diet over many generations and were housed in separate cages in the same room. Interestingly, their external microbiomes (ear, nose, skin) were remarkably similar based on beta-diversity testing and testing for differential bacterial taxa but significantly different based on alpha diversity and richness measurements (which mimicked that of fecal communities). This data further supports that, while environmental factors play a vital role in shaping microbiome structure and composition, underlying host genetics exerts homeostatic pressure for distinct microbiomes between populations.

The most recent phylogenetic analysis of the genus *Sigmodon sp.* found that *S. hispidus* and *S. fulviventer* diverged 5.4 million years ago (39). In the wild, *S. hispidus* and *S. fulviventer* are sympatric species, with *S. fulviventer* being the more dominant animal (40). Separate inbreeding of the two species has made them a useful small animal model in laboratory research (26, 41). Our data show that even when inbred and adapted to a controlled laboratory environment, each species still maintains a unique gut microbiome community structure and composition. This is important to note because the cotton rat is a useful model for respiratory viral infection, pathogenesis, and immunity research (26). Additionally, cotton rats are an excellent small animal model for pre-clinical studies for antiviral and vaccine candidates for viruses e.g., RSV (17), IAV (18, 19), and human rhinovirus (HRV) (26). Recent studies have shown that gut microbiome composition and disruptions by antibiotics may affect some immune responses to flu vaccination (42, 43). Similarly, there are important correlations between vaccination status and the microbiome of the gut and lung (44, 45). Further, it has also been shown that transplantation of the gut microbiome of wild-caught mice to laboratory mice can promote host fitness and improve disease resistance and pathogenesis (46). In light of the above literature suggesting that the microbiome may play a key role in respiratory viral disease exacerbation or remediation and vaccine response (42–45, 47), our findings exemplify the usefulness of this small animal model for understanding viral pathogenesis in the context of different microbial communities. Furthermore, the cotton rat model could be an optimal subject to uncovering and examining the full therapeutic potential of the microbiome by understanding how the host regulates and modulates bacterial communities.

While other small animal models including mice and rats have been used to explore the relationship between genetic patterns and bacterial homeostasis (11, 48–50), the cotton rat could be used in studying the interplay between differing host microbiomes and host immune responses to viral infections. For example, *S. hispidus* had a significantly higher amount of probiotic gut bacteria genera (*Lactobacillus*, *Bifidobacterium*) that have been associated with protection against the severe outcomes from RSV, IAV, and HRV (51–53). Other differentially abundant bacteria between the two species, such as *Escherichia coli* and *Bacillus cereus*, have also been shown to enhance both poliovirus and reovirus replication and pathogenesis (54). With the presence/absence of these bacteria in this animal model, along with our recently published annotated transcriptome (55), host responses in light of the microbiome could be

elucidated. The cotton rat could also be an optimal model for supplementation studies of particular taxa in relation to these viruses.

Here, we have not only comprehensively characterized and established the key differences in microbiome community structure and function between multiple sites from two species of cotton rats housed in same facility for generations and fed with same diet, we also uncovered sex as a factor that can impact the microbiome composition of the gut. Our study has significant strengths compared to the lone study published so far (34), including a large sample size, longitudinal sampling, and shotgun metagenomic sequencing to characterize the gut microbiome at the species level. However, we acknowledge several shortcomings: 1) Due to the lack of an assembled cotton rat genome, we could not examine host genes, genetic patterns, or polymorphisms that may be driving differences in microbial colonization. 2) In both cohorts of cotton rats, samples were taken at different time points, and the cotton rats were followed for different durations. However, we found no evidence of changes in the gut microbiome across 111 days of sampling our first cohort of adult cotton rats. 3) We only conducted shotgun metagenomic sequencing on the gut microbiome from male cotton rats from our first cohort. While there is no cotton rat genome in the public databases, with the *de novo* assembly of the cotton rat transcriptome (55), further studies integrating the microbiome data and gene expression patterns may uncover more relevant information in regards to differences in host and microbiome interactions. We would also like to note that characterizing the microbiome of a unique animal species can be challenging due to the lack of host and/or microbiome sequence databases. The majority of databases and tools are that are commonly available are specifically designed for human microbiome analyses and can result in misclassification of bacteria when used for new animal species. However, until we have better databases, interpretation of data has to be done with caution. Further research is warranted to understand species level microbiome differences and their impact on immune response in all small animal models for better interpretation of preclinical studies of vaccines and anti-microbiological agents. In spite of some limitations, our study creates a steppingstone for future research into these pressing questions of host-microbiome interactions during infection.

Conclusion

Overall, we have comprehensively characterized the cotton rat microbiome, an invaluable small animal model for viral and bacterial infections, and established key differences in microbiome community structure and function between multiple sites from two species of cotton rats (*S. hispidus* and *S. fulviventer*) housed in same facility for generations and fed with same diet. We also uncovered sex as a variable that can impact the microbiome composition of the gut. This foundational study establishes a foundation for future hypothesis testing experiments in understanding the role of microbiome in viral pathogenesis, especially for RSV and Influenza virus. Additionally, this study adds to the small but expanding literature in understanding the role of host genetics on microbiome composition and structure.

Methods

Animals

Four- to six-week-old cotton rats (~ 100 g) were obtained from the inbred colony maintained at Sigmovir Biosystems, Inc. (SBI). Cotton rats in the colony were seronegative by ELISA to adventitious respiratory viruses (i.e. Pneumonia Virus of Mice, Rat parvovirus, Rat coronavirus, Sendai virus). Animals were individually housed in large polycarbonate cages and fed a diet of standard rodent chow and water *ad libitum*.

For rigor and reproducibility, two independent animal experiments were carried out to characterize and establish the healthy cotton rat microbiome structure and composition by comparing two different species, *S. fulviventer and* S. *hispidus.* In the first experimental group, 20 young male cotton rats were examined: *S. fulviventer* (n = 10) and *S. hispidus* (n = 10). Each animal was observed for 111 days, with nose, ear, and skin swabs collected at day 95 and fecal samples collected at days 0, 4, 34, and 111. These samples were used for microbiome characterization. To analyze any sex bias to the microbiome, a second experimental group (at a later time) included 13 young *S. fulviventer* (10 males, 3 females) and 16 young *S. hispidus* (5 males, 9 females). Healthy animals were monitored for 28 days with fecal samples collected on days 0, 7, 13, 21, and 28 and nose, ear, and skin swabs collected on days 7 and 28. To avoid fighting, all the animals were housed individually in large polycarbonate cages (with proper enrichment; nylon bone and glass jar). The cotton rat colony was maintained free of human and rodent viruses. All animal procedures followed NIH and USDA guidelines and were approved by the Sigmovir Biosystems, Inc. IACUC.

Sample Collection

Feces collection: One day prior to feces collection, cage beddings were changed in the late afternoon for each animal. Samples were collected between 10 am and 1 pm with sterile forceps. On average, 10–15 feces pellets were collected from each animal. Immediately after collection, samples were frozen at − 80 °C.

Nose swab: Sterile saline (~ 100 µl) was pipetted into both nostrils of anesthetized cotton rats positioned face down; Fisherbrand Sterile Swabs (Calcium Alginate Fiber Tipped Ultrafine Aluminum Applicator Swab) were then immediately placed in the nostrils to absorb the saline. Swabs were broken into sterile DNase/RNase-free 1.5 ml tubes and stored at − 80 °C.

Ear swab: Sterile saline (~ 100 µl) was pipetted up and down into both ears of each anesthetized cotton rat while the animal was kept in an anesthesia chamber for 1–2 minutes, and residual liquid was absorbed from each ear with Beaver Visitec Ultracell PVA Eye Spears pack of 5 (intended for fluid absorption and tissue manipulation). Swabs were broken into sterile DNase/RNase-free 1.5 ml tubes and stored at − 80 °C.

Skin swab: Sterile saline (~ 200 µl) was put at the back of each anesthetized cotton rat (at approximately the same site for each animal) and rubbed vigorously using Fisherbrand Sterile Swabs (Calcium Alginate

Fiber Tipped wood applicator swab). Swabs were broken in a sterile DNase/RNase-free 1.5 ml tubes and stored at − 80 °C.

Microbiome DNA extraction and 16S rRNA gene sequencing

Genomic DNA was extracted from all samples at Vanderbilt University Medical Center using the Qiagen DNeasy PowerSoil HTP Kit (96-well plates) following the manufacturer's protocol, except the optional 4 °C incubations were skipped. Stool samples were thawed on ice and added directly to the kit plate. Nose, ear, and skin swabs were vortexed in tubes with 800 µL Qiagen PowerBead solution for 5 min; this PowerBead solution was then added to the kit plate. An extraction negative, which did not contain any template but was otherwise processed the same as the rest of the samples, was included on each extraction plate. To mechanically lyse the cells, plates were shaken at 20 Hz in a TissueLyser II system (Qiagen) for 20 minutes. Steps 16–33 of the kit manufacturer's protocol were performed on a QIAcube HT (Qiagen). One-step PCR targeting the V4 region of the 16S rRNA gene was performed using 515F/806R primers (56). MyTaq HS Mix (Bioline) was used to create amplicons, with the following cycling conditions: 95 °C for 2 min; 30 cycles of 95 °C for 20 sec, 50 °C for 15 sec, 72 °C for 5 min; 72 °C for 10 min; 4 °C indefinitely. Positive PCR results were confirmed by the presence of a 400 bp band in 1% agarose gel electrophoresis; all negative controls were verified at this step to not have a visible band. The PCR products were cleaned and normalized using the SequalPrep Normalization Kit (Invitrogen). Samples and complementary controls (extraction negative, PCR negative, and ZymoBIOMICS Microbial Community Standard) were pooled and then cleaned using 1X AMPure XP beads. Sequencing was done on an Illumina MiSeq platform with 2 × 250 bp reads at the Vanderbilt Technologies for Applied Genomics (VANTAGE) core facility.

16S rRNA Gene Data Processing and Statistical Analysis

After sequencing, reads were processed using the mothur SOP (https://mothur.org/wiki/miseq_sop/) (57). Operational taxonomic units (OTUs) were clustered at 97% sequence identity. Non-bacterial sequences, low-quality sequences (1.5% of total reads), and chimeras as identified with UCHIME (58) were removed during data processing. Sequences were aligned to the SILVA database release 128 (59) and taxonomic assignment was made using the Ribosomal Database Project (RDP) database 14 (60). Samples with < 1000 final reads (n = 5) were removed prior to analysis. Statistical analyses were performed using MGSAT [https://bitbucket.org/andreyto/mgsat] (15, 53), which facilitates data analysis by wrapping the R packages as described below.

Alpha- and beta-diversity analyses were performed using the R package *vegan* (61). Prior to alpha- and beta-diversity analysis, counts were rarefied to the lowest library size, and richness, alpha-, and beta-estimates were calculated. This process was repeated 400 times, and the results were averaged. Richness was estimated with the observed OTUs and Chao1 indices; alpha diversity was estimated with the Shannon and Simpson indices, which were converted into their corresponding Hill numbers (62). Statistical testing between site alpha diversity was calculated using Mann-Whitney U or Kruskal-Wallace/Dunn's Post Hoc test where applicable. For beta-diversity analysis, counts were normalized to

simple proportions, and pairwise Bray-Curtis dissimilarities were estimated. The PermANOVA (permutation-based analysis of variance) test as implemented in the *Adonis* function from the R package *vegan* was used to test for significance between overall microbial composition and groups of interest (i.e., *S. hispidus* compared to *S. fulviventer* and males compared to females) over 4000 permutations; results are indicated by “centroid” p-values. Homogeneity of variance within sample groups was tested using *betadisper* function; results are indicated by “dispersion” p-values.

Differential abundance of taxa in association with metadata categories was analyzed using DESeq2 (35). DESeq2 implements a model with raw absolute counts of each taxon with a negative binomial distribution and uses the estimated depth of sequencing of each sample to scale the (unknown) relative abundance that is the parameter of the negative binomial distribution. Compared with using either simple proportion-based normalization or rarefaction for controlling for differential sequencing depth, the DESeq2 approach provides improved sensitivity and specificity (63). Prior to DESeq2 analysis, we eliminated all taxa that were had an average number of < 10 reads, taxa with a minimum quantile mean fraction < 0.25, and taxa with a minimum quantile incidence fraction < 0.25; taxa with a normalized base mean (generated by DESeq2) < 10 were removed. Reported adjusted P values (q) values are the result of a Wald test with the Benjamini and Hochberg correction for multiple comparisons. To build alternative rankings of taxa in regard to their prevalence in one cotton rat species over the other, we also used stabsel and GeneSelector. The stabsel stability selection (64) approach aims to build the relative ranking of the predictor variables (taxa in this case) according to their importance for predicting the outcome. It does so by building multiple “base” models on random subsamples of the data. The elastic net model from the R package glmnet was used as the base feature selection method to be wrapped by the stability protocol. The ranking of taxa and their probability of being selected into the model were reported, as well as the probability cutoff corresponding to the per-family error rate that is controlled by this method. The GeneSelector package (36) was used as a stability feature ranking method that is based on a nonparametric univariate test. In brief, the same ranking method (package function RankingWilcoxon) was applied to multiple random subsamples of the full set of observations (400 replicates, sampling 50% of observations without replacement). RankingWilcoxon ranks features in each replicate according to the test statistic from Wilcoxon rank-sum test with regard to the outcome group (e.g. *S. hispidus* vs. *S. fulviventer*). Consensus ranking between replicates was then found with a Monte Carlo procedure (package function AggregateMC) and the features were reported in the order of that consensus. To account for different sequencing depth, the absolute abundance counts were normalized to simple proportions within each observation. For each feature, we also obtained several types of the effect size, such as common language effect size and rank biserial correlation. These 3 statistical analyses (DESeq2, stabsel, and GeneSelector) allowed for rigorous testing of each particular taxon of interest.

Metagenomic Library Preparation

A subset of fecal samples from 20 total male cotton rats (10 from each species), taken at days 34 and 111 within the first cohort of cotton rats, underwent whole-metagenomic shotgun sequencing. From the same stool samples, genomic DNA was extracted using the Qiagen DNeasy PowerSoil Kit (Cat No./ID:

12888-100) by following the manufacturer's protocol (skipping the optional 4ºC incubations). In addition, a negative sample (which did not contain any template but was otherwise processed the same as the rest of the samples) and a positive control (ZymoBIOMICS Microbial Community Standard) were processed in parallel with samples and sequenced. Samples were normalized to 75 ng/ µL in 1X TE prior to library construction. Metagenomic libraries were prepared using the NEBNext® Ultra™ II FS DNA Library Prep Kit for Illumina® following the manufacturer's protocol for inputs ≤ 100 ng. Samples were fragmented at 37ºC for 12 minutes to yield a fragment size of 200–450 bp. NEBNext Multiplex Adaptors were diluted 10-fold. NEBNext Multiplex Oligos for Illumina (Set 1, NEB #E7335) were used for PCR enrichment of adaptor-ligated DNA, and 5 cycles of PCR were run. Library quality was assessed on an Agilent 2100 Bioanalyzer System using the Agilent High Sensitivity DNA Kit (5067 – 4626). Samples were sequenced via the NovaSeq 6000 2 × 150 platform for Illumina at the Vanderbilt Technologies for Advanced Genomics (VANTAGE) core, aiming for 40 million reads per sample.

Whole Metagenomic Shotgun Sequence Analysis

FastQC (65) followed by MultiQC (66) were used to examine data quality. Trimmomatic (67) was used to remove adaptors and trim low quality reads using the parameters: TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:75. Microbial communities were then profiled using MetaPhlAn2 (68). Differentially abundant bacteria were calculated using MetaPhlAn2's *hclust2.py* function by hierarchical clustering (based on Bray-Curtis dissimilarity) of the top 25 most abundant species according to the 90th percentile of the abundance in each clade as well as DESeq2. Functional, metabolic profiles were analyzed using Humann2, which aligns reads from UniRef (69) cluster abundances to the ChocoPhlAn (37) database. This generates three outputs: UniRef IDs for gene families in reads per million, MetaCyc pathway coverage, and MetaCyc pathway abundances in copies per million (Supplemental File 8). To identify differential pathways between sample groups, associations between cotton rat species were identified by the Humann2.associate script and statistical testing using the Kruskal-Wallis H-test. Data presented (generated by Humann2.barplot script) is from pathway abundances (normalized as relative abundance) within each sample with unmapped/unintegrated pathways removed and was found statistically significant ($p < 0.05$ and $q < 0.05$). Superclasses distribution of identified MetaCyc pathways was manually generated using the online MetaCyc database.

Enumeration of Lactobacillus

Two frozen stool pellets were taken from 20 male cotton rats (10 *S. hispidus*, 10 *S. fulviventer*), weighed, and diluted to 45 mg/mL in sterile 1x PBS. Samples were rocked for 20 min on ice and resuspended manually with pipette mixing. 10^{-1}-10^{-3} serial dilutions were plated on *Lactobacilli* MRS agar (BD 288210) and incubated at 37ºC for 48 h. Colonies on 10^{-2} were counted, and 95 colonies grown from stool from *S. hispidus* and *S. fulviventer* each were randomly picked and inoculated into 1.2 mL MRS broth (BD 288130) in a sterile 96-deep-well plate. The plate was incubated at 37ºC for 20 h with no shaking. Cultures were gently mixed by pipetting, and 20% glycerol stocks were prepared for each culture.

Colony PCR was performed on each isolate by boiling the culture at 95ºC for 10 minutes, then using 10 µL as the template with *Lactobacillus* species-specific primers (70) and MyTaq HS Red (Bioline®) with the following cycling conditions: 95 °C for 2 min; 30 cycles of 95 °C for 20 sec, 50 °C for 15 sec, 72 °C for 1 min; 72 °C for 10 min; 4 °C indefinitely. PCR reactions were spun at 3900 g for 10 min to remove any bacterial debris from the boiled template and run on 1% agarose gel to verify *Lactobacillus*-positive colonies. A new PCR was then repeated using the universal primers Uni331F/Uni797R (71) (following cycling conditions listed above), and purified PCR products were sent for Sanger sequencing. Bacterial isolate identity was determined using NCBI BLAST database.

Determination of bacterial load by qPCR

DNA was extracted from an equal volume of normalized homogenates of cotton rat stool (described in Methods: Enumeration of *Lactobacillus*) using the DNeasy PowerSoil Kit (Qiagen). qPCR reactions were prepared in duplicate using BioRad iQ Supermix with Invitrogen Sybr Green following the manufacture's protocol. Universal eubacteria 16S rRNA primers (UniF340, UniR514) (72) equal volumes of extracted DNA, and targeted standards were used to determine copy number per gram of feces. Each qPCR plate included a corresponding extraction negative and a no-template negative control. A serial dilution of standards containing known bacterial copy numbers specific to the primer pair were used as a standard curve as previously described. PCR reactions were run with a 15 sec 95ºC melting and 1 min 54ºC annealing step for 40 cycles. Cycle threshold (CT) values were plotted against the standard curve to determine copy number, and figures and statistical testing (unpaired T test) were generated using Prism version 8.

Isolation and culture of Lactobacillus strains from cotton rats' stool

Glycerol stocks of identified *Lactobacillus* species were streaked on MRS agar plates, and a single colony was grown in culture using MRS broth. To determine growth parameters, each species was incubated at 37ºC without shaking, and growth efficiency was measured by turbidity. A growth curve was also estimated using a BioTek Synergy HTX plate reader at 37ºC for 24 h; OD_{600} was measured every 10 min following a brief 3 second shake to mix culture. CFU counts were also taken during the log phase by plating a 3-fold serial dilution on MRS agar plates.

Declarations

- # ETHICS APPROVAL AND CONSENT TO PARTICIPATE

No human subjects participated in this study. All animal procedures followed NIH and USDA guidelines and were approved by the Sigmovir Biosystems, Inc. IACUC.

- # CONSENT FOR PUBLICATION

Not Applicable

AVAILABILITY OF DATA AND MATERIALS

The sequencing data will be deposited at the NCBI Short Read Archive (SRA) upon manuscript acceptance. Additionally, please contact corresponding authors for data requests.

COMPETING INTERESTS

JCGB, MCP, AK, WZ, DS, and MSB are employed by Sigmovir Biosystems, Inc. Other authors have no competing interests.

FUNDING

This work was supported by start-up funds from Vanderbilt University Medical Center awarded to SRD and funds from the National Institute of Allergy and Infectious Diseases (under award numbers 1R21AI149262 and 1R21AI154016). SRD is also supported by 1R21AI142321, R21AI142321-01A1S1, U19AI095227, and U19AI110819), and the National Heart, Lung and Blood Institute (1R01HL146401) the Edward P. Evans Foundation, the Vanderbilt Institute for Clinical and Translational Research (grant support from the National Center for Advancing Translational Sciences under award number UL1TR000445), Vanderbilt Technologies for Advanced Genomics Core (grant support from the National Institutes of Health under award numbers UL1RR024975, P30CA68485, P30EY08126, and G20RR030956) and Sigmovir Biosystems Inc's corporate funds.

AUTHORS' CONTRIBUTIONS

SRD, JCGB, BAS conceived the idea. MCP, AK, WZ, DS, and MSB maintained cotton rat colonies and collected samples under the supervision of JCGB. BAS and HHB performed DNA extraction and 16S rRNA gene sequencing library preparation. BAS performed whole metagenomic sequencing library preparation, bacterial cultivation, and qPCR. BAS, MHS, SVR performed sequencing data analysis. MHS, SY, SVR, SRD were consulted during data analysis. CRS assisted in manuscript preparation and data presentation oversight. BAS wrote the manuscript with inputs from SRD, MHS and CRS. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

We thank Dr. Stokes Peebles Jr. for careful reading of the manuscript.

References

1. Unger SA, Bogaert D. 2017. The respiratory microbiome and respiratory infections. J Infect 74 Suppl 1:S84-S88.

2. Atherton JC, Blaser MJ. 2009. Coadaptation of Helicobacter pylori and humans: ancient history, modern implications. J Clin Invest 119:2475-87.
3. Hooper LV, Falk PG, Gordon JI. 2000. Analyzing the molecular foundations of commensalism in the mouse intestine. Curr Opin Microbiol 3:79-85.
4. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling AV, Devlin AS, Varma Y, Fischbach MA, Biddinger SB, Dutton RJ, Turnbaugh PJ. 2014. Diet rapidly and reproducibly alters the human gut microbiome. Nature 505:559-63.
5. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R, Gordon JI. 2008. Evolution of mammals and their gut microbes. Science 320:1647-51.
6. Fierer N, Hamady M, Lauber CL, Knight R. 2008. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. Proc Natl Acad Sci U S A 105:17994-9.
7. Jernberg C, Lofmark S, Edlund C, Jansson JK. 2007. Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. ISME J 1:56-66.
8. Limborg MT, Heeb P. 2018. Special Issue: Coevolution of Hosts and Their Microbiome. Genes (Basel) 9.
9. Foster KR, Schluter J, Coyte KZ, Rakoff-Nahoum S. 2017. The evolution of the host microbiome as an ecosystem on a leash. Nature 548:43-51.
10. O'Brien PA, Webster NS, Miller DJ, Bourne DG. 2019. Host-Microbe Coevolution: Applying Evidence from Model Systems to Complex Marine Invertebrate Holobionts. mBio 10.
11. Korach-Rechtman H, Freilich S, Gerassy-Vainberg S, Buhnik-Rosenblau K, Danin-Poleg Y, Bar H, Kashi Y. 2019. Murine Genetic Background Has a Stronger Impact on the Composition of the Gut Microbiota than Maternal Inoculation or Exposure to Unlike Exogenous Microbiota. Appl Environ Microbiol 85.
12. Turpin W, Espin-Garcia O, Xu W, Silverberg MS, Kevans D, Smith MI, Guttman DS, Griffiths A, Panaccione R, Otley A, Xu L, Shestopaloff K, Moreno-Hagelsieb G, Consortium GEMPR, Paterson AD, Croitoru K. 2016. Association of host genome with intestinal microbial composition in a large healthy cohort. Nat Genet 48:1413-1417.
13. Goodrich JK, Davenport ER, Beaumont M, Jackson MA, Knight R, Ober C, Spector TD, Bell JT, Clark AG, Ley RE. 2016. Genetic Determinants of the Gut Microbiome in UK Twins. Cell Host Microbe 19:731-43.
14. Moeller AH, Caro-Quintero A, Mjungu D, Georgiev AV, Lonsdorf EV, Muller MN, Pusey AE, Peeters M, Hahn BH, Ochman H. 2016. Cospeciation of gut microbiota with hominids. Science 353:380-2.
15. Rosas-Salazar C, Shilts MH, Tovchigrechko A, Chappell JD, Larkin EK, Nelson KE, Moore ML, Anderson LJ, Das SR, Hartert TV. 2016. Nasopharyngeal Microbiome in Respiratory Syncytial Virus Resembles Profile Associated with Increased Childhood Asthma Risk. Am J Respir Crit Care Med 193:1180-3.

16. Ichinohe T, Pang IK, Kumamoto Y, Peaper DR, Ho JH, Murray TS, Iwasaki A. 2011. Microbiota regulates immune defense against respiratory tract influenza A virus infection. Proc Natl Acad Sci U S A 108:5354-9.
17. Prince GA, Hemming VG, Horswood RL, Baron PA, Chanock RM. 1987. Effectiveness of topically administered neutralizing antibodies in experimental immunotherapy of respiratory syncytial virus infection in cotton rats. J Virol 61:1851-4.
18. Blanco JC, Pletneva LM, Wan H, Araya Y, Angel M, Oue RO, Sutton TC, Perez DR. 2013. Receptor characterization and susceptibility of cotton rats to avian and 2009 pandemic influenza virus strains. J Virol 87:2036-45.
19. Ottolini MG, Blanco JCG, Eichelberger MC, Porter DD, Pletneva L, Richardson JY, Prince GA. 2005. The cotton rat provides a useful small-animal model for the study of influenza virus pathogenesis. J Gen Virol 86:2823-2830.
20. Ottolini MG, Porter DD, Hemming VG, Prince GA. 2000. Enhanced pulmonary pathology in cotton rats upon challenge after immunization with inactivated parainfluenza virus 3 vaccines. Viral Immunol 13:231-6.
21. Ottolini MG, Porter DD, Blanco JC, Prince GA. 2002. A cotton rat model of human parainfluenza 3 laryngotracheitis: virus growth, pathology, and therapy. J Infect Dis 186:1713-7.
22. Pfeuffer J, Puschel K, Meulen V, Schneider-Schaulies J, Niewiesk S. 2003. Extent of measles virus spread and immune suppression differentiates between wild-type and vaccine strains in the cotton rat model (Sigmodon hispidus). J Virol 77:150-8.
23. Hamelin ME, Yim K, Kuhn KH, Cragin RP, Boukhvalova M, Blanco JC, Prince GA, Boivin G. 2005. Pathogenesis of human metapneumovirus lung infection in BALB/c mice and cotton rats. J Virol 79:8894-903.
24. Patel MC, Wang W, Pletneva LM, Rajagopala SV, Tan Y, Hartert TV, Boukhvalova MS, Vogel SN, Das SR, Blanco JC. 2016. Enterovirus D-68 Infection, Prophylaxis, and Vaccination in a Novel Permissive Animal Model, the Cotton Rat (Sigmodon hispidus). PLoS One 11:e0166336.
25. Blanco JC, Core S, Pletneva LM, March TH, Boukhvalova MS, Kajon AE. 2014. Prophylactic Antibody Treatment and Intramuscular Immunization Reduce Infectious Human Rhinovirus 16 Load in the Lower Respiratory Tract of Challenged Cotton Rats. Trials Vaccinol 3:52-60.
26. Boukhvalova MS, Prince GA, Blanco JC. 2009. The cotton rat model of respiratory viral infections. Biologicals 37:152-9.
27. Burian M, Rautenberg M, Kohler T, Fritz M, Krismer B, Unger C, Hoffmann WH, Peschel A, Wolz C, Goerke C. 2010. Temporal expression of adhesion factors and activity of global regulators during establishment of Staphylococcus aureus nasal colonization. J Infect Dis 201:1414-21.
28. Carrara AS, Coffey LL, Aguilar PV, Moncayo AC, Da Rosa AP, Nunes MR, Tesh RB, Weaver SC. 2007. Venezuelan equine encephalitis virus infection of cotton rats. Emerg Infect Dis 13:1158-65.
29. Rollin PE, Ksiazek TG, Elliott LH, Ravkov EV, Martin ML, Morzunov S, Livingstone W, Monroe M, Glass G, Ruo S, et al. 1995. Isolation of black creek canal virus, a new hantavirus from Sigmodon hispidus

in Florida. J Med Virol 46:35-9.

30. Holsomback TS, McIntyre NE, Nisbett RA, Strauss RE, Chu YK, Abuzeineh AA, de la Sancha N, Dick CW, Jonsson CB, Morris BE. 2009. Bayou virus detected in non-oryzomyine rodent hosts: an assessment of habitat composition, reservoir community structure, and marsh rice rat social dynamics. J Vector Ecol 34:9-21.
31. Winn WC, Jr., Murphy FA. 1975. Tamiami virus infection in mice and cotton rats. Bull World Health Organ 52:501-6.
32. Dietrich G, Montenieri JA, Panella NA, Langevin S, Lasater SE, Klenk K, Kile JC, Komar N. 2005. Serologic evidence of west nile virus infection in free-ranging mammals, Slidell, Louisiana, 2002. Vector Borne Zoonotic Dis 5:288-92.
33. Ehlen L, Todtmann J, Specht S, Kallies R, Papies J, Muller MA, Junglen S, Drosten C, Eckerle I. 2016. Epithelial cell lines of the cotton rat (Sigmodon hispidus) are highly susceptible in vitro models to zoonotic Bunya-, Rhabdo-, and Flaviviruses. Virol J 13:74.
34. Chaves-Moreno D, Plumeier I, Kahl S, Krismer B, Peschel A, Oxley AP, Jauregui R, Pieper DH. 2015. The microbial community structure of the cotton rat nose. Environ Microbiol Rep 7:929-35.
35. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550.
36. Boulesteix AL, Slawski M. 2009. Stability and aggregation of ranked gene lists. Brief Bioinform 10:556-68.
37. Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, Lipson KS, Knight R, Caporaso JG, Segata N, Huttenhower C. 2018. Species-level functional profiling of metagenomes and metatranscriptomes. Nat Methods 15:962-968.
38. Caspi R, Billington R, Fulcher CA, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Midford PE, Ong Q, Ong WK, Paley S, Subhraveti P, Karp PD. 2018. The MetaCyc database of metabolic pathways and enzymes. Nucleic Acids Res 46:D633-D639.
39. Henson DD, Bradley RD. 2009. Molecular systematics of the genus Sigmodon: results from mitochondrial and nuclear gene sequences. Can J Zool 87:211-220.
40. Petersen MK. 1973 Interactions Between the Cotton Rats, Sigmodon fulviventer and S. hispidus The American Midland Naturalist 90, No. 2, Oct., 1973:319-333.
41. Sadowski W, Semkow R, Wilczynski J, Krus S, Kantoch M. 1987. [The cotton rat (Sigmodon hispidus) as an experimental model for studying viruses in human respiratory tract infections. I. Para-influenza virus type 1, 2 and 3, adenovirus type 5 and RS virus]. Med Dosw Mikrobiol 39:33-42.
42. Ciabattini A, Olivieri R, Lazzeri E, Medaglini D. 2019. Role of the Microbiota in the Modulation of Vaccine Immune Responses. Front Microbiol 10:1305.
43. Hagan T, Cortese M, Rouphael N, Boudreau C, Linde C, Maddur MS, Das J, Wang H, Guthmiller J, Zheng NY, Huang M, Uphadhyay AA, Gardinassi L, Petitdemange C, McCullough MP, Johnson SJ, Gill K, Cervasi B, Zou J, Bretin A, Hahn M, Gewirtz AT, Bosinger SE, Wilson PC, Li S, Alter G, Khurana S,

Golding H, Pulendran B. 2019. Antibiotics-Driven Gut Microbiome Perturbation Alters Immunity to Vaccines in Humans. Cell 178:1313-1328 e13.

44. Harris VC, Armah G, Fuentes S, Korpela KE, Parashar U, Victor JC, Tate J, de Weerth C, Giaquinto C, Wiersinga WJ, Lewis KD, de Vos WM. 2017. Significant Correlation Between the Infant Gut Microbiome and Rotavirus Vaccine Response in Rural Ghana. J Infect Dis 215:34-41.
45. Ding T, Song T, Zhou B, Geber A, Ma Y, Zhang L, Volk M, Kapadia SN, Jenkins SG, Salvatore M, Ghedin E. 2019. Microbial Composition of the Human Nasopharynx Varies According to Influenza Virus Type and Vaccination Status. mBio 10.
46. Rosshart SP, Vassallo BG, Angeletti D, Hutchinson DS, Morgan AP, Takeda K, Hickman HD, McCulloch JA, Badger JH, Ajami NJ, Trinchieri G, Pardo-Manuel de Villena F, Yewdell JW, Rehermann B. 2017. Wild Mouse Gut Microbiota Promotes Host Fitness and Improves Disease Resistance. Cell 171:1015-1028 e13.
47. McAleer JP, Kolls JK. 2018. Contributions of the intestinal microbiome in lung immunity. Eur J Immunol 48:39-49.
48. Hufeldt MR, Nielsen DS, Vogensen FK, Midtvedt T, Hansen AK. 2010. Variation in the gut microbiota of laboratory mice is related to both genetic and environmental factors. Comp Med 60:336-47.
49. Org E, Parks BW, Joo JW, Emert B, Schwartzman W, Kang EY, Mehrabian M, Pan C, Knight R, Gunsalus R, Drake TA, Eskin E, Lusis AJ. 2015. Genetic and environmental control of host-gut microbiota interactions. Genome Res 25:1558-69.
50. Tabrett A, Horton MW. 2020. The influence of host genetics on the microbiome. F1000Res 9.
51. Tomosada Y, Chiba E, Zelaya H, Takahashi T, Tsukida K, Kitazawa H, Alvarez S, Villena J. 2013. Nasally administered Lactobacillus rhamnosus strains differentially modulate respiratory antiviral immune responses and induce protection against respiratory syncytial virus infection. BMC Immunol 14:40.
52. Kumpu M, Kekkonen RA, Korpela R, Tynkkynen S, Jarvenpaa S, Kautiainen H, Allen EK, Hendley JO, Pitkaranta A, Winther B. 2015. Effect of live and inactivated Lactobacillus rhamnosus GG on experimentally induced rhinovirus colds: randomised, double blind, placebo-controlled pilot trial. Benef Microbes 6:631-9.
53. Rosas-Salazar C, Shilts MH, Tovchigrechko A, Schobel S, Chappell JD, Larkin EK, Gebretsadik T, Halpin RA, Nelson KE, Moore ML, Anderson LJ, Peebles RS, Jr., Das SR, Hartert TV. 2018. Nasopharyngeal Lactobacillus is associated with a reduced risk of childhood wheezing illnesses following acute respiratory syncytial virus infection in infancy. J Allergy Clin Immunol 142:1447-1456 e9.
54. Kuss SK, Best GT, Etheredge CA, Pruijssers AJ, Frierson JM, Hooper LV, Dermody TS, Pfeiffer JK. 2011. Intestinal microbiota promote enteric virus replication and systemic pathogenesis. Science 334:249-52.
55. Rajagopala SV, Singh H, Patel MC, Wang W, Tan Y, Shilts MH, Hartert TV, Boukhvalova MS, Blanco JCG, Das SR. 2018. Cotton rat lung transcriptome reveals host immune response to Respiratory

Syncytial Virus infection. Sci Rep 8:11318.

56. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol 79:5112-20.
57. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537-41.
58. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194-200.
59. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glockner FO. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:7188-96.
60. Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM. 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res 37:D141-5.
61. ari Oksanen FGB, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, Eduard Szoecs and Helene Wagner. 2019. vegan: Community Ecology Package. R package version 2.5-6. . https://CRAN.R-project.org/package=vegan. Accessed 09/24/2020.
62. Hill MO. 1973. Diversity and Evenness: A Unifying Notation and Its Consequences. Ecology; Ecological Society of America.
63. McMurdie PJ, Holmes S. 2014. Waste not, want not: why rarefying microbiome data is inadmissible. PLoS Comput Biol 10:e1003531.
64. Nicolai Meinshausen PB. 2010. Stability selection. Journal of the Royal Statistical Society 72:417–473.
65. Andrews S. 2015. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 09/24/2020.
66. Ewels P, Magnusson M, Lundin S, Kaller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047-8.
67. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-20.
68. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. 2012. Metagenomic microbial community profiling using unique clade-specific marker genes. Nat Methods 9:811-4.
69. Suzek BE, Wang Y, Huang H, McGarvey PB, Wu CH, UniProt C. 2015. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. Bioinformatics 31:926-32.

70. Delroisse JM, Boulvin AL, Parmentier I, Dauphin RD, Vandenbol M, Portetelle D. 2008. Quantification of Bifidobacterium spp. and Lactobacillus spp. in rat fecal samples by real-time PCR. Microbiol Res 163:663-70.
71. Nadkarni MA, Martin FE, Jacques NA, Hunter N. 2002. Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. Microbiology (Reading) 148:257-266.
72. Barman M, Unold D, Shifley K, Amir E, Hung K, Bos N, Salzman N. 2008. Enteric salmonellosis disrupts the microbial ecology of the murine gastrointestinal tract. Infect Immun 76:907-15.

Table

Due to technical limitations, table 1 is only available as a download in the Supplemental Files section.

Figures

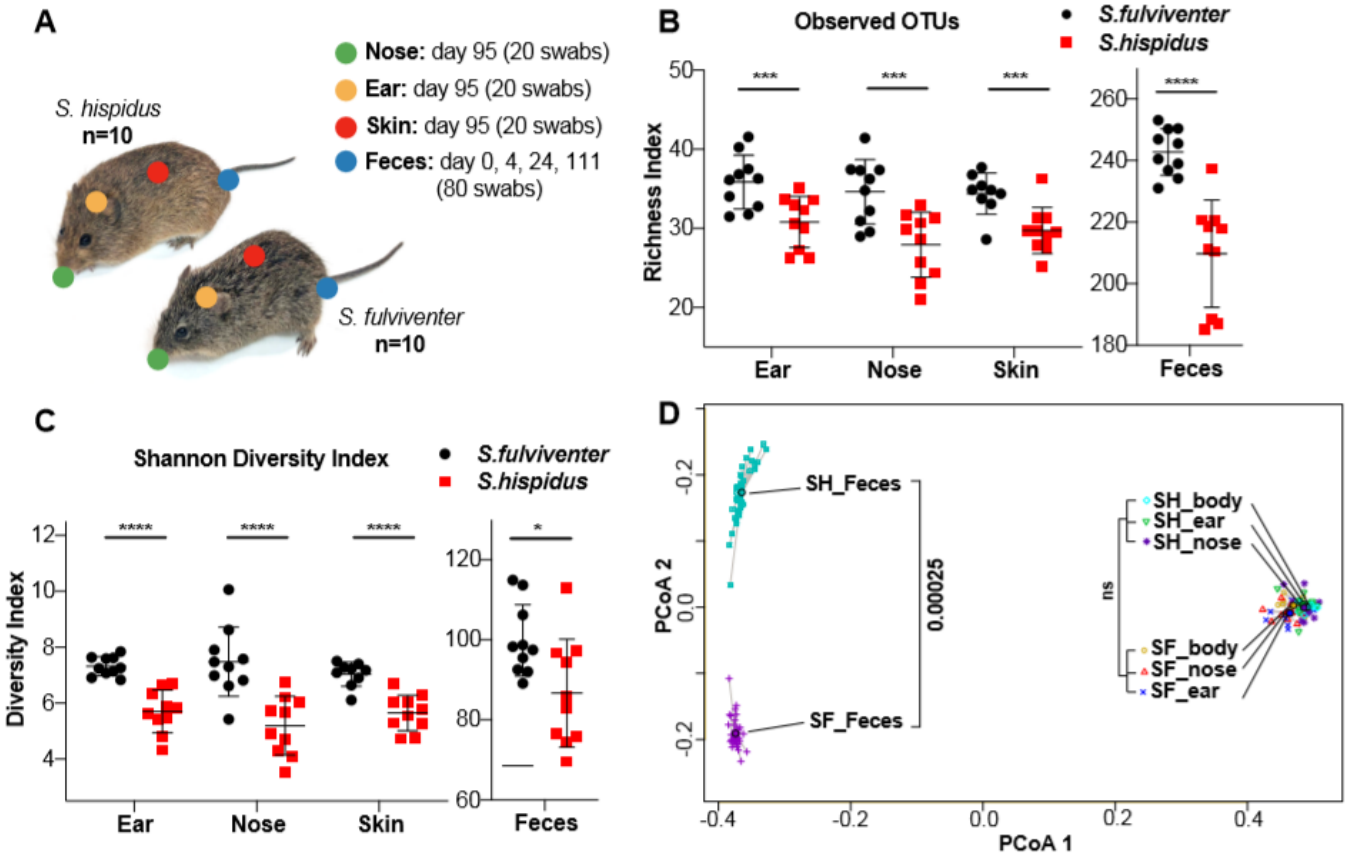

Figure 1

The cotton rat microbiome as examined with 16S rRNA gene sequencing. (A) Study design includes 10 male cotton rats from both S. fulviventer and S. hispidus. Feces were taken across a 111-day period, while sample swabs of nose, ear, skin were taken at day 95. (B) Site richness (Obs. OTUs) and (C) alpha diversity (Shannon index) metrics at the OTU level indicate that the S. fulviventer microbiome was significantly richer and more diverse across all sites. Other alpha-diversity metrics (Chao1, Simpson) are

shown in Figure S1. (D) Ordination of samples (Bray-Curtis dissimilarities, OTU level) reveals distinct microbiota compositions between feces and external sites regardless of species. Fecal microbiomes are also significantly different between species while external site microbiomes were not (SF=S. fulviventer, SH=S. hispidus). Statistical testing was performed using PERMANOVA between species. Each site was plotted separately in Figure S2. ns=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****=P≤0.0001.

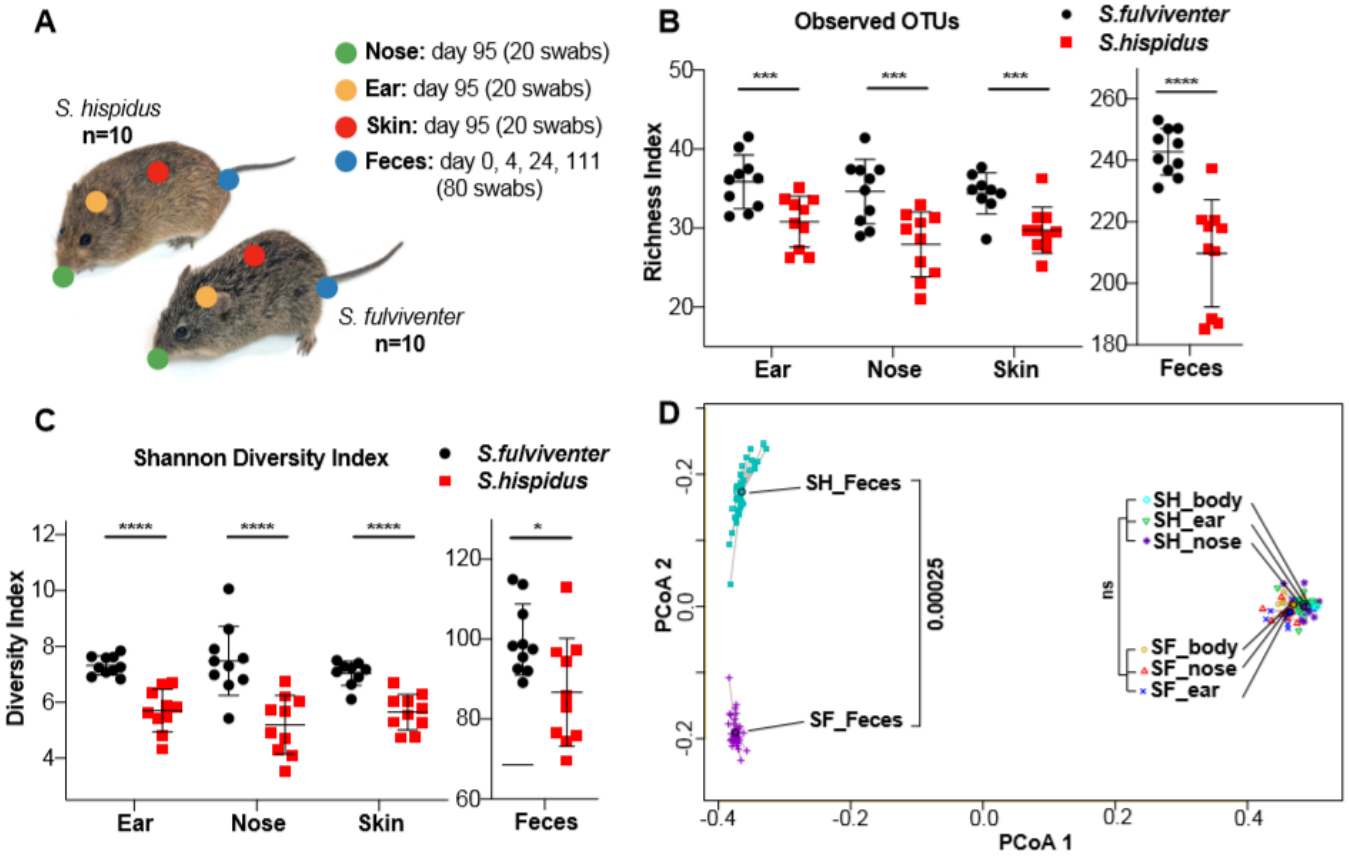

Figure 1

The cotton rat microbiome as examined with 16S rRNA gene sequencing. (A) Study design includes 10 male cotton rats from both S. fulviventer and S. hispidus. Feces were taken across a 111-day period, while sample swabs of nose, ear, skin were taken at day 95. (B) Site richness (Obs. OTUs) and (C) alpha diversity (Shannon index) metrics at the OTU level indicate that the S. fulviventer microbiome was significantly richer and more diverse across all sites. Other alpha-diversity metrics (Chao1, Simpson) are shown in Figure S1. (D) Ordination of samples (Bray-Curtis dissimilarities, OTU level) reveals distinct microbiota compositions between feces and external sites regardless of species. Fecal microbiomes are also significantly different between species while external site microbiomes were not (SF=S. fulviventer, SH=S. hispidus). Statistical testing was performed using PERMANOVA between species. Each site was plotted separately in Figure S2. ns=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****=P≤0.0001.

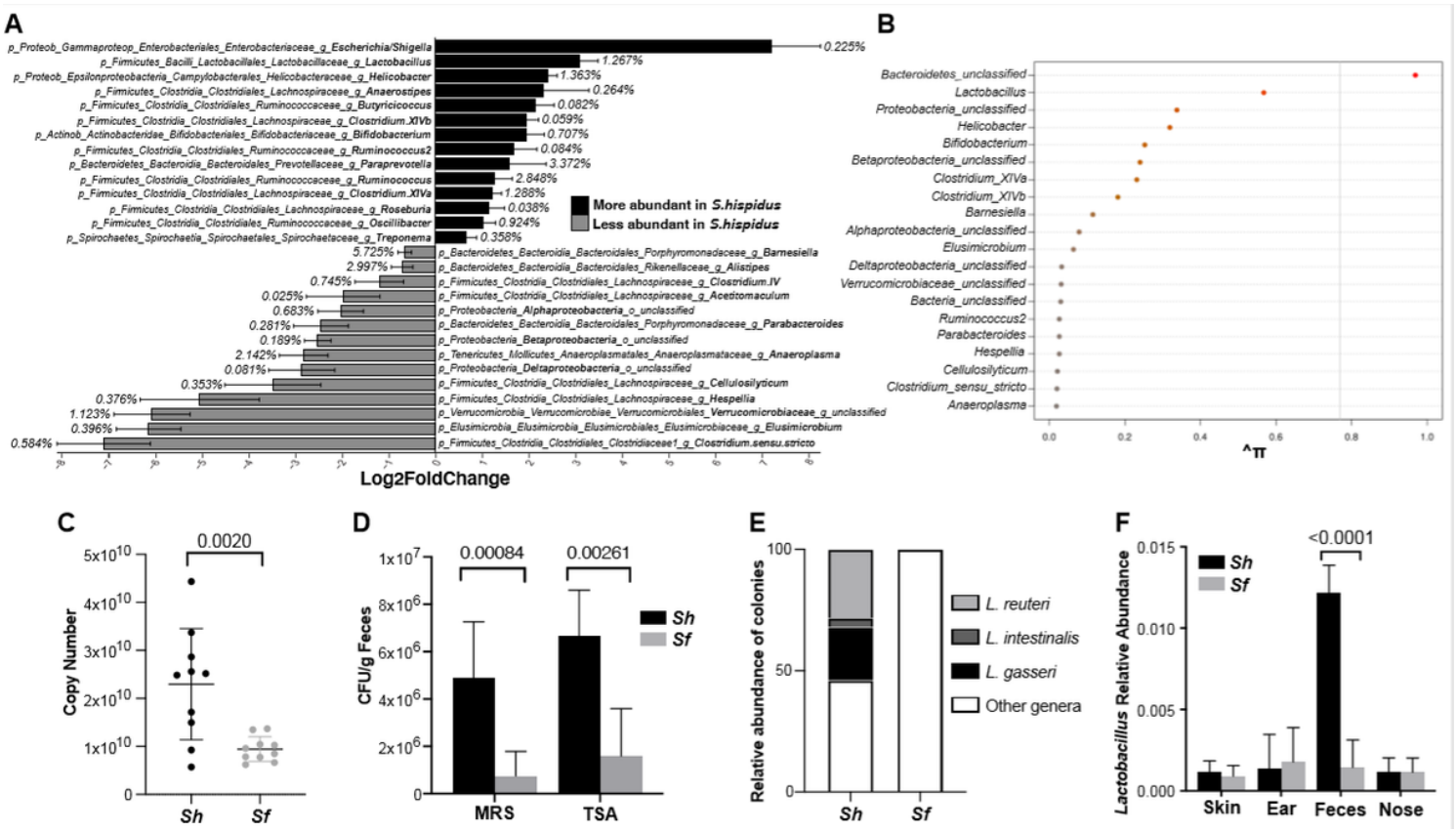

Figure 2

Differential abundance of gut taxa using 16S rRNA gene sequencing. (A) Differential abundance of gut microbial taxa between S. hispidus vs. S. fulviventer that displayed significant differences ($p < 0.05$, $q < 0.05$, l2fc > ±0.65) between host species. The Log2FoldChange is plotted along the x-axis, with genera ranked highest in S. hispidus (black, +l2fc) to highest in S. fulviventer (grey,-l2fc) on the y-axis. Error bars represent the log2 fold change standard error; relative abundances from either S. hispidus or S. fulviventer are denoted next to the corresponding bar. (B) Probability of a gut bacterial genus being selected into a stability selection model distinguishing cotton rat species. The probability of being selected into the model is plotted along the x-axis, with top 20 ranked genera along the y-axis. (C) Bacterial load depicted as copy number/µL of extracted DNA from normalized cotton rat stool. Data were generated by qPCR; and statistics were performed using unpaired T test. (D) Amount of aerobic colony forming units (CFU) per gram of feces on both Lactobacillus-selective (MRS) and general growth (TSA) media. S. hispidus displayed a higher amount of aerobic growth using both methods. Significance was calculated by the student's t-test. (E) Percentage of CFUs with positive detection of Lactobacillus amplicons determined by PCR with primers targeting the Lactobacillus 16S rRNA region. Species-specific identity of colonies was confirmed by Sanger sequencing. "Other genera" include Bacillus, Enterococcus, and Corynebacterium. (F) Relative abundance of Lactobacillus between cotton rat body sites. Significance was calculated by the student's t-test. In the gut, Lactobacillus was one of the higher-abundant taxa with significantly differential abundance between cotton rat species. Figures C-F were generated in Prism 8.

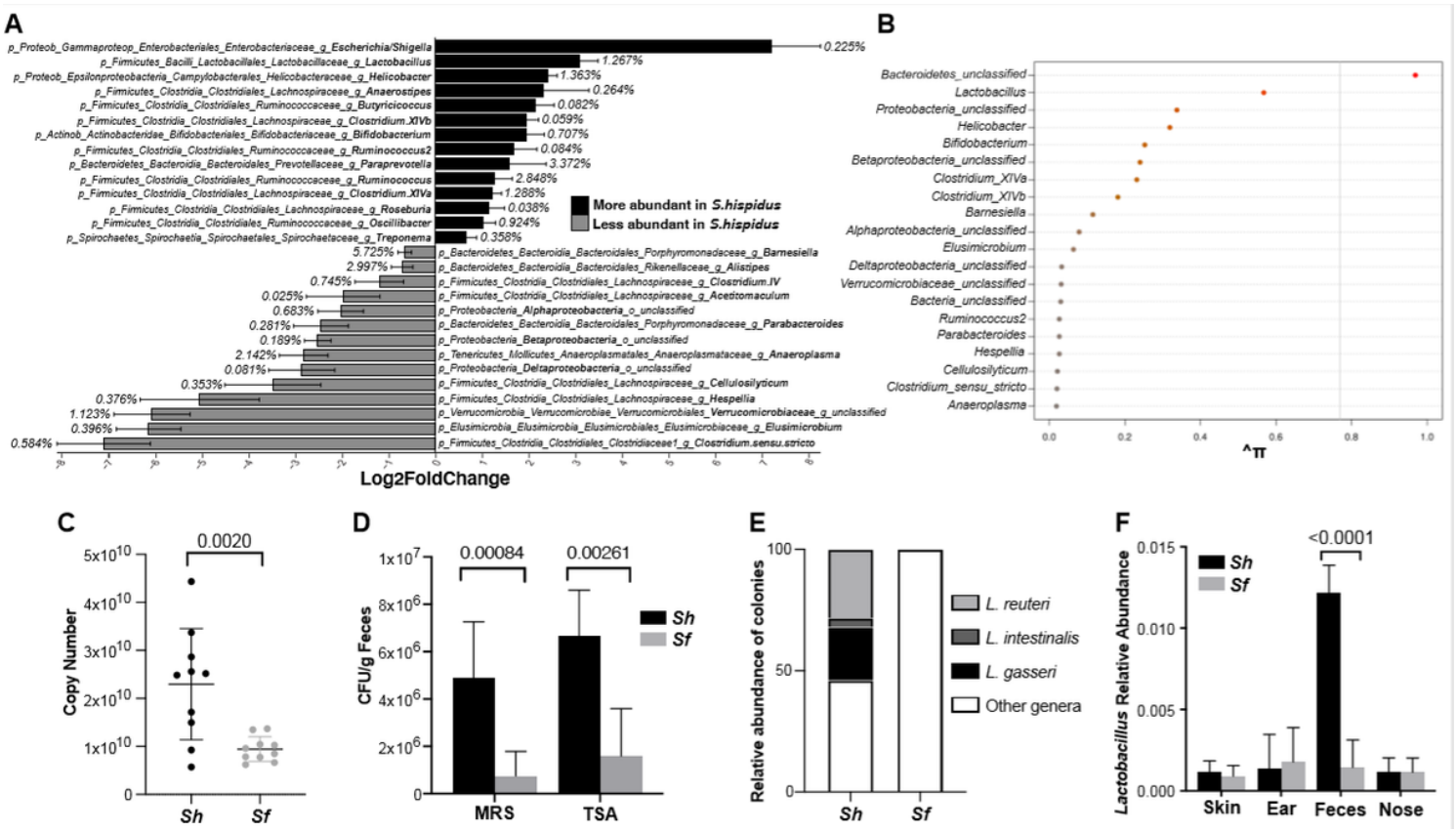

Figure 2

Differential abundance of gut taxa using 16S rRNA gene sequencing. (A) Differential abundance of gut microbial taxa between S. hispidus vs. S. fulviventer that displayed significant differences ($p < 0.05$, $q < 0.05$, l2fc > ±0.65) between host species. The Log2FoldChange is plotted along the x-axis, with genera ranked highest in S. hispidus (black, +l2fc) to highest in S. fulviventer (grey,-l2fc) on the y-axis. Error bars represent the log2 fold change standard error; relative abundances from either S. hispidus or S. fulviventer are denoted next to the corresponding bar. (B) Probability of a gut bacterial genus being selected into a stability selection model distinguishing cotton rat species. The probability of being selected into the model is plotted along the x-axis, with top 20 ranked genera along the y-axis. (C) Bacterial load depicted as copy number/µL of extracted DNA from normalized cotton rat stool. Data were generated by qPCR; and statistics were performed using unpaired T test. (D) Amount of aerobic colony forming units (CFU) per gram of feces on both Lactobacillus-selective (MRS) and general growth (TSA) media. S. hispidus displayed a higher amount of aerobic growth using both methods. Significance was calculated by the student's t-test. (E) Percentage of CFUs with positive detection of Lactobacillus amplicons determined by PCR with primers targeting the Lactobacillus 16S rRNA region. Species-specific identity of colonies was confirmed by Sanger sequencing. "Other genera" include Bacillus, Enterococcus, and Corynebacterium. (F) Relative abundance of Lactobacillus between cotton rat body sites. Significance was calculated by the student's t-test. In the gut, Lactobacillus was one of the higher-abundant taxa with significantly differential abundance between cotton rat species. Figures C-F were generated in Prism 8.

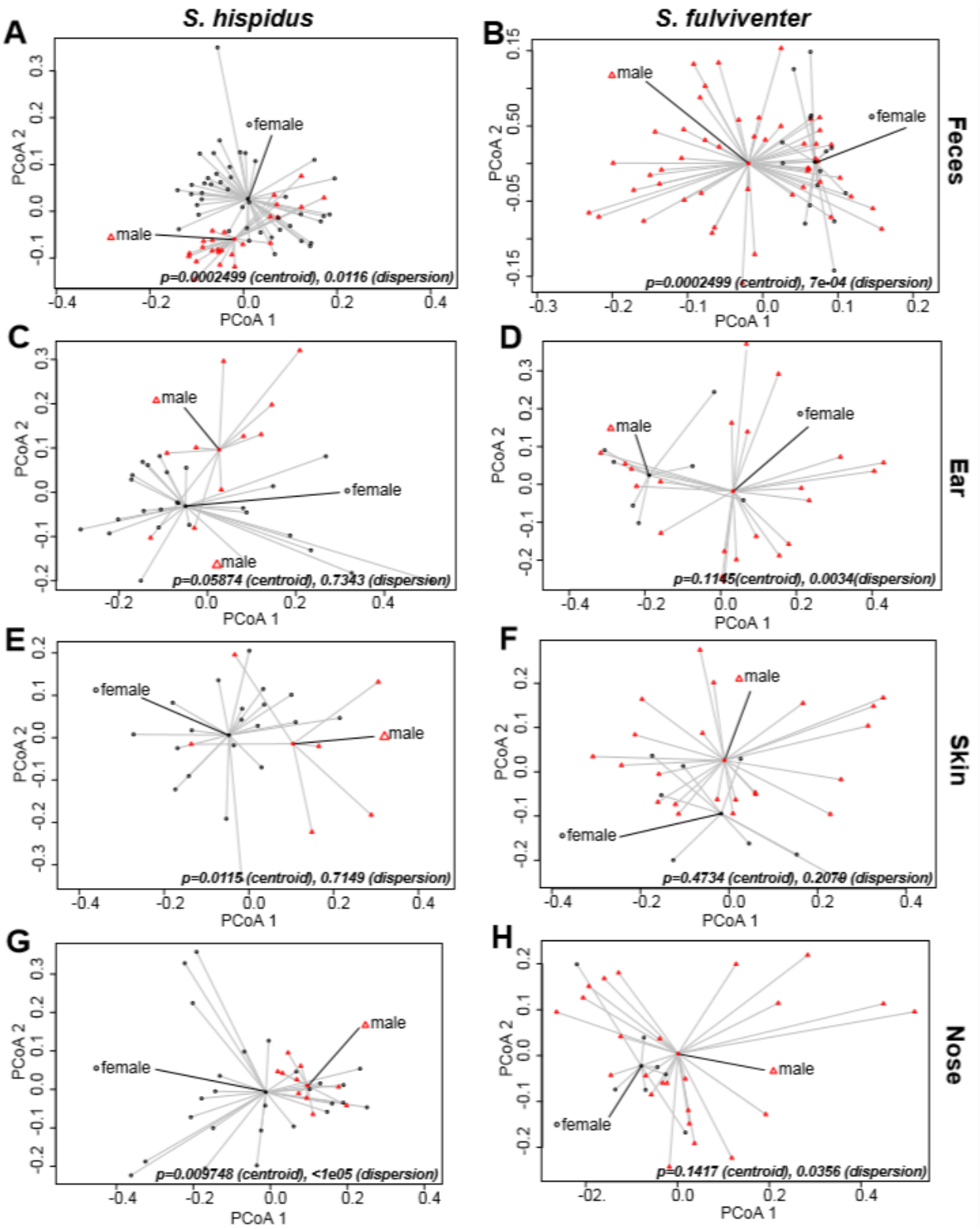

Figure 3

Clustering of site- and species-specific samples (Bray Curtis, OTU level) revealed host sex-dependent communities at most sites. Statistical testing was performed using PERMANOVA for both the geometric mean (or centroid) of the cluster and the dispersion (or variance). (A, B) Gut communities showed significant differences between both S. fulviventer (SF) and S. hispidus (SH) males and females. (C-H)

External sites (ear, skin, and nose) showed sex-based community trends based on both sample mean distances and dispersion.

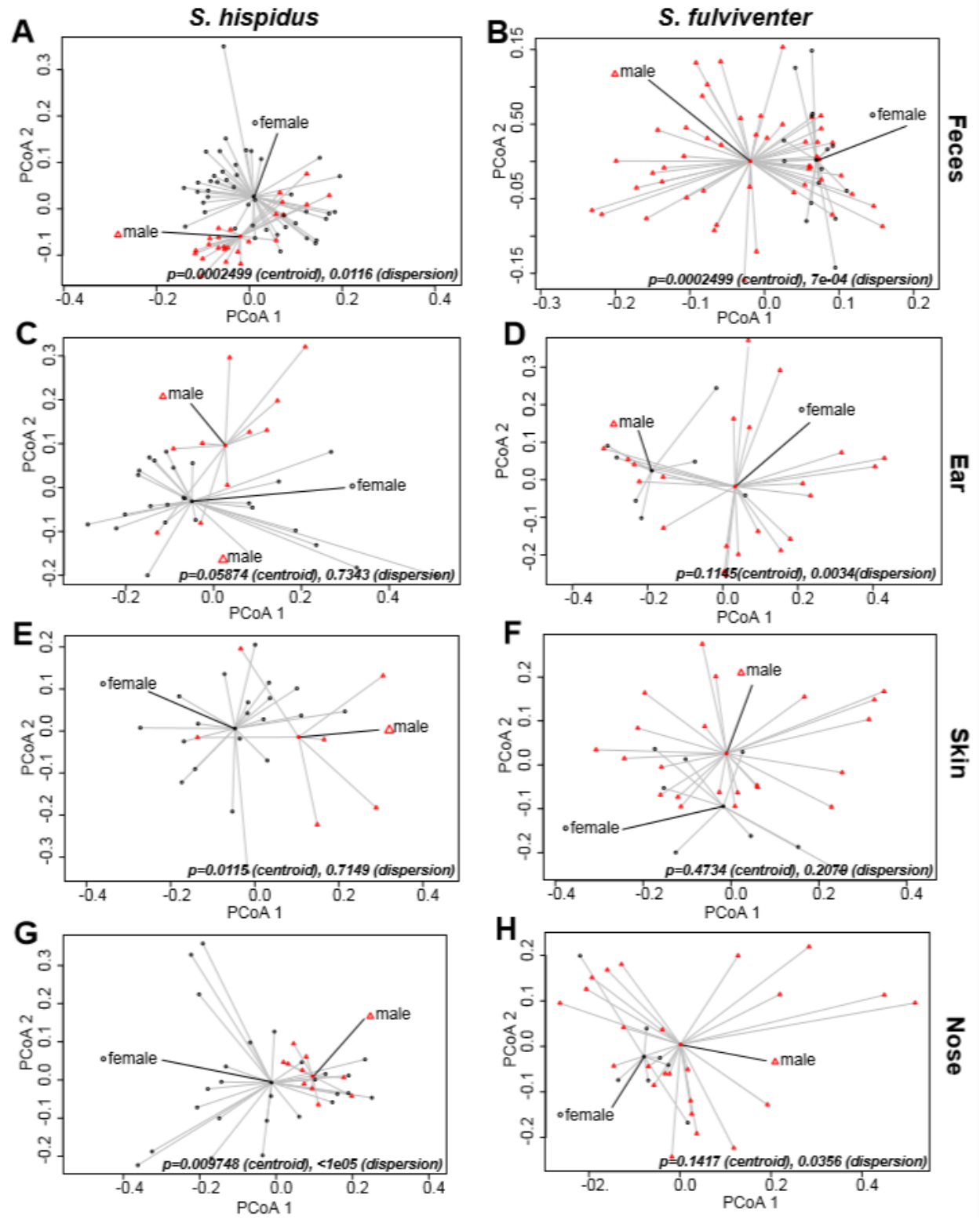

Figure 3

Clustering of site- and species-specific samples (Bray Curtis, OTU level) revealed host sex-dependent communities at most sites. Statistical testing was performed using PERMANOVA for both the geometric mean (or centroid) of the cluster and the dispersion (or variance). (A, B) Gut communities showed

significant differences between both S. fulviventer (SF) and S. hispidus (SH) males and females. (C-H) External sites (ear, skin, and nose) showed sex-based community trends based on both sample mean distances and dispersion.

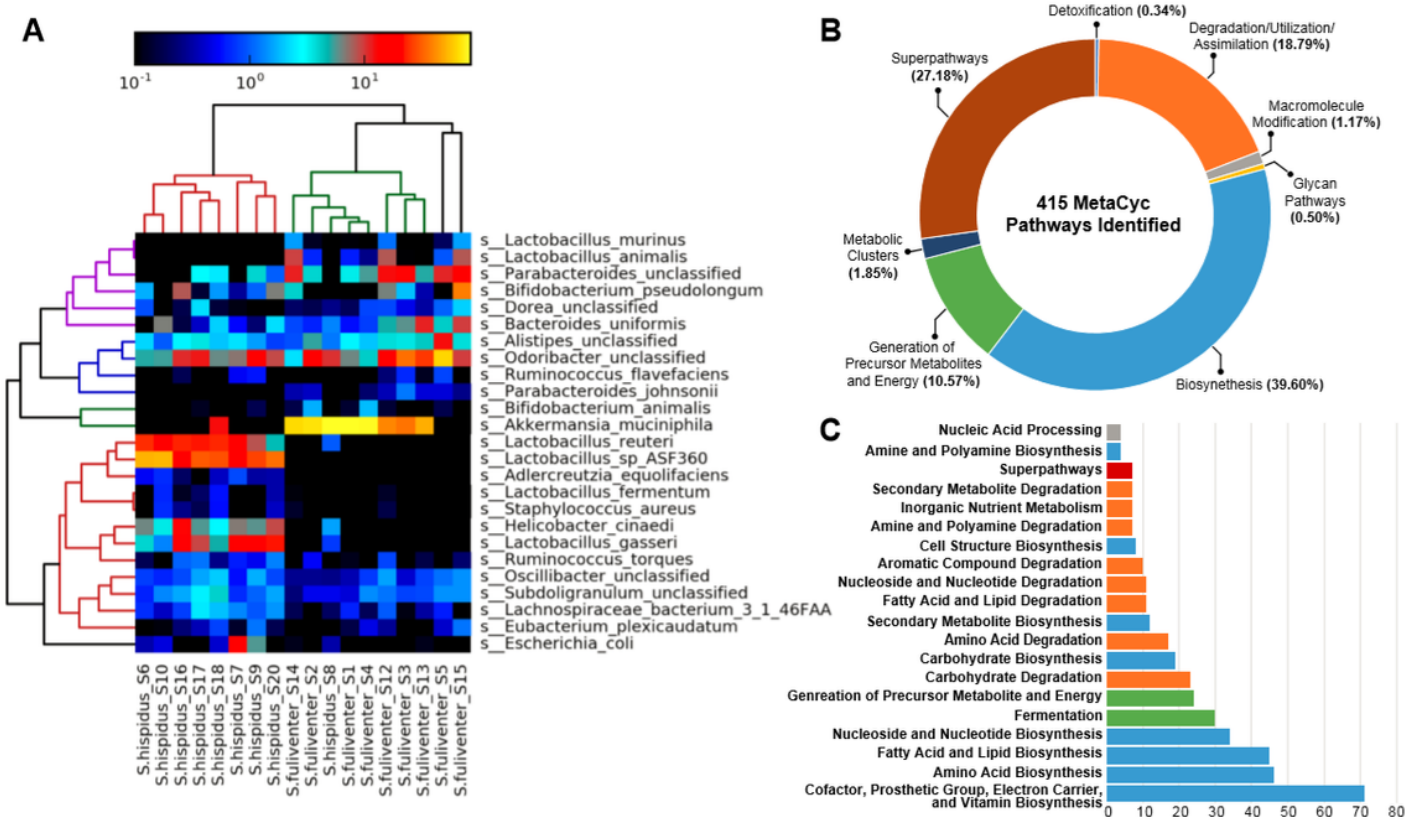

Figure 4

Differential abundance of cotton rat gut taxa and corresponding pathways using whole-genome metagenomic sequencing. (A). Hierarchical clustering (Bray-Curtis) of the top 25 most differentially abundant bacterial species between S. fulviventer vs. S. hispidus. Lactobacillus reuteri and L. gasseri were drastically more abundant in S. hispidus stool, while Akkermansia muciniphila was more abundant in S. fulviventer stool. (B) Distribution of MetaCyc metabolomic pathways predicted from bacterial sequences. 415 unique pathways were identified, and most pathways were classified within the Biosynthesis, Superpathway, and Degradation/Utilization/Assimilation pathway superclasses. (C) Distribution of pathway ontology. Of all the identified pathways, the largest group consisted of Cofactor, Prosthetic Group, Electron Carrier, and Vitamin Biosynthesis, which are often components of host biology sourced solely by commensal bacteria.

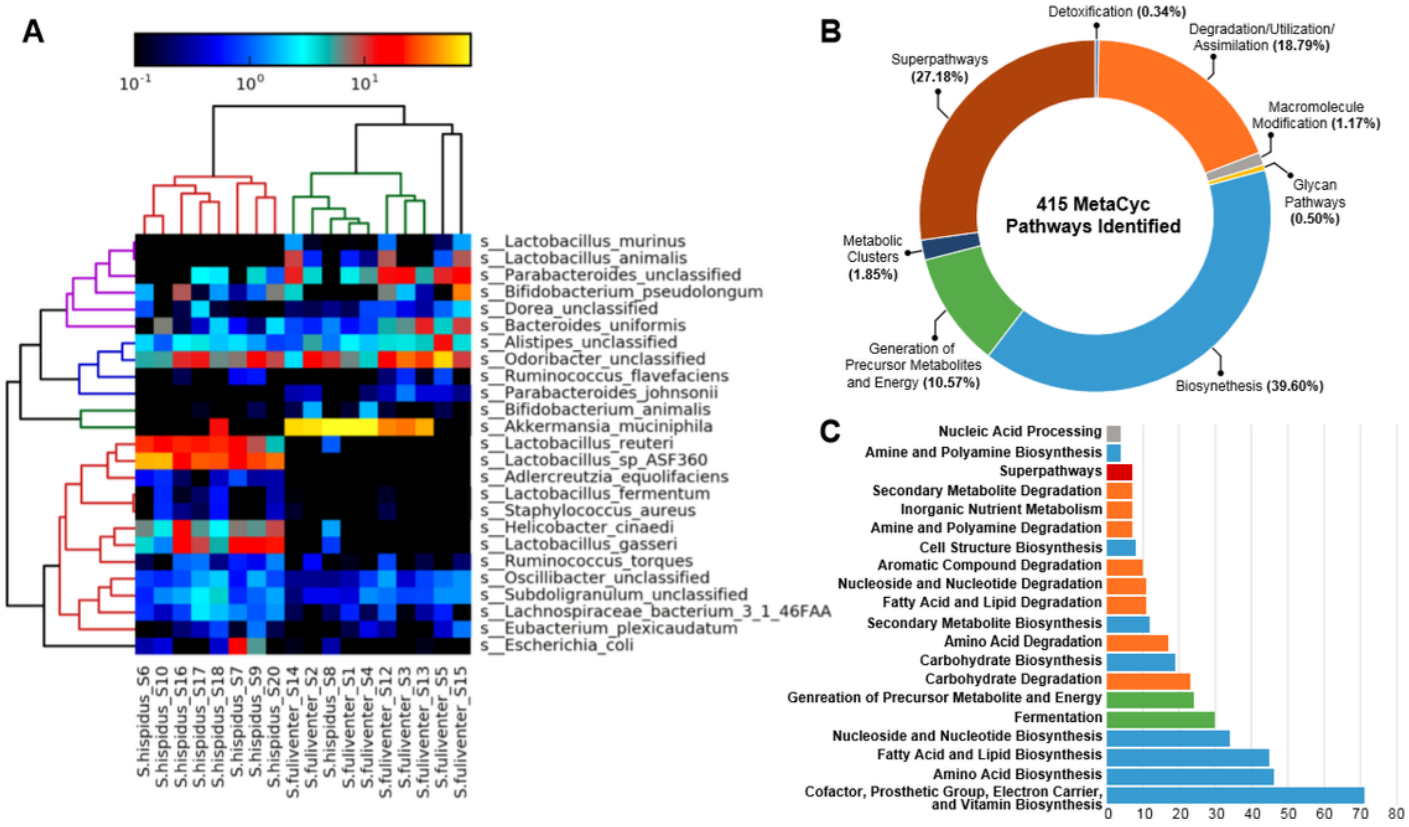

Figure 4

Differential abundance of cotton rat gut taxa and corresponding pathways using whole-genome metagenomic sequencing. (A). Hierarchical clustering (Bray-Curtis) of the top 25 most differentially abundant bacterial species between S. fulviventer vs. S. hispidus. Lactobacillus reuteri and L. gasseri were drastically more abundant in S. hispidus stool, while Akkermansia muciniphila was more abundant in S. fulviventer stool. (B) Distribution of MetaCyc metabolomic pathways predicted from bacterial sequences. 415 unique pathways were identified, and most pathways were classified within the Biosynthesis, Superpathway, and Degradation/Utilization/Assimilation pathway superclasses. (C) Distribution of pathway ontology. Of all the identified pathways, the largest group consisted of Cofactor, Prosthetic Group, Electron Carrier, and Vitamin Biosynthesis, which are often components of host biology sourced solely by commensal bacteria.

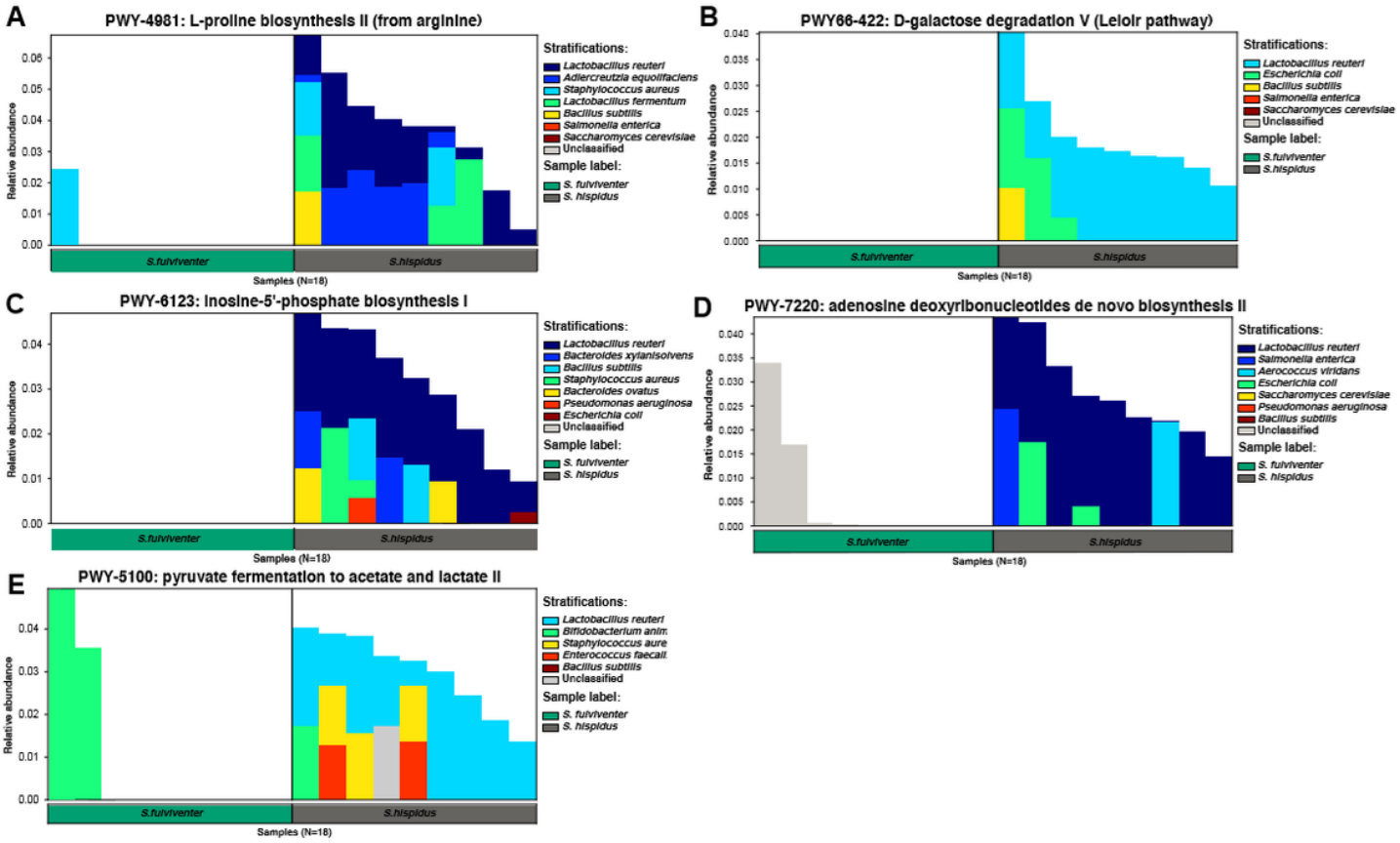

Figure 5

Several pathways that were more active in S. hispidus than S. fulviventer were greatly contributed to by Lactobacillus species. (A) Catalysis of proline biosynthesis by bacterial enzymes (PWY-4981). (B) Catalysis of the conversion of D-galactose to D-glucopyranose 6-phosphate, the more metabolically versatile carbohydrate that can feed directly into glycolysis, by the enzymes of the Leloir pathway (PWY66-422). (C) De novo biosynthesis of purines (PWY-6123). (D) De novo synthesis of ADP for the direct feeding of ATP generation, a pathway that can only accept ribonucleoside diphosphates instead of the mono- or triphosphate forms (PWY-7220). (E) Anaerobic breakdown of glucose to energy (PWY-5100).

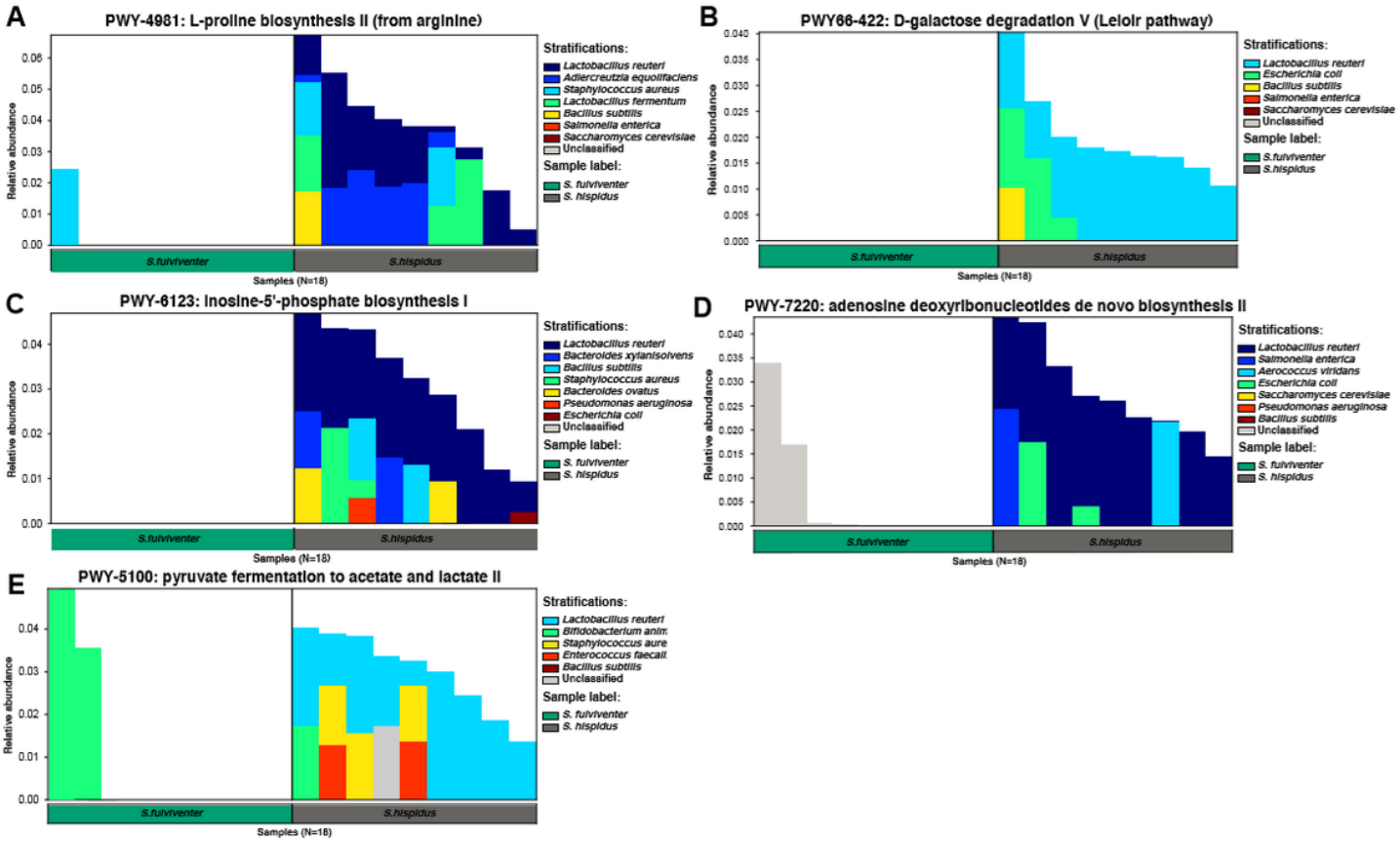

Figure 5

Several pathways that were more active in S. hispidus than S. fulviventer were greatly contributed to by Lactobacillus species. (A) Catalysis of proline biosynthesis by bacterial enzymes (PWY-4981). (B) Catalysis of the conversion of D-galactose to D-glucopyranose 6-phosphate, the more metabolically versatile carbohydrate that can feed directly into glycolysis, by the enzymes of the Leloir pathway (PWY66-422). (C) De novo biosynthesis of purines (PWY-6123). (D) De novo synthesis of ADP for the direct feeding of ATP generation, a pathway that can only accept ribonucleoside diphosphates instead of the mono- or triphosphate forms (PWY-7220). (E) Anaerobic breakdown of glucose to energy (PWY-5100).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.pdf
- Table1.pdf
- FigureS1.pdf
- FigureS1.pdf
- FigureS2.pdf
- FigureS2.pdf

- FigureS3.pdf
- FigureS3.pdf
- FigureS4.pdf
- FigureS4.pdf
- FigureS5.pdf
- FigureS5.pdf
- FigureS6.pdf
- FigureS6.pdf
- FigureS7.pdf
- FigureS7.pdf
- FigureS8.pdf
- FigureS8.pdf
- TableS1.pdf
- TableS1.pdf