

An integrated metagenomics and metabolomics approach implicates the microbiome-gut-brain-axis in the pathogenesis of Huntington's disease transgenic mice

Geraldine Kong

The Florey Institute of Neuroscience and Mental Health

Susan Ellul

University of Melbourne

Vinod Narayana

University of Melbourne

Komal Kanojia

University of Melbourne

Harvey Tran Thai Ha

The Florey Institute of Neuroscience and Mental Health

Shanshan Li

The Florey Institute of Neuroscience and Mental Health

Thibault Renoir

The Florey Institute of Neuroscience and Mental Health

Kim-Anh Le Cao

University of Melbourne

Anthony Hannan (✉ anthony.hannan@florey.edu.au)

Florey Institute of Neuroscience and Mental Health, University of Melbourne <https://orcid.org/0000-0001-7532-8922>

Research

Keywords: Gut microbiome, Huntington's disease, metagenomics, metabolomics, brain disorder, mouse model

Posted Date: February 12th, 2020

DOI: <https://doi.org/10.21203/rs.2.23316/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Huntington's disease (HD) is an autosomal dominant neurodegenerative disorder with onset and severity of symptoms influenced by various environmental factors. Recent discoveries have highlighted the importance of the gastrointestinal microbiome in mediating the bidirectional communication between the central and enteric nervous system via circulating factors. Using shotgun sequencing, we investigated the gut microbiome composition in the R6/1 transgenic mouse model of HD from 4 to 12 weeks of age (early adolescent through to adult stages). Targeted metabolomics was also performed on the blood plasma of these mice (n=9 per group) at 12 weeks of age to investigate potential effects of gut dysbiosis on the plasma metabolome profile. The short-chain fatty acid (SCFA) concentrations in the plasma were validated using LCMS.

Results: Beta diversity differed between HD mice compared to their WT littermates at 12 weeks of age, suggesting that gut dysbiosis occurs prior to overt motor symptoms. Modelled time profiles of each species, KEGG Orthologs and bacterial genes, revealed heightened volatility in the HD mice, indicating potential early effects of HD mutation in the gut prior to significant cognitive and motor dysfunction. In addition to gut dysbiosis at 12 weeks of age, we also found functional differences between the WT and HD mice. The butanoate metabolism pathway, which leads to the production of the protective short-chain fatty acid, butyrate, was increased in the gut. However, the plasma concentration of butyrate and propionate were both decreased in the HD mice, as determined by the SCFA validation. The concentrations of ATP were increased by over 4-fold in the HD plasma, which was contrary to expectations. The statistical integration of the metagenomics and metabolomics unraveled several *Bacteroides* species that were positively correlated with butyrate levels, and negatively correlated with ATP and piperolic acid in the plasma. *Blautia producta* and *Prevotella scopos* were found to be negatively correlated with the butyrate and ATP respectively.

Conclusions: Our study has revealed a previously unknown relationship between the gut bacteria and plasma metabolome, suggesting the potential role of gut in modulating the pathogenesis of HD via specific altered plasma metabolites which mediate gut-brain signalling.

Introduction

Huntington's disease (HD) is a progressive neurodegenerative disorder caused by a trinucleotide expansion (CAG) in the huntingtin gene, which is expressed ubiquitously throughout the brain and peripheral tissues (1). Despite being a genetic disease, the onset and severity of the symptoms is affected by a combination of environmental factors including diet, physical activity and stress (2). There is increasing evidence implicating the gut microbiome as a factor influencing neurodegenerative diseases including Alzheimer's disease and Parkinson's disease (3–7). Recently, two independent studies in mouse models have shown evidence of gut dysbiosis in HD (8,9). This was not entirely surprising, as gut dysbiosis is tightly associated with gastrointestinal motility and HD patients have bowel complications (10).

One important function of the gut microbiome is the production of metabolites for the maintenance of the host's metabolic homeostasis as well as the shaping of the gut-brain-axis (11). A significant part of the metabolites circulating in mammalian blood are derived from the intestinal microbial community, some of which are important in regulating brain health (12,13). In particular, short-chain fatty acids (SCFA) are notable microbial-derived metabolites from the fermentation of dietary fiber in the gut (14). Manipulation of SCFA levels in the plasma via oral ingestion have shown improvements in the severity of symptoms in various metabolic and neurodegenerative diseases (15–17).

Since HD is a genetic autosomal dominantly inherited disease, the presence of the mutant protein in the gut could have an effect on the gut microbial structure in early life which may then modulate the onset and severity of other HD symptoms. Hence, in this study, we sought to determine the microbial structure of HD from young to adult stages of life, prior to onset of overt motor deficits modeling clinical symptoms, relative to wild-type littermate controls.

We profiled the gut microbiome of the R6/1 transgenic mouse model of HD at five timepoints: 4, 6, 8, 10 and 12 weeks of age, using shotgun sequencing and metagenomic analyses. The timepoint of 12 weeks was the age at which we recently discovered imbalance using 16S rRNA amplicon sequencing of fecal DNA (8). We performed targeted metabolomics on the plasma of these mice at 12 weeks of age, followed by multi-omics integration analysis, to determine any potential effects of gut dysbiosis on the plasma metabolome profile.

Methods

Animal husbandry

Hemizygous R6/1 transgenic male mice were crossed with F1 (CBA x C57Bl/6) females to generate male wild-type (WT) and R6/1 littermates, herein termed HD mice (n = 9 per group). The mice were co-housed according to genotype due to the coprophagic nature of mice, and they were distributed among 4 cages per group, with 2-3 mice per cage. The animals had *ad libitum* access to water and chow in a temperate (22 °C) and humidity (45%) controlled room, with a 12-hour light / dark cycle (lights on at 0700 h). Cage changes were performed weekly and body weight was monitored. All experiments were approved by the Florey Institute of Neuroscience and Mental Health Animal Ethics Committee and performed in accordance with research guidelines of the National Health and Medical Research Council.

Fecal samples and plasma collection

Fecal samples were collected at 4, 6, 8, 10 and 12 weeks of age as previously described (8). Briefly, individual mice were placed in clean cages for 5-10 min. Fresh pellets were collected and immediately frozen with liquid nitrogen prior to storage at -80 °C until use. Blood was collected at 12 weeks of age via terminal cardiac puncture in EDTA tubes. The plasma was obtained after centrifugation at 5000g for 10 mins.

Fecal DNA extraction and shotgun sequencing

The 90 fecal samples were randomized prior to DNA extraction to ensure the samples were equally represented across sequencing runs so as to minimize batch effects. DNA was extracted from 50-200mg of raw sample with preliminary step of bead beating using 0.1 diameter glass beads (BioSpec Products #11079101) on the Powerlyser 24 homogenizer (Mo-Bio #13155). Each sample was added to a bead tube filled with 750mL of Bead Solution (Qiagen #12855-100-BS) and 60mL of solution C1 was added to sample tube and vortexed to mix. The tubes were heated at 65°C for 10 minutes. The sample was then bead beat for 5 mins at 2000rpm, and then centrifuged for 1 min at 10000g. The resulting lysate was transferred to a new collection tube. Extraction was as per Qiagen DNeasy Powersoil Kit (#12888-100). Final elution volume was? 50mL.

Libraries were prepared according to the manufacturer's protocol using Nextera XT Library Preparation Kit (Illumina #FC-131-1096). The only alterations to the protocol as outlined was the reduction of the total reaction volume for processing in 96 well plate format. Library preparation and bead clean-up was run on the Mantis Liquid Handler (Formulatrix) and Epmotion (Eppendorf #5075000301) automated platform. These programs cover "Tagment Genomic DNA" to "Amplify DNA" in the protocol (Mantis-Nextera XT library prep protocol) and "Clean Up Libraries" in the protocol (Epmotion – Library Clean Up protocol). On Completion of the library prep protocol, each library was quantified and QC was performed using the Quant-iT™ dsDNA HS Assay Kit (Invitrogen) and Agilent D1000 HS tapes (#5067-5582) on the TapeStation 4200 (Agilent #G2991AA) as per the manufacturer's protocol.

Nextera XT libraries were pooled at equimolar amounts of 2nM per library to create a sequencing pool. The library pool was quantified in triplicates using the Qubit™ dsDNA HS Assay Kit (Invitrogen). Library QC is performed using the Agilent D1000 HS tapes (#5067-5582) on the TapeStation 4200 (Agilent #G2991AA) as per the manufacturer's protocol. The library was prepared for sequencing on the NextSeq500 (Illumina) using NextSeq500/550 High Output v2 2 x 150bp paired end chemistry in the Australian Centre for Ecogenomics (University of Queensland, Australia) according to the manufacturer's protocol. The reads obtained were between 6 to 8 million reads for each sample.

Taxonomic assignment and microbial structure analysis

The reads were de-replicated and the host contaminated reads were removed for the subsequent analysis. The contaminant-free dataset has been made publicly available at NCBI (uploading in progress) and the R code for the following analysis have been uploaded onto https://github.com/gkong1/longitudinal_metagenomics_analysis. The quality of the reads was checked using FASTQC. To identify the taxonomic composition, the reads were classified using Kaiju v1.6.3 against the NCBI BLAST nr+euk database, under Greedy mode with MEM 11 parameter to maximize the classification rate (18).

For alpha diversity analysis, reads were rarefied to 25,000 reads to calculate several alpha diversity metrics including species richness, Shannon and Inverse Simpson metrics using the Phyloseq v1.28 R package (19). Two-way ANOVA was performed to determine the effect of genotype and age in any of these metrics while Kruskal-Wallis rank test was performed to determine the significance of the difference at each timepoint.

For subsequent analysis, the data was normalized to their relative abundance and only features above the relative abundance threshold of 0.001 were included, resulting in a table of 2041 species. Using Phyloseq, Bray-Curtis dissimilarity distances were calculated and used for ordination in Principal Coordinates Analysis (PCoA). To determine whether the visually observed differences were statistically significant, Adonis PERMANOVA (Permutational Multivariate Analysis of Variance Using Distance Matrices) from the 'vegan' R package was performed with 999 permutations (20). Multivariate sparse Partial Least Square Discriminant analysis (sPLS-DA), a supervised method implemented in mixOmics v6.8.5 R package, was used to identify the microbial signature distinguishing the two genotypes (21). Leave-one-out cross-validation method was used to assess the performance of the method in terms of classification error rate.

Metagenomic reads functional analysis

Metagenomic reads were annotated using MG-RAST Version v4.0.3 (<https://www.mg-rast.org>) (22). The Kyoto Encyclopedia of Genes and Genomes (KEGG) database was used to explore the microbial functions in the samples. Similarity search between protein reads and the SEED/KEGG databases was conducted by using an E-value cutoff of 10^{-5} minimum identity cutoff of 60%, and minimum alignment length of 15aa. The annotated reads were sorted into Level 3 and Level 4 subsystems for further analysis. Similar to the taxonomic data analysis, sPLS-DA was used to determine the signature of enzyme commission (EC) numbers of KEGG orthologues (KOs) which could distinguish between the two genotypes.

Identifying cage and sequencing run effects

To assess potential cage and sequencing run effects, we used linear mixed models to test for cage and run effects with cage or sequencing run as random effects (significance level 0.05). We also examined the PCA sample plots from the centred log ratio (CLR) transformed relative abundance data across all time points, and at each time point to assess potential age effects.

Time course analysis

We used the timeomics R pipeline from (23) to perform the following analysis, focused on assessing relative stability or volatility over time. To cluster bacterial features, such as species/KO/gene profiles across time, we modelled each feature across the samples of the same genotype using Linear Mixed Effect Model Splines (LMMS) (24). LMMS fits one of the best possible models for each species: a linear model which assumes that the response is not affected by individual variation (Model 1), a model of nonlinear response patterns (Model 2), a model fitted with subject-specific random intercept (Model 3) or a model fitted with both subject specific intercept and slope (Model 4). This method accounts for between and within individual variability and irregular time sampling whilst avoiding under- or over-smoothing.

Straight line modelling (Model 1) may occur when the inter-individual variability is very high. An additional filtering step based on Breusch-Pagan test and mean square error were used to assess homoskedasticity of residuals to remove profiles that were considered noisy (23). Time profiles of species/KO/ gene profiles were clustered based on PCA, as proposed by (23). Sparse PCA was applied to obtain a smaller cluster signature.

As correlations can be spurious for compositional microbiome data, we used the proportionality distance between species identified as relevant in our method, as a measure of association (23). A small value indicates that in proportion, the pair of profiles are strongly associated (0 represents the strongest). Median distances within clusters were calculated and compared to all the median distance between all species (Supplemental data).

Metabolomics sequencing

Quenching and extraction of polar metabolites

Polar metabolites were extracted in mouse plasma from different conditions using acetonitrile: methanol: milli-Q water (2:2:1 v/v/v, 180mL) containing internal standards ($^{13}\text{C}_{10}$, $^{15}\text{N}_5$ -AMP, ^{13}C , ^{15}N -UMP, $^{13}\text{C}_6$ -

Sorbitol and $^{13}\text{C}_5$, ^{15}N -Valine) at 2mM concentration per sample and vortexed. The homogenate was incubated for 10 mins at 4°C on the thermomixer to ensure lysis of all cell membranes and then centrifuged to maximum speed 14000 rpm at 4°C to remove the cell debris. Around 150mL of the clear supernatant was transferred into a HPLC vials with inserts for LC/MS analysis.

LC/MS analysis of polar metabolites

Polar metabolites were analysed by LC/MS. Chromatography was performed on Agilent 1200 series LC system (Agilent Technologies). Samples were stored in an autosampler at 4°C. Metabolite separation was performed by injecting a 7mL sample onto a SeQuant ZIC-pHILIC column (150mm x 4.6mm, 5mm, Merck Millipore) maintained at 30°C using Solvent A (20mM $(\text{NH}_4)^2\text{CO}_3$, pH 9.0, Sigma-Aldrich) and Solvent B (100% acetonitrile, Hypergrade for LCMS LiChrosolv, Merck) at a flow rate of 300mL/min. The gradients used were: time (t) = 0 min, 80% B; t = 0.5 min, 80% B; t = 15.5 min, 50% B; t = 17.5 min, 30% B; t = 18.5 min, 5% B, t = 21.0 min, 5% B; t = 23 min, 80% B, t = 33 min, 80% B.

The mass spectrometry analysis was performed on an Agilent 6545B series quadrupole time-of-flight mass spectrometer (QTOF MS) (Agilent Technologies). The LC flow was directed to an electrospray ionisation source (ESI), where metabolite ionisation in negative mode was performed with a capillary voltage of 2500V, a drying gas (N_2) pressure of 20 psi with a gas flow rate of 10.0 L/min, a gas temperature in the capillary of 300°C and fragmentor and skimmer cap voltages of 125V and 45V, respectively. Data was collected in centroid mode with a mass range of 60-1200 m/z and an acquisition rate of 1.5 spectra/s in all-ion fragmentor (AIF) mode, which included three collision energies (0, 10 and 20V). Prior to analysis, mass calibration was performed for negative mode to 0.5ppm accuracy of the m/z value. Internal mass calibration was performed using Agilent ESI-TOF Reference Mass Solution containing purine (119.036320) and hexakis (^1H , ^1H , ^3H - tetrafluoropropoxy) phosphazine (981.99509) (API-TOF Reference Mass Solution Kit, Agilent technologies), which was continuously infused into the ESI source at a flow rate of 200mL/min.

Targeted data matrices were generated using MassHunter Quantitative Analysis software (version B.09.00, Agilent Technologies) with metabolite identification (Metabolomics Standard Initiative (MSI) level 1) based on the retention time and molecular masses matching an authentic standard (25).

Metabolomics analysis

The missing values (2 in total) were imputed and the metabolomic dataset was median-normalized. Similarly, sPLS-DA was performed to obtain a signature of metabolites which discriminate the two groups.

Multi-omics data integration

We applied the DIABLO multi-omics integration method from the mixOmics v6.8.5 R package to integrate the shotgun sequencing data, both bacterial species and genes, and the plasma metabolomics data at 12 weeks of age, as well as to discriminate between the WT mice and HD littermates (26). The method is supervised and allows variable selection. A design matrix of 0.1 was used to focus primarily on the discrimination between the genotype groups leading to a selection of 30 species and 15 metabolites, as detailed in (26). The resulting network was analyzed further using Cytoscape v3.7.2 (27).

SCFA extraction, derivatization and LCMS analysis

SCFAs were extracted and analysed by LCMS in accordance with the previously described protocol (28). Briefly, 380 mL of acetonitrile:water (1:1, v/v) containing 4 mM of 4-methylvaleric acid as internal standard was added to the Eppendorf tubes holding 20 mL mouse plasma and the tubes were vortexed for 30 seconds before incubating for 10 minutes on a thermomixer (Eppendorf) at 10 °C with 950 rpm. Insoluble material was removed by centrifugation for 5 minutes at 4 °C at 16,200 g (Beckman Coulter Microfuge 22R refrigerated microcentrifuge) and the supernatant transferred to a fresh tube. SCFA standards were prepared from stock solutions (1mM) in acetonitrile. Calibrators (0.1 to 50 mM) were made in acetonitrile:water (1:1, v/v) from serial dilutions of the standards. A small portion (40 mL) of the supernatant and calibrators was derivatized using 3-nitrophenylhydrazine (NPH) and 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide(EDC), which converts SCFA to their 3-nitrophenylhydrazones. The reaction was stopped using 20 mL of 200 mM quinic acid dissolved in acetonitrile:water (1:1, v/v). The samples were then spiked with 20 mL of 10 mM isotope-labelled internal standard mix prepared in accordance with the previously described protocol (28). Finally, the samples (1 mL) were then injected into an Agilent 6490 triple quadruple mass spectrometer using the conditions previously described (28).

One-way ANOVA was performed to test whether the concentration of acetate, propionate and butyrate differed between WT and HD mice plasma. Significance threshold was set at 0.1.

Results

We present longitudinal fecal microbiome data for WT and HD mice at 4, 6, 8, 10 and 12 weeks of age to determine the onset of gut dysbiosis. A total of 20928 species were detected, and those above a relative abundance of 0.001% were included for further analysis. The remaining 2024 species were not affected by cage or sequencing run effects.

As expected, *Bacteroidetes* and *Firmicutes* were the two most abundant bacteria phylum in the mouse fecal microbiome followed by *Proteobacteria*, *Actinobacteria*, and *Verrucomicrobia*. We examined the phylum composition at 5 different time points from the shotgun metagenomics sequencing data (Fig. 1A-D). Two-way repeated ANOVA revealed significant effects of age on the relative abundances of *Bacteroidetes* (Age:p=0.0019, Geno:p=0.5865), *Firmicutes* (Age:p=0.0017, Geno:p=0.9378), and *Proteobacteria* (Age:p=0.0001, Geno:p=0.7592). Overall, there were no significant differences in the relative abundances of *Bacteroidetes* and *Firmicutes* between WT mice and HD mice at any of the time points. Similarly, we observed no significant differences in the relative abundances of different bacterial families when comparing WT and HD mice at any of the time points (data not shown).

Subtle differences in bacterial species composition at 12 weeks of age

To further characterize how the phylogenetic and functional differences observed above change with age, we performed a principal coordinates analysis (PCoA) of the obtained taxonomic and gene profiles respectively (Fig. 2). We did not observe any strong differences between samples according to their genotype at 4, 6, 8, and 10 weeks of age. However, at 12 weeks of age, which is prior to overt motor symptom onset, PERMANOVA testing revealed significant effects of genotype in the bacterial species composition based on Bray-Curtis distance (p=0.029).

For alpha diversity measures, we observed no significant effects of age or genotype on the number of bacterial species observed (Two-way repeated-measures ANOVA, Fig. 1E). Age had a significant effect (p=0.006) but not genotype based on the Inverse Simpson index (Fig. 1F).

In terms of beta diversity analysis, Bray-Curtis index indicated significant differences in HD mouse gut microbial structure when compared to their WT littermates at 12 weeks of age (Fig. 2). Thus, we performed sPLS-DA to determine the specific bacterial species that discriminate between the two groups. In total, 50 bacterial species were selected as a signature (classification error rate of 0.3). The differentially abundant bacterial species were found to be *Clostridium mt 5*, *Treponema phagedenis*, *Clostridium leptum CAG:27*, *Desulfatirhabdium butyrativorans*, *Plasmodium chabaudi*, *Defulfuribacillus alkaliarsenatis*, *Plasmodium yoelii* and *Chlamydia abortus*.

The known commensal butyrate producers, *Faecalibacterium praunitizii* and *Eubacterium hallii*, were more abundant in the gut of HD mice than the WT mice at 12 weeks of age. *Akkermansia muciniphila*, another well-known species of commensal bacteria was less abundant in the HD gut only at 12 weeks of

age. There were no differences in the abundances of other butyrate producers including *Roseburia intestinalis*, *Clostridium symbiosum* and *Eubacterium rectale*.

The remainder of the bacterial species we identified were not well annotated, and thus we were unable to identify their potential effects on gut dysbiosis.

The gut microbiome of HD mice is functionally distinct from their WT littermates at 12 weeks of age

One advantage of whole genome shotgun sequencing is that it allows the profiling of microbial genes to interrogate the function of the gut microbiome, thus enabling us to determine the specific functional effects of the gut microbial composition alteration in the HD mouse. We identified 333 genes above the cut-off relative abundance of 0.1%, corresponding to 245 KEGG orthologs (KOs). We found no effects of cage and sequencing run on the bacterial genes and KOs, so we proceeded to the subsequent analysis without further filtering.

For the KO analysis, the most abundant were “aminoacyl-tRNA biosynthesis”, followed by “alanine, aspartate and glutamate metabolism” and “ABC transporters and purine metabolism”. Focusing on the 12-week time point, sPLS-DA identified a KO signature that characterized the two genotype groups. Five pathways were identified which included galactose metabolism, sulfur metabolism, lysine degradation, glutathione metabolism and butanoate metabolism (classification error rate of 0.2, Fig. 3).

For the gene analysis, sPLS-DA identified 20 genes which form the signature to discriminate the WT and the HD mouse gut (classification error rate of 0.27, Fig. 4). The most discriminant genes were *pyre*, *pepD*, *oadB*, *gmd* which were also identified in the pathways from the sPLS-DA KO analysis (Fig. 4).

Modelling of bacterial species and KO trajectories revealed significant volatility in the gut of HD mice

As HD is a progressive neurodegenerative disease, we sought to investigate whether there were any potential effects in the stability of gut microbial structure as well as the function from early adolescence (4 weeks) to the adult stage of life (12 weeks).

LMMS models categorized each species or KO into one of the four different models denoted as 0, indicating a relatively flat profile, 1 indicating either increasing or decreasing over time, 2 and 3 indicating complex curves across time (Fig. 5). After filtering of the noisy profiles, models for 1708 species, 124 KOs and 1228 genes were able to be fitted for both HD and WT mice.

This analysis revealed that in the gut of WT mice, the temporal profile of the bacterial species, genes and KOs remains largely stable throughout adolescence and adult stage: 98%, 97.6% and 100% respectively were categorized into Model 0 (Fig. 5). In the gut of HD mice, the majority of the temporal profiles of bacterial species, genes and KOs were unstable: only 23.5%, 41% and 39.5% respectively were categorized into Model 0 whilst the rest were largely categorized into Model 1 (74.9%, 57.3%, and 58.1% respectively) and the remainder were categorized into Model 2 (Fig. 3). The median proportionality distance showed that most of the members within each model were strongly associated (distance close to 0, Supplementary Table 1).

Targeted metabolomic profiling of plasma revealed distinct differences between HD and WT mice at 12 weeks of age

We further investigated the potential functional differences of the gut microbiome composition in the circulatory system which could be one of the pathways of bidirectional communication between the brain and the gut microbiome. Due to the nature of the amount of blood required, and the progressive nature of the gut dysbiosis, only the plasma metabolites at 12 weeks of age were examined, following terminal cardiac blood collection. In total, the platform identified 221 metabolites, which belong to the following broad categories: amino acids, carbohydrates, lipids, nucleotides, peptides and xenobiotics. sPLS-DA selected 15 metabolites as a signature which could distinguish the samples based on their genotype (Fig. 6A). The top metabolites selected included adenosine triphosphate (ATP), 3-Methylhistidine, urocanic acid, carnosine, threonic acid, homocitrulline, orotic acid, ADP, p-Aminobenzoic acid, 2-methylbutyrylglycine, trigonelline, alpha-hydroxyisobutyric acid, propionic acid, butyric acid, pipercolic acid, 2-hydroxybutyric acid, isobutyrylglycine, 2-hydroxy-3-methylbutyric acid and ribitol (Fig. 6B).

Significant reduction of propionate and butyrate in the plasma of HD mice compared to their WT littermates

The gut function analysis revealed that butanoate metabolism was affected in the HD gut, hence, we sought to determine a possible change of SCFAs in the blood of the HD mice. Even though SCFA data was detected by the targeted metabolomics, the majority of the SCFAs in the circulatory system would

have been metabolized by the liver and a minimal amount would have passed through to the bloodstream. Therefore, the targeted metabolomics approach may not accurately determine the concentration of these SCFAs. Hence, we conducted an independent SCFA assay to validate the SCFA levels in the plasma. Decreases in the plasma concentration of propionate ($p = 0.09$) and butyrate ($p = 0.09$, one-way ANOVA) were observed in the HD mice when compared to their WT littermates, but no differences in the concentration of acetate were observed (Fig. 6C).

Integration of plasma metabolomics and shotgun metagenomic sequencing data

Our integrative analysis with DIABLO identified a signature of 30 bacterial species, 20 genes and 8 metabolites that were highly associated (correlation > 0.7 after variable selection, Fig. 7A). Further visualization in Cytoscape revealed that many microbes were found to be highly correlated with butyrate, homocitrulline, ATP, L-asparagine, 3-methylhistidine, orotic acid, and isobutyrylglycine.

We further narrowed our analysis to the metabolites with known functions in the nervous system, including ATP, butyrate and pipercolic acid, and the associated bacterial species which have been classified. Closer inspection of each network revealed many overlaps in the bacterial species between the three networks. Several *Bacteroides* species including *B.pyogenes*, *B.ihuae*, *B.oleiciplenus* and *B.timonensis* were positively correlated with butyrate but negatively correlated with ATP and pipercolic acid (Fig. 7B). Similarly, we identified *Prevotella ruminicola* and *Odoribacter laneus* to be positively correlated with butyrate levels and negatively correlated with ATP and pipercolic acid.

Blautia producta was uniquely identified to be negatively correlated with butyrate. *Prevotella scopos* was identified to be negatively correlated with ATP levels and no associations were found with butyrate and pipercolic acid.

Discussion

The gut microbiome is often noted for its ecological stability (29–31). Individuals may carry different microbial species from each other, however any one individual tends to carry the same key set of species for long periods (32). This stability is considered critical for host health and well-being because it ensures that beneficial symbionts and their associated functions are maintained over time (29). Similar to humans, the gut microbiome in mice achieves a stable state 15 days post weaning (33). Our study allows for deeper investigation into the temporal stability of the gut microbiome via longitudinal shotgun

sequencing, and the host gene-microbiome interaction from adolescence to adulthood in both HD and WT mice. Our HD mice displayed a high intra-temporal instability in the microbiome structure based on LMMS splines analysis, and even more strikingly, the functional attributes of the microbiome also fluctuated compared to the WT littermates. High volatility of the gut microbiome has been reported in patients with irritable bowel syndrome (34) and colorectal cancer (35). Fluctuations in the microbiome function during the post-weaning developmental period (adolescence) may have irreversible detrimental effects on neurodevelopment during growth.

At 12 weeks of age in the HD mice, prior to overt motor dysfunction, we found significant changes in microbial composition as indicated by Bray-Curtis dissimilarity distance. In addition, functional metagenomics analysis revealed downregulation of galactose metabolism whose temporal profile is similar to the weight-loss trajectory of the mice (data not shown), indicating that decreased capacity of the gut in breaking down food sources into absorbable components by the host. Interestingly, butanoate metabolism was increased at 12 weeks of age, which is mainly driven by the overexpression of the butyrate kinase, an enzyme catalyzing the conversion of butyryl-CoA to butyrate (36), in the HD gut. The known commensal butyrate producers: *Faecalibacterium prausnitzii* and *Eubacterium hallii*, were more abundant in the HD gut than the WT counterparts at 12 weeks of age, although *Akkermansia muciniphila* was slightly decreased at 12 weeks of age in the HD mice (data not shown).

Butyrate is the preferred source of energy by the colonocytes with a plethora of roles in promoting gut integrity through tight junction regulation, cell proliferation and stimulation of mucin production by Goblet cells (37) as well as anti-inflammatory effects (38). It is possible that increased SCFA concentration in the gut lumen is toxic to the host. One study has shown that excessive butyrate is cytotoxic and may induce severe epithelial cell apoptosis and disrupt the intestinal barrier using (39).

In addition, commensal butyrogenic bacteria typically utilizes pyruvate as a source for butyrate production whereas pathogenic ones utilize other pathways including lysine, glutarate and 4-aminobutyrate as sources for butyrate production. *Fusobacterium* is known as an oral pathogen but its role in the gut is rather unclear, as it produces butyrate by using lysine as a source and ammonia is released as a by-product (40–42). This by-product could change the gut lumen pH and could be potentially damaging to gut health. Indeed, we found that the main *Fusobacterium* species, *F.nucleatum* and *F.ulcerans* were both more abundant in the HD gut at 12 week of age (data not shown).

However, the HD blood plasma does not reflect this increase in butanoate metabolism in the gut, as we found butyrate concentration to be lower in the HD plasma. Excess butyrate not taken up by the

colonocytes will be taken up by the liver through the hepatic portal vein (43). Perhaps there is a malfunction in the absorption by either the liver or the colonocytes, which would have system-wide effects throughout the host.

We also found some similarities as well as dissimilarities in our plasma metabolome data with other published findings. Our plasma metabolomics data and gut functional profiling in HD mice corroborated with a report by Underwood et al. (44), showing pro-catabolic phenotype in early HD prior to symptom onset. In a human HD plasma metabolomics study, they reported a significant difference in levels of serotonin which was not differentially abundant in these HD mice at 12 weeks of age (data not shown) (45). ATP was increased by over 4-fold in the plasma, which was opposite to expectations since mitochondrial dysfunction is prominent in HD and hence, impaired ATP production, suggesting that the gut microbiome could contribute to the bulk of the changes found in the plasma (46,47). Indeed, bacteria could release ATP via mechanosensitive channels or stimulate ATP production by providing SCFA as food source for the colonocytes (48,49).

In our multi-omics integration analysis, we focused on ATP and pipercolic acid, in addition to butyrate, since there is evidence supporting their role in CNS and they can be microbially derived (12,48–54). Matsumoto and colleagues have shown evidences of the contribution of gut microbiome to the cerebral levels of pipercolic acid and ATP using germ-free mice, however it is unclear which specific bacterial species contributed to those observations (12). DIABLO analysis highlighted several key species overlapping between the three networks. *Bacteroides* species including *B.pyogenes*, *B.timonensis*, *B.oleiciplenus* as well as *Odoribacter laneus* and *Prevotella ruminicola*, were found to be highly correlated with all three metabolites. *Odoribacter laneus* is a common constituent of the human fecal microbiome although no functions, to this date, have been tied to it (55). *Prevotella ruminicola* has known functions in SCFA production and was found to be reduced in our fecal samples (56). On the other hand, we found negative correlations between *Blautia producta* and plasma butyrate levels. One report had shown that supplementation of *Blautia producta* together with other potential commensal species in germ-free rats increased the fecal concentration of butyrate (57). It is possible that the distal gut's functional output is not entirely reflected in the plasma despite a high correlation between the two datasets (correlation coefficient of 0.76). Nevertheless, this approach could help in untangling the complicated cross-talk between the host and the microbiome as well as identifying potential candidates for a probiotic intervention as part of future treatments for HD.

Our data demonstrated that the gut dysbiosis in HD occurs prior to overt motor and cognitive deficits and may play a role, via metabolite release into the circulatory system, in modulating pathogenesis and

disease progression. Furthermore, these novel findings identify potential therapeutic targets that can be pursued in future studies.

Conclusions

In the present study, we have discovered a previously unknown relationship between the gut bacteria and plasma metabolome, suggesting the potential role of gut in modulating the pathogenesis of HD via specific altered plasma metabolites which mediate gut-brain signaling. Specific changes in the gut microbiome of the HD mice may contribute to the onset of pathogenesis and disease progression. Both the metagenomic findings from the gut microbiome and the metabolomics, as well as their integrated analyses, have implications for our understanding of peripheral and central pathology in HD, and may pave the way for future development of novel therapeutic approaches.

Declarations

Acknowledgments

We thank members of the Hannan Laboratory for useful discussions regarding experimental design and analyses. We would also like to acknowledge Metabolomics Australia which is funded by Bioplatforms Australia through the Australian Government National Collaborative Research Infrastructure Strategy (NCRIS).

ETHICS APPROVAL

All experiments were approved by the Florey Institute of Neuroscience and Mental Health Animal Ethics Committee and performed in accordance with research guidelines of the National Health and Medical Research Council.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

All data will be made publicly available on publication.

COMPETING INTERESTS

The authors declare that they have no competing interests.

FUNDING

The shotgun sequencing was funded by a Computational Biology Research Initiative (CBRI) Grant to KALC and AJH, from the University of Melbourne. KALC is an NHMRC Career Development Fellow (GNT1159458). TR is an NHMRC-ARC Dementia Research Fellow. AJH is an NHMRC Principal Research Fellow. None of the funding bodies had any role in the design of the study or collection, analysis, and interpretation of data, or in writing the manuscript.

AUTHORS' CONTRIBUTIONS

GK carried out the mouse studies, fecal and blood plasma collections, and associated experimental work, as well as the bioinformatics, and drafted the manuscript. SE contributed to the bioinformatic analyses. VKN and KK contributed to the metabolomic studies. HTTC and SL contributed to mouse husbandry and handling, tissue and sample collections. TR, KALC and AJH conceived of the study, and participated in its design, coordination and implementation, and helped to draft the manuscript. All authors read and approved the final manuscript.

References

1. Moffitt H, McPhail GD, Woodman B, Hobbs C, Bates GP. Formation of polyglutamine inclusions in a wide range of non-CNS tissues in the HdhQ150 knock-in mouse model of Huntington's disease. *PLoS One*. 2009;4(11):e8025.
2. Gubert C, Kong G, Renoir T, Hannan AJ. Exercise, diet and stress as modulators of gut microbiota: Implications for neurodegenerative diseases. *Neurobiol Dis*. 2020 Feb 1;134:104621.
3. Giongo A, Gano KA, Crabb DB, Mukherjee N, Novelo LL, Casella G, et al. Toward defining the autoimmune microbiome for type 1 diabetes. *ISME J*. 2011;5(1):82.
4. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature*. 2006;444(7122):1027.
5. Sampson TR, Debelius JW, Thron T, Janssen S, Shastri GG, Ilhan ZE, et al. Gut microbiota regulate motor deficits and neuroinflammation in a model of Parkinson's disease. *Cell*. 2016;167(6):1469–80.
6. Scheperjans F, Aho V, Pereira PA, Koskinen K, Paulin L, Pekkonen E, et al. Gut microbiota are related to Parkinson's disease and clinical phenotype. *Mov Disord*. 2015;30(3):350–8.

7. Vogt NM, Kerby RL, Dill-McFarland KA, Harding SJ, Merluzzi AP, Johnson SC, et al. Gut microbiome alterations in Alzheimer's disease. *Sci Rep*. 2017 Oct 19;7(1):13537.
8. Kong G, Lê Cao K-A, Judd LM, Li S, Renoir T, Hannan AJ. Microbiome profiling reveals gut dysbiosis in a transgenic mouse model of Huntington's disease. *Neurobiol Dis*. 2018;
9. Radulescu CI, Garcia-Miralles M, Sidik H, Bardile CF, Yusof NABM, Lee HU, et al. Manipulation of microbiota reveals altered callosal myelination and white matter plasticity in a model of Huntington disease. *Neurobiol Dis*. 2019;127:65–75.
10. Kobal J, Matej K, Koželj M, Podnar S. Anorectal dysfunction in presymptomatic mutation carriers and patients with Huntington's disease. *J Huntingt Dis*. 2018;7(3):259–67.
11. Forsythe P, Bienenstock J, Kunze WA. Vagal pathways for microbiome-brain-gut axis communication. In: *Microbial Endocrinology: The Microbiota-Gut-Brain Axis in Health and Disease*. Springer; 2014. p. 115–33.
12. Matsumoto M, Kibe R, Ooga T, Aiba Y, Sawaki E, Koga Y, et al. Cerebral low-molecular metabolites influenced by intestinal microbiota: a pilot study. *Front Syst Neurosci*. 2013;7:9.
13. Wikoff WR, Anfora AT, Liu J, Schultz PG, Lesley SA, Peters EC, et al. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc Natl Acad Sci*. 2009 Mar 10;106(10):3698.
14. Cummings J, Pomare E, Branch W, Naylor C, Macfarlane G. Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut*. 1987;28(10):1221–7.
15. Kidd SK, Schneider JS. Protection of dopaminergic cells from MPP⁺-mediated toxicity by histone deacetylase inhibition. *Brain Res*. 2010 Oct 1;1354:172–8.
16. Laurent RS, O'Brien LM, Ahmad ST. Sodium butyrate improves locomotor impairment and early mortality in a rotenone-induced *Drosophila* model of Parkinson's disease. *Neuroscience*. 2013;246:382–90.
17. Kilgore M, Miller CA, Fass DM, Hennig KM, Haggarty SJ, Sweatt JD, et al. Inhibitors of class 1 histone deacetylases reverse contextual memory deficits in a mouse model of Alzheimer's disease. *Neuropsychopharmacology*. 2010;35(4):870–80.
18. Menzel P, Ng KL, Krogh A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun*. 2016;7:11257.
19. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS One*. 2013;8(4):e61217.
20. Dixon P. VEGAN, a package of R functions for community ecology. *J Veg Sci*. 2003;14(6):927–30.
21. Le Cao K-A, Costello M-E, Lakis VA, Bartolo F, Chua X-Y, Brazeilles R, et al. MixMC: a multivariate statistical framework to gain insight into microbial communities. *PloS One*. 2016;11(8).
22. Glass EM, Wilkening J, Wilke A, Antonopoulos D, Meyer F. Using the metagenomics RAST server (MG-RAST) for analyzing shotgun metagenomes. *Cold Spring Harb Protoc*. 2010;2010(1):pdb-prot5368.

23. Bodein A, Chapleur O, Droit A, Lê Cao K-A. A generic multivariate framework for the integration of microbiome longitudinal studies with other data types. *Front Genet.* 2019;10.
24. Straube J, Gorse A-D, Huang BE, Le Cao K-A. A linear mixed model spline framework for analysing time course 'omics' data. *PLoS One.* 2015;10(8):e0134540.
25. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, et al. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics Off J Metabolomic Soc.* 2007 Sep;3(3):211–21.
26. Rohart F, Gautier B, Singh A, Lê Cao K-A. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol.* 2017;13(11):e1005752.
27. Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics.* 2010;27(3):431–2.
28. Han J, Lin K, Sequeira C, Borchers CH. An isotope-labeled chemical derivatization method for the quantitation of short-chain fatty acids in human feces by liquid chromatography–tandem mass spectrometry. *Anal Chim Acta.* 2015;854:86–94.
29. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience of the human gut microbiota. *Nature.* 2012 Sep 13;489(7415):220–30.
30. Voreades N, Kozil A, Weir TL. Diet and the development of the human intestinal microbiome. *Front Microbiol.* 2014;5:494.
31. Claesson MJ, Cusack S, O'Sullivan O, Greene-Diniz R, de Weerd H, Flannery E, et al. Composition, variability, and temporal stability of the intestinal microbiota of the elderly. *Proc Natl Acad Sci.* 2011;108(Supplement 1):4586–91.
32. Vanhoutte T, Huys G, de Brandt E, Swings J. Temporal stability analysis of the microbiota in human feces by denaturing gradient gel electrophoresis using universal and group-specific 16S rRNA gene primers. *FEMS Microbiol Ecol.* 2004 Jun 1;48(3):437–46.
33. Schloss PD, Schubert AM, Zackular JP, Iverson KD, Young VB, Petrosino JF. Stabilization of the murine gut microbiome following weaning. *Gut Microbes.* 2012;3(4):383–93.
34. Halfvarson J, Brislawn CJ, Lamendella R, Vázquez-Baeza Y, Walters WA, Bramer LM, et al. Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat Microbiol.* 2017 Feb 13;2:17004–17004.
35. Scanlan PD, Shanahan F, Clune Y, Collins JK, O'Sullivan GC, O'Riordan M, et al. Culture-independent analysis of the gut microbiota in colorectal cancer and polyposis. *Environ Microbiol.* 2008 Mar 1;10(3):789–98.
36. Louis P, Duncan SH, McCrae SI, Millar J, Jackson MS, Flint HJ. Restricted distribution of the butyrate kinase pathway among butyrate-producing bacteria from the human colon. *J Bacteriol.* 2004;186(7):2099–106.
37. Morrison DJ, Preston T. Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. *Gut Microbes.* 2016;7(3):189–200.

38. den Besten G, van Eunen K, Groen AK, Venema K, Reijngoud D-J, Bakker BM. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J Lipid Res.* 2013;54(9):2325–40.
39. Peng L, He Z, Chen W, Holzman IR, Lin J. Effects of Butyrate on Intestinal Barrier Function in a Caco-2 Cell Monolayer Model of Intestinal Barrier. *Pediatr Res.* 2007 Jan 1;61(1):37–41.
40. Vital M, Howe AC, Tiedje JM. Revealing the Bacterial Butyrate Synthesis Pathways by Analyzing (Meta)genomic Data. Moran MA, editor. *mBio.* 2014 May 1;5(2):e00889-14.
41. Anand S, Kaur H, Mande SS. Comparative In silico Analysis of Butyrate Production Pathways in Gut Commensals and Pathogens. *Front Microbiol.* 2016;7:1945.
42. Barker HA, Kahn JM, Hedrick L. Pathway of lysine degradation in *Fusobacterium nucleatum*. *J Bacteriol.* 1982 Oct 1;152(1):201.
43. van der Beek CM, Bloemen JG, van den Broek MA, Lenaerts K, Venema K, Buurman WA, et al. Hepatic uptake of rectally administered butyrate prevents an increase in systemic butyrate concentrations in humans. *J Nutr.* 2015;145(9):2019–24.
44. Underwood BR, Broadhurst D, Dunn WB, Ellis DI, Michell AW, Vacher C, et al. Huntington disease patients and transgenic mice have similar pro-catabolic serum metabolite profiles. *Brain.* 2006 Feb 7;129(4):877–86.
45. Rosas HD, Doros G, Bhasin S, Thomas B, Gevorkian S, Malarick K, et al. A systems-level “misunderstanding”: the plasma metabolome in Huntington’s disease. *Ann Clin Transl Neurol.* 2015 Jul 1;2(7):756–68.
46. Cui L, Jeong H, Borovecki F, Parkhurst CN, Tanese N, Krainc D. Transcriptional repression of PGC-1 α by mutant huntingtin leads to mitochondrial dysfunction and neurodegeneration. *Cell.* 2006;127(1):59–69.
47. Jenkins BG, Koroshetz WJ, Beal MF, Rosen BR. Evidence for impairment of energy metabolism in vivo in Huntington’s disease using localized ¹H NMR spectroscopy. *Neurology.* 1993;43(12):2689–2689.
48. Firmansyah A, Penn D, Lebenthal E. Isolated colonocyte metabolism of glucose, glutamine, n-butyrate, and β -hydroxybutyrate in malnutrition. *Gastroenterology.* 1989;97(3):622–9.
49. Donohoe DR, Garge N, Zhang X, Sun W, O’Connell TM, Bunger MK, et al. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab.* 2011;13(5):517–26.
50. Ivanova EP, Alexeeva YV, Pham DK, Wright JP, Nicolau DV. ATP level variations in heterotrophic bacteria during attachment on hydrophilic and hydrophobic surfaces. *Int Microbiol.* 2006;9(1):37.
51. Takahama K, Hashimoto T, Wang M-W, Akaike N, Hitoshi T, Okano Y, et al. Pipecolic acid enhancement of GABA response in single neurons of rat brain. *Neuropharmacology.* 1986;25(3):339–42.
52. Matsumoto S, Yamamoto S, Sai K, Maruo K, Adachi M, Saitoh M, et al. Pipecolic acid induces apoptosis in neuronal cells. *Brain Res.* 2003 Aug 8;980(2):179–84.

53. Edwards FA, Gibb AJ, Colquhoun D. ATP receptor-mediated synaptic currents in the central nervous system. *Nature*. 1992;359(6391):144–7.
54. Proietti M, Perruzza L, Scribano D, Pellegrini G, D'Antuono R, Strati F, et al. ATP released by intestinal bacteria limits the generation of protective IgA against enteropathogens. *Nat Commun*. 2019 Jan 16;10(1):250.
55. Nagai F, Morotomi M, Watanabe Y, Sakon H, Tanaka R. *Alistipes indistinctus* sp. nov. and *Odoribacter laneus* sp. nov., common members of the human intestinal microbiota isolated from faeces. *Int J Syst Evol Microbiol*. 2010;60(6):1296–302.
56. Strobel HJ. Vitamin B12-dependent propionate production by the ruminal bacterium *Prevotella ruminicola* 23. *Appl Environ Microbiol*. 1992 Jul 1;58(7):2331.
57. Becker N, Kunath J, Loh G, Blaut M. Human intestinal microbiota: characterization of a simplified and stable gnotobiotic rat model. *Gut Microbes*. 2011;2(1):25–33.

Figures

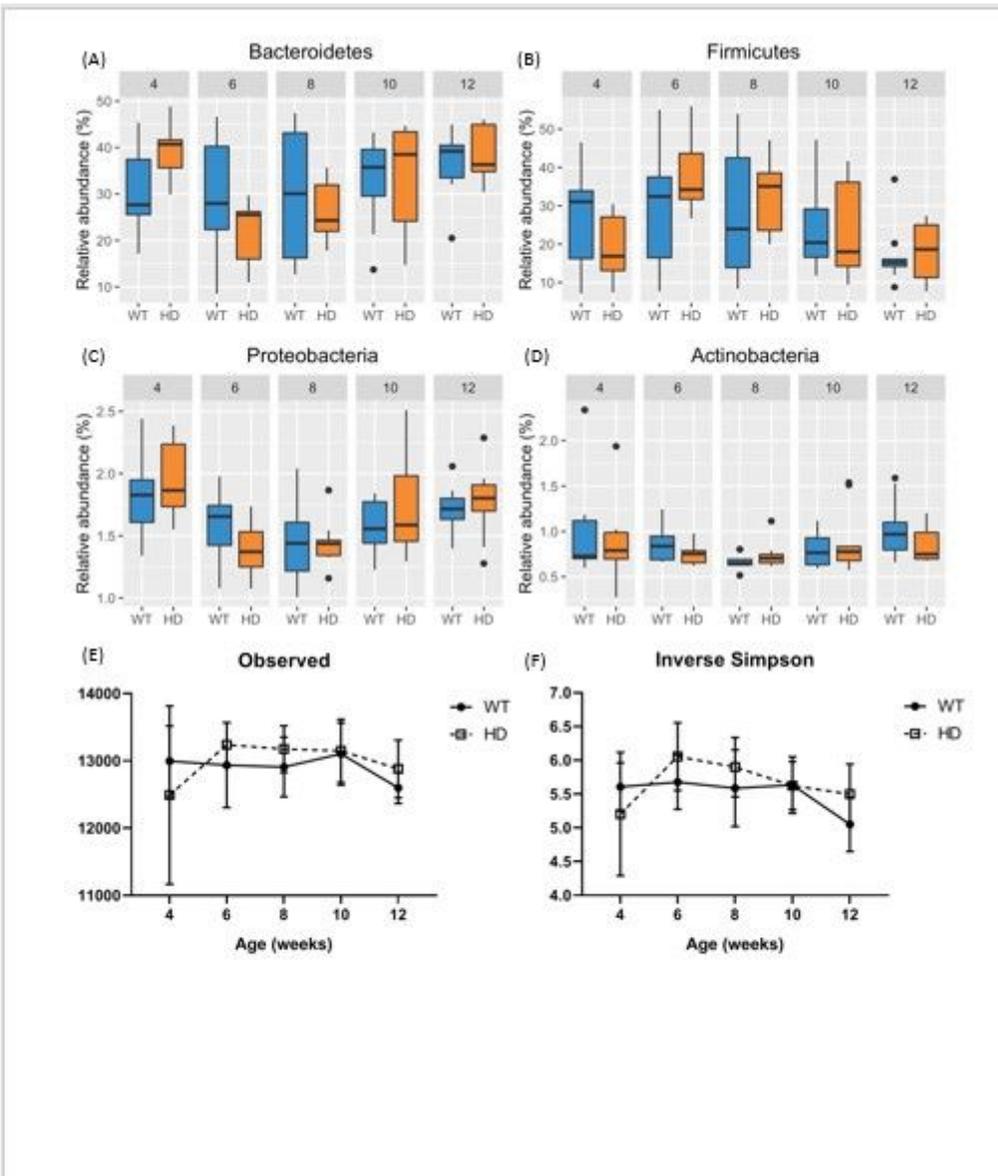


Figure 1

Overview of the gut microbiome in the HD mice and their WT littermates from 4 to 12 weeks of age. Boxplots of the relative abundance of the major phyla present in the mouse gut microbiome: (A) Bacteroidetes, (B) Firmicutes, (C) Proteobacteria and (D) Actinobacteria. The alpha diversity indices, including the (E) number of bacterial species observed and (F) Inverse Simpson index, a measure of evenness, was accounted for from 4 to 12 weeks of age in these mice. There were no significant differences in any of these measures at any timepoints (n = 9 per genotype and time point).

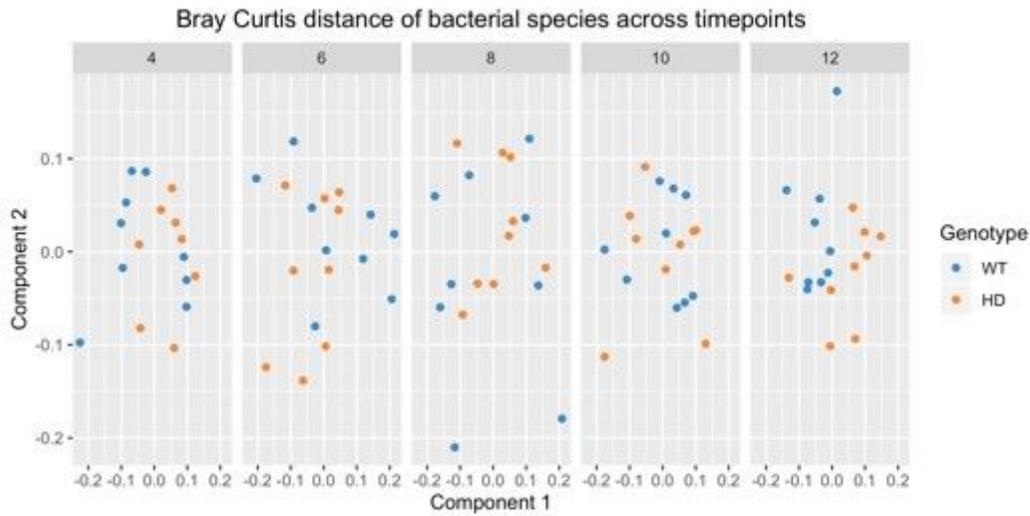


Figure 2

Principal Coordinate Analysis sample plots based on Bray-Curtis dissimilarity distance of bacterial species data on HD mouse gut and their WT littermates per time point (4 to 12 weeks). Adonis PERMANOVA testing revealed significant differences in microbial structure at 12 weeks of age ($p=0.029$) ($n = 9$ per genotype and time point).

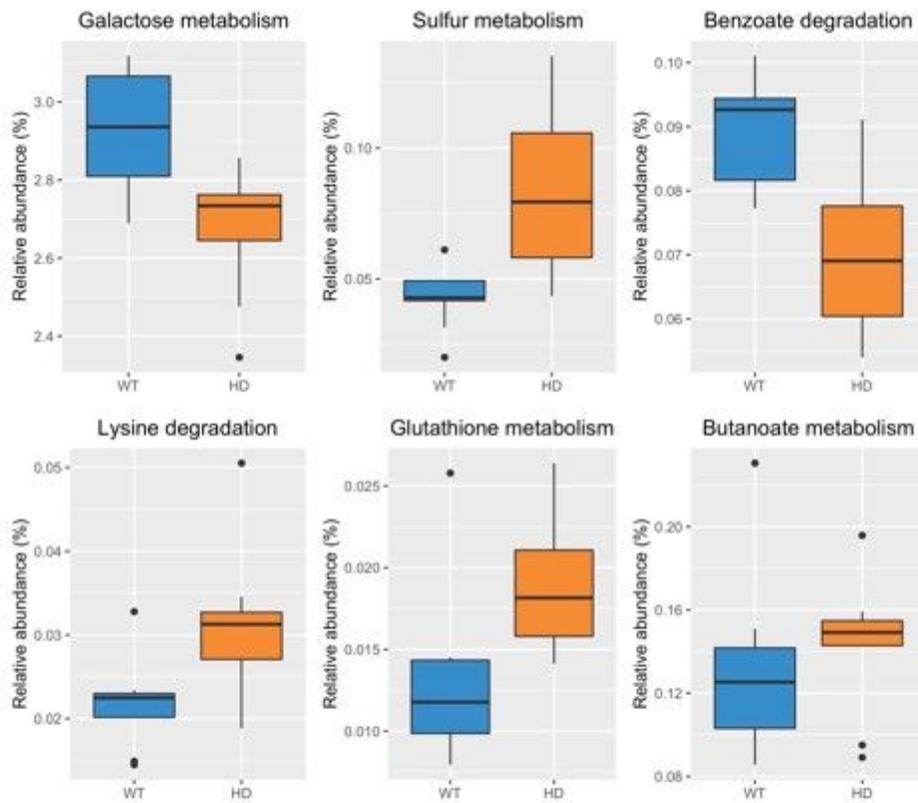


Figure 3

At time point 12 weeks, sPLS-DA analysis was performed on the KO pathways data to determine the signature capable of discriminating the HD mice from their WT littermates. The relative abundances of the top KO pathways selected are shown above. The galactose metabolism and benzoate degradation were decreased in the gut of HD mice whereas sulfur metabolism, lysine degradation, glutathione metabolism and butanoate metabolism were increased in the gut of HD mice (n = 9 per genotype).

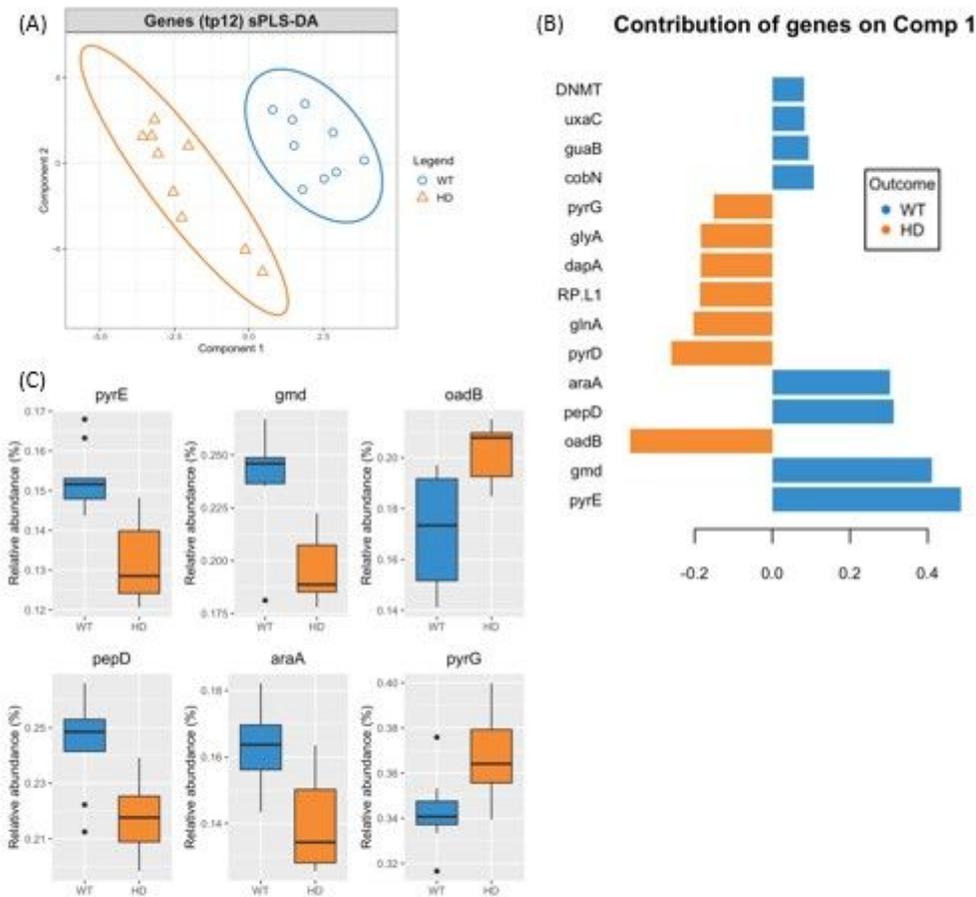


Figure 4

sPLS-DA analysis of genes data HD mice and their WT littermates at 12 weeks of age. (A) Sample plot with 95% confidence ellipses. Strong stratification of samples were observed using the signature of genes determined by sPLS-DA. (B) Contribution of each gene selected by sPLS-DA, color indicates the genotype in which the median relative abundance is maximal, value indicate the contribution to the first component that discriminates the genotypes. (C) Relative abundances of the top genes selected by sPLS-DA. There was a strong stratification of samples observed on the sample plot when only taking into account the abundance of the signature genes. The relative abundances of *pyrE*, *gmd*, *pepD* and *araA* were decreased in the HD gut whilst *oadB* and *pyrG* were increased in HD gut when comparing to WT littermates (n = 9 per genotype).

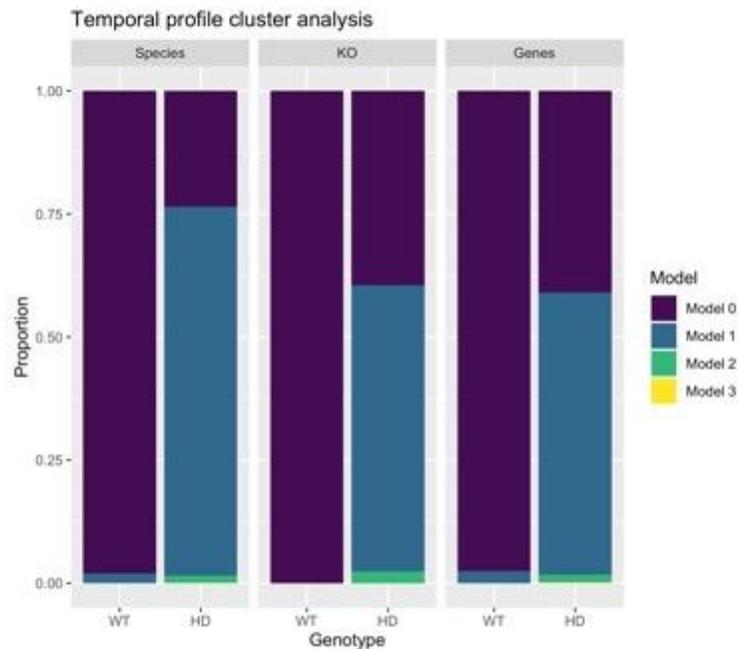


Figure 5

Temporal stability based on LMMS-splines time profile modelling on the bacterial species, KEGG Orthologs (KO) and genes data on R6/1 (HD) mouse gut and their wild-type (WT) littermates. LMMS fits the best model on each species or KOs: Model 0 indicates a relatively flat profile, Model 1 either increasing or decreasing over time, whereas Model 2 indicates complex curves across time. In WT mice, the temporal profile of species, KOs and genes remained largely stable (Model 0: 98%, 97.6%, 100% respectively). In the HD mice, the majority of the temporal profiles of species, KOs and genes were categorized into Model 1 (Model 0: 23.5%, 41%, 39.5% respectively, Model 1: 74.9%, 57.3%, 58.1% respectively) (n=9 per genotype and time point).

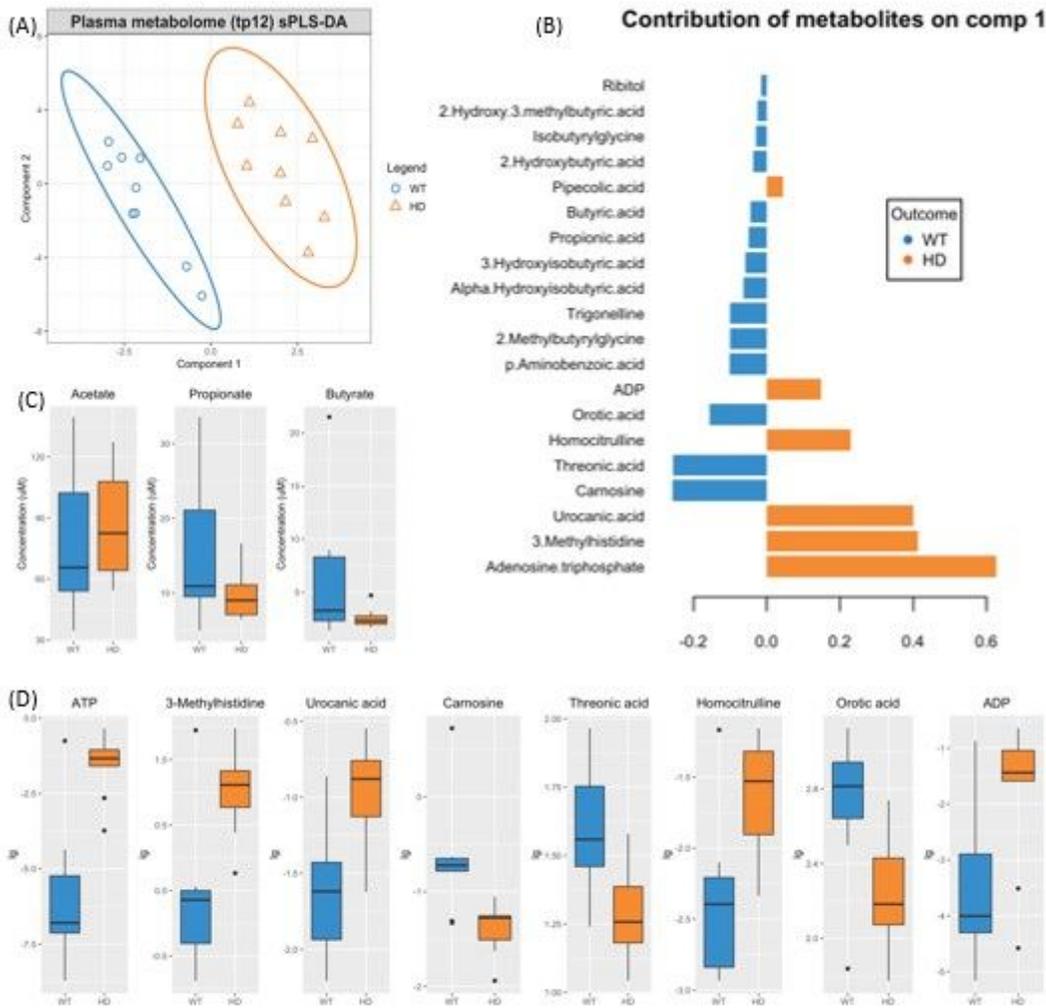


Figure 6

sPLS-DA analysis of plasma metabolites of R6/1 (HD) mice and their wild-type (WT) controls at 12 weeks of age. Plot (A) shows individual samples points classified by sPLS-DA with 95% confidence interval shown by the ellipse. (B) The loading plots of each metabolites selected by sPLS-DA and the amount of contribution made by each gene towards component 1, which is the component clustering the two genotypes. (C) Concentrations of the short-chain fatty acids, acetate, propionate and butyrate. The log abundances of the top metabolites selected by sPLS-DA is shown in (D). Concentration of short-chain fatty acid assay in WT and HD mice validated independently at 12 weeks of age. One-way ANOVA revealed significant decrease in the concentration of propionate and butyrate in the HD plasma when compared to their WT littermates (n = 9 per genotype).

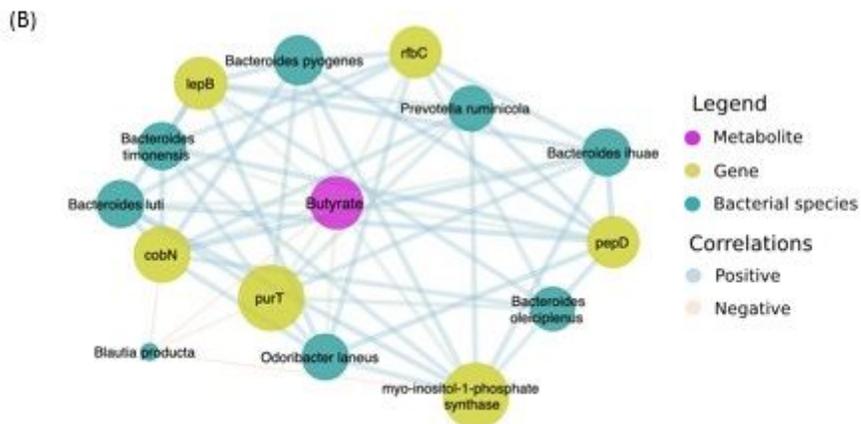
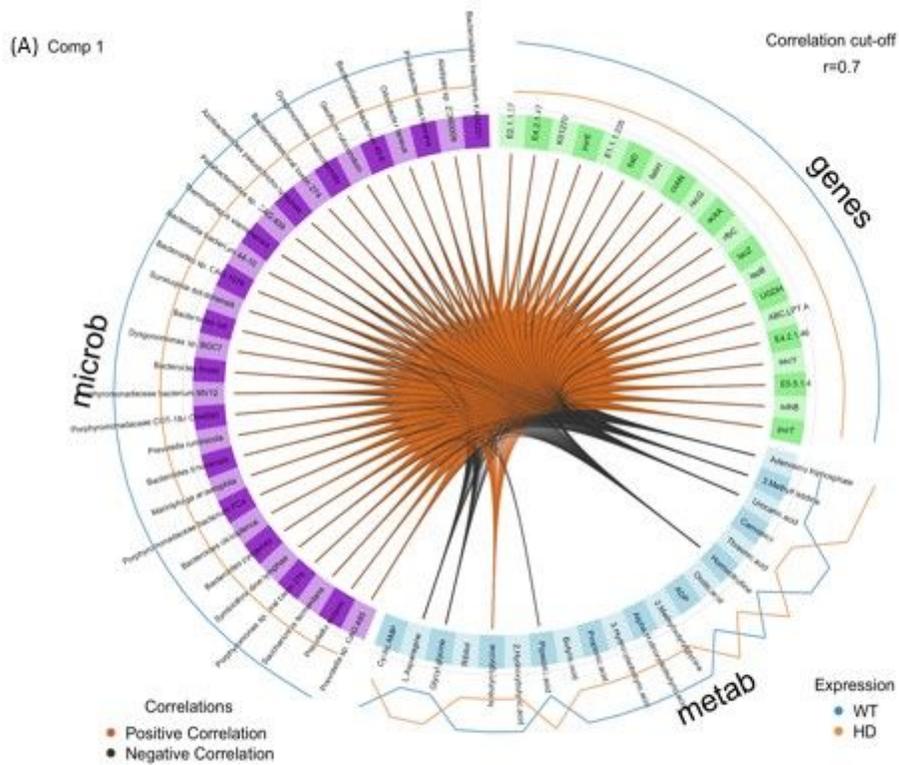


Figure 7

Integration of shotgun metagenomics data, functional metagenomics data and plasma metabolomics data using DIABLO. (A) Circos plot output from DIABLO. The most outer lines show the expression levels in HD mice in orange and WT mice in blue. Each quadrant indicates the type of molecular features, i.e. bacterial species (purple), genes (green), and metabolites (light blue), and lines indicate measure of association (correlation), either positive or negative. (B) Further visualization of the network from DIABLO using Cytoscape, focusing on butyrate and their relationship with classified bacterial species and genes which may contribute to their abundance. The size of the nodes indicate the number of interacting

partners within the network. Many *Bacteroides* species, as well as *Odoribacter laneus* and *Prevotella ruminicola*, were found to correlate positively with butyrate. *Blautia producta* was found to correlate negatively with butyrate.

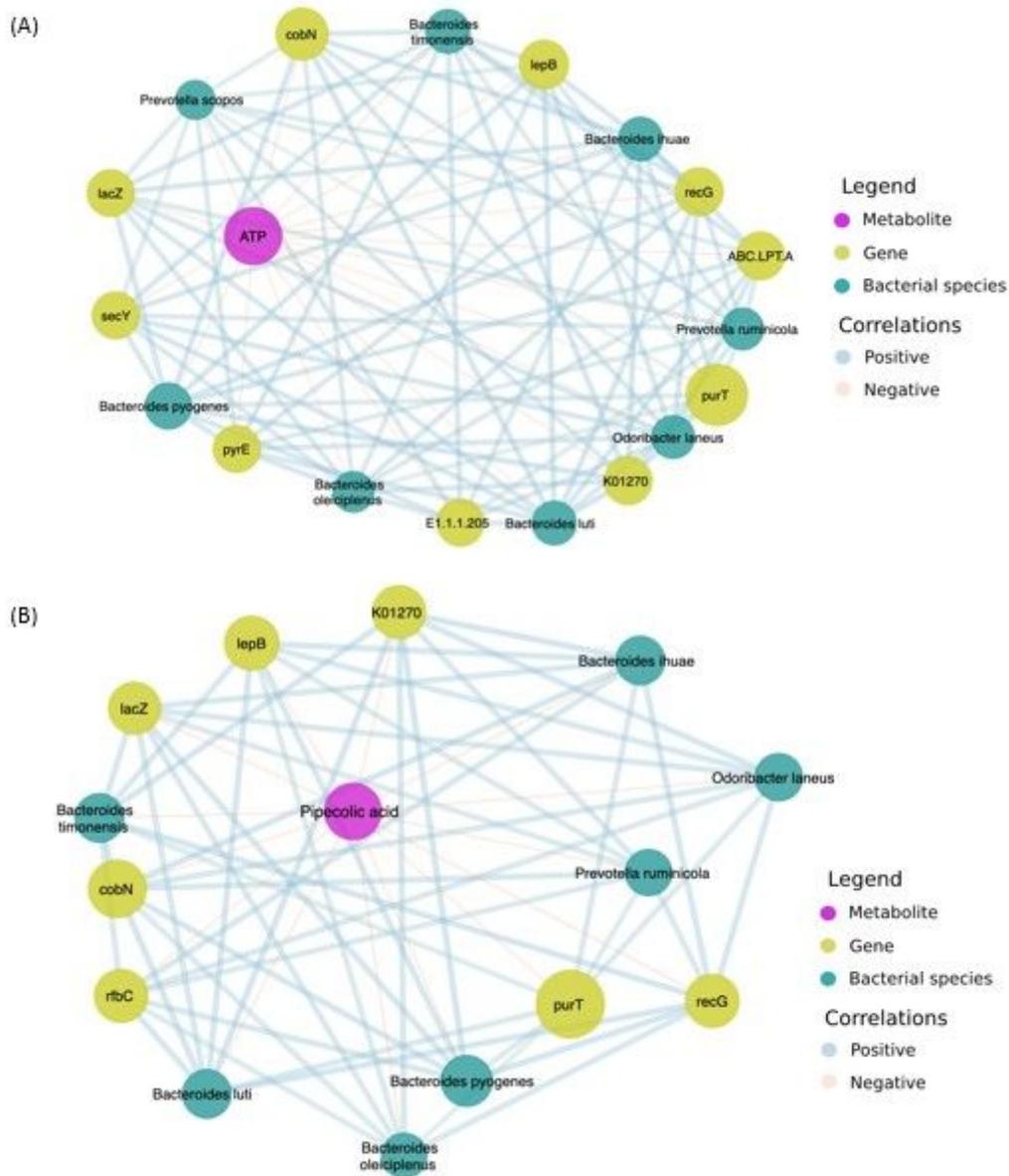


Figure 8

Network analysis from DIABLO using Cytoscape, focusing on (A) ATP and (B) Pipecolic acid, and their relationship with classified bacterial species and genes which may contribute to their abundance. The size of the nodes indicate the number of interacting partners within the network. There were many overlaps between the members of the two networks shown. Many *Bacteroides* species, as well as

Odoribacter laneus and *Prevotella ruminicola*, correlated negatively with ATP and pipecolic acid. In addition, *Prevotella scopos* also correlated negatively with ATP.

Supplementary Files

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

- [SuppTable1WTproportionalitydistanceallgenes0220.htm](#)