ACIBADEM MEHMET ALI AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

MULTI-OMICS INTEGRATION AND ANALYSIS OF METAGENOMICS AND METABOLOMICS DATA FROM RAT MODEL OF ACQUIRED EPILEPSY

ERAY ŞAHİN

PH.D. THESIS

DEPARTMENT OF BIOSTATISTICS AND BIOINFORMATICS

SUPERVISOR
Prof. Uğur Sezerman

ISTANBUL-2024

ACIBADEM MEHMET ALI AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

MULTI-OMICS INTEGRATION AND ANALYSIS OF METAGENOMICS AND METABOLOMICS DATA FROM RAT MODEL OF ACQUIRED EPILEPSY

ERAY ŞAHİN

PH.D. THESIS

DEPARTMENT OF BIOSTATISTICS AND BIOINFORMATICS

SUPERVISOR
Prof. Uğur Sezerman

ISTANBUL-2024

Department:	Biostatistics and Bioinformatics
Program:	Biostatistics and Bioinformatics Ph.D. Program
Thesis Title:	Multi-Omics Integration and Analysis of Metagenomics and Metabolomics Data from Rat Model of Acquired Epilepsy
Student's name and Surname:	Eray Şahin
Date of Defence:	/ /

This is to certify that I have examined this copy of Ph.D thesis. I have found that she/he prepared after fulfilling the specified requirements in the associated legislations before the final examining committee whose signatures are below.

Jury Member (Head of the Defense)	Title, Name Surname Instituition	Signature
Jury Member (Thesis Supervisor)	Title, Name Surname Instituition	Signature
Jury Member	Title, Name Surname Instituition	Signature
Jury Member	Title, Name Surname Instituition	Signature
Jury Member	Title, Name Surname Instituition	Signature

DECLARATION

I declare that this thesis work is my own work, I had no unethical behavior at any stages from the planning to the writing of the thesis, I obtained all the information in this thesis in accordance with academic and ethicel rules, I cited all the information and comments that were not obtained with this thesis work, and I provided resources in the list of references. I also declare that there was no violation of any patents and copyrights during the study and writing of this thesis. Part of this study was published in May 2024 with DOI of 10.1016/j.nbd.2024.106469.

27/06/2024

Eray Şahin

PREFACE AND ACKNOWLEDGEMENT

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. O. Uğur Sezerman, for his invaluable guidance, endless patience, and continuous support throughout my research journey. I would also like to extend my sincere thanks to the jury members, Assoc. Prof. Öznur Taştan, Prof. Günseli Bayram Akçapınar, Prof. Dr. Emel Timuçin, and Assoc. Prof. Burcu Bakır Güngör, Prof. Dr. Nerses Bebek, and Assoc. Prof. Ahmet Can Timuçin for their insightful feedback and valuable contributions that have greatly enhanced the quality of my work. I am profoundly grateful to Prof. Dr. Annamaria Vezzani and Dr. Teresa Ravizza from Milan, Prof. Dr. Pasquale Striano, Dr. Antonella Riva, Greta Volpedo, Andrea Petretto, and Chiara Lavarello from Genova for conducting the pivotal study, having a fruitful collaboration and granting permission to use their data in my thesis. Without their generous support, completing my thesis would not have been possible.

A special thanks to my lab friends, Narod, Hazal, Berk, Nogay, Tayyip, Ali, Tugce, Huseyin Avni and all other Sezerman Lab members, old and new. I am deeply grateful to my current colleagues at CosmosID, but especially BIS Team members, Kelly Moffat, Dana Walsh, Jon Jeun, and Andrew Aradi for their unwavering support, encouragement, and for such a great teamwork. To my friends who are like family, Merve Erdem, Basar Cetin, Yagiz Baris, Ogulcan Habiboglu, Basak Arslan and Gizem Saribiyik, your constant support and encouragement have been invaluable. Thank you for always being there for me, especially during the most trying times.

My heartfelt thanks go to my family members for their unconditional love and support. But especially my aunt and my grandmother, thank you for always believing in me and encouraging me to pursue my dreams, and welcoming me at your home.

I am also immensely grateful to TUBITAK for the scholarship from 2244 project no.118C039 between 2019 and 2023, and for the travel grant to visit University of Copenhagen for six months based on 2214-A International Research Fellowship Programme for PhD Students with application no. 1059B142000810.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

SE	Status epilepticus
KA	Kainic acid
CNS	Central nervous system
DR	Drug-resistant
DS	Drug-sensitive
KD	Ketogenic diet
GABA	Gamma-aminobutyric acid
No-Epi	Kainic acid induced epilepsy rats without seizure development
Epi	Kainic acid induced epilepsy rats with seizure development
MetaPhlAn4	METAgenomic PHyLogenetic Analysis 4
RefSeq	Reference Sequence
MAGs	Metagenome-assembled genomes
GTBD	Genome Taxonomy Database
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	KEGG Orthology
GIFTs	Genome-Inferred Functional Traits
iTOL	Interactive Tree Of Life
KW	Kruskal-Wallis
PERMANOVA	Permutational multivariate analysis of variance
PCoA	Principal coordinate analysis
DA	Differential abundance
MS	Mass spectrometry
HILIC	Hydrophilic interaction liquid chromatography
C18-Pos	C18 positive
C18-Neg	C18 negative
HILIC-Pos	HILIC positive
HILIC-Neg	HILIC negative

ANOVA	Analysis of variance
FDR	False discovery rate
HSD	Honestly significant difference
RF	Random forest
EN	Elastic Net
MOFA	Multi-Omics Factor Analysis
LF	Latent factor
clr	Centered-log ratio
Faith's PD	Faith's phylogenetic diversity
LPC	Lysphosphatidylcholin
abs	Absolute
LPE	Lysophosphatidylethanolamine
SCFA	Short-chain fatty acid

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Multi-Omics Integration and Analysis of Metagenomics and Metabolomics Data from Rat Model of Acquired Epilepsy

Epilepsy is a neurological disorder characterized by repetitive seizures observed due to abnormal alterations in the electrical functioning of the brain with excessive and hypersynchronous neuronal activity. Crosstalk between gut and brain has been revealed to be an important factor in development of the central nervous system. Up to date, numerous studies conducted to reveal role and importance of gut microbiome on variety of human neurological disorders including epilepsy. In this study, shotgun metagenomics sequencing data from faecal samples and untargeted metabolomics data from serum samples obtained from kainic acid-induced pediatric rat model of epilepsy are analyzed to reveal omics fingerprints associated with epilepsy development. A total of three rat subgroups were defined as Sham controls, kainic acid administered rats with and without seizure development as Epi and No-Epi. Untargeted serum metabolomics revealed highest number of significantly different metabolites between Epi vs Sham comparison, which was specifically associated with lipid metabolism. Shotgun metagenomics data analysis showed dysbiosis in both Epi and No-Epi groups compared to Sham controls based on beta diversity, while there was no significant community-wise difference between Epi and No-Epi gut microbiota. A bacterial catalogue of 324 metagenome assembled genomes was established by this study. Taxonomic analyses in phylum, family, genus, and species levels pointed out significant differences in abundances of several bacteria that play role in short chain fatty acids production. Finally, multi-omics integration revealed significant correlations between two bacterial species, annotated as *CAG-873 sp009775195* and unknown species of genus *Angelakisella*, and serum N-Acetylornithine in opposite directions. Taken together, this study provides specific alterations in gut microbiota and serum metabolomics in epilepsy development in a pediatric rat model.

Keywords: Epilepsy, Multi-omics, Metagenomics, Metabolomics

ÖZET

Kazanılmış Epilepsi Sıçan Modelinden Metagenomik ve Metabolomik Verilerin Çoklu Omik Analizi ve Entegrasyonu

Epilepsi, aşırı ve hipersenkron nöronal aktivite ile beynin elektriksel işleyişindeki anormal değişikliklere bağlı, tekrarlayan nöbetlerle karakterize nörolojik bir hastalıktır. Bugüne kadar bağırsak mikrobiyotasının, epilepsi de dahil olmak üzere çeşitli insan nörolojik bozukluklarındaki rolünü ve önemini ortaya çıkaran çok sayıda çalışma yapılmıştır. Bu çalışmada, kainik asitle indüklenen pediatrik sıçan epilepsi modelinden toplanan dışkı örneklerinden elde edilen shotgun metagenomik ve serum örneklerinden elde edilen hedeflenmemiş metabolomik verileri, epilepsi gelişimi ile ilişkili bağırsak mikrobiyotası etkilerini ortaya çıkarmak için analiz edilmiştir. Bu bağlamda, Sham kontrolleri, kainik asit uygulanan sıçanlarda nöbet geliştirmeyen No-Epi grubu ile nöbet gelişimi olan Epi grubu olmak üzere toplam üç grup tanımlanmıştır. Özellikle Epi ve Sham karşılaştırmasında, lipid metabolizması ile ilişkili, en yüksek sayıda ve anlamlı derecede farklı serum metabolitleri ortaya çıkarmıştır. Metagenomik beta çeşitliliğine dayalı analizler, Sham kontrollerine kıyasla hem Epi hem de Epi olmayan gruplarda disbiyoz varlığını gösterirken; Epi ve No-Epi grupları arasında anlamlı bir fark elde edilememiştir. Bu çalışma ile toplam 324 adet metagenom verileri ile birleştirilmiş genom kataloğu oluşturulmuştur. Taksonomik analizlerde, kısa zincirli yağ asitleri sentezi ile ilişkilendirilmiş bir çok bakteri seviyelerinde istatiksel olarak anlamlı farklar elde edilmiştir. Son olarak, çoklu omik verileri entegrasyonu sonrası, *CAG-873 sp009775195* ve *Angelakisella* cinsinin bilinmeyen türü ile serum N-Asetillornitin seviyeleri arasında zıt yönlerde anlamlı korelasyonlar elde edilmiştir. Bütün bu bulgular ışığında bu çalışma ile, pediatrik bir sıçan modelinde epilepsi gelişiminde bağırsak mikrobiyotasında ve serum metabolomiklerinde spesifik değişiklikler olduğu ortaya koyulmuştur.

Anahtar Sözcükler: Epilepsi, Multi-omik, Metagenomik, Metabolomik

1 INTRODUCTION

1.1 Introduction

Epilepsy is a globally prevalent neurological disorder affecting approximately 50 million people globally (World Health Organization - (1)), characterized by occurrence of seizures that result from abnormal synchronized activities in brain (2). These seizures lead to various cognitive, psychological, and social complications (3). Despite intensive research over the past two decades into the molecular mechanisms underlying epilepsy, understanding its complex pathology remains challenging (3). Integration across different molecular layers through omics studies is crucial for unravelling these complexities. Multi-omics approaches, which combine clinical data with knowledge of disease-associated molecular pathways, offer hope for better diagnostic and prognostic tools for epilepsy (4).

The importance of multi-omics data in studying complex diseases like epilepsy lies in its ability to provide a comprehensive understanding that single omics data cannot achieve alone (5). While individual omics data types can identify differences associated with the disease and suggest involved biological pathways, their analysis often reveals correlations that are more reflective of reactive processes than causative ones (5). By integrating various omics data, researchers can uncover potential causative changes and identify treatment targets, offering a multidimensional perspective on disease mechanisms that enhances our understanding of disease development and progression.

Appreciation of the role of microbiome in development and progress of different disease mechanisms led to involvement of numerous omics assays in the study design to be able to explore the microbiome by several omics layers (6). This may aim to identify microbial features (taxa, metabolic activities or metabolic products themselves) as biomarkers associated with different stages of disease pathogenesis (6). On the other hand, despite the exciting potential of multi-omics data to study the microbiome and its role in host health, one of the biggest challenges is the bottleneck

in methodologies that enable comprehensive analysis of multi-omics data to achieve a systems-level understanding of the microbiome.

While meta-omics analysis techniques enable researchers to examine the microbial communities of interest in genomic, transcriptomic, proteomic or metabolomic levels, combination of two or more multi-omics layers opens new horizons by capturing the dynamics in the complex microbial community. Coupling meta-omics data with a host omics analysis leads better understanding of the mechanical basis behind bidirectional communication between microbiome and the host. In a recent example to the gut-brain axis studies within the context of epilepsy, researchers combined 16S-amplicon based sequencing data from the gut bacterial community and both blood and brain metabolomics in epileptic mice and established gut microbial-based regulation of seizure susceptibility (7).

1.2 Aim of the Study

The main goal of this study is the multi-omics investigation of the specific gut microbiota and blood metabolomics signatures in a rat model of kainic acid induced epilepsy development.

2 BACKGROUND

2.1 Epilepsy

Epilepsy is a neurological disorder characterized by repetitive, unprovoked seizures, which can be observed due to abnormal alterations in the electrical functioning of the brain with excessive and hypersynchronous neuronal activity with diverse underlying factors (8). It is ranked as the second most burdensome neurological disorder globally based on a study led by WHO (9). It affects roughly more than 70 million people globally (10), and the development of about three quarters of the epileptic cases starts during childhood (8).

The current clinical classification of epilepsy, according to the 2017 International League Against Epilepsy Classification, relies on a complex and multilevel assessment through seizure types (based on onset or beginning: generalized seizures (entire brain), focal seizures (one hemisphere of the brain), and unknown which includes epileptic spasms), epilepsy types (a clinical diagnosis with help of electroencephalography; focal, generalized, combination of both, and unknown where the accurate classification cannot be carried out by physician due to lack of information), and epilepsy syndrome as third level (encompassing different features including type of seizure and clinical findings from electroencephalography and imaging techniques, which are correlated with age; idiopathic (cause in unknown) generalized epilepsies - also named as genetic generalized epilepsies- and self-limited focal epilepsies generally beginning in childhood) (11,12).

Typically, antiseizure drugs are prescribed to the epileptic patients for the treatment targeting decrease and elimination of the seizures. However, a group of patience, diagnosed with refractory epilepsy, cannot respond (failure in control of seizure occurrences) to two or more antiepileptic medications and constitutes approximately 30% of the epilepsy patients (13).

2.1.1 Animal models of epilepsy

There are numerous animal models developed up to date. One of the main groups are involving genetic animal models which can be further divided into two categories; the ones with naturally occurring epilepsy while epilepsy development is achieved via genetic engineering in the second group (14). In another main group constitute healthy animals for which the epilepsy or seizure development is induced by employing different techniques, such as electrically or chemically induced seizure models (15). Most studies have focused on the animal models that can mimic the physiology of human epilepsy cases, and rodents constitute the majority of the model organisms used (16).

Chemoconvulsants consist of wide range of substances with different characteristics, and can induce seizures when they administered into animals via injection (17). The development of seizures after a latent period is a procedure resulted from the initial injury of the brain, *status epilepticus* (SE), a term used to describe as long lasting seizures (more than 5 minutes) or multiple seizures with loss of consciousness between the attacks (18–20). There are two widely used chemoconvulsants; pilocarpine and kainic acid (21). Pilocarpine functions as an agonist of muscarinic acetylcholine receptors and induce SE (21). It has been showed that pilocarpine treatment of hippocampal neurons leads to a disruption in the equilibrium between signals that stimulate and those that suppress neuronal activity *in vitro* (22), and pilocarpine administration mediated seizure is further associated with NMDA receptor activity, and elevated extracellular glutamate levels were reported (23).

Kainic acid (KA) is another widely used agent in epilepsy animal model. It is a cyclic analogue of L-glutamate, and classified as a neuroexcitotoxic chemical employed in the development of temporal lobe epilepsy, most frequent partial epilepsy type in adult humans, and characterized with recurrent spontaneous seizures (20,24). Even the full molecular mechanism of KA in SE development is not clear yet, it has been reported that KA induces several destructive events in neurons in the CA3 region

of hippocampus, such as production of reactive oxygen species, elevated intracellular CA^{2+} levels, and mitochondrial dysfunction, and glial activation by acting as an agonist of the ionotropic glutamate receptors (25,26).

2.2 Gut Microbiota

Gut microbiota consists of several types of microorganisms including bacteria, fungi and viruses. In the last two decades, great advancements in next generation sequencing technologies eliminated the need of culture of microorganisms in laboratory conditions, and expanded the knowledge about their structures, functions, and relationships with the environment they live in (27). The importance of the microbiome in well-being has been stated by ancient Greek physician, Hippocrates, as he claimed that "All disease begins in the gut" more than 2000 years ago (28). In recent years several important studies investigated the interplay between the gut microbiome and the host system in relation to well-being and disease, and eventually led the gut to be recognized as a new 'organ' (29). While it has been stated that the bacterial cell mass is outnumbering the host cells in humans, a study carried by Sender and colleagues in 2016 revised the ratio as 1:1 (30).

Formation of the gut microbiome begins after birth, but initial bacteria localizations starts in later stages of pregnancy (31). After birth, lactation in first six months, and later involvement of nutrition with solid foods leads formation of microbiota similar to the adults (31). Nevertheless, the microbiota composition during infancy fluctuates a lot and stabilized during childhood (31). Initial establishment of gut microbiota has critical role on immune system maturation in early stages of the host's life. Previous studies point out that commensal gut microorganisms help mucosal immune system to gain an ability to differentiate the gut residents and the pathogenic invaders (32). A bidirectional regulation is present between microbiota and host's innate immune system, and any perturbation in one of those may have negative impacts on other (33).

Throughout the life of the host, the microbial community of the gut is affected by the several external factors including the diet and medications used. In turn, it shapes the host physiology through several metabolic pathways. Several classes of metabolites such as amino acid and derivatives, vitamins, lipids and carbohydrates were produced and released through gut microbial pathways and are key factors in systemic regulation of the host metabolic pathways (34). A previous study highlighted the importance of the metabolic capacity of the gut microbiome on the host-gut microbiome interplay for the regulation of the 71% of faecal and 15% of blood metabolites, rather than focusing on the specific taxa based on a large twin cohort (35). In this regard, impact of the gut microbiota on development and progression on several diseases such as obesity (36), liver diseases (37), cancer (38), inflammatory bowel disease (39), and neurological disorders (40) were revealed. Considering the huge impact on the bidirectional communication between gut and several organs, variety of specialized studies focusing on different axes such as gut-liver, gut-kidney and gut-brain has been established (41).

2.2.1 Gut-brain axis

Several internal and peripheral factors play key roles in development of the central nervous system (CNS), and crosstalk between gut and brain has been revealed after several studies carried on animals which are germ-free or treated with antibiotics (42). Microglia are the macrophage-like, neuronal support cells which reside in central nervous system and acting as immune mediators (43). Defects in microglial maturation in germ-free and antibiotic-treated normal mice have been reported by Erny and colleagues, and short-chain fatty acids produced by microbiota have been shown to have regulatory role on microglia homeostasis (44). Another member of the CNS, astrocytes are highly populated glial cells, and they regulate numerous activities including synapse formation and function, control of the regional blood flow in CNS and blood-brain barrier, regulation of ionic homeostasis for synaptic transmission, and metabolic activities (45). Astrocytes are involved in the pathogenesis of MS through staging into a reactive state which results in dysregulation in BBB function, increased inflammation, and glial scar formation which may inhibit remyelination and axonal

outgrowth (46). Tryptophan metabolization by gut microbiome produce an agonist of aryl hydrocarbon receptor, which has been shown to be partly activated by type I interferon in experimental mouse model of MS, and reduced inflammation and disease score have been reported (47). Apart from those, a recently published study revealed a gut-brain circuit in which regulatory effect of gut microbiota in expression of cFos transcription factor of extrinsic enteric neurons (48). Additionally, a set of brainstem sensory nuclei have been shown to have an ability to control gut sympathetic activity and to sense dysbiosis (48).

2.2.2 Gut-brain axis in epilepsy

Up to date, numerous studies conducted to reveal role and importance of gut microbiome on variety of human neurological disorders; Parkinson's Disease, Alzheimer's Disease, mood disorders, and epilepsy (40). Emerging evidence highlights a significant and intricate interplay between epilepsy and gut microbiota. Study carried on rats to explore the relationship between gut microbiota and the stress-induced epilepsy revealed microbiota-mediated seizure development and activity through cross-faecal microbiota transplantations between treatment and control groups (49). Longitudinal gut microbiome investigations on WAG/Rij rats, as an animal model of generalized absence epilepsy and absence epileptogenesis accompanied by chronic low-grade depression (50), and age-matched controls to track changes in gut community composition and potential relationship with seizure development showed remarkable differentiation starting from 1 month of age (before the seizure development), and decreased *Bacteroidetes*/*Firmicutes* ratio in epileptic rats (51). They further showed the microbial community transfer from healthy control rats or epileptic rats treated with anti-epileptic drug were shown to be effective in weakening of seizure frequency and durations (51).

Even there are variety of medications developed and used for the treatment of the epilepsy, one third of the patients may develop drug resistance, and it has been of great interest to highlight potential role of gut microbiome in development and progress of this phenotype (52). On a human-based exploration of the gut community structures

of drug-resistant (DR) epilepsy and drug-sensitive (DS) epilepsy patients with their matched healthy relatives, 16S amplicon sequencing showed that DR group showed remarkable dysbiosis compared to both DS patients and healthy controls, and the differences between DS patients and healthy samples were much lower (53). In another study focusing on adult patients grouped into DR or DS showed significant alterations in abundance of different taxa even the alpha and beta diversity measurements were similar (54).

Apart from the anti-epileptic drug treatments, diet-based studies have also been carried out to investigate the effect on treatment of epileptic patients -especially ones with refractory epilepsy- while modulating the gut microbiome. Among them, ketogenic diet (KD) is highly popular and proven to be beneficial by several studies (reviewed in (55)). As one of those studies, positive effect of the KD on epileptic patients were reported; significant drop in seizure frequency and modification of the gut microbiome such as increased *Bacteroides* levels (56). In a very detailed study carried on mice revealed the act of mechanism initiated by the KD against seizure susceptibility through two commensals of the gut. Accordingly, KD elevates the *Akkermensia muciniphila* levels in gut, which supports the growth of another species, *Parabacteroides spp.* that plays role in lower gamma-glutamylation of amino acids (7). Decreased gamma-glutamyl amino acids reach to brain throughout the blood stream, and elevates gamma-aminobutyric acid (GABA)/glutamate ratio in brain and hence decreasing seizure susceptibility (7).

3 MATERIALS AND METHODS

3.1 Study Design and Data Generation

All the data used in this thesis were generated as a part of a project from two research groups led by Prof. Teresa Ravizza and Prof. Annamaria Vezzani from Istituto di Ricerche Farmacologiche Mario Negri IRCCS (Milano/Italy), Dr. Antonella Riva and Prof. Pasquale Striano from IRCCS Istituto Giannina Gaslini (Genova/Italy).

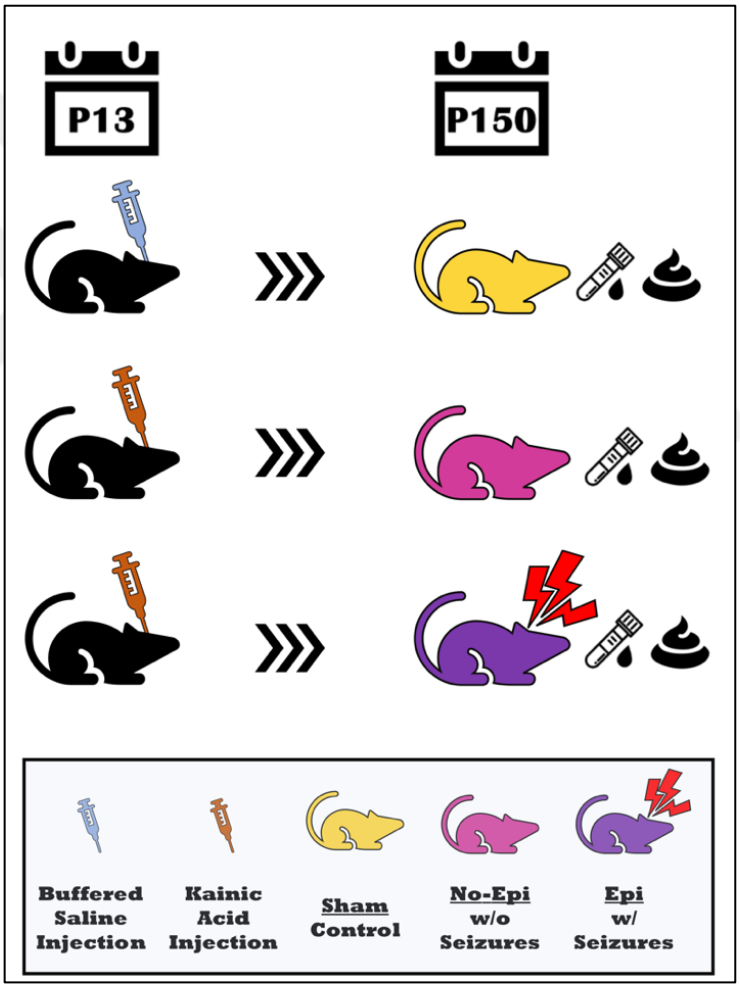

Figure 1. Illustration of the study design, development of three rat groups, and collection of faecal and blood samples.

All detailed information of the methods in animal handling, treatments and sample collection can be found in our article (57). Briefly, a total of 66 (72 rats at the

beginning) male Sprague-Dawley rats were used starting from age of postnatal day 13 (P13), which is the equivalent of preschool age in humans. 54 of them were injected with kainic acid (KA) at P13 for the development of status epilepticus, and seizure recordings were carried out between P110 and P135. The rats developing epilepsy with spontaneous seizures were named as Epi (n = 26), and the rats which did not show seizure development were classified as No-Epi group (n = 20). To be used as a sham control group, 20 rats were injected with saline and treated under similar conditions including electrode implantation. Faecal samples were collected at P150 to be used in shotgun metagenomics, and blood samples were obtained from tail vein of anaesthetized rats at the same day and plasma portion was isolated to be used in untargeted metabolomics (Figure 1).

3.2 Analysis of Shotgun Metagenomics Data

Shotgun metagenomics sequencing of the DNA extracted from total of 66 faecal samples using Zymo Quick-DNA/RNA Mag Bead Kit (Zymo Research) were carried out by CeGaT GmbH (Tübingen/Germany). Library preparation was performed using Illumina Nextera XT DNA Library Preparation Kit with 1 ng of sample DNA, and they were sequenced on NovaSeq 6000 platform (Illumina) with 2x100bp reading. Demultiplexed and adapter trimmed .fastq files were delivered by the facility, and used in this study. Briefly, a total of 3,982,592,302 reads were obtained from shotgun metagenomics sequencing of 66 samples. Minimum, median, and maximum number of paired-end reads obtained were 38.322 million, 54.413 million, and 99.630 million reads, respectively.

3.2.1 Methods used in optimization of bioinformatic analysis of shotgun metagenomics data

In this part of the methods, different tools with specific algorithms were employed for taxonomic annotation of the shotgun sequencing data.

3.2.1.1 Pre-processing

Pre-processing of the raw reads is composed of three steps; quality control and trimming, removing read artifacts, and host read removal as the last step. To begin with quality control, FastQC (ver. 0.11.9) (58) and MultiQC (ver. 1.12) (59) tools were used to create a quality control summary for each sample and overall single summary report, respectively. Parameters such as read length distributions, quality scores of bases in each position, GC contents, duplicate sequences, and presence of adapters were checked. Quality filtering of the reads was carried out using fastp (ver. 0.23.2) (60). Minimum Phred quality score required was set to 20. By this way, reads with more than 40% of their bases having a score below Q20 will be discarded. Second additional filtering option was used for N base content. Any reads having N bases more than 5 were filtered. For trimming, polyG and polyX sequences were set to be removed from tail regions of the reads. No additional adapter removal was performed as the reads were reported to be adapter free in MultiQC summary. Optical duplicates were removed using clumpify tool of BBTools (BBMap version 38.57) (61) with "dedupe optical" parameter and by setting dupedist to 12000 (for NovaSeq runs). Then, another tool from BBTools, filterbytile (BBMap version 38.57) (61), was ran to get rid of reads with low-quality because of analytical issues arisen from location on Illumina flow cell. In order to get rid of any potential vector contamination, "UniVec_Core" dataset was downloaded from NCBI UniVec Database (62) on 2022-07-12. Fasta files were indexed using Bowtie 2 (ver. 2.4.2) (63). Quality trimmed reads were aligned to vector contaminant database using Bowtie 2, and then aligned reads were filtered out using samtools (ver. 1.11) (64). Next step was the removal of sequences with host origin. Reference genome assembly for the host, Rattus norvegicus, was obtained from NCBI RefSeq library (mRatBN7.2 GCF_015227675.2). Genome indexing was carried out using Bowtie 2 (63). Artifact removed reads were aligned to indexed genome using Bowtie 2, and aligned reads were filtered out using samtools (64). Lastly, sequencing artifacts and phiX control sequences were discarded using BBDuk (BBMap version 38.57) (61).

3.2.1.2 Taxonomic analysis using assembly-free methods

Pre-processes reads were analysed using three assembly-free methods using individual algorithms; Kraken2 (65), MetaPhlAn4 (66), and Kaiju (67).

Kraken algorithm relies on taxonomic classification of the reads using k-mers (65). The tool first divides the query sequence into defined k-mers, subsequence in length of k, and then assigns the taxonomic label using lowest common ancestor method based on the matched k-mers of reference genomes in the database (65). For taxonomic assignments of the pre-processed reads through Kraken algorithm, Kraken 2 tool (ver. 2.1.2) was run. The indexed standard database (ver. 20220926), which includes NCBI RefSeq genomes of bacteria, archaea, viral, and human, was obtained from (68), and used in the analysis. As a custom parameter confidence threshold level was set to 5e-06. Classified and unclassified percentages of the input reads were manually obtained from log file of each sample.

Second choice of the algorithm, METAgenomic PHyLogenetic Analysis 4 (MetaPhlAn4) (66), relies on marker-based method for the classification of the short reads. It maps the short reads to the database of markers specific to species-level genome bins comprising both *de novo* metagenomic assemblies and reference sequences, and based on the homology and coverage rates, annotation and abundance calculation are carried out (66). For the taxonomic annotation of the pre-processed reads, MetaPhlAn4 (ver. 4.0.4) was run against the marker gene database (ver. mpa_vJan21_CHOCOPhlAnSGB_202103) by excluding eukaryotes and including viruses via additional parameters of "--ignore_eukaryotes --add_viruses". In order to obtain relative taxonomic abundances with unclassified group, a second run was carried out using the alignment files from previous run "--unclassified_estimation", and then all sample quantifications were merged into one abundance table via "merge_metaphlan_tables.py" script of the tool.

Last algorithm Kaiju (67) utilizes taxonomic classification on the basis of DNA-to-protein. For that purpose, the short query reads are first translated into respective

amino acid sequences, and then performs amino acid alignment to the protein sequences in the database to find maximum exact matches with an option of allowed mismatches defined by the user (67). Pre-built "nr_euk" database, which contains indices of archaeal, bacterial, microbial eukaryotic, and viral protein sequences from NCBI Reference Sequence (RefSeq) and RefSeq non-redundant databases, was obtained from Kaiju server (69) on 10-03-2022. Kaiju tool (ver. 1.9.0) was run using the pre-processed short reads as input and by setting the E-value to 0.00001. After obtaining output files for each sample, classification summary files were created in different taxonomic levels using the "kaiju2table" program.

3.2.2 Analysis workflow of the short reads by generating metagenome-assembled genomes (MAGs)

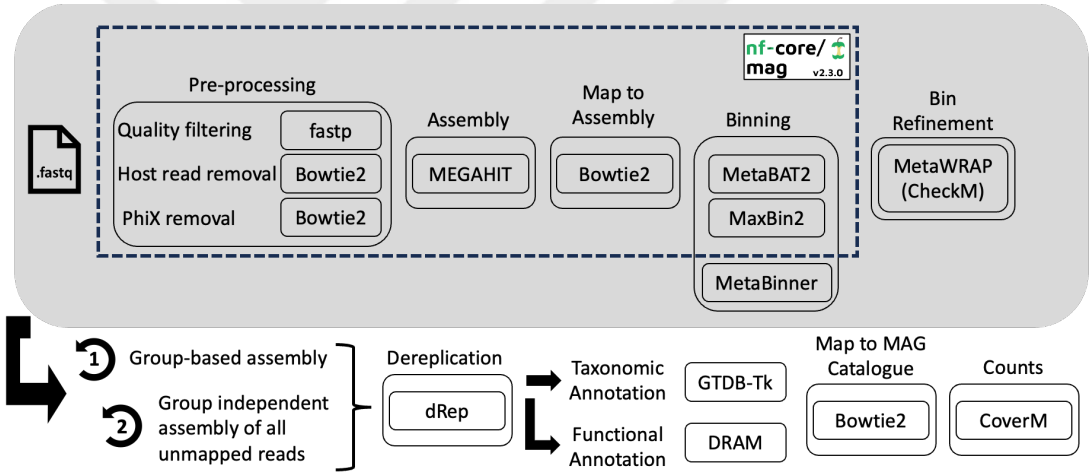

Figure 2. The bioinformatic analysis pipeline used for the process of the shotgun metagenomics data.
The flow in the grey area describes the steps from input processing to bin refinement. These steps were repeated twice with (1) rat group-based assembly and (2) rat group-independent assembly of the reads. All MAGs obtained from two iterations were dereplicated and final MAG catalogue is established. Following annotations in taxonomic and functional levels, short reads were mapped to the MAG catalogues and count information for the MAGs and predicted gene sequences were obtained. Steps encircled by a dashed line are carried out using the nf-core mag (ver. 2.2.1) (70) pipeline.

Following testing the accuracies and the performance of different strategies outlined above, it was decided to perform analysis of the reads by generating *de novo*

assemblies into contigs and then by binning them to produce draft genome sequences that are called metagenome-assembled genomes (MAGs). This section involves the final pipeline employed; starting from pre-processing of the raw reads and ends with quantification of the taxa and predicted genes of the generated MAG catalogue (Figure 2).

3.2.2.1 Generation of the MAG catalogue

Pre-processing, hybrid-assembly and binning steps were carried out using nf-core/mag analysis pipeline (ver. 2.2.1) (70). Briefly, fastp (ver. 0.23.2) (60) was used for quality filtering with additional parameters "--reads_minlength 70 --fastp_qualified_quality 20 --fastp_cut_mean_quality 20", host reads were removed via Bowtie2 (ver. 2.4.2) (63) using *rattus norvegicus* reference genome downloaded from NCBI RefSeq with accession ID of GCF_015227675.2, and as a last step of pre-processing, PhiX, a bacteriophage used in Illumina sequencing runs as a control sequence and may present as a contaminant in sample reads, removal were carried. Sample group-based co-assembly was performed on pre-processed reads using MEGAHIT (ver. 1.2.9) (71) with parameters "coassemble_group megahit_options '—k-list 21,29,39,59,79'", and pre-processed reads were mapped to these assemblies by Bowtie2 (63). Lastly, MetaBAT2 (ver. 2:2.12.1) (72) and MaxBin2 (ver. 2.2.7) (73) were used for binning of the contigs. All following steps were carried out independent from the "nf-core/mag" pipeline. As a third metagenomic binning tool, MetaBinner (ver. 1.4.4) (74) was used. All the bins produced by three binning tools were combined, and processed by bin refinement module of MetaWRAP (ver. 1.3.0) (75) using CheckM (76), based on minimum completeness rate of 70% and maximum contamination of 10%. After obtaining first set of MAGs, pre-processed reads were mapped to them, and unmapped reads were used for second iteration of MAG generation by following the steps above, except, all the unmapped reads were pooled and used for co-assembly. Two sets of MAGs were combined, and dereplicated via dRep tool (ver. 3.4.0) (77) with default ANI threshold parameters.

3.2.2.2 Taxonomic and functional annotation of the MAGs, and obtaining counts

Taxonomic annotation of the final set of MAGs was carried out by Genome Taxonomy Database (GTBD)-Tk (ver. 2.1.1) (78) using reference database version of R207_v2. For functional annotation of the MAGs, DRAM pipeline (ver. 1.3.5) (79) was employed. Pre-processed reads were mapped to MAGs using Bowtie2, and PCR duplicated were removed by samtools (ver. 1.11). The counts of the properly mapped paired reads to each MAG were obtained by CoverM (ver. 0.6.1) (80) with additional flag of "--proper-pairs-only". In DRAM annotation pipeline, amino acid sequences were predicted using Prodigal (ver. 2.6.3) (81), and they are annotated using Pfam (82), Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology (KO) (83), UniProt (84), CAZy (85), MEROPS (86), and VOGDB (87) databases. For the quantification of predicted genes, pre-processed reads were mapped to predicted gene sequences, PCR duplicates were removed, and raw counts of the properly mapped pairs were obtained.

3.2.2.3 Metabolic capacity prediction of MAGs

To infer metabolic capacities of the MAGs through the presence of annotated genes, Genome-Inferred Functional Traits (GIFTs) were obtained by distilling raw annotations obtained from DRAM using a custom R package, called "distillR" (88). The package contains several biosynthesis and degradation pathways/modules of several biomolecules, and metabolic capacities of bacterial communities were inferred via presence of genes necessary to accomplish the relevant metabolic activity. Estimated GIFTs were corrected using MAG completeness values to minimize bias of completeness on functional capacity (89).

3.2.2.4 Data analysis

All the downstream steps were carried in R (ver. 4.1.3) (90). As suggested by previous studies (91,92), raw MAG counts and counts of the predicted genes were

normalized using normalization factor, calculated via DESeq2 package (ver. 1.34.0) (93).

3.2.2.4.1 Taxonomic data analysis

For alpha and beta diversity analysis, phyloseq package (ver. 1.38.0) (94) was used by creating a phyloseq object using normalized MAG counts, metadata, taxonomic information of the MAGs, and phylogenetic tree was pruned using ape package (ver. 5.7) (95) to include 324 MAGs only. Interactive Tree Of Life (iTOL) (96) was used for illustration of phylogenetic tree. Rarefication was performed using the minimum sample size of 15,374,826. All available alpha diversity metrics were calculated by "estimate_richness" function of phyloseq (94). For this study, among the calculated metrics, Chao1 and Shannon Diversity indices were used. As the third metric, Faith's Phylogenetic Diversity (PD) index was calculated using picante package (ver. 1.8.2) (97). Kruskal-Wallis (KW) test was used for testing any significant difference in alpha diversity indices among rat groups. For beta diversity analysis, Bray-Curtis and Jaccard distances between samples were calculated. Plots were drawn based on first two axes of principal coordinate analysis (PCoA) ordination. Permutational multivariate analysis of variance (PERMANOVA) was used through "adonis2" function in vegan package (ver. 2.6.4) (98) to test whether there are any significant differences between clusters based on rat test groups. Additionally, pairwise PERMANOVA was performed using "pairwise.adonis2" function in the pairwiseAdonis package (ver. 0.4) (99). For both PERMANOVA tests, the number of permutations was set to 9999.

For differential abundance (DA) analysis of taxonomic composition, normalized MAG counts were aggregated into three taxonomic ranks; phylum, genus, and species. For each taxonomic rank, relative abundances were calculated. Any taxa with relative abundance less than 0.01% for phylum, and less than 0.001% for genus and species levels in less than 1:5 of the samples in every three rat groups were filtered out prior to downstream analysis. DA was performed using Kruskal-Wallis test followed by post hoc Dunn's test with Bonferroni correction of raw p values. For each pairwise

comparison, taxa with at least 2-fold change and with adjusted p value less than 0.05 were reported. Before visualization of significant taxa using heatmap, zero values in relative abundance matrices for each phylum, genus, and species level taxa were replaced with half of minimum relative abundance values, then clr transformation was applied via decostand function of vegan (98). Significant taxa were filtered and aggregated into one matrix, and values were scaled between 0 and 1 to enhance visualization. Heatmap was constructed using the ComplexHeatmap package (ver. 2.10.0) (100).

3.2.2.4.2 Functional data analysis

For functional changes across groups, DA analysis of the KO's was carried out. Normalized gene counts were aggregated based on unique KO IDs, and relative abundances were calculated. Before proceeding to DA, a filtering step was applied; if a KO annotated gene is present in at least 1/5 of the samples in at least one of the rat groups, it was retained for downstream DA analysis. Kruskal Wallis and post hoc Dunn's test with Bonferroni p value correction were performed. KO changes of each pairwise comparison were illustrated using EnhancedVolcano R package (ver. 1.12.0) (101).

3.2.2.4.3 Functional pathway enrichment analysis

KO's with adjusted p value less than 0.05 and with at least two fold change were used for pathway enrichment. KO ID's were mapped to pathways via KEGG Orthology Database web page (102). After obtaining pathways and KO's mapped to those pathways, total number of KO's on KO Database (as of 12-02-2023), total number of KO's linked to the relevant pathway, and total number of KO's in query list obtained from DA were manually curated, and hypergeometric test was performed. Obtained p values were corrected using Bonferroni method, and significance threshold of 0.05 was applied.

3.3 Analysis of Processed Untargeted Serum Metabolomics Data

Procedures from mass spectrometry (MS) analysis to annotation and data processing were all carried out by Dr. Chiara Lavarello and Dr. Andrea Petretto from IRCCS Istituto Giannina Gaslini (Genova/Italy). Four different metabolomics data tables were produced by combining data from two chromatography approaches, with reversed-phase C18 or hydrophilic interaction liquid chromatography (HILIC) columns, and two data acquisition modes in positive or negative ionization. Lastly, the abundances were quantile-normalized, log2-transformed, and median scaled by them. Those processed four mode tables, C18 positive (C18-Pos), C18 negative (C18-Neg), HILIC positive (HILIC-Pos), HILIC negative (HILIC-Neg) in ready-to-analyse format were obtained and used in this study (Table 1).

Table 1. Number of blood samples from each rat groups and number of metabolites present in each four mode data tables

	C18-Pos	**C18-Neg**	**HILIC-Pos**	**HILIC-Neg**
Sham	17	18	18	18
No-Epi	16	18	18	18
Epi	22	23	24	24
Metabolites	393	341	90	294

3.3.1 Statistical analysis

Each of the four modes were independently analysed using R (90). The normal distribution was evaluated using Shapiro-Wilk test, and the majority of the data was satisfying the normality ($p > 0.05$) together with other analysis of variance (ANOVA) assumptions. Accordingly, metabolites significantly changing between rat groups were determined using one-way ANOVA, followed by post-hoc Tukey's honestly significant difference (HSD) based on false discovery rate (FDR) corrected p value threshold of 0.05. Metabolite changes obtained from each pairwise comparison were illustrated using EnhancedVolcano package (101).

3.3.2 Pathway enrichment

For each pairwise comparison (Sham vs No-Epi, Sham vs Epi, and No-Epi vs Epi), metabolites with FDR<0.05 from each of four modes were combined, and KEGG compound IDs were retrieved from metabolite names or HMDB ID's either by "Compound ID Conversion" utility of MetaboAnalyst 5.0 (103) web interface (www.metaboanalyst.ca) or by manual inspection. Unique list of KEGG compound ID's were used as a query for pathway enrichment analysis by MBROLE 2.0 (104) web tool (https://csbg.cnb.csic.es/mbrole2) with selection of "KEGG Pathways" option in annotations, and *rattus norvegicus* as a background set. The results were downloaded, and pathways with FDR-adjusted p value less than 0.05 were used.

3.4 Multi-Omics Data Integration

3.4.1 Feature selection

3.4.1.1 Feature selection by machine learning

All machine learning models were trained using the "caret" package (ver. 6.0.94) (105) in R. For all the model development and testing applications, datasets were divided into train and test datasets in 80% and 20% ratios, respectively. For training of both random forest (RF) and elastic net (EN) models; the number of combinations was set to 10 using "tuneLength" which combines different parameters for hypertuning, "trainControl" function was used for 2-fold cross-validation, and hyperparameter search was set to random. Additionally, kappa was assigned as the parameter used in selection of the optimal model. For random forest, 1001 decision trees were used. Trained model was used for classification on test data set, and the accuracy and kappa values were calculated by comparing the actual and predicted rat group labels. In order to check the effect of different sample selection to the train and test datasets during splitting, seed number was set to 100 different models and 100 different partitioning, model training, and performance measurement on test data set were obtained. Performance distributions of the 100 trained models on test data

classification were visualized using histogram plots using kappa values, for each algorithm. Median values of the performances of 100 models were shown on plots with vertical lines.

3.4.1.2 Feature selection by Multi-Omics Factor Analysis (MOFA)

Multi-Omics Factor Analysis (MOFA) (106) was used to obtain important features from both omics data. MOFA introduces latent factors (LFs) that are used to explain variance in the omics datasets, in an unsupervised fashion (106). Prior the MOFA, all four modes of the metabolomics data were merged based on matched samples, and final composition of the groups were as 17, 16, and 21 samples for Sham, No-Epi, and Epi group rats, respectively. For metagenomics data, normalized MAG counts were transformed using a centred log ratio (clr). After importing both datasets into R, and downstream steps were carried out using MOFA2 R package (ver. 1.4.0). First, the optimum number of LFs to train a MOFA model was determined to be 15 based on highest rate of variance explained in both omics data. MOFA model was trained using 15 LFs and without the metadata information, and LFs that explains at least 2% of variance were selected for metabolomics and metagenomics, independently. Top 10 features weighted by the selected LFs were recorded for each omics dataset.

3.4.1.3 Feature selection by differential abundance analysis

For this part, unique MAGs and metabolites significantly changing in at least one of the three pairwise DA analyses (No-Epi vs Sham, Epi vs Sham, and Epi vs No-Epi) were collected. Of note, as distinct from the DA analysis carried out in different taxonomic levels as described above, relative abundances of MAGs were calculated and used for DA analysis irrespective of any specific taxonomic aggregation.

3.4.2 Correlations between selected features

Spearman correlation analysis was carried between selected features from metabolomic and metagenomic datasets, and obtained p values were adjusted by Bonferroni method.

4 RESULTS

4.1 Analysis Results of Shotgun Metagenomics Sequencing Data

4.1.1 Optimizations for bioinformatic analysis of shotgun metagenomics data

Summary tables of the three algorithms used for the taxonomic classifications of the pre-processed reads were prepared, and used for the evaluation of the performances (Figure 3). When the annotations are aggregated into two categories based on classification rates, for Kraken2; median rates are obtained as 17.46% and 82.53% for classified and unclassified reads, respectively. In case of annotation using MetaPhlAn4, the classified and unclassified portions of the reads are 56.76% and 43.24%, respectively.

Lastly, as distinct from Kraken2 and MetaPhlAn4, Kaiju classification summaries prepared for individual samples show percentage of the classified reads in five major categories. First two categories belong to “Bacteria” and “Archaea”, and individual taxa of those two kingdoms are shown in desired taxonomic level by a separate row. Other three categories show the percentage of the reads belonging to “Viruses”, “cannot be assigned to a (non-viral) phylum”, and “unclassified”. As a classification category specific to Kaiju, “cannot be assigned to a (non-viral) phyla” defines the reads that are assigned to a taxonomy level above the target taxonomy level, which is phylum in this case, and also the ones that cannot be assigned to a specific kingdom and hence assigned to “cellular organisms”. For phylum level, the median levels for each category are 49.4% for bacteria, 0.016% for archaea, 0.66% for viruses, 9.487% for “cannot be assigned to a (non-viral) phylum”, and 39.19% unassigned. Besides phylum level, species level summary tables were also prepared to be able to see the resolution differences in highest and lowest taxonomic level Kaiju allows. In case of species level; median levels are 24.98% for bacteria, 0.013% for archaea, 0.66% for viruses, 34.89% for “cannot be assigned to a (non-viral) species”, and 39.11% unassigned.

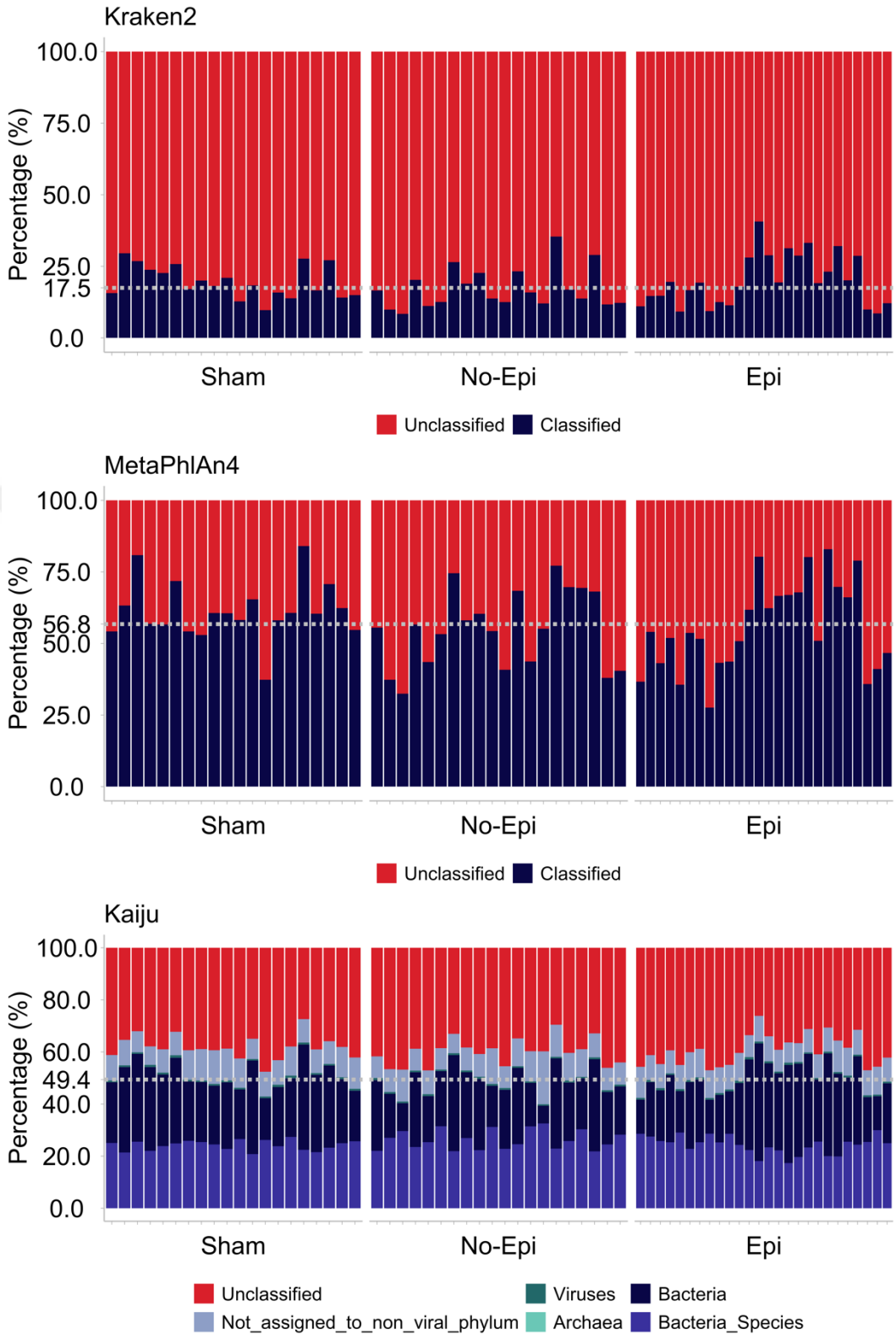

Figure 3. Plots summarizing the taxonomy-based classification rates for each assembly-free method.

Figure 3. Plots summarizing the taxonomy-based classification rates for each assembly-free method (continued)
Classification rates of reads using Kraken2 (top), MetaPhlAn4 (middle), Kaiju (bottom) tools. Plots show the stacked bars illustrating the sample-wise percentages of the classified reads. Rat groups were shown on x-axis, and percentages are shown on y-axis. The white dashed line belongs to the median percentage shown on the y-axis.

The difference between rates of reads that cannot be assigned to phylum and species levels is approximately 25.4%, indicating the loss of annotation rates in -mainly- bacteria and archaea.

4.1.2 Analysis results of the shotgun metagenomics data by generation of MAGs

4.1.2.1 Establishment of MAG catalogue

As detailed in the methods section, two different workflows were employed for generation of the raw MAG sets. In the first round, reads from the three rat group samples were co-assembled separately (group-based co-assembly), and this workflow yielded 176, 179, and 175 MAGs from Sham, No-Epi, and Epi group sample reads, respectively. In the second round, all sample reads that are not mapped to those three MAG groups were pooled and assembled, and 109 MAGs were produced. After dereplication of all sets of MAGs, a final MAG catalogue composed of 324 unique MAGs were obtained. When the pre-processed reads were mapped to the MAG catalogue, a median mapping rate of 52.42% was obtained for pre-processed reads of 66 samples based on unique and concordant mapping settings (Figure 4).

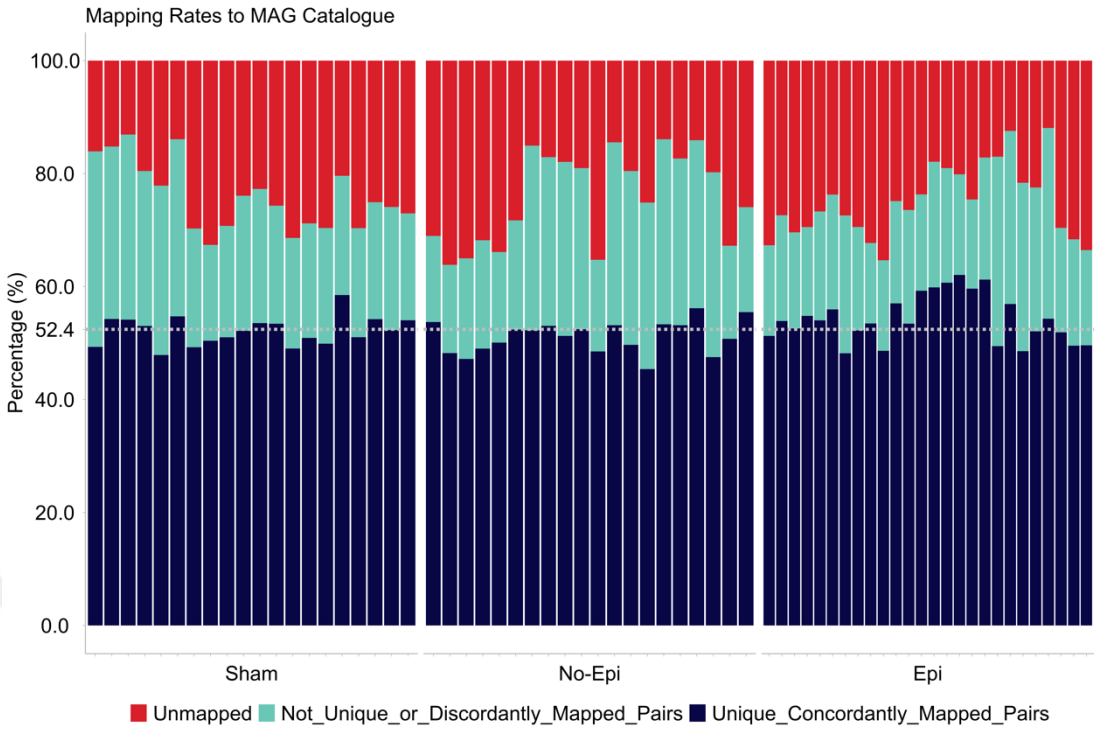

Figure 4. Mapping rate summary of the pre-processed reads of 66 samples to the MAG catalogue.
Plot shows the stacked bars illustrating the sample-wise percentages of the reads that show overall alignment in two subclasses based on unique and concordant mapping, and unmapped percentages. Rat groups were shown on x-axis, and percentages are shown on y-axis. The white dashed line belongs to the median percentage shown on the y-axis.

4.1.2.2 Taxonomic analysis results

4.1.2.2.1 Overview of the taxonomic annotations of MAGs

For objective taxonomic classification, all the MAGs were annotated using the Genome Taxonomy Database (GTBD)-Tk v2.11 (78). As a result, 100% of the MAGs were efficiently annotated at the phylum, class, order, and family level, while 97.5% and 62% of them were annotated at the genus and species level based on GTDB-Tk database classification (Appendix1). Phylogenetic relationships of the MAGs and their median abundance distributions across each rat group were illustrated in Figure 5.

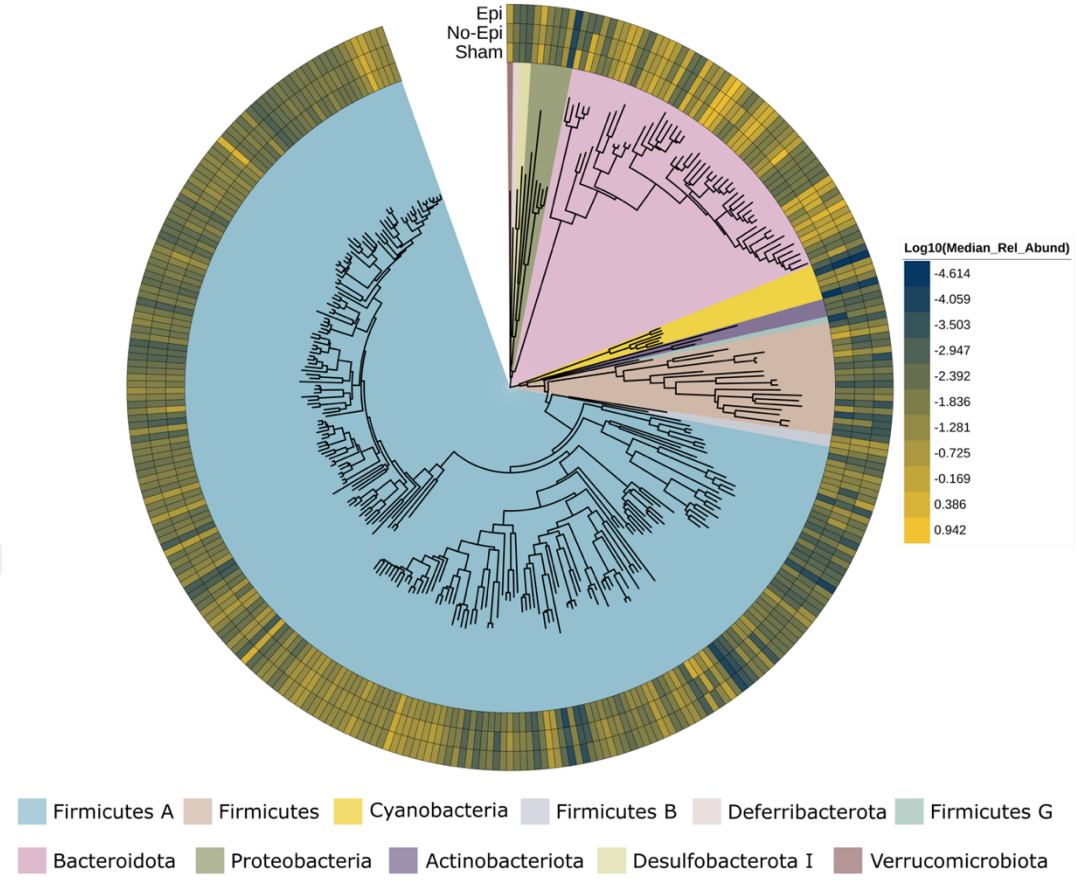

Figure 5. Phylogenetic tree of the 324 MAGs annotated using GTDB-Tk database. Phylogenetic tree was coloured based on the 11 phyla each MAG belongs to. Outer circle layers are the heatmaps showing log10 median abundance of the respective MAG in each of the three rat groups.

4.1.2.2.2 Alpha diversity analysis results

For alpha diversity analysis, instead of taxonomic abundance in different phylogenetic levels, MAG abundances and compositions in each rat sample were evaluated through three different indices (Figure 6). Richness evaluating index, Chao1, was similar across three rat groups with same median value of 322, suggesting all three group rat guts have similar number of MAGs. When both richness and evenness are taken into consideration via Shannon index, higher value was obtained from No-Epi rats compared to control Sham and Epi group ($Sham_{median}$ = 3.82, $No\text{-}Epi_{median}$ = 3.97, Epi_{median} = 3.79). Even it indicates increased evenness in No-Epi group samples as Chao1 reveals similar richness in all groups, the difference is not statistically different (KW p value = 0.076). Lastly, as the third alpha index, Faith's PD was employed to

check if there is any difference by means of the differences in phylogenetic relationships of the MAGs in the community. Similar to Shannon, comparably higher value was obtained in No-Epi rats ($Sham_{median}$ = 83.4, $No\text{-}Epi_{median}$ = 83.7, Epi_{median} = 83.4), indicating presence of more phylogenetically distant MAGs compared to other two rat groups; however, the difference was not significant either (KW p value = 0.081). Altogether, three different alpha diversity measurements agreeably show similar bacterial composition across rats in each group.

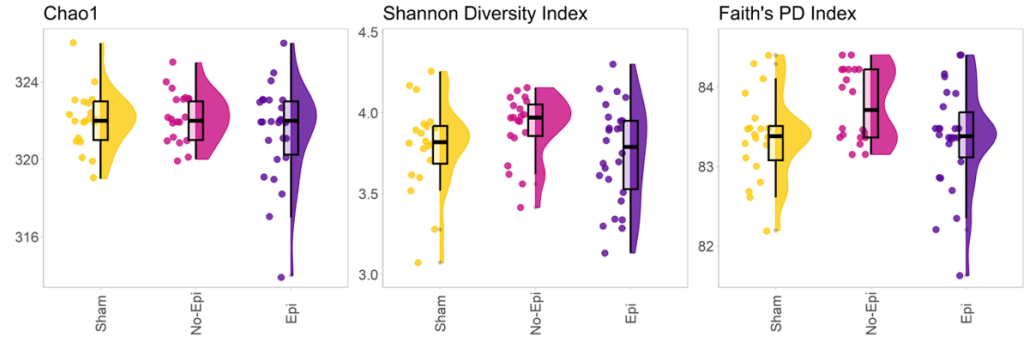

Figure 6. Alpha diversity analysis of three rat groups based on MAG abundance and distributions.
Changes in the alpha diversity of each rat group were tested using Chao1 (left), Shannon Diversity (middle), and Faith's Phylogenetic Diversity (right) indices. Individual sample data of calculated alpha diversity indices represented by points, and the distributions in each group were illustrated by violin and boxplots (with black middle line representing median, and boxes between interquartile ranges of the 25th and 75th percentiles).

4.1.2.2.3 Beta diversity analysis results

After comparison of the three rat groups by means of the sample-wise MAG community compositions, similarity/dissimilarity of the communities was compared using beta diversity analysis (Figure 7). PERMANOVA tests on both Bray-Curtis and Jaccard dissimilarity indices showed statistically significant differences between rat group communities (Bray-Curtis: F = 2.1, p value = 0.0078; Jaccard: F = 1.8052, p value = 0.0044). Pairwise PERMANOVA testing clarified which of the pairwise comparisons carry the significance. For both Bray-Curtis and Jaccard dissimilarity indices, KA-injected rat MAG communities can be distinguishable from control rat communities (Bray-Curtis: No-Epi vs Sham – F = 2.6688, p value = 0.007; Epi vs

Sham – F = 2.2676, p value = 0.019; Jaccard: No-Epi vs Sham – F = 2.2253, p value = 0.006; Epi vs Sham – F = 1.927, p value = 0.016). On the other hand, neither Bray-Curtis (F = 1.4253, p value = 0.162), nor Jaccard (F = 1.3112, p value = 0.156) showed significant difference between Epi and No-Epi groups.

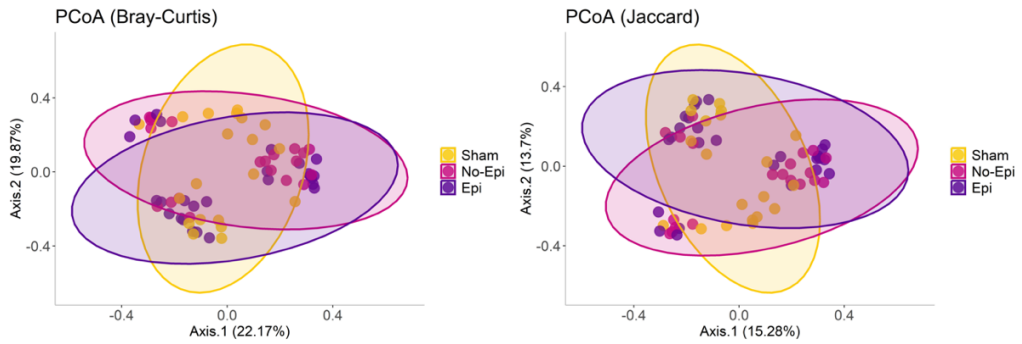

Figure 7. Beta diversity analysis of three rat groups based on MAG compositions. PCoA plots were drawn based on the Bray-Curtis (left) and Jaccard (right) dissimilarity index values. Points represent each sample and ellipses correspond to 95% confidence intervals for each of the rat groups.

4.1.2.2.4 Taxonomic differential abundance analysis results

Following community-level analyses, individual taxa were compared between rat groups through differential abundance analysis. For that purpose, abundances were aggregated into phylum, family, genus, and species levels, and differential abundance analysis was carried out by comparing each group with each other, and fold change threshold is set to 2 for evaluation besides p.adj value threshold of 0.05. Relative abundance distributions of the significant taxa obtained from pairwise comparisons carried out in each taxonomic levels of interest were illustrated in Figure 8.

4.1.2.2.4.1 Phylum level

In phylum level, two phyla were obtained from comparison of either No-Epi or Epi rats with Sham group. One of them, *Firmicutes* relative abundances were detected lower in both No-Epi (median = 0.585%) and Epi rats (median = 0.631%) compared to Sham (median = 1.934%). Specific to Epi vs Sham comparison, *Desulfobacterota_I*

is the second phylum detected in lower relative abundance compared to control group ($Sham_{median}$ = 0.0184%, Epi_{median} = 0.0033%). Another dominant phylum, *Bacteroidota*, is not obtained from any of the three pairwise comparisons. On the other hand, the ratio of *Bacteroidota*/*Firmicutes* was significantly increasing in both No-Epi (median = 59.457) and Epi rats (median = 79.376) compared to Sham (median = 27.523). None of the phyla tested are significantly different between two epilepsy rat groups. All the taxa passing the significance thresholds were summarized in Table 2.

Table 2. Significantly different phyla with at least 2-fold change for each pairwise comparison of three rat groups obtained from taxonomic DA analysis

No-Epi vs Sham				
Phylum	**Sham median**	**No-Epi median**	**No-Epi/Sham Log2(Fold Change)**	**p.adj**
Firmicutes	1.93444	0.58526	-1.72476	0.00655
Bacteroidetes*/*Firmicutes	27.52305	59.45680	1.11120	0.02458
Epi vs Sham				
Phylum	**Sham median**	**Epi median**	**Epi/Sham Log2(Fold Change)**	**p.adj**
Desulfobacterota I	0.01839	0.00333	-2.46609	0.00465
Firmicutes	1.93444	0.63136	-1.61539	0.00142
Bacteroidetes*/*Firmicutes	27.52305	79.37575	1.52806	0.00885
Epi vs No-Epi				
Phylum	**Epi-median**	**No-Epi median**	**Epi/No-Epi Log2(Fold Change)**	**p.adj**
No Significant Change				

4.1.2.2.4.2 Family level

Among the 46 annotated families, total of 10 different families were found to be significantly different based on pairwise comparisons (Table 3). Six families were obtained from comparison of No-Epi and Sham rats. Families detected lower in No-Epi group rats are; *Lactobacillaceae* ($Sham_{median}$ = 1.3617%, $No\text{-}Epi_{median}$ = 0.2879%) and *Acutalibacteraceae* ($Sham_{median}$ = 4.4639%, $No\text{-}Epi_{median}$ = 1.8776%). *Anaerotignaceae* ($Sham_{median}$ = 0.0157%, $No\text{-}Epi_{median}$ = 0.095%), *CAG-552* ($Sham_{median}$ = 0.1265%, $No\text{-}Epi_{median}$ = 0.5427%), *UBA5755* ($Sham_{median}$ = 0.0571%, $No\text{-}Epi_{median}$ = 0.1697%), and *Butyricicoccaceae* ($Sham_{median}$ = 0.1582%, $No\text{-}Epi_{median}$ = 0.4382%) abundances were detected higher in No-Epi compared to control.

Most of the families obtained from Epi vs Sham comparison were detected in lower abundances. Five the six families were less abundant in Epi group; *Turicibacteraceae* ($Sham_{median}$ = 0.0061%, Epi_{median} = 0.0002%), *Peptostreptococcaceae* ($Sham_{median}$ = 0.0268%, Epi_{median} = 0.0029%), *Desulfovibrionaceae* ($Sham_{median}$ = 0.0184%, Epi_{median} = 0.0033%), *Lactobacillaceae* ($Sham_{median}$ = 1.3617%, Epi_{median} = 0.2560%), and *Acutalibacteraceae* ($Sham_{median}$ = 4.4639%, Epi_{median} = 1.8412%). Family of Anaerotignaceae is the only taxa detected in higher relative abundance in Epi group compared to Sham ($Sham_{median}$ = 0.0157%, Epi_{median} = 0.0595%).

Epi vs No-Epi comparison in family level yielded two families with lower relative abundance in Epi group; *CAG-274* ($No\text{-}Epi_{median}$ = 0.0622%, Epi_{median} = 0.0101%) and *Butyricicoccaceae* ($No\text{-}Epi_{median}$ = 0.4382%, Epi_{median} = 0.2027%).

Table 3. Significantly different families with at least 2-fold change for each pairwise comparison of three rat groups obtained from taxonomic DA analysis

No-Epi vs Sham				
Family	**Sham median**	**No-Epi median**	**No-Epi/Sham Log2(Fold Change)**	**p.adj**
Anaerotignaceae	0.01573	0.09495	2.59338	0.00036
Lactobacillaceae	1.36168	0.28786	-2.24192	0.00067
CAG-552	0.12647	0.54266	2.10120	0.03600
UBA5755	0.05707	0.16975	1.57257	0.00692
Butyricicoccaceae	0.15818	0.43821	1.47001	0.00586
Acutalibacteraceae	4.46390	1.87758	-1.24943	0.00207
Epi vs Sham				
Family	**Sham median**	**Epi median**	**Epi/Sham Log2(Fold Change)**	**p.adj**
Turicibacteraceae	0.00611	0.00018	-5.05414	0.00880
Peptostreptococcaceae	0.02683	0.00295	-3.18525	0.03601
Desulfovibrionaceae	0.01839	0.00333	-2.46609	0.00465
Lactobacillaceae	1.36168	0.25603	-2.41100	0.00154
Anaerotignaceae	0.01573	0.05951	1.91935	0.01040
Acutalibacteraceae	4.46390	1.84121	-1.27765	0.00032
Epi vs No-Epi				
Family	**Epi median**	**No-Epi median**	**Epi/No-Epi Log2(Fold Change)**	**p.adj**
CAG-274	0.01011	0.06218	-2.62025	0.04764
Butyricicoccaceae	0.20274	0.43821	-1.11200	0.04325

4.1.2.2.4.3 Genus level

In genus level analysis, 16, 12, and 4 genera were detected to be significantly altered between No-Epi vs Sham, Epi vs Sham, and Epi vs Sham groups (Table 4).

Compared to Sham rats, relative abundance of 7 genera *Ruminococcus_E* ($Sham_{median}$ = 0.9730%, $No\text{-}Epi_{median}$ = 0.0020%), *Caccenecus* ($Sham_{median}$ = 0.0762%, $No\text{-}Epi_{median}$ = 0.0061%), *Limosilactobacillus* ($Sham_{median}$ = 0.0947%, $No\text{-}Epi_{median}$ = 0.0119%), *Lactobacillus* ($Sham_{median}$ = 0.4030%, $No\text{-}Epi_{median}$ = 0.0554%), *Ligilactobacillus* ($Sham_{median}$ = 0.7594%, $No\text{-}Epi_{median}$ = 0.1868%), *UBA1394* ($Sham_{median}$ = 0.1885%, $No\text{-}Epi_{median}$ = 0.0600%), and *CAG-273* ($Sham_{median}$ = 0.2547%, $No\text{-}Epi_{median}$ = 0.1145%) decreased, whereas 9 genera of *Anaerotignum* ($Sham_{median}$ = 0.0158%, $No\text{-}Epi_{median}$ = 0.0950%), *Avoscillospira_A* ($Sham_{median}$ = 0.0042%, $No\text{-}Epi_{median}$ = 0.0189%), *CAJFPI01* ($Sham_{median}$ = 0.0489%, $No\text{-}Epi_{median}$ = 0.2223%), *Butyricicoccus* ($Sham_{median}$ = 0.0060%, $No\text{-}Epi_{median}$ = 0.0247%), *Angelakisella* ($Sham_{median}$ = 0.0666%, $No\text{-}Epi_{median}$ = 0.2323%), *QXXE01* ($Sham_{median}$ = 0.1175%, $No\text{-}Epi_{median}$ = 0.4032%), *Avidehalobacter* ($Sham_{median}$ = 0.0573%, $No\text{-}Epi_{median}$ = 0.1704%), *Dysosmobacter* ($Sham_{median}$ = 0.3137%, $No\text{-}Epi_{median}$ = 0.8101%), and *Lawsonibacter* ($Sham_{median}$ = 0.0269%, $No\text{-}Epi_{median}$ = 0.0591%) increase in No-Epi group.

Table 4. Significantly different genera with at least 2-fold change for each pairwise comparison of three rat groups obtained from taxonomic DA analysis

No-Epi vs Sham				
Genus	**Sham median**	**No-Epi median**	**No-Epi/Sham Log2(Fold Change)**	**p.adj**
Ruminococcus_E	0.97298	0.00201	-8.91657	0.02708
Caccenecus	0.07622	0.00609	-3.64631	0.01492
Limosilactobacillus	0.09472	0.01185	-2.99863	0.00907
Lactobacillus	0.40304	0.05542	-2.86248	0.00087
Anaerotignum	0.01577	0.09502	2.59091	0.00038
Avoscillospira_A	0.00415	0.01893	2.18858	0.00342
CAJFPI01	0.04890	0.22230	2.18467	0.00731
Butyricicoccus	0.00601	0.02474	2.04092	0.02910
Ligilactobacillus	0.75944	0.18676	-2.02376	0.01010
Angelakisella	0.06656	0.23227	1.80316	0.00162
QXXE01	0.11749	0.40318	1.77893	0.00496
UBA1394	0.18854	0.06003	-1.65122	0.00907
Avidehalobacter	0.05728	0.17038	1.57270	0.00692
Dysosmobacter	0.31371	0.81010	1.36866	0.00793
CAG-273	0.25467	0.11448	-1.15348	0.04532
Lawsonibacter	0.02692	0.05914	1.13552	0.00395
Epi vs Sham				
Genus	**Sham median**	**Epi median**	**Epi/Sham Log2(Fold Change)**	**p.adj**
Scybalousia	0.02421	0.00044	-5.77775	0.02259
Turicibacter	0.00612	0.00018	-5.05093	0.00905
Caccenecus	0.07622	0.00726	-3.39177	0.01918
Choladocola	0.54230	0.05924	-3.19435	0.01796
Romboutsia	0.02692	0.00297	-3.18068	0.03446
Ligilactobacillus	0.75944	0.10933	-2.79627	0.00094
Limosilactobacillus	0.09472	0.01391	-2.76759	0.02018

Table 4. Significantly different genera with at least 2-fold change for each pairwise comparison of three rat groups obtained from taxonomic DA analysis (continued)

Lactobacillus	0.40304	0.07809	-2.36781	0.02502
Schaedlerella	0.29564	0.07794	-1.92348	0.01551
Anaerotignum	0.01577	0.05965	1.91937	0.01070
Emergencia	0.13611	0.05379	-1.33946	0.00093
UBA1394	0.18854	0.08801	-1.09908	0.01887
Epi vs No-Epi				
Genus	**Epi median**	**No-Epi median**	**Epi/No-Epi Log2(Fold Change)**	**p.adj**
Scybalousia	0.00044	0.01944	-5.46161	0.00819
MGBC114079	0.01013	0.06247	-2.62431	0.04764
RGIG8607	0.00490	0.02678	-2.44980	0.04035
Pseudobutyricicoccus	0.00768	0.02041	-1.40921	0.01466

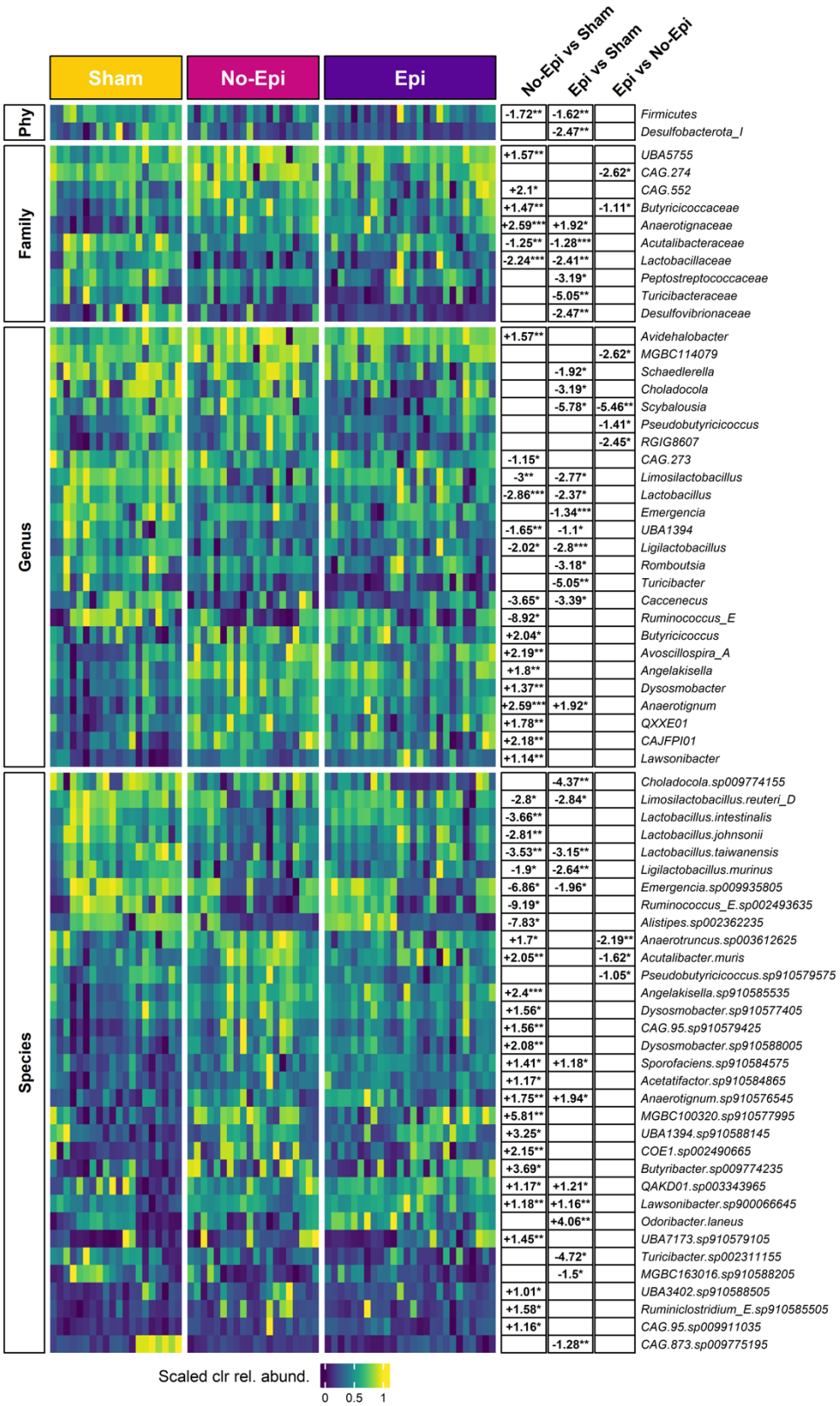

Figure 8. Differential abundance analysis of the taxonomic compositions of samples in phylum, family, genus, and species levels.

Figure 8. Differential abundance analysis of the taxonomic compositions of samples in phylum, family, genus, and species levels (continued)
Relative abundances of each taxon were clr-transformed, and scaled between 0 and 1 to increase contrast and better discrimination the compositional differences across samples and groups. The annotation boxes on the left of heatmap shows the taxonomic level the heatmap parts belong to (Phy = Phylum). Three columns on the right side of the heatmap belong to taxa which were significantly changed between two rat groups tested. Bacteria with at least two-fold change between compared groups, and with p.adj < 0.05 obtained from Kruskal Wallis test followed by Dunn's Test with Bonferroni correction were shown only. Significance stars represent; "*" for p.adj<0.05, "**" for p.adj < 0.01, and "***" for p.adj < 0.001. "+" and "-" indicate direction of the changes in relative abundance of the respective taxa between compared rat groups.

From the comparison of Epi and Sham rats, amongst the 12 genera; *Anaerotignum* ($Sham_{median}$ = 0.0158%, Epi_{median} = 0.0597%) is the only genus detected in higher abundance in Epi, while the rest, namely, *Scybalousia* ($Sham_{median}$ = 0.0242%, Epi_{median} = 0.0004%), *Turicibacter* ($Sham_{median}$ = 0.0061%, Epi_{median} = 0.0002%), *Caccenecus* ($Sham_{median}$ = 0.0762%, Epi_{median} = 0.0073%), *Choladocola* ($Sham_{median}$ = 0.5423%, Epi_{median} = 0.0592%), *Romboutsia* ($Sham_{median}$ = 0.0269%, Epi_{median} = 0.0030%), *Ligilactobacillus* ($Sham_{median}$ = 0.7594%, Epi_{median} = 0.1093%), *Limosilactobacillus* ($Sham_{median}$ = 0.0947%, Epi_{median} = 0.0139%), *Lactobacillus* ($Sham_{median}$ = 0.4030%, Epi_{median} = 0.0781%), *Schaedlerella* ($Sham_{median}$ = 0.2956%, Epi_{median} = 0.0779%), *Emergencia* ($Sham_{median}$ = 0.1361%, Epi_{median} = 0.0538%), and *UBA1394* ($Sham_{median}$ = 0.1885%, Epi_{median} = 0.0880%) were in lower abundance in Epi group.

Lastly, Epi vs No-Epi comparison showed decreased relative abundance of 4 genera in Epi group; *Scybalousia* ($No\text{-}Epi_{median}$ = 0.0194%, Epi_{median} = 0.0004%), *MGBC114079* ($No\text{-}Epi_{median}$ = 0.0625%, Epi_{median} = 0.0101%), *RGIG8607* ($No\text{-}Epi_{median}$ = 0.0268%, Epi_{median} = 0.0049%), and *Pseudobutyricicoccus* ($No\text{-}Epi_{median}$ = 0.0204%, Epi_{median} = 0.0077%).

4.1.2.2.4.4 Species level

DA analysis in species level revealed 27, 13, and 3 species for No-Epi vs Sham, Epi vs Sham, and Epi vs No-Epi comparisons, respectively (Table 5).

Decreased levels of *Ruminococcus_E sp002493635* ($Sham_{median}$ = 1.5973%, No-Epi_{median} = 0.0027%), *Alistipes sp002362235* ($Sham_{median}$ = 1.2113%, No-Epi_{median} = 0.0053%), *Emergencia sp009935805* ($Sham_{median}$ = 0.0565%, No-Epi_{median} = 0.0005%), *Lactobacillus intestinalis* ($Sham_{median}$ = 0.0101%, No-Epi_{median} = 0.0008%), *Lactobacillus taiwanensis* ($Sham_{median}$ = 0.0181%, No-Epi_{median} = 0.0016%), *Lactobacillus johnsonii* ($Sham_{median}$ = 0.3394%, No-Epi_{median} = 0.0482%), *Limosilactobacillus reuteri_D* ($Sham_{median}$ = 0.1320%, No-Epi_{median} = 0.0190%), and *Ligilactobacillus murinus* ($Sham_{median}$ = 0.8685%, No-Epi_{median} = 0.2328%), and increased abundances of *MGBC100320 sp910577995* ($Sham_{median}$ = 0.0022%, No-Epi_{median} = 0.1223%), *Butyribacter sp009774235* ($Sham_{median}$ = 0.0054%, No-Epi_{median} = 0.0694%), *UBA1394 sp910588145* ($Sham_{median}$ = 0.0019%, No-Epi_{median} = 0.0182%), *Angelakisella sp910585535* ($Sham_{median}$ = 0.0149%, No-Epi_{median} = 0.0783%), *COE1 sp002490665* ($Sham_{median}$ = 0.0280%, No-Epi_{median} = 0.1245%), *Dysosmobacter sp910588005* ($Sham_{median}$ = 0.0678%, No-Epi_{median} = 0.2874%), *Acutalibacter muris* ($Sham_{median}$ = 0.0471%, No-Epi_{median} = 0.1950%), *Anaerotignum sp910576545* ($Sham_{median}$ = 0.0135%, No-Epi_{median} = 0.0454%), *Anaerotruncus sp003612625* ($Sham_{median}$ = 0.0141%, No-Epi_{median} = 0.0458%), *Ruminiclostridium_E sp910585505* ($Sham_{median}$ = 0.0030%, No-Epi_{median} = 0.0089%), *CAG-95 sp910579425* ($Sham_{median}$ = 0.0099%, No-Epi_{median} = 0.0293%), *Dysosmobacter sp910577405* ($Sham_{median}$ = 0.0306%, No-Epi_{median} = 0.0902%), *UBA7173 sp910579105* ($Sham_{median}$ = 0.0049%, No-Epi_{median} = 0.0134%), *Sporofaciens sp910584575* ($Sham_{median}$ = 0.0207%, No-Epi_{median} = 0.0551%), *Lawsonibacter sp900066645* ($Sham_{median}$ = 0.0329%, No-Epi_{median} = 0.0744%), *QAKD01 sp003343965* ($Sham_{median}$ = 0.0411%, No-Epi_{median} = 0.0924%), *Acetatifactor sp910584865* ($Sham_{median}$ = 0.0244%, No-Epi_{median} = 0.0549%), *CAG-95 sp009911035* ($Sham_{median}$ = 0.0056%, No-Epi_{median} = 0.0124%), and *UBA3402 sp910588505* ($Sham_{median}$ = 0.0148%, No-Epi_{median} = 0.0299%) were observed in No-Epi compared to Sham group.

Table 5. Significantly different species with at least 2-fold change for each pairwise comparison of three rat groups obtained from taxonomic DA analysis

No-Epi vs Sham				
Species	**Sham median**	**No-Epi median**	**No-Epi/Sham Log2(Fold Change)**	**p.adj**
Ruminococcus_E sp002493635	1.59735	0.00274	-9.18607	0.02981
Alistipes sp002362235	1.21131	0.00532	-7.83156	0.01122
Emergencia sp009935805	0.05650	0.00049	-6.85807	0.02644
MGBC100320 sp910577995	0.00218	0.12235	5.80774	0.00570
Butyribacter sp009774235	0.00538	0.06938	3.68829	0.01921
Lactobacillus intestinalis	0.01013	0.00080	-3.66247	0.00721
Lactobacillus taiwanensis	0.01814	0.00158	-3.52530	0.00510
UBA1394 sp910588145	0.00191	0.01825	3.25324	0.03950
Lactobacillus johnsonii	0.33938	0.04824	-2.81468	0.00815
Limosilactobacillus reuteri_D	0.13205	0.01898	-2.79837	0.01492
Angelakisella sp910585535	0.01487	0.07831	2.39688	0.00040
COE1 sp002490665	0.02797	0.12451	2.15433	0.00144
Dysosmobacter sp910588005	0.06779	0.28740	2.08395	0.00162

Table 5. Significantly different species with at least 2-fold change for each pairwise comparison of three rat groups obtained from taxonomic DA analysis (continued)

Acutalibacter muris	0.04709	0.19501	2.04993	0.00655
Ligilactobacillus murinus	0.86854	0.23282	-1.89935	0.01122
Anaerotignum sp910576545	0.01348	0.04543	1.75283	0.00772
Anaerotruncus sp003612625	0.01409	0.04582	1.70111	0.03277
Ruminiclostridium_E sp910585505	0.00297	0.00889	1.57886	0.04330
CAG-95 sp910579425	0.00993	0.02929	1.56102	0.00731
Dysosmobacter sp910577405	0.03060	0.09018	1.55935	0.03950
UBA7173 sp910579105	0.00490	0.01335	1.44637	0.00907
Sporofaciens sp910584575	0.02069	0.05506	1.41214	0.01921
Lawsonibacter sp900066645	0.03287	0.07436	1.17791	0.00342
QAKD01 sp003343965	0.04108	0.09238	1.16921	0.02458
Acetatifactor sp910584865	0.02442	0.05490	1.16869	0.02399
CAG-95 sp009911035	0.00557	0.01242	1.15794	0.01652
UBA3402 sp910588505	0.01480	0.02985	1.01174	0.03052

Table 5. Significantly different species with at least 2-fold change for each pairwise comparison of three rat groups obtained from taxonomic DA analysis (continued)

Epi vs Sham				
Species	**Sham median**	**Epi median**	**Epi/Sham Log2(Fold Change)**	**p.adj**
Turicibacter sp002311155	0.00741	0.00028	-4.71716	0.01460
Choladocola sp009774155	0.57281	0.02764	-4.37346	0.00375
Odoribacter laneus	0.02073	0.34542	4.05840	0.00977
Lactobacillus taiwanensis	0.01814	0.00205	-3.14750	0.00621
Limosilactobacillus reuteri_D	0.13205	0.01849	-2.83624	0.02850
Ligilactobacillus murinus	0.86854	0.13964	-2.63687	0.00212
Emergencia sp009935805	0.05650	0.01453	-1.95939	0.04440
Anaerotignum sp910576545	0.01348	0.05156	1.93535	0.04591
MGBC163016 sp910588205	0.00292	0.00104	-1.49521	0.03010
CAG-873 sp009775195	0.01889	0.00779	-1.27695	0.00178
QAKD01 sp003343965	0.04108	0.09529	1.21404	0.02492
Sporofaciens sp910584575	0.02069	0.04694	1.18197	0.03908
Lawsonibacter sp900066645	0.03287	0.07329	1.15701	0.00734

Table 5. Significantly different species with at least 2-fold change for each pairwise comparison of three rat groups obtained from taxonomic DA analysis (continued)

Epi vs No-Epi				
Species	**Epi median**	**No-Epi median**	**Epi/No-Epi Log2(Fold Change)**	**p.adj**
Anaerotruncus sp003612625	0.01006	0.04582	-2.18691	0.00730
Acutalibacter muris	0.06325	0.19501	-1.62450	0.02923
Pseudobutyricicoccus sp910579575	0.01259	0.02605	-1.04851	0.01519

Epi vs Sham comparison revealed increased *Odoribacter laneus* ($Sham_{median}$ = 0.0207%, Epi_{median} = 0.3454%), *Anaerotignum sp910576545* ($Sham_{median}$ = 0.0135%, Epi_{median} = 0.0516%), *QAKD01 sp003343965* ($Sham_{median}$ = 0.0411%, Epi_{median} = 0.0953%), *Sporofaciens sp910584575* ($Sham_{median}$ = 0.0207%, Epi_{median} = 0.0469%), and *Lawsonibacter sp900066645* ($Sham_{median}$ = 0.0329%, Epi_{median} = 0.0733%), conversely, decreased abundances of *Turicibacter sp002311155* ($Sham_{median}$ = 0.0074%, Epi_{median} = 0.0003%), *Choladocola sp009774155* ($Sham_{median}$ = 0.5728%, Epi_{median} = 0.0276%), *Lactobacillus taiwanensis* ($Sham_{median}$ = 0.0181%, Epi_{median} = 0.0020%), *Limosilactobacillus reuteri_D* ($Sham_{median}$ = 0.1320%, Epi_{median} = 0.0185%), *Ligilactobacillus murinus* ($Sham_{median}$ = 0.8685%, Epi_{median} = 0.1396%), *Emergencia sp009935805* ($Sham_{median}$ = 0.0565%, Epi_{median} = 0.0145%), *MGBC163016 sp910588205* ($Sham_{median}$ = 0.0029%, Epi_{median} = 0.0010%), and *CAG-873 sp009775195* ($Sham_{median}$ = 0.0189%, Epi_{median} = 0.0078%) for Epi rats.

Species level comparison of Epi and No-Epi revealed lower abundance of *Anaerotruncus sp003612625* ($No\text{-}Epi_{median}$ = 0.0458%, Epi_{median} = 0.0101%), *Acutalibacter muris* ($No\text{-}Epi_{median}$ = 0.1950%, Epi_{median} = 0.0632%), and *Pseudobutyricicoccus sp910579575* ($No\text{-}Epi_{median}$ = 0.0260%, Epi_{median} = 0.0126%) in Epi than No-Epi.

4.1.2.3 Functional analysis results

4.1.2.3.1 Functional differential abundance analysis

A total of 874,616 protein-coding gene sequences were predicted, and 364,330 (41.66 %) of them were annotated to 5179 unique KEGG orthologs (KO's). 5075 KO's were retained after manual filtering prior differential abundance analysis. As in taxonomic DA analysis, KO's with at least 2-fold change between the compared groups, and with adjusted p value less than 0.05 were used in downstream analysis (Figure 9). Accordingly, total of 276 KO's from No-Epi vs Sham (59 are higher, and 217 are lower in No-Epi compared to Sham), 195 KO's from Epi vs Sham (16 are higher, and 179 are lower in Epi compared to Sham), and 36 KO's from Epi vs No-Epi (18 are higher, and 18 are lower in Epi compared to No-Epi) comparisons (Table 6).

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis

No-Epi vs Sham				
KO ID	**Sham median**	**No-Epi median**	**No-Epi/Sham Log2(Fold Change)**	**p.adj**
K18285	0.00368	0.00001	-9.30324	0.00355
K11784	0.00171	0.00001	-8.19973	0.01482
K11782	0.00120	0.00001	-7.68557	0.01598
K11785	0.00101	0.00001	-7.43268	0.00862
K12206	0.00080	0.00001	-7.10667	0.00567
K11783	0.00051	0.00001	-6.43836	0.00354
K12217	0.00046	0.00001	-6.30614	0.00870
K12203	0.00029	0.00001	-5.62294	0.01109
K12218	0.00028	0.00001	-5.57099	0.00568
K12209	0.00027	0.00001	-5.54924	0.04567
K22330	0.00021	0.00001	-5.15617	0.00844

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K25588	0.00018	0.00001	-4.91622	0.01853
K24915	0.00016	0.00001	-4.80960	0.00028
K05304	0.00001	0.00016	4.74184	0.00770
K12204	0.00015	0.00001	-4.71238	0.00910
K12213	0.00014	0.00001	-4.57059	0.03674
K00966	0.00001	0.00010	4.15251	0.02466
K11707	0.00001	0.00010	4.11202	0.03276
K13727	0.00022	0.00001	-3.98915	0.00232
K09979	0.00026	0.00002	-3.96321	0.00219
K20391	0.00008	0.00001	-3.85078	0.02968
K15976	0.00024	0.00002	-3.79022	0.00007
K05020	0.00001	0.00008	3.78096	0.03709
K05296	0.00008	0.00001	-3.76420	0.02736
K22394	0.00021	0.00002	-3.31096	0.00162
K20811	0.00087	0.00010	-3.17977	0.00055
K12244	0.00006	0.00053	3.14652	0.01186
K10530	0.00100	0.00012	-3.11167	0.00012
K03693	0.00161	0.00019	-3.10191	0.00026
K02081	0.00124	0.00015	-3.08962	0.00797
K09155	0.00084	0.00010	-3.05977	0.00114
K25641	0.00047	0.00006	-3.04262	0.00347
K01907	0.00003	0.00028	3.04214	0.01623
K23248	0.00001	0.00012	3.04012	0.03813
K03095	0.00026	0.00003	-3.02283	0.00189
K20112	0.00022	0.00003	-3.01328	0.00900
K10353	0.00037	0.00005	-3.00942	0.00427
K25113	0.00005	0.00001	-3.00548	0.03917
K22935	0.00003	0.00020	2.98601	0.00283
K00158	0.00121	0.00015	-2.97236	0.00069

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K22580	0.00105	0.00014	-2.91557	0.00143
K22393	0.00022	0.00003	-2.91522	0.00204
K03182	0.00285	0.00038	-2.89160	0.01803
K08987	0.00037	0.00005	-2.85313	0.00101
K18673	0.00067	0.00010	-2.81878	0.00214
K09694	0.00128	0.00018	-2.81685	0.00554
K17073	0.00123	0.00018	-2.80793	0.00121
K18104	0.00067	0.00010	-2.80641	0.01447
K06994	0.00222	0.00032	-2.79684	0.00223
K17215	0.00003	0.00020	2.78441	0.02630
K02840	0.00097	0.00015	-2.72172	0.00141
K17677	0.00095	0.00014	-2.71972	0.01207
K02788	0.00156	0.00024	-2.70025	0.00135
K01361	0.00871	0.00138	-2.66215	0.00016
K15770	0.00132	0.00021	-2.65995	0.00287
K10041	0.00041	0.00007	-2.65841	0.00068
K20344	0.00190	0.00030	-2.65819	0.00539
K07393	0.00110	0.00017	-2.65765	0.00860
K21567	0.00062	0.00010	-2.59654	0.02228
K03647	0.00035	0.00006	-2.58875	0.00018
K03328	0.00106	0.00018	-2.57455	0.01122
K02786	0.00013	0.00002	-2.56586	0.00235
K02243	0.00120	0.00020	-2.56167	0.00093
K13929	0.00009	0.00051	2.54894	0.01758
K08152	0.00148	0.00025	-2.53901	0.02399
K05796	0.00003	0.00001	-2.53085	0.02168
K08728	0.00018	0.00003	-2.52892	0.00431
K05685	0.00021	0.00004	-2.51740	0.01662
K01597	0.00104	0.00018	-2.50509	0.00153

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K01641	0.00148	0.00026	-2.49059	0.00054
K25157	0.00212	0.00038	-2.48803	0.00042
K01220	0.00120	0.00021	-2.48173	0.00201
K08996	0.00032	0.00006	-2.46483	0.00750
K01635	0.00169	0.00031	-2.43667	0.00510
K22116	0.00001	0.00003	2.43023	0.00933
K08724	0.00362	0.00067	-2.42758	0.00052
K02248	0.00033	0.00006	-2.42314	0.00089
K10810	0.00125	0.00023	-2.41809	0.00201
K03713	0.00009	0.00002	-2.40043	0.03475
K16169	0.00210	0.00040	-2.39691	0.00082
K02859	0.00019	0.00004	-2.38887	0.00173
K01047	0.00076	0.00015	-2.33814	0.02639
K01575	0.00067	0.00013	-2.33765	0.00169
K02076	0.00041	0.00008	-2.32375	0.00786
K13953	0.00139	0.00028	-2.31621	0.00342
K02244	0.00099	0.00020	-2.31465	0.00165
K23265	0.00045	0.00009	-2.30011	0.01302
K19540	0.00032	0.00158	2.29788	0.01122
K00086	0.00034	0.00007	-2.29646	0.04083
K07351	0.00003	0.00001	-2.28750	0.02859
K18692	0.00170	0.00036	-2.25916	0.00207
K01916	0.00140	0.00029	-2.25879	0.00153
K16925	0.00073	0.00015	-2.25734	0.00109
K23978	0.00053	0.00011	-2.24902	0.00489
K03716	0.00043	0.00009	-2.24692	0.02112
K03697	0.00292	0.00062	-2.24336	0.00096
K00938	0.00087	0.00018	-2.23504	0.00105
K16323	0.00162	0.00035	-2.23353	0.00178

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K18918	0.00003	0.00001	-2.23066	0.02681
K09976	0.00008	0.00002	-2.22891	0.01463
K00841	0.00219	0.00047	-2.22084	0.00263
K18891	0.00224	0.00048	-2.21859	0.00088
K02794	0.00222	0.00048	-2.21248	0.00213
K00867	0.00131	0.00028	-2.20586	0.00482
K21910	0.00001	0.00003	2.20581	0.00674
K23269	0.00205	0.00044	-2.20468	0.01152
K10550	0.00003	0.00001	-2.19720	0.02198
K02077	0.00127	0.00028	-2.19489	0.00115
K07128	0.00026	0.00006	-2.18521	0.02494
K23977	0.00003	0.00001	-2.17924	0.01756
K16509	0.00035	0.00008	-2.17099	0.00055
K06164	0.00018	0.00004	-2.16526	0.02299
K03322	0.00163	0.00037	-2.15366	0.00226
K18892	0.00275	0.00062	-2.14716	0.00077
K00869	0.00093	0.00021	-2.14651	0.00240
K23149	0.00029	0.00125	2.13539	0.00883
K05338	0.00003	0.00001	-2.13345	0.00138
K23304	0.00064	0.00015	-2.12671	0.01245
K09940	0.00091	0.00021	-2.11944	0.02341
K11733	0.00163	0.00038	-2.11223	0.01037
K04086	0.00279	0.00065	-2.11094	0.00255
K05311	0.00123	0.00028	-2.10987	0.00418
K02530	0.00079	0.00018	-2.10860	0.00255
K01728	0.02415	0.00561	-2.10553	0.03771
K01274	0.00127	0.00030	-2.10066	0.00313
K19005	0.00789	0.00184	-2.09714	0.00074
K09685	0.00225	0.00053	-2.09283	0.00019

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K22212	0.00128	0.00031	-2.07022	0.00702
K02245	0.00012	0.00003	-2.06495	0.00498
K07693	0.00031	0.00007	-2.06409	0.00361
K01595	0.00144	0.00035	-2.05336	0.00137
K03489	0.00074	0.00018	-2.04768	0.00331
K22130	0.00003	0.00011	2.04560	0.01748
K03179	0.00124	0.00030	-2.03313	0.02160
K03078	0.00046	0.00011	-2.02654	0.01779
K11102	0.00080	0.00020	-2.02500	0.03683
K06216	0.00116	0.00029	-2.01345	0.00384
K09699	0.00095	0.00024	-2.00282	0.00860
K12339	0.00076	0.00019	-2.00021	0.00236
K02240	0.00187	0.00047	-1.98932	0.00055
K02531	0.00073	0.00018	-1.98476	0.00112
K03048	0.00083	0.00021	-1.98420	0.00932
K06191	0.00004	0.00001	-1.98375	0.01026
K07586	0.00112	0.00029	-1.95701	0.00883
K01911	0.00156	0.00040	-1.95349	0.00620
K13788	0.00020	0.00005	-1.94941	0.04560
K01819	0.00053	0.00014	-1.93342	0.00503
K09698	0.00298	0.00078	-1.92826	0.00195
K02532	0.00108	0.00029	-1.92098	0.01329
K16235	0.00132	0.00035	-1.91643	0.01737
K02798	0.00019	0.00005	-1.91322	0.04596
K14205	0.00279	0.00074	-1.91193	0.00711
K04025	0.00002	0.00001	-1.91083	0.03099
K20109	0.00002	0.00001	-1.86481	0.04116
K02796	0.00337	0.00093	-1.86286	0.01214
K02679	0.00002	0.00001	-1.86145	0.04151

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K22489	0.00002	0.00001	-1.86145	0.04847
K25261	0.00002	0.00001	-1.86038	0.02628
K22432	0.00029	0.00008	-1.85482	0.02255
K06162	0.00038	0.00010	-1.85449	0.04196
K19789	0.00090	0.00025	-1.84379	0.00074
K16209	0.00530	0.00148	-1.83873	0.00248
K25158	0.00075	0.00021	-1.83248	0.00207
K22103	0.00047	0.00013	-1.82947	0.00123
K10037	0.00288	0.00082	-1.80901	0.00751
K03366	0.00075	0.00021	-1.80348	0.03683
K06956	0.00112	0.00032	-1.79917	0.02340
K21417	0.00004	0.00015	1.79648	0.00518
K12270	0.00143	0.00041	-1.79566	0.00570
K19784	0.00226	0.00066	-1.77840	0.01382
K20374	0.00046	0.00013	-1.76984	0.04954
K16928	0.00005	0.00016	1.76304	0.04296
K18983	0.00058	0.00195	1.75747	0.00883
K24967	0.00099	0.00029	-1.75441	0.00031
K07467	0.00039	0.00130	1.74047	0.00586
K02069	0.00339	0.00102	-1.73760	0.00148
K07683	0.00002	0.00001	-1.72279	0.01337
K18926	0.00891	0.00270	-1.71971	0.00637
K06286	0.00249	0.00076	-1.70788	0.02019
K23078	0.00175	0.00558	1.67051	0.00342
K12555	0.00214	0.00067	-1.66798	0.00731
K24163	0.00253	0.00080	-1.66432	0.00255
K18118	0.00010	0.00031	1.66431	0.01413
K00054	0.00159	0.00050	-1.66318	0.00123
K09116	0.00097	0.00306	1.66068	0.02708

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K03492	0.00215	0.00068	-1.65317	0.02341
K02302	0.00159	0.00497	1.64570	0.02708
K25308	0.00042	0.00013	-1.63386	0.02704
K19134	0.00086	0.00267	1.63134	0.01214
K02246	0.00050	0.00016	-1.63063	0.00395
K09136	0.00214	0.00069	-1.62377	0.02981
K16182	0.00011	0.00004	-1.59482	0.00184
K01039	0.00008	0.00025	1.57559	0.02609
K09946	0.00046	0.00136	1.55957	0.00263
K07978	0.00005	0.00016	1.53593	0.02897
K06948	0.00161	0.00056	-1.53015	0.01010
K04047	0.00168	0.00061	-1.47541	0.00183
K21739	0.00873	0.00314	-1.47496	0.00112
K10240	0.00584	0.01611	1.46507	0.00189
K11912	0.00002	0.00001	-1.46330	0.02450
K12541	0.00002	0.00001	-1.46330	0.02128
K01281	0.00421	0.00153	-1.45812	0.00109
K05340	0.00190	0.00069	-1.45178	0.00045
K01751	0.00399	0.01082	1.43838	0.00082
K17108	0.01808	0.00669	-1.43369	0.00469
K18578	0.00326	0.00121	-1.42855	0.00116
K25227	0.00424	0.01130	1.41455	0.04136
K23059	0.00308	0.00116	-1.41089	0.03435
K10537	0.00014	0.00038	1.40895	0.01970
K07223	0.00106	0.00040	-1.39479	0.00207
K01048	0.00069	0.00180	1.37161	0.01346
K22736	0.00131	0.00051	-1.35215	0.01531
K07570	0.00023	0.00009	-1.34961	0.00538
K07000	0.00934	0.00367	-1.34539	0.01970

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K00036	0.00457	0.00180	-1.34533	0.00603
K18986	0.00159	0.00400	1.33249	0.00637
K00441	0.00035	0.00087	1.32046	0.00384
K03346	0.00212	0.00085	-1.31725	0.01152
K01247	0.00345	0.00138	-1.31543	0.02399
K02622	0.00298	0.00120	-1.31254	0.00655
K05781	0.00012	0.00005	-1.31031	0.03813
K25181	0.00102	0.00042	-1.30087	0.01122
K00887	0.00080	0.00032	-1.30020	0.00395
K09167	0.00018	0.00007	-1.28165	0.00246
K15554	0.00251	0.00606	1.27425	0.00079
K00613	0.00001	0.00001	-1.25361	0.01096
K07533	0.00200	0.00084	-1.24389	0.00158
K01604	0.00024	0.00057	1.24024	0.00731
K02791	0.00228	0.00097	-1.23703	0.04042
K01026	0.00185	0.00434	1.22835	0.00007
K03724	0.02068	0.04839	1.22650	0.02285
K18940	0.00210	0.00091	-1.21624	0.01738
K10039	0.00342	0.00148	-1.21178	0.00167
K20608	0.00587	0.00253	-1.21134	0.02708
K02825	0.00260	0.00113	-1.20611	0.00304
K15896	0.00058	0.00025	-1.20573	0.00711
K23541	0.00017	0.00040	1.20413	0.02122
K01531	0.00478	0.00208	-1.19730	0.02399
K00712	0.00442	0.00193	-1.19654	0.01611
K14977	0.00021	0.00047	1.18496	0.00362
K07494	0.00922	0.02091	1.18112	0.00226
K03740	0.00226	0.00101	-1.16644	0.01182
K07652	0.00564	0.00253	-1.15790	0.00313

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K05823	0.00129	0.00058	-1.15134	0.04532
K07014	0.00330	0.00149	-1.14843	0.01970
K02020	0.00207	0.00456	1.14092	0.00096
K10241	0.00202	0.00444	1.13816	0.03685
K00087	0.00740	0.01617	1.12740	0.02285
K18898	0.00219	0.00101	-1.11880	0.01570
K01421	0.01508	0.00701	-1.10588	0.00015
K13480	0.00088	0.00189	1.10197	0.00384
K19337	0.00181	0.00386	1.09445	0.02175
K01464	0.00309	0.00657	1.08628	0.00793
K02481	0.00474	0.00223	-1.08521	0.03126
K12269	0.00314	0.00149	-1.07856	0.04232
K02781	0.00026	0.00012	-1.07242	0.03435
K09961	0.00112	0.00235	1.06801	0.02910
K05306	0.00167	0.00080	-1.06117	0.01346
K00342	0.00031	0.00064	1.05702	0.00932
K01807	0.00296	0.00143	-1.05326	0.00692
K18119	0.00038	0.00080	1.05099	0.00034
K02795	0.00263	0.00127	-1.04972	0.02580
K11537	0.00221	0.00107	-1.04793	0.01382
K01295	0.00134	0.00276	1.04143	0.00069
K09118	0.00672	0.00326	-1.04123	0.01782
K04031	0.00051	0.00106	1.03982	0.01010
K03831	0.00187	0.00383	1.03535	0.00148
K02241	0.00348	0.00713	1.03463	0.01652
K00363	0.00001	0.00001	-1.03061	0.02146
K03739	0.00238	0.00117	-1.02636	0.01214
K18928	0.00231	0.00114	-1.02061	0.00603
K07141	0.00186	0.00377	1.01862	0.00072

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K01846	0.00009	0.00018	1.01761	0.04959
K02488	0.00289	0.00583	1.01369	0.00287
K02019	0.00055	0.00110	1.00392	0.00373
K03491	0.00150	0.00301	1.00152	0.01214
Epi vs Sham				
KO ID	**Sham median**	**Epi median**	**Epi/Sham Log2(Fold Change)**	**p.adj**
K22684	0.00197	0.00001	-8.39829	0.02240
K12203	0.00029	0.00001	-5.62294	0.00600
K12218	0.00028	0.00001	-5.57099	0.02295
K07057	0.00001	0.00027	5.53328	0.04468
K16922	0.00329	0.00008	-5.41773	0.00564
K22330	0.00021	0.00001	-5.15617	0.00256
K25588	0.00018	0.00001	-4.91622	0.04994
K12217	0.00046	0.00002	-4.90234	0.01945
K12204	0.00015	0.00001	-4.71238	0.00151
K06607	0.00014	0.00001	-4.55259	0.00852
K12206	0.00080	0.00004	-4.42336	0.02631
K00261	0.00048	0.00002	-4.40834	0.00036
K01533	0.00009	0.00001	-3.99431	0.00201
K03713	0.00009	0.00001	-3.97398	0.00166
K13727	0.00022	0.00001	-3.87241	0.01218
K05296	0.00008	0.00001	-3.76420	0.01046
K00809	0.00001	0.00019	3.73144	0.01889
K09667	0.00007	0.00001	-3.52671	0.01284
K24915	0.00016	0.00001	-3.48543	0.00316
K08153	0.00028	0.00003	-3.27996	0.03108
K08969	0.00015	0.00002	-3.09525	0.02541
K25113	0.00005	0.00001	-3.00548	0.01363

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K21567	0.00062	0.00008	-2.97968	0.00193
K09694	0.00128	0.00016	-2.96903	0.00084
K12205	0.00004	0.00001	-2.93879	0.04244
K15976	0.00024	0.00003	-2.88177	0.01063
K00086	0.00034	0.00005	-2.84319	0.03609
K02076	0.00041	0.00006	-2.82821	0.00016
K23265	0.00045	0.00006	-2.80392	0.00387
K02651	0.00004	0.00001	-2.79133	0.02128
K23269	0.00205	0.00030	-2.78502	0.00137
K09729	0.00004	0.00001	-2.76239	0.00829
K01597	0.00104	0.00016	-2.71048	0.00121
K17677	0.00095	0.00015	-2.69942	0.00680
K06191	0.00004	0.00001	-2.69873	0.00268
K03328	0.00106	0.00016	-2.68579	0.00250
K00869	0.00093	0.00014	-2.68196	0.00090
K02859	0.00019	0.00003	-2.67592	0.00348
K15770	0.00132	0.00021	-2.65713	0.00164
K22212	0.00128	0.00020	-2.65579	0.00236
K10353	0.00037	0.00006	-2.58218	0.01437
K02798	0.00019	0.00003	-2.56825	0.00073
K08152	0.00148	0.00025	-2.55437	0.00248
K22580	0.00105	0.00018	-2.54995	0.02211
K16869	0.00003	0.00001	-2.53428	0.04413
K03366	0.00075	0.00013	-2.49340	0.00096
K06878	0.00006	0.00001	-2.49188	0.01175
K18673	0.00067	0.00012	-2.48692	0.02186
K01047	0.00076	0.00014	-2.48458	0.00455
K09155	0.00084	0.00015	-2.47799	0.01193
K09699	0.00095	0.00017	-2.47048	0.00050

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K06994	0.00222	0.00040	-2.46820	0.02462
K17073	0.00123	0.00023	-2.43578	0.01437
K25308	0.00042	0.00008	-2.43569	0.00134
K14205	0.00279	0.00052	-2.42449	0.00131
K07778	0.00076	0.00014	-2.42279	0.00100
K10041	0.00041	0.00008	-2.41526	0.01864
K10037	0.00288	0.00054	-2.40611	0.00174
K16323	0.00162	0.00031	-2.37865	0.00083
K20374	0.00046	0.00009	-2.37598	0.00147
K00938	0.00087	0.00017	-2.35561	0.00498
K07393	0.00110	0.00022	-2.33937	0.00401
K04086	0.00279	0.00056	-2.32346	0.00110
K02840	0.00097	0.00019	-2.31972	0.00179
K22935	0.00003	0.00013	2.31516	0.03926
K16235	0.00132	0.00027	-2.31260	0.00246
K00841	0.00219	0.00045	-2.29607	0.00117
K03697	0.00292	0.00060	-2.28948	0.00226
K09167	0.00018	0.00004	-2.28761	0.00122
K01728	0.02415	0.00496	-2.28516	0.00375
K16925	0.00073	0.00015	-2.27652	0.00125
K18891	0.00224	0.00046	-2.26664	0.00305
K08987	0.00037	0.00008	-2.26264	0.00683
K12270	0.00143	0.00030	-2.26039	0.00069
K20811	0.00087	0.00018	-2.25615	0.02538
K16180	0.00013	0.00003	-2.24865	0.01069
K03489	0.00074	0.00016	-2.23862	0.00300
K03048	0.00083	0.00018	-2.23321	0.00072
K03693	0.00161	0.00034	-2.23243	0.02681
K25157	0.00212	0.00045	-2.21822	0.00201

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K01220	0.00120	0.00026	-2.20505	0.00758
K10530	0.00100	0.00022	-2.19478	0.01916
K19784	0.00226	0.00050	-2.19308	0.00148
K08724	0.00362	0.00080	-2.17773	0.00124
K01819	0.00053	0.00012	-2.17750	0.01388
K18892	0.00275	0.00061	-2.17501	0.00396
K03095	0.00026	0.00006	-2.17290	0.00483
K00158	0.00121	0.00027	-2.15940	0.01569
K02530	0.00079	0.00018	-2.15356	0.00505
K05338	0.00003	0.00001	-2.13345	0.00178
K02244	0.00099	0.00023	-2.12610	0.00146
K02240	0.00187	0.00044	-2.10082	0.00100
K09136	0.00214	0.00050	-2.09849	0.00102
K25641	0.00047	0.00011	-2.08519	0.03223
K02788	0.00156	0.00037	-2.08498	0.01009
K18692	0.00170	0.00040	-2.07152	0.00169
K02243	0.00120	0.00029	-2.05179	0.00469
K13929	0.00009	0.00036	2.05051	0.01735
K25158	0.00075	0.00018	-2.04564	0.00121
K07693	0.00031	0.00008	-2.04172	0.00085
K23304	0.00064	0.00016	-2.04100	0.00099
K16509	0.00035	0.00009	-2.03176	0.00139
K01635	0.00169	0.00041	-2.03161	0.01036
K02077	0.00127	0.00031	-2.01607	0.00582
K17213	0.00002	0.00010	2.01229	0.04222
K01575	0.00067	0.00017	-2.01168	0.00823
K02531	0.00073	0.00018	-2.00821	0.01070
K01274	0.00127	0.00032	-1.99026	0.00843
K01916	0.00140	0.00035	-1.98471	0.00512

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K16169	0.00210	0.00053	-1.98452	0.00299
K12268	0.00316	0.00081	-1.96860	0.00218
K06216	0.00116	0.00030	-1.95697	0.00460
K01911	0.00156	0.00041	-1.92864	0.00361
K07586	0.00112	0.00029	-1.92805	0.00392
K12555	0.00214	0.00056	-1.92649	0.01696
K03322	0.00163	0.00043	-1.91838	0.01745
K01430	0.00011	0.00003	-1.91529	0.00432
K08728	0.00018	0.00005	-1.91288	0.04603
K19005	0.00789	0.00214	-1.88587	0.00253
K13953	0.00139	0.00038	-1.88091	0.00181
K20109	0.00002	0.00001	-1.86481	0.00087
K25261	0.00002	0.00001	-1.86038	0.00012
K10810	0.00125	0.00034	-1.86020	0.01297
K11733	0.00163	0.00045	-1.84442	0.01590
K06948	0.00161	0.00045	-1.84306	0.00376
K00867	0.00131	0.00037	-1.82685	0.02014
K22130	0.00003	0.00010	1.82460	0.00407
K07645	0.00022	0.00078	1.81740	0.04738
K00054	0.00159	0.00045	-1.81622	0.00073
K02248	0.00033	0.00009	-1.80869	0.01038
K24163	0.00253	0.00073	-1.79981	0.00231
K02246	0.00050	0.00014	-1.78003	0.00449
K23978	0.00053	0.00015	-1.77887	0.00553
K12269	0.00314	0.00092	-1.76948	0.00049
K06286	0.00249	0.00074	-1.75840	0.00544
K03647	0.00035	0.00011	-1.72486	0.00333
K01641	0.00148	0.00045	-1.71965	0.00254
K02794	0.00222	0.00068	-1.71152	0.00674

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K25651	0.00170	0.00052	-1.71021	0.00851
K23264	0.00008	0.00002	-1.70869	0.01970
K03191	0.00064	0.00020	-1.70779	0.00089
K02796	0.00337	0.00103	-1.70440	0.00684
K18926	0.00891	0.00276	-1.69029	0.00219
K08956	0.00016	0.00051	1.68421	0.01088
K03492	0.00215	0.00067	-1.67914	0.00094
K11102	0.00080	0.00025	-1.67690	0.02021
K00712	0.00442	0.00139	-1.67347	0.00098
K07570	0.00023	0.00007	-1.67025	0.00274
K22432	0.00029	0.00009	-1.65117	0.01562
K09698	0.00298	0.00096	-1.64128	0.00496
K24967	0.00099	0.00032	-1.62560	0.00294
K01595	0.00144	0.00047	-1.60726	0.04888
K00836	0.00254	0.00084	-1.58988	0.00300
K05823	0.00129	0.00043	-1.58174	0.00666
K02245	0.00012	0.00004	-1.57599	0.02304
K05311	0.00123	0.00042	-1.55904	0.00539
K01281	0.00421	0.00146	-1.52626	0.00417
K09976	0.00008	0.00003	-1.51063	0.01061
K13665	0.00015	0.00042	1.50822	0.01251
K01772	0.00175	0.00062	-1.49753	0.00364
K07652	0.00564	0.00204	-1.46589	0.00005
K11912	0.00002	0.00001	-1.46330	0.00012
K12541	0.00002	0.00001	-1.46330	0.00011
K17466	0.00007	0.00020	1.45668	0.03386
K03346	0.00212	0.00078	-1.45045	0.00225
K18940	0.00210	0.00077	-1.44272	0.00228
K00590	0.00108	0.00040	-1.41445	0.04159

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K09740	0.00009	0.00024	1.40632	0.02226
K22103	0.00047	0.00018	-1.40465	0.02288
K07097	0.00134	0.00051	-1.40256	0.00133
K20541	0.00059	0.00153	1.36670	0.03893
K03739	0.00238	0.00093	-1.36139	0.00131
K23245	0.00043	0.00017	-1.33873	0.02106
K23078	0.00175	0.00442	1.33284	0.02562
K11144	0.00088	0.00036	-1.28767	0.01696
K16181	0.00011	0.00005	-1.27306	0.00610
K14089	0.00019	0.00008	-1.27083	0.02459
K01286	0.00523	0.00217	-1.26889	0.00885
K03740	0.00226	0.00094	-1.26397	0.00446
K03841	0.00032	0.00077	1.25608	0.02850
K00613	0.00001	0.00001	-1.25361	0.00515
K02622	0.00298	0.00127	-1.23223	0.02567
K14088	0.00050	0.00021	-1.21615	0.02231
K14292	0.00001	0.00001	-1.17495	0.04097
K03709	0.00079	0.00036	-1.13511	0.00031
K07120	0.00028	0.00061	1.12569	0.03361
K11312	0.00003	0.00005	1.11304	0.04182
K06927	0.00036	0.00017	-1.10279	0.00616
K00887	0.00080	0.00038	-1.08844	0.00818
K16013	0.00383	0.00180	-1.08621	0.01336
K13678	0.00109	0.00052	-1.07283	0.01391
K02621	0.00570	0.00276	-1.04352	0.01922
K06895	0.00057	0.00028	-1.01981	0.01822
K05340	0.00190	0.00094	-1.00825	0.01926
K03311	0.00731	0.00364	-1.00371	0.00103

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

Epi vs No-Epi				
KO ID	**No-Epi median**	**Epi median**	**Epi/No-Epi_ Log2(Fold Change)**	**p.adj**
K01907	0.00028	0.00002	-3.93400	0.03991
K05020	0.00008	0.00001	-3.78096	0.04038
K06311	0.00004	0.00001	-2.68355	0.03556
K23149	0.00125	0.00027	-2.24008	0.03465
K21910	0.00003	0.00001	-2.20581	0.00397
K00752	0.00031	0.00007	-2.11187	0.01513
K10561	0.00003	0.00011	2.08111	0.03163
K16650	0.00650	0.00217	-1.58215	0.01003
K01361	0.00138	0.00410	1.57611	0.04028
K22994	0.01629	0.00549	-1.56973	0.00539
K06996	0.00007	0.00020	1.56397	0.00346
K19244	0.00187	0.00065	-1.53454	0.00289
K00899	0.00041	0.00014	-1.48720	0.03531
K04835	0.00082	0.00032	-1.36442	0.00582
K19268	0.00076	0.00030	-1.33960	0.01294
K21741	0.00047	0.00118	1.31827	0.00506
K07115	0.00033	0.00083	1.31562	0.02228
K05908	0.00047	0.00110	1.22545	0.00635
K00844	0.00125	0.00293	1.22502	0.04120
K01846	0.00018	0.00008	-1.22372	0.00608
K08093	0.00044	0.00020	-1.17597	0.01414
K06039	0.00001	0.00001	1.17223	0.01610
K19793	0.00034	0.00076	1.15562	0.03466
K00318	0.00031	0.00068	1.13597	0.04206
K01903	0.00073	0.00159	1.12590	0.00737
K02443	0.00179	0.00083	-1.11896	0.00625

Table 6. Significantly different KO's with at least 2-fold change for each pairwise comparison of three rat groups obtained from functional DA analysis (continued)

K02302	0.00497	0.00230	-1.11155	0.01535
K10716	0.00057	0.00121	1.09198	0.01136
K18702	0.00116	0.00055	-1.05890	0.03051
K15876	0.00032	0.00067	1.05380	0.04285
K01531	0.00208	0.00427	1.03541	0.01291
K13015	0.00305	0.00150	-1.02790	0.01435
K15726	0.00789	0.01598	1.01895	0.02923
K18330	0.00067	0.00134	1.01413	0.01297
K23121	0.00159	0.00322	1.01190	0.02562
K16055	0.00127	0.00255	1.00842	0.00998

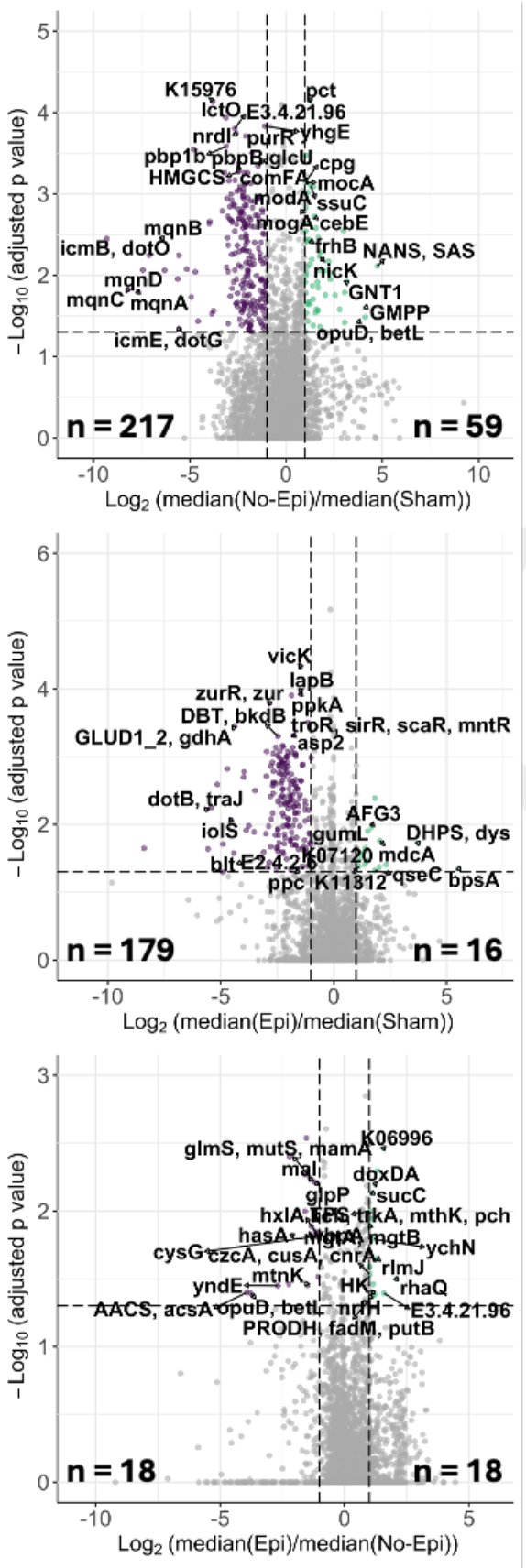

Figure 9. Differential abundance analysis of the KO annotated bacterial genes.

Figure 9. Differential abundance analysis of the KO annotated bacterial genes (continued)
Volcano plots are used to illustrate the KO abundance changes obtained from comparison of No-Epi vs Sham (left), Epi and Sham (middle), and Epi and No-Epi (right) groups. Plots were generated using fold changes obtained from median relative abundances of KO's in two groups, and p.adj values. Genes are represented by the data points, and dashed lines represent the p.adj and fold-change thresholds. KO's with at least two-fold change and with p.adj less than 0.05 were coloured, and decreased and increased KO's are submitted in purple and green colours, respectively. Numbers of the significantly changing bacteria were written at the bottom of each plot by bold text for each direction.

4.1.2.3.2 Pathway enrichment

KEGG pathway enrichment using significantly changing KO's for each comparison yielded pathways dominated with pathways classified under "Carbohydrate metabolism"; "Galactose metabolism" (No-Epi vs Sham, and Epi vs Sham), "Butanoate metabolism" (No-Epi vs Sham only), "Fructose and mannose metabolism" (No-Epi vs Sham only), "Pyruvate metabolism" (No-Epi vs Sham only), "Amino sugar and nucleotide sugar metabolism" (No-Epi vs Sham only), "C5-Branched dibasic acid metabolism" (Epi vs No-Epi only), and "Glyoxylate and dicarboxylate metabolism" (Epi vs No-Epi only). There were two "Glycan biosynthesis and metabolism" class pathways; "Teichoic acid biosynthesis" (No-Epi vs Sham, and Epi vs Sham), and "Peptidoglycan biosynthesis" (No-Epi vs Sham, and Epi vs Sham). Additionally, "Ubiquinone and other terpenoid-quinone biosynthesis" (No-Epi vs Sham only) classified under "Metabolism of cofactors and vitamins", "Terpenoid backbone biosynthesis" (No-Epi vs Sham, and Epi vs Sham) classified under "Metabolism of terpanoids and polyketides", and "Lysine biosynthesis" (Epi vs Sham only) under class "Amino acid metabolism" were other metabolism-associated pathways. Lastly, two "Membrane transport" class pathways; "Phosphotransferase system" (No-Epi vs Sham, and Epi vs Sham), and "ABC transporters" (No-Epi vs Sham only) were significantly enriched (Table 7 and Figure 10).

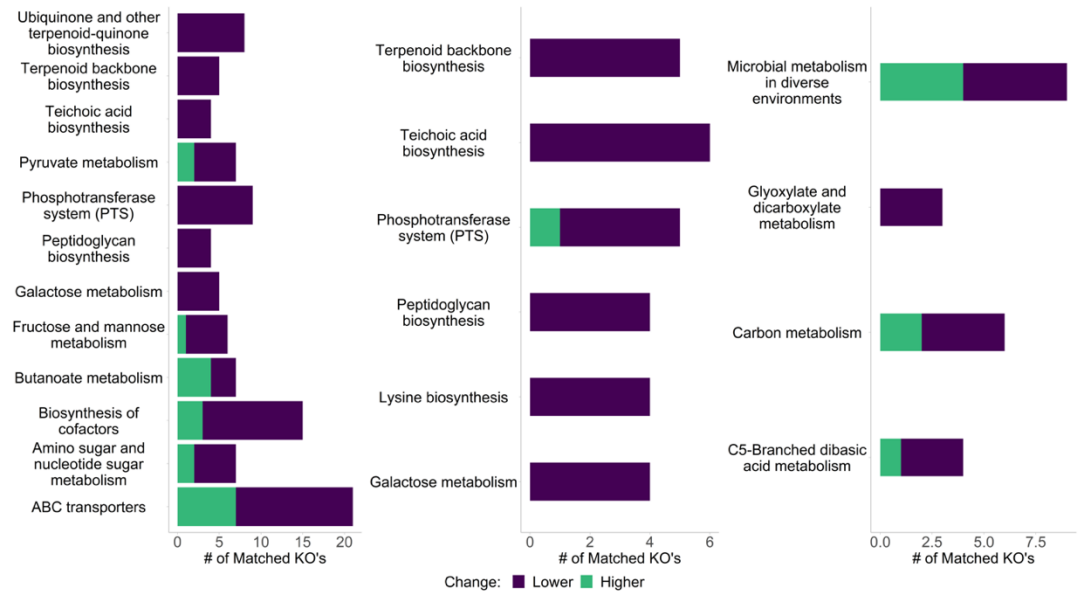

Figure 10. Stacked bar plots of KEGG Orthology pathways significantly enriched using differentially abundant KO's obtained from No-Epi vs Sham, Epi vs Sham, and Epi vs No-Epi comparisons.
X-axis represents the number of matched KO's to the background KO's associated with respective pathway, purple and green colours represent the decreased or increased KO's between compared groups, respectively.

Table 7. Significantly enriched pathways obtained by using significantly changing KO's from each pairwise comparison of three rat groups as query

No-Epi vs Sham				
Pathway ID	**Pathway name**	**Pathway size**	**Matched KO's**	**p.adj**
ko01100	Metabolic pathways	4702	87	0.000004
ko02060	Phosphotransferase system (PTS)	72	9	0.000004
ko02010	ABC transporters	515	21	0.000005
ko00130	Ubiquinone and other terpenoid-quinone biosynthesis	59	8	0.000005
ko01240	Biosynthesis of cofactors	375	15	0.000286

Table 7. Significantly enriched pathways obtained by using significantly changing KO's from each pairwise comparison of three rat groups as query (continued)

ko00650	Butanoate metabolism	114	7	0.003775
ko00620	Pyruvate metabolism	133	7	0.007206
ko00900	Terpenoid backbone biosynthesis	63	5	0.007206
ko00552	Teichoic acid biosynthesis	41	4	0.010277
ko00520	Amino sugar and nucleotide sugar metabolism	156	7	0.012273
ko00051	Fructose and mannose metabolism	112	6	0.012273
ko00052	Galactose metabolism	78	5	0.012273
ko00550	Peptidoglycan biosynthesis	53	4	0.018606
Epi vs Sham				
Pathway ID	**Pathway name**	**Pathway size**	**Matched_ KO's**	**p.adj**
ko00552	Teichoic acid biosynthesis	41	6	0.000057
ko00900	Terpenoid backbone biosynthesis	63	5	0.005225
ko02060	Phosphotransferase system (PTS)	72	5	0.006562
ko00300	Lysine biosynthesis	48	4	0.010709
ko00550	Peptidoglycan biosynthesis	53	4	0.012518
ko01100	Metabolic pathways	4702	52	0.035463
ko00052	Galactose metabolism	78	4	0.037602

Table 7. Significantly enriched pathways obtained by using significantly changing KO's from each pairwise comparison of three rat groups as query (continued)

Epi vs No-Epi				
Pathway ID	**Pathway name**	**Pathway size**	**Matched KO's**	**p.adj**
ko00660	C5-Branched dibasic acid metabolism	29	4	0.000003
ko01100	Metabolic pathways	4702	19	0.000064
ko01200	Carbon metabolism	365	6	0.000143
ko01120	Microbial metabolism in diverse environments	1281	9	0.000512
ko00630	Glyoxylate and dicarboxylate metabolism	104	3	0.003335

4.2 Analysis Results of Shotgun Metagenomics Sequencing Data

4.2.1 Differential abundance analysis

Among all the pairwise comparisons (Figure 11), major shift in the composition of serum metabolites, based on the number of metabolites that are significantly different between the two groups compared, were obtained from Epi group samples compared to the control group; 175 metabolites were detected to be significantly different. Among those, 109 were higher, and 66 were lower in Epi group in comparison to control. Epi vs No-Epi comparison ranked in second position with a total of 86 metabolites, 31 of which are enriched in Epi group, while 55 of them are detected in lower concentrations. Contrast to the metagenomic results, No-Epi vs Sham comparison yielded the lowest number of metabolites; a total of 51 metabolites with 17 metabolites are higher and 34 of them are lower in No-Epi group (Table 8).

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis

No-Epi vs Sham					
Metabolite ID	**KEGG ID**	**Sham mean**	**No-Epi mean**	**No-Epi/Sham Log2(Fold Change)**	**p.adj**
c18neg.m30	C00026	4.25035	3.74975	-0.50061	0.02018
hilicpos.m13	C00064	9.59208	9.33756	-0.25452	0.03919
hilicpos.m38	C00065	7.18518	6.91247	-0.27271	0.03316
hilicpos.m30	C00079	8.51874	8.20115	-0.31759	0.00526
hilicpos.m42	C00148	1.11557	0.55516	-0.56042	0.01912
c18neg.m105	C00245	3.40169	3.82681	0.42511	0.02412
hilicneg.m36	C00275	-4.86540	-3.33752	1.52788	0.02253
hilicpos.m28	C00327	7.29016	7.01543	-0.27473	0.03403
hilicneg.m85	C00407	5.37346	4.63172	-0.74174	0.00053
hilicneg.m117	C00755	-2.10958	-2.65140	-0.54182	0.04507
c18neg.m36	C00785	2.31578	1.27759	-1.03819	0.02983
hilicneg.m1	C01087	-1.74414	-0.44184	1.30230	0.03471
hilicpos.m44	C01879	8.14362	7.83573	-0.30789	0.01866

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

hilicpos.m21	C02218	3.38453	3.12238	-0.26215	0.04078
hilicpos.m4	C05127	2.66589	2.25823	-0.40765	0.02649
c18neg.m169	C06954	-1.68703	-3.28430	-1.59728	0.01671
hilicneg.m124	C07203	0.72208	0.39384	-0.32823	0.01936
c18neg.m195	C10596	-2.63802	-3.90812	-1.27010	0.02282
c18pos.m381	C10861	-6.90845	-8.07075	-1.16231	0.02362
c18pos.m201	C10935	-0.13103	-0.74978	-0.61874	0.00592
c18neg.m6	C11301	2.01182	2.97165	0.95983	0.01592
hilicneg.m54	C11301	-0.48367	0.69869	1.18236	0.03687
c18pos.m2	C14691	0.77037	-0.75915	-1.52952	0.00022
hilicpos.m40	C14691	0.65674	0.22332	-0.43342	0.01888
hilicneg.m232		-2.44910	-1.74787	0.70123	0.01354
c18neg.m241		-0.58257	-1.07271	-0.49014	0.02300
c18pos.m305		-5.58572	-4.08749	1.49823	0.00002
hilicneg.m144		3.63670	2.87792	-0.75878	0.00929
c18neg.m133		1.22738	0.82790	-0.39947	0.01044

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

c18pos.m5		-4.34865	-1.68248	2.66618	0.04966
hilicneg.m249		-1.63746	-1.10933	0.52814	0.01233
hilicneg.m219		-9.07929	-7.58987	1.48942	0.00050
hilicneg.m234		0.01858	0.55432	0.53574	0.00182
c18pos.m39		-2.10888	-0.90733	1.20154	0.00123
c18neg.m271		-3.60104	-2.76223	0.83880	0.01092
hilicpos.m67		7.13121	6.81650	-0.31471	0.02128
hilicneg.m248		-1.64659	-1.89574	-0.24916	0.03486
hilicneg.m291		-4.54287	-5.88170	-1.33883	0.00025
c18neg.m170		-2.27136	-2.63012	-0.35876	0.04052
c18pos.m99		1.72108	1.28074	-0.44034	0.04292
c18neg.m41		-1.00947	-1.51911	-0.50964	0.03138
c18pos.m170		1.55883	1.26491	-0.29392	0.04514
c18neg.m209		-1.92286	-1.63551	0.28736	0.03154
hilicneg.m293		0.28517	0.04879	-0.23639	0.03275

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

hilicneg.m294		-0.49749	-1.08485	-0.58736	0.00182
c18neg.m216		-3.40406	-5.18535	-1.78129	0.01969
c18pos.m16		-1.01191	0.29158	1.30349	0.00168
hilicneg.m202		-4.52564	-3.11485	1.41079	0.00471
hilicneg.m167		-8.97250	-7.79011	1.18239	0.02433
hilicneg.m289		-1.45972	-2.13900	-0.67928	0.03992
hilicpos.m52		6.21839	5.90535	-0.31305	0.01773
Epi vs Sham					
Metabolite ID	**KEGG ID**	**Sham mean**	**No-Epi mean**	**No-Epi/Sham Log2(Fold Change)**	**p.adj**
c18pos.m233	C00037	-2.10449	-1.67069	0.43380	0.00281
c18pos.m17	C00065	1.30607	1.78914	0.48307	0.04202
hilicpos.m22	C00078	8.60988	8.08668	-0.52320	0.00226
hilicneg.m8	C00078	7.54027	7.08376	-0.45652	0.00479
hilicpos.m30	C00079	8.51874	8.18929	-0.32946	0.00181
hilicpos.m20	C00099	8.07336	7.75883	-0.31453	0.01581

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

hilicpom18	C00152	6.71488	6.45030	-0.26458	0.03602
hilicneg.m75	C00152	6.85148	6.60906	-0.24242	0.02060
c18neg.m124	C00157	-2.28933	-1.31465	0.97467	0.00235
hilicneg.m3	C00158	5.43193	6.06619	0.63426	0.01440
hilicneg.m81	C00186	6.61493	6.32240	-0.29253	0.03295
c18neg.m128	C00186	2.03346	1.50969	-0.52377	0.02086
hilicneg.m99	C00188	-1.99488	-2.70163	-0.70675	0.02661
c18neg.m86	C00219	-0.49491	0.07054	0.56544	0.01633
hilicneg.m128	C00221	-0.84745	-0.24163	0.60582	0.04888
hilicneg.m190	C00249	-3.02731	-3.54695	-0.51964	0.01721
hilicneg.m36	C00275	-4.86540	-3.51018	1.35523	0.03219
hilicneg.m84	C00300	5.97643	5.53055	-0.44588	0.01511
hilicpos.m61	C00318	3.12989	1.10879	-2.02110	0.01415
hilicpos.m28	C00327	7.29016	6.99602	-0.29414	0.01303
hilicneg.m44	C00334	2.70985	2.36022	-0.34964	0.01191

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

hilicneg.m85	C00407	5.37346	4.58602	-0.78744	0.00009
hilicpos.m25	C00437	4.53569	4.12828	-0.40741	0.00884
hilicneg.m107	C00489	1.06832	0.49589	-0.57243	0.04164
c18neg.m211	C00627	1.53649	1.97411	0.43762	0.03649
c18pos.m78	C00791	4.56448	4.80157	0.23710	0.01567
hilicneg.m14	C00811	-1.56332	-2.25251	-0.68920	0.00324
hilicneg.m67	C00870	-2.68329	-3.12083	-0.43754	0.03999
hilicneg.m119	C01042	-5.95560	-3.50926	2.44635	0.00116
c18neg.m100	C01595	-1.52691	-0.69324	0.83367	0.03065
c18pos.m354	C02530	-4.44299	-5.51688	-1.07390	0.00080
hilicpos.m24	C02997	3.39042	2.89683	-0.49359	0.02058
c18pos.m133	C04230	-3.71732	-4.62896	-0.91163	0.00124
c18pos.m136	C04230	2.09736	1.48271	-0.61465	0.00008
c18pos.m135	C04230	0.41979	-0.22160	-0.64140	0.00006

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

hilicneg.m274	C04230	1.84615	1.00896	-0.83719	0.00000
c18neg.m341	C04230	-3.50824	-4.15858	-0.65034	0.00136
c18pos.m138	C04230	0.37552	0.10176	-0.27376	0.01184
c18pos.m105	C04230	6.51432	6.28513	-0.22919	0.00121
c18pos.m131	C04230	1.04538	0.20879	-0.83659	0.00641
c18neg.m318	C04230	-0.44068	-0.09917	0.34151	0.03482
c18pos.m110	C04230	5.79934	5.12265	-0.67669	0.00079
c18neg.m121	C04230	5.22137	4.86563	-0.35574	0.01691
hilicneg.m58	C04230	4.10254	3.41863	-0.68391	0.03856
c18pos.m113	C04230	5.72397	5.47278	-0.25119	0.03288
c18neg.m22	C04230	-2.09347	-2.59427	-0.50080	0.01789
c18pos.m174	C04230	-2.56288	-3.05845	-0.49556	0.00596
c18pos.m178	C04230	-0.30324	-0.92105	-0.61781	0.01597
c18neg.m24	C04230	1.36521	0.68806	-0.67715	0.00265

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

c18pos.m165	C04230	7.48081	7.23947	-0.24134	0.00408
c18pos.m164	C04230	5.30282	4.94035	-0.36247	0.00099
c18pos.m385	C04230	-1.43273	-1.71744	-0.28471	0.01593
hilicneg.m284	C04230	3.84033	3.40556	-0.43478	0.00040
hilicneg.m60	C04230	5.80327	5.10164	-0.70163	0.00051
c18pos.m167	C04230	0.93339	0.29423	-0.63915	0.03760
c18pos.m150	C04230	-3.34182	-3.69826	-0.35644	0.00444
c18neg.m31	C04230	0.44997	-0.10344	-0.55341	0.00016
c18pos.m151	C04230	0.02389	-0.62729	-0.65118	0.00145
c18neg.m32	C04230	1.61310	0.94934	-0.66376	0.00006
c18pos.m149	C04230	-2.30063	-2.84676	-0.54614	0.00857
c18pos.m172	C04230	4.40908	3.84992	-0.55916	0.00326
c18neg.m327	C04230	-2.04372	-2.51633	-0.47261	0.01513
c18neg.m33	C04230	2.00653	1.62922	-0.37731	0.03312

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

c18pos.m59	C05497	0.71438	1.28743	0.57305	0.02463
c18pos.m11	C05548	-0.50441	-0.12330	0.38111	0.00599
c18neg.m78	C06001	4.78945	5.52931	0.73986	0.03054
hilicneg.m45	C06001	6.29268	6.92044	0.62776	0.00199
c18pos.m195	C06413	4.50388	4.69282	0.18894	0.00460
c18pos.m193	C06423	5.63867	5.82579	0.18712	0.00430
hilicneg.m124	C07203	0.72208	0.31978	-0.40229	0.00161
hilicneg.m129	C10448	1.87714	0.95774	-0.91940	0.01738
c18pos.m381	C10861	-6.90845	-8.41604	-1.50760	0.00107
c18pos.m201	C10935	-0.13103	-0.74270	-0.61167	0.00318
c18pos.m7	C12288	-0.25888	0.17334	0.43222	0.02162
c18pos.m2	C14691	0.77037	-0.68588	-1.45624	0.00015
c18neg.m53	C16202	-0.13897	0.38234	0.52132	0.03208
hilicneg.m26	C16679	0.22077	-1.06904	-1.28981	0.00951
c18pos.m301	C17431	1.69102	1.21397	-0.47705	0.00469

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

hilicneg.m264	C18043	2.13604	1.20741	-0.92863	0.00146
c18pos.m139	C21484	3.13999	2.86504	-0.27495	0.00657
hilicneg.m186	D00386	-2.38902	-2.93401	-0.54499	0.01787
hilicneg.m232		-2.44910	-1.84771	0.60139	0.02562
c18+m226		-0.17422	-0.89567	-0.72145	0.00010
c18neg.m259		1.55222	1.91761	0.36539	0.04318
c18pos.m251		-1.89813	-1.13044	0.76769	0.02821
c18pos.m206		2.43768	2.65509	0.21741	0.00940
hilicneg.m177		-0.74226	-0.36123	0.38103	0.03664
c18neg.m156		-2.41184	-1.93812	0.47372	0.00076
c18neg.m300		-1.67345	-1.36583	0.30762	0.03884
c18neg.m245		1.39944	1.95065	0.55121	0.00123
c18pos.m101		-0.10928	-0.70414	-0.59487	0.00767
hilicneg.m159		-0.47943	-0.11841	0.36102	0.03961
c18neg.m333		1.31863	1.64093	0.32229	0.03086

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

c18neg.m187		-3.85044	-6.27525	-2.42481	0.00002
c18posm31		1.20039	0.93227	-0.26812	0.00403
c18pos.m305		-5.58572	-4.56447	1.02125	0.00149
c18neg.m253		-2.90437	-2.62102	0.28335	0.00865
c18pos.m307		-1.11028	-1.79057	-0.68029	0.00051
c18neg.m117		1.77673	1.19061	-0.58612	0.00011
c18neg.m254		5.87340	6.42028	0.54688	0.00057
c18pos.m355		0.09657	1.14081	1.04424	0.00626
c18neg.m332		-0.54082	-0.28540	0.25543	0.04505
c18pos.m293		-1.21781	-1.93560	-0.71779	0.00016
hilicneg.m288		0.13488	-0.47446	-0.60934	0.00103
c18pos.m306		-1.68401	-2.42192	-0.73791	0.00116
c18pos.m187		2.25312	2.63773	0.38460	0.03458
c18pos.m132		3.28251	2.99028	-0.29223	0.03933
c18neg.m12		0.41234	-0.92160	-1.33394	0.00063

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

c18pos.m5		-4.34865	-0.32843	4.02023	0.00073
c18neg.m126		-3.00938	-2.56447	0.44491	0.01618
c18neg.m167		-2.85498	-1.98801	0.86697	0.03585
hilicneg.m249		-1.63746	-1.19605	0.44141	0.02816
c18neg.m227		-4.28525	-3.36207	0.92318	0.00019
hilicneg.m219		-9.07929	-7.98660	1.09269	0.00715
c18pos.m210		-4.30792	-3.84358	0.46433	0.03430
c18neg.m246		-2.68834	-2.23236	0.45598	0.02773
hilicneg.m234		0.01858	0.47523	0.45666	0.00479
hilicneg.m137		5.89929	5.53837	-0.36092	0.02802
c18pos.m120		-1.68568	-3.20146	-1.51578	0.00028
c18neg.m181		-0.34541	-1.56045	-1.21504	0.00571
c18pos.m283		0.49954	-0.13338	-0.63292	0.00618
c18pos.m299		-2.42587	-3.25791	-0.83203	0.00123
c18pos.m213		0.81384	1.49545	0.68161	0.00930

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

c18neg.m240		-3.35476	-4.70253	-1.34777	0.00364
hilicneg.m230		0.06482	-1.59810	-1.66292	0.00112
c18pos.m204		-0.28080	-0.15181	0.12899	0.02595
c18pos.m304		-0.10675	-0.80844	-0.70169	0.00012
c18neg.m271		-3.60104	-2.71865	0.88239	0.00411
c18neg.m13		0.03599	-1.13483	-1.17082	0.00144
c18neg.m247		-2.84368	-2.19371	0.64997	0.00137
hilicneg.m248		-1.64659	-1.87834	-0.23176	0.03610
hilicneg.m291		-4.54287	-5.43484	-0.89196	0.01061
c18pos.m350		-1.13271	-2.47161	-1.33889	0.00793
hilicpos.m26		7.39407	6.87593	-0.51814	0.00254
c18neg.m239		-2.56814	-2.08583	0.48230	0.02257
c18pos.m43		-1.99685	-2.84676	-0.84990	0.01262
hilicpom71		-1.59815	-0.45781	1.14034	0.02086
c18pos.m99		1.72108	1.04998	-0.67110	0.00046

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

c18negm41		-1.00947	-2.03171	-1.02224	0.00000
hilicneg.m261		3.00247	2.64704	-0.35543	0.00849
c18neg.m286		-4.28846	-5.29509	-1.00662	0.00004
hilicneg.m276		2.53952	1.78351	-0.75601	0.00091
c18neg.m297		-1.79493	-2.61527	-0.82033	0.00955
hilicneg.m281		4.64369	4.05819	-0.58551	0.00000
c18pos.m170		1.55883	1.05865	-0.50018	0.00011
c18pos.m121		-0.05838	0.13461	0.19299	0.00499
c18neg.m267		0.35036	0.61118	0.26082	0.02497
c18neg.m215		-0.26820	-1.11437	-0.84617	0.00551
c18neg.m255		1.70894	1.97714	0.26820	0.01768
c18pos.m280		-1.78891	-1.61460	0.17430	0.04288
hilicneg.m123		-0.69553	-2.50100	-1.80547	0.00119
c18pos.m323		-1.26857	-1.86380	-0.59524	0.00936
hilicneg.m203		-1.83757	0.29514	2.13271	0.01747

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

c18neg.m207		-2.54931	-0.65463	1.89468	0.01377
c18pos.m370		-0.90051	-0.22081	0.67970	0.04179
c18neg.m209		-1.92286	-1.49951	0.42336	0.00045
hilicneg.m293		0.28517	0.01738	-0.26779	0.00771
hilicneg.m287		2.24082	1.02249	-1.21834	0.00019
c18posm58		-3.35153	-3.11793	0.23360	0.04922
hilicneg.m194		-4.78830	-7.22147	-2.43318	0.00131
c18neg.m206		0.07700	1.55384	1.47684	0.03184
c18pos.m297		0.69520	-0.02605	-0.72125	0.00570
hilicneg.m202		-4.52564	-3.36915	1.15649	0.01479
c18pos.m382		-0.81684	-1.40663	-0.58979	0.00014
c18pos.m262		-4.39598	-3.18209	1.21389	0.01188
c18pos.m308		-1.96897	-2.92757	-0.95861	0.00406
c18neg.m293		0.11991	-1.66422	-1.78413	0.04184
c18neg.m334		-2.38621	-1.36756	1.01865	0.03486

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

hilicneg.m289		-1.45972	-2.64181	-1.18208	0.00006
c18neg.m329		0.25605	-0.37735	-0.63340	0.01656
hilicneg.m206		-1.06011	-0.62168	0.43844	0.01002
c18neg.m289		-3.29710	-4.74464	-1.44754	0.00733
c18neg.m336		-3.19670	-4.05614	-0.85944	0.02284
c18pos.m189		1.91707	1.62354	-0.29353	0.03146
c18pos.m209		-1.51664	-1.24963	0.26701	0.00954
Epi vs No-Epi					
Metabolite ID	**KEGG ID**	**Sham mean**	**No-Epi mean**	**No-Epi/Sham Log2(Fold Change)**	**p.adj**
c18neg.m76	C00065	1.69091	1.90839	0.21748	0.04769
c18neg.m211	C00627	1.53770	1.97411	0.43641	0.03712
c18pos.m191	C00785	-0.40469	0.41271	0.81740	0.04748
c18pos.m78	C00791	4.56556	4.80157	0.23601	0.01845
hilicneg.m14	C00811	-1.70259	-2.25251	-0.54992	0.02238
hilicneg.m67	C00870	-2.66724	-3.12083	-0.45359	0.03195

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

hilicneg.m119	C01042	-5.15897	-3.50926	1.64972	0.03675
c18neg.m100	C01595	-1.49520	-0.69324	0.80196	0.03910
hilicneg.m38	C01595	4.18057	4.74884	0.56827	0.04166
c18pos.m324	C03242	-8.41901	-9.02611	-0.60710	0.04332
hilicneg.m150	C03519	-3.74682	-4.93440	-1.18758	0.02695
c18neg.m117	C04230	1.56517	1.19061	-0.37456	0.01595
c18pos.m133	C04230	-4.03140	-4.62896	-0.59756	0.04891
c18pos.m136	C04230	1.88121	1.48271	-0.39850	0.01379
c18pos.m135	C04230	0.18880	-0.22160	-0.41040	0.01314
hilicneg.m274	C04230	1.59039	1.00896	-0.58142	0.00021
c18neg.m341	C04230	-3.59146	-4.15858	-0.56712	0.00566
c18pos.m110	C04230	5.80476	5.12265	-0.68211	0.00089
c18neg.m121	C04230	5.16760	4.86563	-0.30197	0.04942
c18neg.m24	C04230	1.17099	0.68806	-0.48293	0.04085

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

c18pos.m177	C04230	1.00834	0.32660	-0.68175	0.04328
hilicneg.m284	C04230	3.68779	3.40556	-0.28223	0.02728
c18negm31	C04230	0.24153	-0.10344	-0.34496	0.02298
c18pos.m151	C04230	-0.17883	-0.62729	-0.44846	0.03975
c18neg.m32	C04230	1.30197	0.94934	-0.35263	0.04204
c18neg.m82	C04661	3.93994	1.50474	-2.43521	0.01131
c18neg.m280	C04661	0.74564	-1.33254	-2.07819	0.02030
c18pos.m195	C06413	4.51437	4.69282	0.17844	0.00902
c18pos.m193	C06423	5.65696	5.82579	0.16883	0.01242
hilicneg.m37	C06427	1.10928	1.72020	0.61092	0.01765
c18neg.m169	C06954	-3.28430	-1.91872	1.36558	0.03322
c18pos.m76	C07091	-0.79613	-0.11152	0.68461	0.00487
c18pos.m35	C07432	-3.20823	-2.95493	0.25330	0.03922
c18neg.m322	C07544	-3.96684	-5.98810	-2.02127	0.00054
hilicneg.m33	C08362	-0.35484	0.30341	0.65825	0.04987

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

c18pos.m366	C08445	-3.44387	-3.81042	-0.36655	0.02675
c18neg.m195	C10596	-3.90812	-2.60820	1.29992	0.01245
c18neg.m6	C11301	2.97165	1.46917	-1.50248	0.00004
hilicneg.m54	C11301	0.69869	-1.30026	-1.99895	0.00008
hilicneg.m175	C13297	-0.17656	0.43369	0.61024	0.03409
c18pos.m329	C13828	2.28950	-0.13282	-2.42231	0.02567
c18pos.m327	C16921	-4.00241	-5.86160	-1.85918	0.03702
c18neg.m233		-0.93978	-0.68644	0.25334	0.01440
c18pos.m211		0.46555	0.58917	0.12362	0.03802
c18neg.m245		1.57594	1.95065	0.37471	0.03603
c18neg.m241		-1.07271	-0.61692	0.45579	0.02553
c18neg.m187		-4.79827	-6.27525	-1.47698	0.01141
c18neg.m253		-2.86396	-2.62102	0.24294	0.02789
c18neg.m254		6.05420	6.42028	0.36608	0.02656
c18neg.m95		-1.53718	-1.12877	0.40840	0.03145

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

hilicneg.m92		4.07436	4.87899	0.80462	0.03712
c18neg.m332		-0.60074	-0.28540	0.31535	0.01026
c18pos.m293		-1.43248	-1.93560	-0.50312	0.01047
c18negm12		0.02228	-0.92160	-0.94388	0.01908
c18pos.m64		0.14311	-2.83546	-2.97857	0.01081
c18pos.m123		4.61503	2.40296	-2.21207	0.02284
c18pos.m62		2.85610	0.58513	-2.27096	0.03878
c18pos.m162		-0.93381	-4.74966	-3.81584	0.02957
c18neg.m85		-0.72555	-3.96691	-3.24137	0.01243
c18neg.m167		-2.84498	-1.98801	0.85697	0.03851
hilicneg.m238		-3.19537	-4.18649	-0.99112	0.01007
c18neg.m246		-2.68856	-2.23236	0.45620	0.02765
c18pos.m120		-2.04299	-3.20146	-1.15847	0.00695
c18neg.m328		-0.81916	-0.31680	0.50235	0.04903
c18neg.m305		-1.21523	-3.70767	-2.49244	0.00359

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

c18neg.m137		1.15660	1.72213	0.56554	0.02681
c18pos.m304		-0.38646	-0.80844	-0.42198	0.02856
c18pos.m365		-4.28227	-6.17192	-1.88965	0.00138
c18neg.m247		-2.91060	-2.19371	0.71689	0.00040
c18pos.m70		-0.13573	-3.37697	-3.24124	0.00953
c18pos.m43		-2.06753	-2.84676	-0.77923	0.02700
c18negm41		-1.51911	-2.03171	-0.51260	0.02038
hilicneg.m261		2.99117	2.64704	-0.34413	0.01115
c18neg.m286		-4.51880	-5.29509	-0.77628	0.00153
hilicneg.m276		2.27844	1.78351	-0.49494	0.03910
hilicneg.m281		4.37066	4.05819	-0.31248	0.01479
c18pos.m364		-2.46352	-4.84424	-2.38072	0.01583
c18pos.m277		-3.33793	-5.85395	-2.51601	0.00403
c18pos.m321		-3.13585	-5.61968	-2.48383	0.02786
c18neg.m256		-5.19970	-6.57171	-1.37202	0.00765

Table 8. Significantly different metabolites for each pairwise comparison of three rat groups obtained from serum metabolomics DA analysis (continued)

hilicneg.m32		1.78889	2.61119	0.82230	0.02605
c18pos.m297		0.60190	-0.02605	-0.62795	0.02037
c18pos.m382		-1.03985	-1.40663	-0.36679	0.02424
c18neg.m293		0.76186	-1.66422	-2.42608	0.00378
c18neg.m289		-2.99464	-4.74464	-1.75000	0.00102
c18neg.m336		-3.25851	-4.05614	-0.79763	0.03728

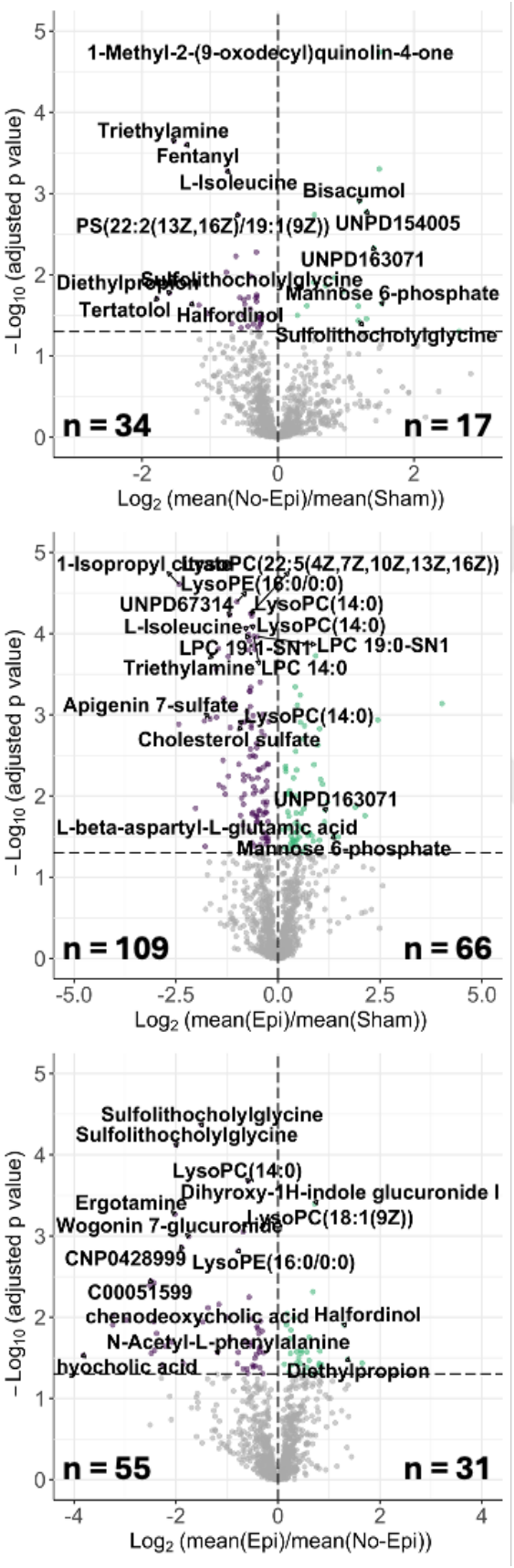

Figure 11. Differential abundance analysis of serum metabolomics data.

Figure 11. Differential abundance analysis of serum metabolomics data (continued) Volcano plots summarizing differentially abundant blood metabolites between No-Epi and Sham (left), Epi and Sham (middle), and Epi and No-Epi groups (right). Plots were generated using mean abundance values between two groups and adjusted p values. Metabolites are represented by the data points, and dashed lines represent the p.adj and fold-change thresholds. Purple and green points belong to decreased and increased metabolites, respectively, with adjusted p value less than 0.05. Numbers of the significantly changing metabolites were written at the bottom of each plot by bold text for each direction.

4.2.2 Pathway enrichment

After conversion of annotated metabolites to respective KEGG ID's, 47, 22, and 26 unique KEGG ID's were obtained from Epi vs Sham, No-Epi vs Sham, and Epi vs No-Epi comparisons, respectively. After pathway enrichment analysis using those converted IDs, total of ten unique KEGG pathways were obtained from all the comparisons (Table 9 and Figure 12). Predominantly, several amino acid metabolism pathways were mapped by the metabolites changing in both epileptic rat groups compared to control. Interestingly, the arginine and proline metabolism, as well as the alanine, aspartate and glutamate metabolic pathways were generally downregulated both in the No-Epi and Epi groups, compared to the Sham controls, indicating a TBI-specific metabolic profile. "Aminoacyl-tRNA biosynthesis" and "ABC transporters" were two other pathways shared by Epi vs Sham and No-Epi vs Sham comparisons. Remarkably, Epi group rats were observed to be differentiated from the others by significant changes in several lysphosphatidylcholines (LPCs), and pathways related to the lipid metabolism. Furthermore, linoleic acid metabolism was significantly upregulated in the Epi vs. Sham, but remained similar in the No-Epi vs. Sham group. Comparison of Epi vs No-Epi yielded only one pathway, "Biosynthesis of unsaturated fatty acids", which was upregulated in the Epi vs. No-Epi comparison, suggesting that lipid metabolism reprogramming might be specific to epileptogenesis.

Table 9. Significantly enriched pathways obtained by using significantly changing serum metabolites from each pairwise comparison of three rat groups as query

No-Epi vs Sham				
Pathway ID	**Pathway Name**	**Pathway Size**	**Matched metabolites**	**p.adj**
rno02010	ABC transporters	13	6	0.00003
rno00970	Aminoacyl-tRNA biosynthesis	13	5	0.00014
rno00340	Histidine metabolism	13	3	0.00783
rno00471	D-Glutamine and D-glutamate metabolism	13	2	0.00881
rno04964	Proximal tubule bicarbonate reclamation	13	2	0.01435
rno00250	Alanine, aspartate and glutamate metabolism	13	2	0.02045
rno00330	Arginine and proline metabolism	13	3	0.02045
Epi vs Sham				
Pathway ID	**Pathway Name**	**Pathway Size**	**Matched metabolites**	**p.adj**
rno00970	Aminoacyl-tRNA biosynthesis	33	7	0.00057
rno00260	Glycine, serine and threonine metabolism	33	5	0.00407
rno00591	Linoleic acid metabolism	33	3	0.02212
rno00330	Arginine and proline metabolism	33	5	0.02212
rno02010	ABC transporters	33	5	0.02212
rno01100	Metabolic pathways	33	24	0.02212
rno00250	Alanine, aspartate and glutamate metabolism	33	3	0.02212

Table 9. Significantly enriched pathways obtained by using significantly changing serum metabolites from each pairwise comparison of three rat groups as query (continued)

Epi vs No-Epi				
Pathway ID	**Pathway Name**	**Pathway Size**	**Matched metabolites**	**p.adj**
rno01040	Biosynthesis of unsaturated fatty acids	13	3	0.02601

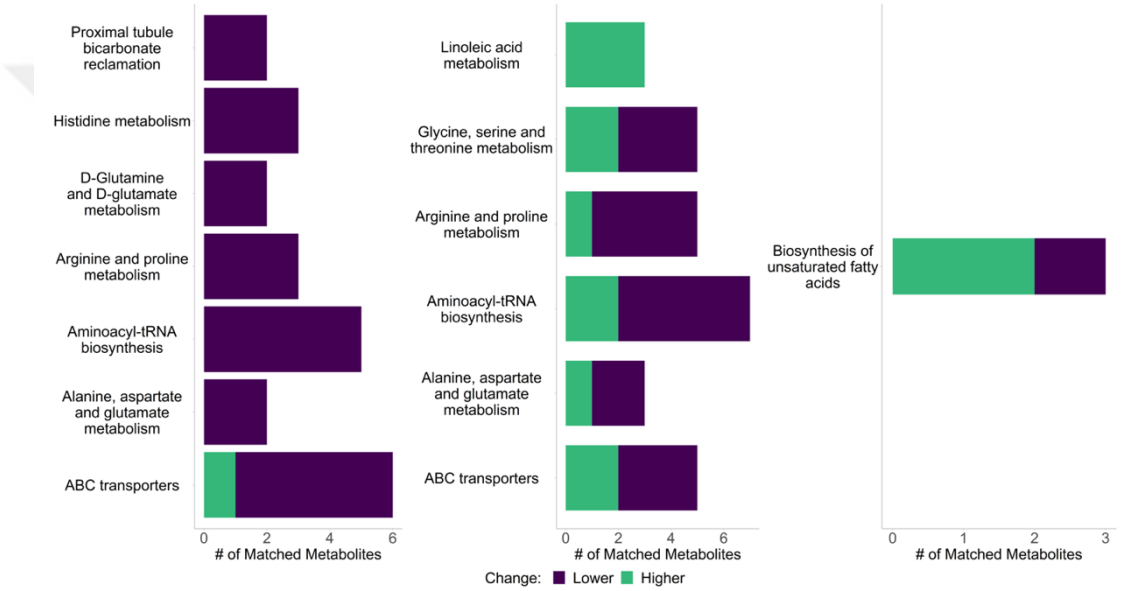

Figure 12. Stacked bar plots of KEGG pathways significantly enriched using differentially abundant serum metabolites obtained from No-Epi vs Sham, Epi vs Sham, and Epi vs No-Epi comparisons.
X-axis represents the number of matched metabolites to the background metabolites associated with respective pathway, purple and green colours represent the decreased or increased metabolites between compared groups, respectively.

4.3 Multi-Omics Integration of Shotgun Metagenomics and Serum Metabolomics Data

4.3.1 Feature selections

4.3.1.1 Feature selection by machine learning

Neither random forest nor random forest gave satisfactory performance, especially on the metagenomics dataset. The median kappa values obtained from classification performance of 100 trained models were 0.19 and 0.40 for elastic net and random forest algorithms, respectively (Figure 13). For metabolomics data, higher performances were obtained. Median kappa value was 0.54 for both elastic net and random forest (Figure 13). Considering unsatisfactory performance from both algorithms, feature selection was not performed.

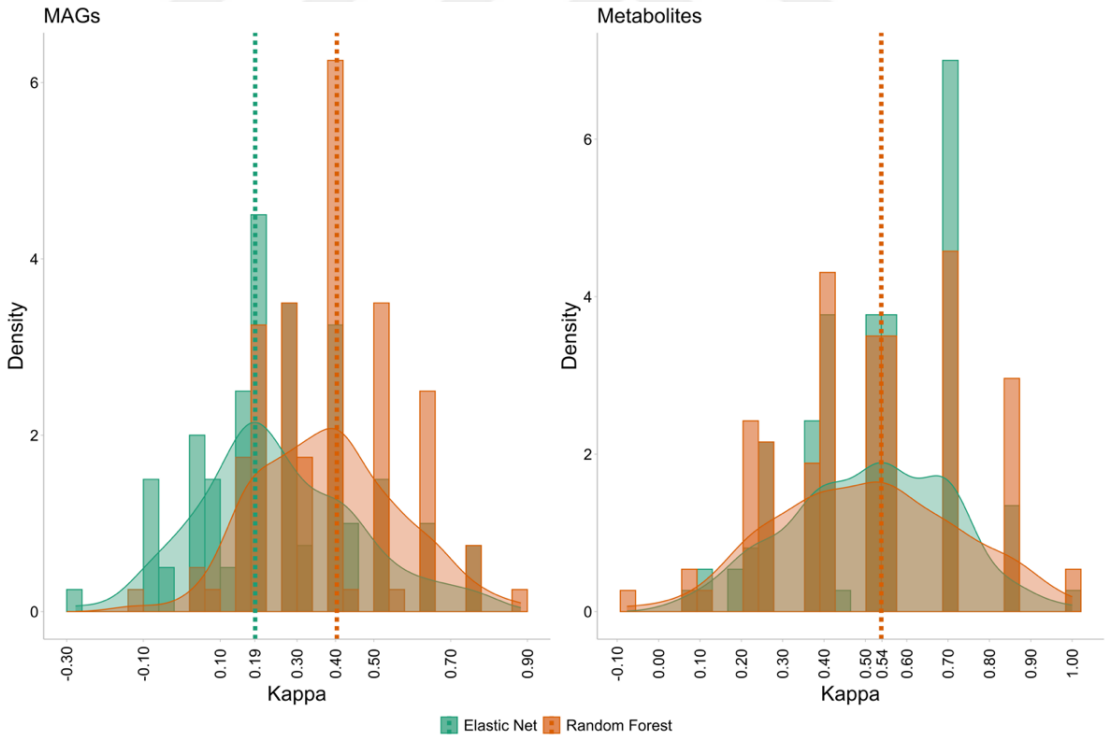

Figure 13. Histograms illustrating the classification performance of the trained machine learning models on test data.
Kappa values obtained from performance of 100 different models were summarized by the histogram plots for metagenomics (left) and metabolomics (right) datasets.

Figure 13. Histograms illustrating the classification performance of the trained machine learning models on test data (continued)
Distribution of the frequencies were additionally shown by density plots. Median kappa values for elastic net and random forest algorithms were shown by vertical dashed lines with respective score on the x-axis.

4.3.1.2 Feature selection by MOFA

A total of 1118 metabolites and 324 MAGs were used for MOFA analysis (Figure 14-A). MOFA model was built using 15 latent factors (LF's), which explains a total biological variation of 49.47% and 55.29% for metabolomic and metagenomic data, respectively (Figure 14-B). For each omics data, LFs explaining at least 2% biological variation were selected. Accordingly, LFs 1, 3, 4, and 7 for metagenomics, and LFs 1, 2, 3, 5, 6, 8, and 9 for metabolomics were selected. The classification capacity of the LFs were visualized by pairing each of the 9 LFs with each other (Figure 14-C). After selecting top 10 weighted features associated with each filtered LF, 37 unique MAGs and 70 unique metabolites were obtained.

4.3.1.3 Feature selection by DA Analysis

With response to low efficiencies obtained from supervised and unsupervised feature selection methods in different stages, comparatively simpler approach was tested. As the ultimate goal is to explore potential role of gut microorganisms in epilepsy development, and differentiate the rat groups between each other to identify specific taxa and functions for different rat groups, the MAGs and metabolites that are differentially abundant when at least two of the three groups were compared. Thereby, features obtained from DA analysis were used for downstream correlation analysis. Since, MAGs were aggregated into taxonomic levels of interest for previous taxonomic DA analysis, an independent DA analysis was carried out using the normalized MAG abundances directly. A total of 56 MAGs and 239 distinct serum metabolites from at least one of the No-Epi vs Sham, Epi vs Sham, and Epi vs No-Epi comparisons were listed to be used in correlation.

4.3.2 Correlations between selected features

4.3.2.1 Correlations between features selected by MOFA

Among all the correlations performed, top 31 significant correlations (absolute (abs) (Spearman's rho) ≥ 0.65 and p.adj < 0.05) were obtained between 11 MAGs and 11 metabolic features (Figure 15). Those MAGs are *CAG-873 sp009775195*, *CAG-873 sp001689665*, *CAG-873 sp910578095*, *Coproplasma sp910587005*, *Paramuribaculum intestinale*, *Paramuribaculum sp910579675*, *Turicimonas muris*, *UBA1394 sp910588145*, *UBA7173 sp001689685*, and two unknown species of *MGBC100320* and *UBA3789*.

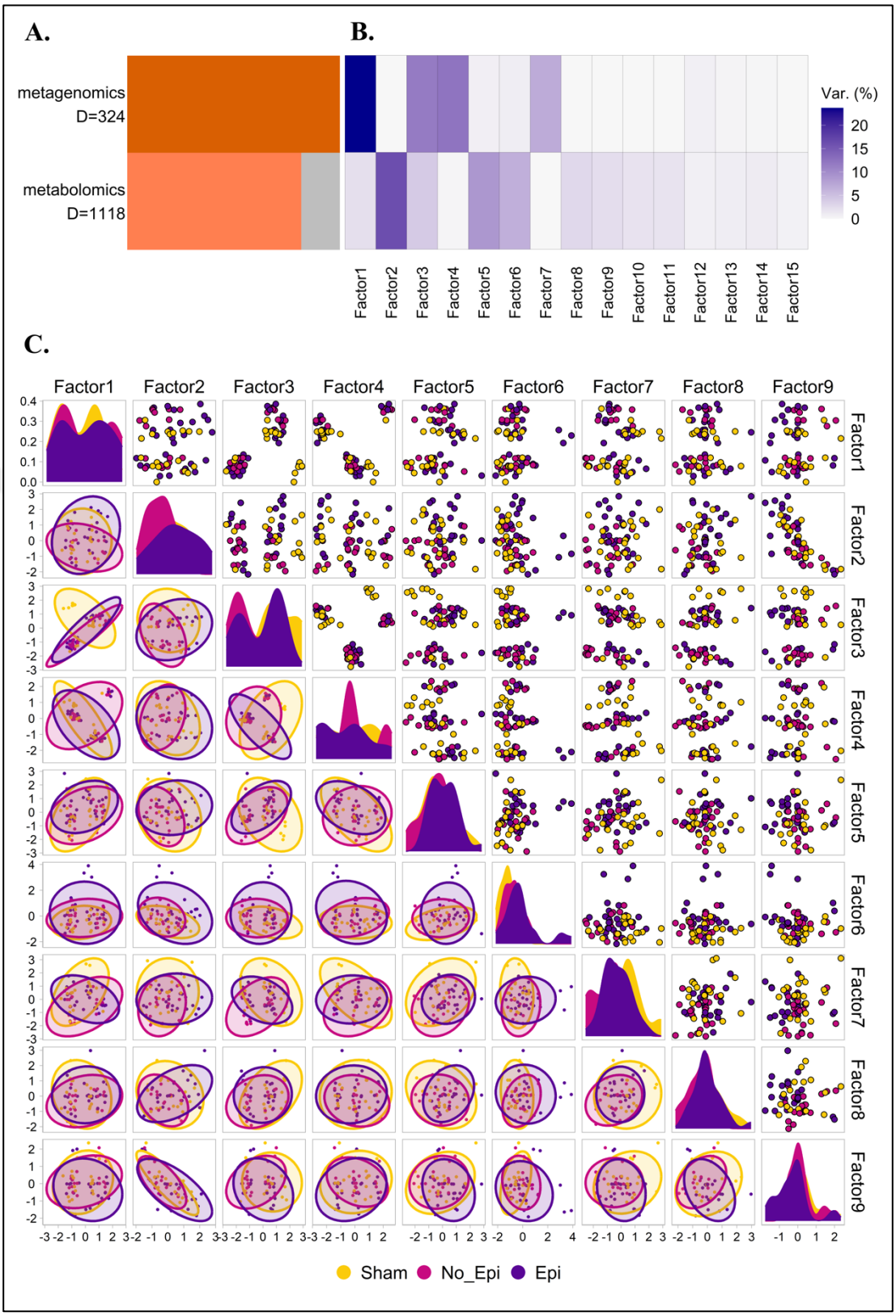

Figure 14. Feature selection using MOFA.
A. Distribution of features (D values) and samples (length of orange bar plot with total length of 66) in each omic layers used to train MOFA model.

Figure 14. Feature selection using MOFA (continued)
B. 15 latent factors are used to train the model, and heatmap illustrates individual contributions to explain variance % in respective dataset. C. Paired factor plot generated by pairwise coupling of first 9 LFs. Points in upper and lower triangle plots represent each sample used, and coloured by respective rat group. Ellipses in the lower triangle plots correspond to 95% confidence intervals for each of the rat groups. The density plots on the diagonal show the densities of factor values of the 3 different clusters.

On the other hand, only 3 of 11 metabolic features were annotated, and they were all derivatives of lysophosphatidylethanolamines (LPEs), namely, LPE O-16:1, LPE O-18:1, and LPE O-18:2. Two unknown species of *MGBC100320* and *UBA3789,* and *Coproplasma sp910587005* were positively correlated with all three forms of LPE's. However, the abundance of taxa, *CAG-873 sp910578095* and *Paramuribaculum sp910579675* were inversely correlated with all three LPE's. *CAG-873 sp009775195* was determined to be negative correlating with LPE O-18:1, and LPE O-18:2. Lastly, *Turicimonas muris* and *UBA7173 sp001689685* are negatively correlated with LPE O-16:1 only.

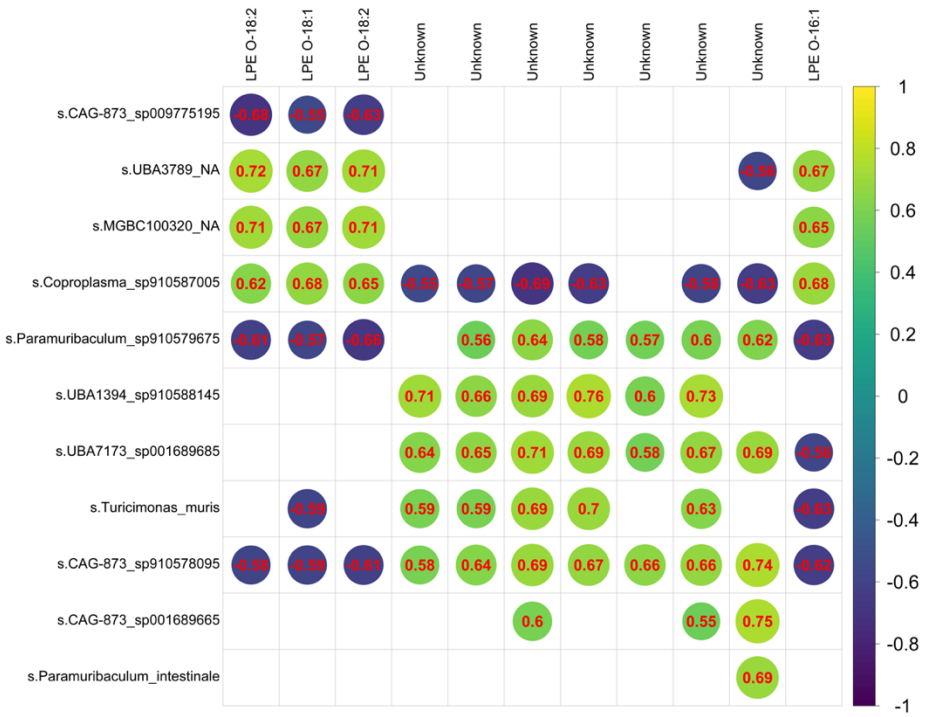

Figure 15. Correlations between metabolic and metagenomic features selected by MOFA.

Figure 15. Correlations between metabolic and metagenomic features selected by MOFA (continued)
Correlation plot summarizing the results for strongest correlations, filtered by abs(Spearman's rho) ≥ 0.65, among all the significant correlations tested. Correlations with p.adj < 0.05 are shown only, and the size of the circles are proportional to the value of correlation coefficient.

4.3.2.2 Correlations between features selected by DA analysis

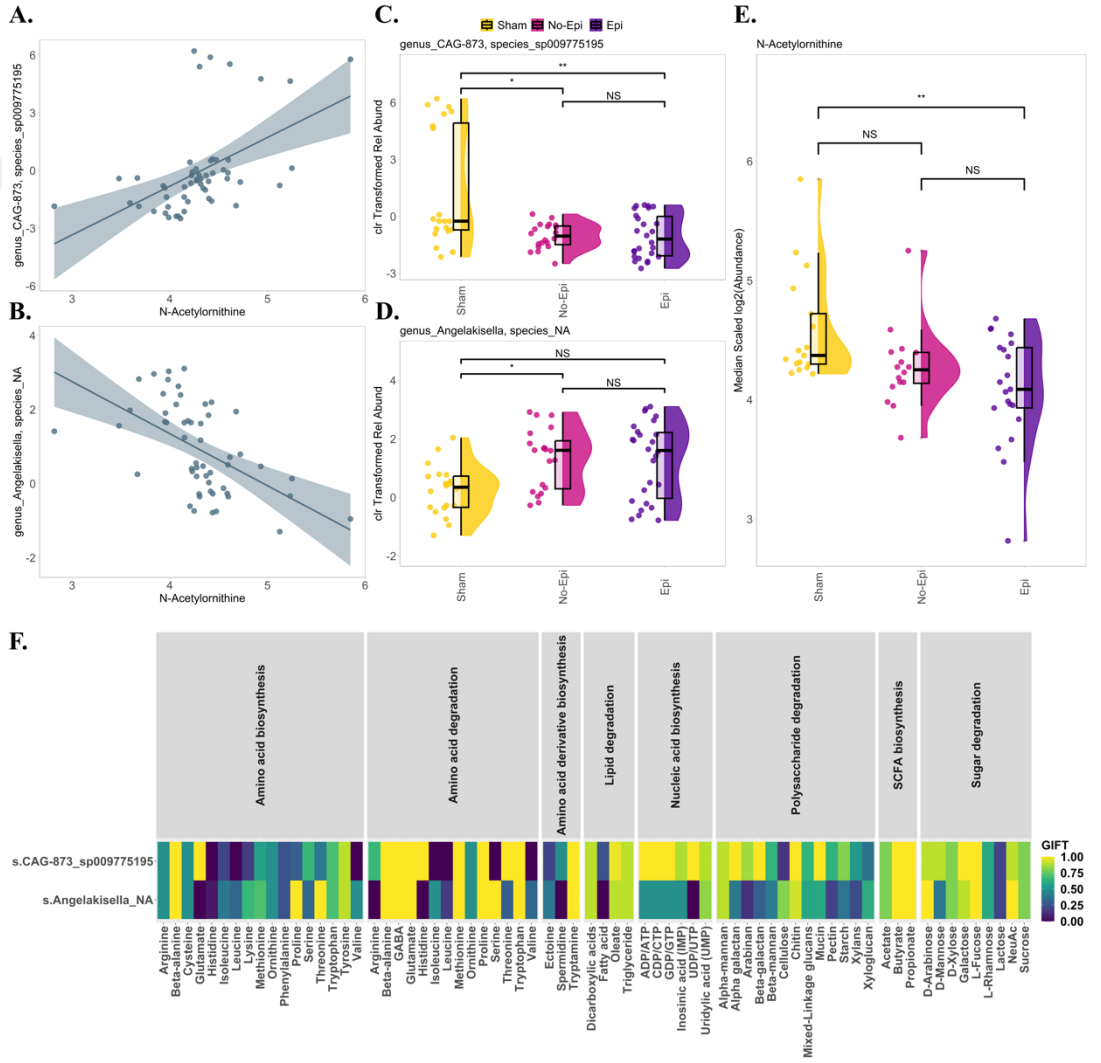

Figure 16. Multi-omics integration of serum metabolomics and metagenomics.
Correlation analysis between MAGs and metabolites which were determined to be significantly changing in at least one of No-Epi vs Sham, Epi vs Sham, and Epi vs No-Epi pairwise comparisons revealed two significant correlations based on Spearman's rho ≥ 0.5 and p.adj < 0.05.
Figure 16. Multi-omics integration of serum metabolomics and metagenomics (continued)

A. Positive correlation between MAG annotated in species level as *CAG-873 sp009775195* and the metabolite annotated as theanine. **B.** Negative correlation between MAG annotated in genus level as *Angelakisella* and the same metabolite, theanine. **C.** Heatmap for *CAG-873 sp009775195* and *Angelakisella* MAGs is illustrating their metabolic capacities (GIFT scores) to synthesize or degrade certain elements. Elements on x-axis were grouped based on different functions. Data of MAG and metabolites in A and B were presented by points, violin and boxplots (with black middle line representing median, and boxes between interquartile ranges of the 25th and 75th percentiles). "**", "*", and "NS" represent p.adj < 0.01, p.adj < 0.05, and p.adj ≥ 0.05 obtained from Kruskal Wallis test followed by Dunn's Test with Bonferroni correction.

After performing a Spearman correlation analysis between abundance of these two omics features, compared to correlation results obtained from MOFA-selected features, relatively lower correlation coefficients were obtained. Therefore, the absolute correlation coefficient threshold was decreased to 0.5. Accordingly, 6 significant correlations (abs(Spearman's rho) ≥ 0.5 and p.adj < 0.05) were obtained between 4 MAGs and 5 metabolites. Taxonomic annotations of those 4 MAGs have been as *Alistipes sp910587675*, *Odoribacter laneus*, *CAG-873 sp009775195* and unknown species of *Angelakisella*. Unfortunately, only one of the 5 metabolites could have been annotated, which is N-acetylornithine. Based on the final selected correlations which were filtered by successful annotations, there are two significant high MAG-metabolite correlations, that are in opposite directions. First of them is a positive correlation (rho = 0.594, p.adj = 0.029) between *CAG-873 sp009775195* and N-acetylornithine (Figure 16-A). An opposite relationship was obtained between MAG annotated to unknown species of *Angelakisella* and N-acetylornithine (rho = -0.597, p.adj = 0.025; Figure 16-B). In order to reveal potential metabolic capacity differences between those two MAGs to be able to explain oppsite correlation with N-acetylornithine, element-wise metabolic capacities (GIFTs) calculated using the presence/absence information of the genes present on the two MAGs against different pathway and modules associated with biosynthesis or catabolism of different biomolecules selected (Figure 16-F). Different GIFT scores were obtained for synthesis of amino acids including glutamate, histidine, isoleucine, leucine, lysine, methionine, proline, serine, threonine, and valine; for degradation of amino acids including arginine, histidine, isoleucine, leucine, serine, threonine, and valine; for synthesis of amino acid derivatives including ectoine and spermidine; for degradation

of lipids including fatty acid and oleate; for biosynthesis of nucleic acids including ADP/ATP, CDP/CTP, GDP/GTP, inosinic acid, UDP/UTP, and uridylic acid; for degradation of polysaccharides including alpha galactan, arabinan, beta-galactan, beta-mannan, cellulose, chitin, mixed-linkage glucans, mucin, pectin, starch, xylans, and xyloglucan; and lastly for degradation of sugars including D-arabinose, D-mannose, galactose, and sucrose. On the other hand, two MAGs were sharing the same GIFT score for synthesis of short-chain fatty acids (SCFAs) (acetate, butyrate, and propionate); for synthesis of amino acids including arginine, beta-alanine, cysteine, ornithine, phenylalanine, tryptophan, and tyrosine; for the degradation of amino acids including beta-alanine, GABA, glutamate, methionine, ornithine, proline, and tryptophan; for biosynthesis of amino acid derivative tryptamine; for degradation of lipids including dicarboxylic acids and triglyceride; for degradation of polysaccharides including alpha-mannan; and for degradation of sugars including D-xylose, L-rhamnose, lactose, and sucrose.

5 DISCUSSION

The findings in this thesis study represent unique gut microbial and serum metabolomics changes in a rat model of kainic acid-induced epilepsy. Using multi-omics analyses, the changes in seizure developments were investigated by comparing two epileptic groups with and without seizure development, and also the changes in KA-induction of the epilepsy in epileptic rats by comparing them with the Sham controls.

Composition of the gut microbiota and the functional metabolic capacities of the gut microbiomes in three rat groups were examined by analysis of deep shotgun metagenomics sequencing data. Shotgun metagenomic sequencing is an informative but very complex method exhibiting several challenges depending on the sample type, variation between subjects analyzed, volume of the data necessary to capture the diversity, host and environmental contaminations, and analysis strategies (107). There is no gold standard established for the bioinformatic analysis of the shotgun metagenomics data, and existing methods and pipelines expose varying advantages and disadvantages. Researchers decide on the analysis methodology considering different aspects such as the sensitivity they are looking for depending on the research hypothesis, usage simplicity of the tools or the computational power and time necessary to perform whole analysis. In this study different analysis tools and methods were employed and taxonomic classifications were compared. There are two main approaches tested, assembly-free read mapping-based algorithms and *de novo* assembly of the reads for the metagenome construction. The assembly-free approaches present some advantages like elimination of assembly-based problems, lower requirements of the computational power and presents higher speed, and classification of rare taxa that may not be captured by assembly-based methods due to low coverage (108). On the other hand, those methods have some main limitations such as necessity of a reference database, limited capacity to identify uncharacterized microorganisms and risk of high false positive rates (108,109). Since this study aims to identify novel characteristic changes in epilepsy, relatively conservative approaches were followed by choosing more restrictive parameters to eliminate false-positives as much as

possible such as setting the E-value to 0.00001 for Kaiju analysis or setting the confidence level threshold to 5e-06 for Kraken2. Accordingly, lower classification rates with decreased resolution in lower taxonomic ranks were obtained compared to default setting classifications. Among them, MetaPhlAn4 showed relatively better classification performance, however, for the functional annotation of the sequenced reads, another tool from the same developers, called HUMAnN (110) is necessary, and when it tested for one sample, the running time was exceeding more than three days (data not shown), therefore this approach was abandoned.

Next, the second approach, generation of MAGs were carried out and a catalogue composed of 324 MAGs is established. No comparable study found in the literature as most of the extensive MAG catalogues studies rely on human (111) and mouse (112) metagenome, while 16S amplicon based sequencing is widely employed in rat studies. There is one study used assembly to produce contigs using reads of 98 faecal samples collected from 49 male Sprague-Dawley rats at different time points, without binning and generating MAGs, and they then performed both taxonomic and functional annotation using the predicted gene sequences (113). Nevertheless, the 324 medium- and high-quality bacterial MAGs and the 874,616 predicted gene sequences will be a good source for future studies to be used as a reference dataset.

Following taxonomic annotation of the MAGs and counting, taxonomic compositions of the three rat groups were compared. As a starting point, within and between sample gut community similarities/dissimilarities were examined through significance testing of alpha and beta diversity indices. Instead of aggregating into a specific taxonomic level, normalized MAG counts were used for both alpha and beta diversity analysis. None of the three alpha diversity indices tested gave significance between the three rat groups, suggesting similar MAG richness and even distribution within samples. Based on community-wise differences through quantitative Bray-Curtis and presence/absence information based Jaccard indices, both Epi and No-Epi MAG compositions are significantly different than Sham group. On the other hand, there is no significant difference obtained from comparison of No-Epi and Epi comparisons. These results suggest that kainic acid injection and induction of epilepsy

caused a shift in bacterial community composition compared to controls, but the seizure development has not made a big difference statistically.

Besides community level change assessments, bacterial changes were investigated in individual taxonomic levels. Phylum level analyses revealed decrease in *Firmicutes* and *Desulfobacterota I* levels in both Epi and No-Epi and in Epi only, respectively, compared to Sham controls. Increased *Firmicutes* levels were found in pediatric patients with refractory epilepsy compared to controls (114). Another study performed on both pediatric and adult epilepsy patients with drug-sensitive and drug-resistant subgroups and healthy controls also reported relatively higher *Firmicutes* abundance in drug resistant group while it was similar among drug sensitive and healthy control samples (53). Similarly, elevated *Firmicutes* abundance was observed in pentylenetetrazol-administered epilepsy rats compared to controls (115). In a small study, where febrile seizure development in mice was induced by hot air conditions, they cannot observe significant change between treatment and control groups (116). It requires further investigations to understand the reason behind the contradictory results obtained in this study.

We detected decreased *Lactobacillus* levels for both No-Epi and Epi group as compared to the Sham rats, and in species level, decreased *Lactobacillus taiwanensis* in both No-Epi and Epi, *Lactobacillus intestinalis* and *Lactobacillus johnsonii* in No-Epi only. *Lactobacillus* species have been attributed to several beneficial effects on host including healthy brain functions through regulation of GABA (117). *Lactobacillus johnsonii* is another beneficial species with some strains used as a probiotic supplementation and regulate SCFA production (118). In a mouse study of pentylenetetrazole (GABA receptor antagonist)-induced epilepsy, treatment with a widely used epilepsy treatment medicine, topiramate, elevated *Lactobacillus johnsonii* levels in the gut (119). Additionally, when mice were ingested with both topiramate and *L. johnsonii* increased butyrate levels and cortex GABA/glutamate levels which resulted in gain of resistance against seizures (119). In another pentylenetetrazol-induced epilepsy model study, ingestion of arbutin showed beneficial effects by decreased the seizures by alleviating the inflammatory mediators (120), and

Lactobacillus intestinalis has been shown to be promoted in gut community through arbutin in another study (121), indicating beneficial impact of arbutin is strongly associated with *Lactobacillus intestinalis*. G-protein-coupled receptors are involved in regulation of several pathways including modulation of acquired epilepsy and seizures (122). In a study, reduced abundance of *Lactobacillus taiwanensis* detected in mice that were treated with Designer Receptors Exclusively Activated by Designer Drugs, which are synthetic ligands activating heterotrimeric G proteins (123). As a member of the commensal gut family of *Lactobacillaceae*, *Limosilactobacillus reuteri* is another potential regulator of gut-brain axis in several nervous system disorders, and we detected lower abundance in both epileptic rat microbiota. Reduced level of *L. reuteri* was detected in mouse model of autism, and they have been shown to regulate the GABA levels not only in RNA levels but also in protein levels (124). In another study, *L. reuteri* strain 17938 were associated with changes in N-acetylated amino acid levels (125) which can improve memory deficiency and delayed seizures in KA-induced epileptic animals (126). Lastly, *Ligilactobacillus murinus*, previously classified as *Lactobacillus murinus*, was detected in reduced levels in both Epi and No-Epi groups. This species was detected in elevated levels in Alzheimer's Disease model rats upon treatment with a herbal formula named Kai-Xin-San which alleviates neuroinflammation and decrease neuronal injuries (127).

Epi group rats showed significant alterations in genus level compared to the control rats. One of those genera, *Turicibacter* were in significantly lower abundance. These bacteria are associated with regulation of variety of fat metabolites in the host (reviewed in (128)). Our investigations in serum metabolomics and downstream pathway enrichment analysis revealed changes in polyunsaturated fatty acid linoleic acid levels, and the only enriched "Biosynthesis of unsaturated fatty acids" pathway specific to Epi group compared to the No-Epi counterparts. The potential involvement of *Turicibacter* species in the regulation of fat metabolism in KA-induced epilepsy development can be further examined through targeted lipidomic assays. Epi group rats had decreased *Romboutsia* genus bacteria compared to Sham controls. Contradictory to our findings, these bacteria have been reported to be in higher levels in epileptic patients (129). In another study on mouse model of Dravet syndrome, a

form of genetic epilepsy observed in infants or in early childhood, a positive correlation between *Rombutsia* levels and phenotypical measurements was reported and proposed to be a potential biomarker in Dravet syndrome (130). Specific to Epi group, we detected lower *Emergencia* levels compared to control group. In species level DA analysis, both Epi and No-Epi group showed decrease in *Emergencia Emergencia sp009935805*, which is not classified in NCBI taxonomy. The only classified member of this genus, *Emergencia timonensis* plays role in production of trimethylamine in human gut microbiota (131). Even there is no established association of trimethylamine with epilepsy, there are three cases reporting co-occurrences of trimethylaminuria, a condition with excess trimethylamine secretion in the body, and epilepsy (132). Our results show another reduction in *Schaedlerella* levels specific to Epi group. The knowledge about its role in gut is very limited in the literature, but described as one of the beneficial bacteria in a previous report (133). In another study, dietary intervention that lowers the GABA a receptor levels in amygdala has been shown to be increasing the *Schaedlerella* genus abundance in rats (134). This study also revealed increased *Odoribacter laneus* relative abundance in Epi group compared to Sham rat group. This species has been shown to decrease circulating succinate levels in obesity mouse models (135). Regarding findings of another report in which the authors proposed potential causality between neuronal degeneration in KA-induced status epilepticus and succinate accumulation through elevated oxidative stress (136), increase in the *O. laneus* levels may indicate the response to deal with KA-induced epilepsy. Even we cannot detect succinate from metabolomics DA analysis, "Alanine, aspartate and glutamate metabolism" pathway has been one of the enriched pathways from Epi vs Sham comparison. L-glutamate to succinate conversion via GABA takes place in this pathway, and we detected lower GABA levels in Epi group compared to Sham, which may be an indirect indicator of lessened succinate levels, but needs further investigations.

Specific to No-Epi vs Sham comparisons, *Dysosmobacter* levels were observed in higher levels in No-Epi rats. Even the role of this genus in the gut has not been well-known, one of the strains, *Dysosmobacter welbionis* J115^{T} has been linked to butyrate production and shown to be playing health-promoting role in obesity (137). Another

No-Epi vs Sham comparison specific genus detected in elevated abundance is *Angelakisella*. Those bacteria were reported to be potentially impacting metabolism of glutathione and glycerophospholipids by a study performed on samples from patients diagnosed with adenoma and colorectal cancer (138). We detected higher *Butyricicoccus* abundance in No-Epi group. A member of this genus, *Butyricicoccus pullicaecorum* were reported to be in decreased levels in patients diagnosed with neurodevelopmental disorders (139). Potential biomarker role was described for stroke in a study carried on patients, and higher levels were reported compared to healthy individuals (140). In an epilepsy study, anti-seizure medicine, topiramate administration was shown to enhance abundance of *Butyricicoccus* in mice (119). Additionally, *Lawsonibacter* was another genus detected in elevated levels specific to No-Epi when compared with Sham rats. *Lawsonibacter asaccharolyticus*, a butyrate producer identified recently (141), was in lower abundance in elderly people with with mild cognitive impairment (142). In contrast to those genera, *Ruminococcus-E* bacteria were detected in significantly lower levels in No-Epi group compared to Sham. *Ruminococcus* bacteria has been linked to either health promoting or disease-associated roles in previous studies because of their metabolic activities such as SCFA production, degeneration of intestinal mucosa that cause intestinal inflammation, and metabolism of dopamine which can modulate epileptic activities in bidirectional way depending on the dopamine receptor types involved (reviewed in (143)). In a metagenomic study carried on epileptic patients, changes in opposite directions were reported for the three species of this genus; higher levels of *R.* gnavus, *R. torques* and lower levels of *R. bicirculans* were detected in patients compared to their healthy relatives (144). Besides, a positive correlation was found between *R.* gnavus and epilepsy in the same study (144). Another clinical study in which both drug-resistant and drug-sensitive epilepsy patients and healthy controls were involved, investigated the changes differences in gut microbiome through 16S amplicon sequencing, and increased *Ruminococcus* abundances were found in resistant group compared to control (53). Another interesting study compared the gut composition of two epileptic mice groups with and without seizures (145). Epilepsy was induced via administration of Theiler murine encephalomyelitis virus, and *Ruminococcus* along with four other genera were associated with epileptic group without seizures (145).

We next sought for functional differences between rat groups by comparing the annotated KO gene abundances, followed by pathway enrichment using significantly changing KOs from each pairwise comparisons. Even the differences in KO relative abundance differences in compared groups is not a direct indication of whether the associated metabolic pathway has an enhanced or suppressed activity potential, majority of the KO's associated with those pathways are in lower abundances in former groups of Epi vs Sham, No-Epi vs Sham, and Epi vs No-Epi comparisons.

Both No-Epi vs Sham and Epi vs Sham comparisons share a few metabolic pathways, mainly related to the growth and signaling functions amongst gut bacterial communities. Among them, there are two pathways that may be involved in the regulation of the host immune system. First one is "Teichoic acid biosynthesis" and all associated KO's were in lower abundance in epileptic rats. Teichoic acids are one of the molecules found on the cell walls of the gram-positive bacteria and play role in the microbe-host communication and affect host immunity (146). The second one, "Peptidoglycan biosynthesis" is related to regulation of another important cell wall molecule, peptidoglycan, which is involved in stimulation of immune response in host (147).

Regarding particular importance of SCFA metabolisms, "Butanoate metabolism" enrichment from No-Epi vs Sham comparison specifically is really important and has a potential to enlighten the lack of seizure development in No-Epi group to be investigated further in future. Interestingly, higher number of enrichments of metabolic pathways of different sugar molecules were obtained in No-Epi vs Sham compared to Epi vs Sham comparison, and might indicate differences in energy consumption and regulations for seizure susceptibility, as energy homeostasis is an important factor linked to seizure development and activity (148).

Lastly, only one pathway was enriched from significantly differentially abundant KO's obtained from comparison of epileptic group rats; "C5-Branched dibasic acid metabolism", and showing potential differences between No-Epi and Epi group

bacteria to modulate branched-chain amino acids, which can regulate the seizure activity (149).

In contrast to the findings from shotgun metagenomics data where the main differentiation observed between No-Epi and Sham groups and the Epi vs No-Epi comparison yielded the minimum number of taxa and KO genes, serum metabolomics comparison showed that Epi vs Sham group comparison exhibits the highest number of changes in serum metabolites while the No-Epi vs Sham yielded the minimum number of different metabolites. It indicates development of status epilepticus and spontaneous seizures, rather than kainic acid injection, causes unique metabolite changes in blood. Within that frame, enrichment of metabolites of lipid class in Epi group rats is a crucial finding to see unique changes in response to the seizure development. In our metabolomics analysis, lineoleic acid was detected in significantly higher concentrations in serum samples of Epi group rats compared to both Sham and to No-Epi group. Besides, alpha-linolenic acid levels were found to be also significantly higher in Epi group compared to No-Epi group. The third metabolite associated with linoleic acid metabolism is arachidonate, which was detected in relatively higher concentrations in Epi group compared to Sham. Increased serum arachidonate levels were found in pediatric epilepsy patients with drug resistance in response to ketogenic diet which reduced the seizure frequency (150). In contrast, another study reported decreased amount of serum arachidonate in paediatric epilepsy who responded to the ketogenic diet compared to non-response group (151).

Apart from lipid metabolism, metabolite and pathway level comparisons showed substantial changes in several glucose and amino acid related pathways in both Epi and No-Epi groups compared to Sham rats.

The first attempt to obtain features from metagenomic and metabolomics datasets was carried out by employing two supervised algorithms, random forest and elastic net, aiming to obtain features that could differentiate the rat groups successfully. Considering low sample size for a machine learning, potential bias arising from random selection of the samples to the train and test datasets was checked by training

100 different models that uses 100 different pools of samples involved in train and test datasets, and by comparing the model performance distributions. As expected, for both omics datasets, the performances were highly variable, and a small portion of them could give satisfactory accuracy in the classification. Besides, the performances were much lower on MAG dataset compared to metabolite data. Therefore, this approach was abandoned, and an unsupervised method, MOFA was tested. Most of the latent factors produced could explain the variance in the two datasets individually. Among them, factor 3 could explain the highest percentage of biological variance in both datasets. Different pairing of the 15 LF's could not fully separate the three groups, and the best clustering pattern obtained from pairing of LF 3 and LF 4 was visually similar to the PCoA plots obtained from beta diversity analysis. Nevertheless, the correlation coefficients obtained from MAGs and metabolites selected through MOFA are the highest, and identified potential relationships between different MAGs and LPE derivatives . At the end, this approach was also abandoned since majority of the features in the high correlation group are not significantly different between the rat groups, hence could not be used in explaining the changes in epilepsy and seizure development. Lastly, MAGs and metabolites obtained from differential abundance analysis were used in correlation analysis, as they were detected to be significantly changing by at least one of the pairwise comparisons. Compared to correlations from MOFA-selected features, this approach yielded lower correlation coefficients, and only one of the metabolites could have been annotated among the selected metabolites in high correlation group; N-acetylornithine. For metagenomic features, two MAGs were obtained in high correlation group, annotated to unclassified species of *Angelakisella* and another unclassified bacteria *CAG-873 sp009775195* and the correlations with N-acetylornithine are in opposite direction.

Previous studies established significant correlations between N-acetylornithine and liver (152) and kidney (153) functions. Metabolic investigations on rat kidney samples revealed positive correlation between N-acetylornithine and arginine metabolism (154). Faecal metabolite analysis from patients suffering from constipation showed positive correlation between arginine biosynthesis pathway and disease progression, and even they could not detect significant change in arginine

levels, lower N-acetylornithine levels were found in disease group compared to healthy controls (155). 16S amplicon analysis from the same study showed correlation between five taxa which are A*listipes*, *Ruminiclostridium 5*, *Klebsiella*, and the *[Eubacterium] ventriosum group* (155), which may show possible involvement of the gut microbiome in the regulation of this metabolite in host.

When the metabolic capacities of the two MAGs were examined in our study, both bacteria have same GIFT score for arginine biosynthesis, indicating presence of similar gene repertoire and functional capacity, and requires additional omics analyses to be able to assess whether there is any functional difference that could explain the opposite correlations with N-acetylornithine.

There is no information found specifically investigating the role or changes in N-acetylornithine in epilepsy. A pilot study performed on serum samples from schizophrenia patients and healthy controls reported slight but insignificant decrease in patients, while significant changes were observed two other derivatives, higher N-acetylglutamine and lower N^6-acetyl-L-lysine were reported in patient blood samples (156). In another study focusing on the causal effects on human metabolites on five patient groups diagnosed with different psychiatric disorders, significantly lower serum levels of N-acetylornithine were found in patients diagnosed with schizophrenia or bipolar disorder (157). Decreased amount of N-acetylornithine were found in brain tissue of mice treated with an antipsychotic drug used in schizophrenia treatment (158). It required further investigations to understand role of N-acetylornithine in function of central nervous system, and how it can participate in epilepsy development and progression.

6 CONCLUSION

By analysing deep shotgun metagenomics and blood metabolomics data, significant fingerprints associated with the development of epilepsy have been discovered in this study. However, there are some limitations in this study that need to be addressed in future research to fully understand the role of gut microbiota in epilepsy pathogenesis. These include (i) the absence of additional meta-omics data, such as gut metabolomics, to understand how metagenomic capacity is translated and to identify which serum metabolites can be traced back to the gut; (ii) the lack of targeted metabolomics analyses to examine well-established gut-brain axis mediators, such as short-chain fatty acids; (iii) the absence of brain transcriptomics and/or metabolomics data to establish the impact of these changes on epilepsy-related pathways; and (iv) the use of single time-point measurements that cannot capture the dynamic nature of gut microbiota changes during epilepsy development, necessitating a longitudinal study design. Nevertheless, these findings will guide future investigations that seek to establish causal relationships between gut microbiota and epilepsy, with potential implications for diagnostics and therapy.

7 REFERENCES

1. WHO - Epilepsy [Internet]. [cited 2024 May 19]. Available from: https://www.who.int/news-room/fact-sheets/detail/epilepsy
2. Upadhya D, Kodali M, Gitai D, Castro OW, Zanirati G, Upadhya R, et al. A Model of Chronic Temporal Lobe Epilepsy Presenting Constantly Rhythmic and Robust Spontaneous Seizures, Co-morbidities and Hippocampal Neuropathology. Aging Dis. 2019;10(5):915.
3. Du Y, Li R, Fu D, Zhang B, Cui A, Shao Y, et al. Multi-omics technologies and molecular biomarkers in brain tumor-related epilepsy. CNS Neurosci Ther. 2024 Apr;30(4):e14717.
4. Boeri L, Donnaloja F, Campanile M, Sardelli L, Tunesi M, Fusco F, et al. Using integrated meta-omics to appreciate the role of the gut microbiota in epilepsy. Neurobiol Dis. 2022 Mar;164:105614.
5. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. Genome Biol. 2017 Dec;18(1):83.
6. Muller E, Shiryan I, Borenstein E. Multi-omic integration of microbiome data for identifying disease-associated modules. Nat Commun. 2024 Mar 23;15(1):2621.
7. Olson CA, Vuong HE, Yano JM, Liang QY, Nusbaum DJ, Hsiao EY. The Gut Microbiota Mediates the Anti-Seizure Effects of the Ketogenic Diet. Cell. 2018 Jun;173(7):1728-1741.e13.
8. Stafstrom CE, Carmant L. Seizures and Epilepsy: An Overview for Neuroscientists. Cold Spring Harb Perspect Med. 2015 Jun 1;5(6):a022426–a022426.
9. Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet. 2012 Dec;380(9859):2197–223.
10. Thijs RD, Surges R, O'Brien TJ, Sander JW. Epilepsy in adults. The Lancet. 2019 Feb;393(10172):689–701.
11. Scheffer IE, Berkovic S, Capovilla G, Connolly MB, French J, Guilhoto L, et al. ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. Epilepsia. 2017 Apr;58(4):512–21.
12. Harris L, Angus-Leppan H. Epilepsy: diagnosis, classification and management. Medicine (Baltimore). 2020 Aug;48(8):522–8.
13. Bialer M, White HS. Key factors in the discovery and development of new antiepileptic drugs. Nat Rev Drug Discov. 2010 Jan;9(1):68–82.
14. Lerche H, Petrou S. Genetic Animal Models of Epileptic Seizures. In: Panayiotopoulos CP, editor. Atlas of Epilepsies [Internet]. London: Springer London; 2010 [cited 2024 May 15]. p. 295–308. Available from: http://link.springer.com/10.1007/978-1-84882-128-6_39
15. Löscher W. Critical review of current animal models of seizures and epilepsy used in the discovery and development of new antiepileptic drugs. Seizure. 2011 Jun;20(5):359–68.
16. Devinsky O, Vezzani A, O'Brien TJ, Jette N, Scheffer IE, De Curtis M, et al. Epilepsy. Nat Rev Dis Primer. 2018 May 3;4(1):18024.

17. Henshall DC. Focally Applied Chemoconvulsants. In: Models of Seizures and Epilepsy [Internet]. Elsevier; 2017 [cited 2024 May 15]. p. 513–27. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780128040669000365
18. Cherian A, Thomas S. Status epilepticus. Ann Indian Acad Neurol. 2009;12(3):140.
19. Betjemann JP, Lowenstein DH. Status epilepticus in adults. Lancet Neurol. 2015 Jun;14(6):615–24.
20. Lévesque M, Avoli M, Bernard C. Animal models of temporal lobe epilepsy following systemic chemoconvulsant administration. J Neurosci Methods. 2016 Feb;260:45–52.
21. Kandratavicius L, Balista P, Lopes-Aguiar C, Ruggiero R, Umeoka E, Garcia-Cairasco N, et al. Animal models of epilepsy: use and limitations. Neuropsychiatr Dis Treat. 2014 Sep;1693.
22. Priel MR, Albuquerque EX. Short-Term Effects of Pilocarpine on Rat Hippocampal Neurons in Culture. Epilepsia. 2002 Jun;43(s5):40–6.
23. Smolders I, Khan GM, Manil J, Ebinger G, Michotte Y. NMDA receptor-mediated pilocarpine-induced seizures: characterization in freely moving rats by microdialysis. Br J Pharmacol. 1997 Jul;121(6):1171–9.
24. Tatum WO. Mesial Temporal Lobe Epilepsy. J Clin Neurophysiol. 2012 Oct;29(5):356–65.
25. Rusina E, Bernard C, Williamson A. The Kainic Acid Models of Temporal Lobe Epilepsy. eneuro. 2021 Mar;8(2):ENEURO.0337-20.2021.
26. Zheng XY, Zhang HL, Luo Q, Zhu J. Kainic Acid-Induced Neurodegenerative Model: Potentials and Limitations. J Biomed Biotechnol. 2011;2011:1–10.
27. Sadowsky MJ, Staley C, Heiner C, Hall R, Kelly CR, Brandt L, et al. Analysis of gut microbiota – An ever changing landscape. Gut Microbes. 2017 May 4;8(3):268–75.
28. Lyon L. 'All disease begins in the gut': was Hippocrates right? Brain. 2018 Mar 1;141(3):e20–e20.
29. Baquero F, Nombela C. The microbiome as a human organ. Clin Microbiol Infect. 2012 Jul;18:2–4.
30. Sender R, Fuchs S, Milo R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. PLOS Biol. 2016 Aug 19;14(8):e1002533.
31. Tanaka M, Nakayama J. Development of the gut microbiota in infancy and its impact on health in later life. Allergol Int. 2017 Oct;66(4):515–22.
32. Lazar V, Ditu LM, Pircalabioru GG, Gheorghe I, Curutiu C, Holban AM, et al. Aspects of Gut Microbiota and Immune System Interactions in Infectious Diseases, Immunopathology, and Cancer. Front Immunol. 2018 Aug 15;9:1830.
33. Jiao Y, Wu L, Huntington ND, Zhang X. Crosstalk Between Gut Microbiota and Innate Immunity and Its Implication in Autoimmune Diseases. Front Immunol. 2020 Feb 21;11:282.
34. Cox TO, Lundgren P, Nath K, Thaiss CA. Metabolic control by the microbiome. Genome Med. 2022 Dec;14(1):80.
35. Visconti A, Le Roy CI, Rosa F, Rossi N, Martin TC, Mohney RP, et al. Interplay between the human gut microbiome and host metabolism. Nat Commun. 2019 Oct 3;10(1):4505.

36. 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 Dec;444(7122):1027–31.
37. Tilg H, Cani PD, Mayer EA. Gut microbiome and liver diseases. Gut. 2016 Dec;65(12):2035–44.
38. Gopalakrishnan V, Helmink BA, Spencer CN, Reuben A, Wargo JA. The Influence of the Gut Microbiome on Cancer, Immunity, and Cancer Immunotherapy. Cancer Cell. 2018 Apr;33(4):570–80.
39. 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(5):17004.
40. Cryan JF, O'Riordan KJ, Sandhu K, Peterson V, Dinan TG. The gut microbiome in neurological disorders. Lancet Neurol. 2020 Feb;19(2):179–94.
41. Ahlawat S, Asha, Sharma KK. Gut–organ axis: a microbial outreach and networking. Lett Appl Microbiol. 2021 Jun;72(6):636–68.
42. Ma Q, Xing C, Long W, Wang HY, Liu Q, Wang RF. Impact of microbiota on central nervous system and neurological diseases: the gut-brain axis. J Neuroinflammation. 2019 Dec;16(1):53.
43. Bachiller S, Jiménez-Ferrer I, Paulus A, Yang Y, Swanberg M, Deierborg T, et al. Microglia in Neurological Diseases: A Road Map to Brain-Disease Dependent-Inflammatory Response. Front Cell Neurosci. 2018 Dec 18;12:488.
44. Erny D, Hrabě De Angelis AL, Jaitin D, Wieghofer P, Staszewski O, David E, et al. Host microbiota constantly control maturation and function of microglia in the CNS. Nat Neurosci. 2015 Jul;18(7):965–77.
45. Sofroniew MV, Vinters HV. Astrocytes: biology and pathology. Acta Neuropathol (Berl). 2010 Jan;119(1):7–35.
46. Kamermans A, Planting KE, Jalink K, Van Horssen J, De Vries HE. Reactive astrocytes in multiple sclerosis impair neuronal outgrowth through TRPM7-mediated chondroitin sulfate proteoglycan production. Glia. 2019 Jan;67(1):68–77.
47. Rothhammer V, Mascanfroni ID, Bunse L, Takenaka MC, Kenison JE, Mayo L, et al. Type I interferons and microbial metabolites of tryptophan modulate astrocyte activity and central nervous system inflammation via the aryl hydrocarbon receptor. Nat Med. 2016 Jun;22(6):586–97.
48. Muller PA, Schneeberger M, Matheis F, Wang P, Kerner Z, Ilanges A, et al. Microbiota modulate sympathetic neurons via a gut–brain circuit. Nature. 2020 Jul 16;583(7816):441–6.
49. Medel-Matus J, Shin D, Dorfman E, Sankar R, Mazarati A. Facilitation of kindling epileptogenesis by chronic stress may be mediated by intestinal microbiome. Epilepsia Open. 2018 Jun;3(2):290–4.
50. Russo E, Citraro R, Constanti A, Leo A, Lüttjohann A, Van Luijtelaar G, et al. Upholding WAG/Rij rats as a model of absence epileptogenesis: Hidden mechanisms and a new theory on seizure development. Neurosci Biobehav Rev. 2016 Dec;71:388–408.

51. Citraro R, Lembo F, De Caro C, Tallarico M, Coretti L, Iannone LF, et al. First evidence of altered microbiota and intestinal damage and their link to absence epilepsy in a genetic animal model, the WAG/Rij rat. Epilepsia. 2021 Feb;62(2):529–41.
52. Holmes M, Flaminio Z, Vardhan M, Xu F, Li X, Devinsky O, et al. Cross talk between drug-resistant epilepsy and the gut microbiome. Epilepsia. 2020 Dec;61(12):2619–28.
53. Peng A, Qiu X, Lai W, Li W, Zhang L, Zhu X, et al. Altered composition of the gut microbiome in patients with drug-resistant epilepsy. Epilepsy Res. 2018 Nov;147:102–7.
54. Lee H, Lee S, Lee DH, Kim DW. A comparison of the gut microbiota among adult patients with drug-responsive and drug-resistant epilepsy: An exploratory study. Epilepsy Res. 2021 May;172:106601.
55. D'Andrea Meira I, Romão TT, Pires Do Prado HJ, Krüger LT, Pires MEP, Da Conceição PO. Ketogenic Diet and Epilepsy: What We Know So Far. Front Neurosci. 2019 Jan 29;13:5.
56. Xie G, Zhou Q, Qiu CZ, Dai WK, Wang HP, Li YH, et al. Ketogenic diet poses a significant effect on imbalanced gut microbiota in infants with refractory epilepsy. World J Gastroenterol. 2017 Sep 7;23(33):6164–71.
57. Riva A, Sahin E, Volpedo G, Petretto A, Lavarello C, Di Sapia R, et al. Identification of an epilepsy-linked gut microbiota signature in a pediatric rat model of acquired epilepsy. Neurobiol Dis. 2024 May;194:106469.
58. FastQC [Internet]. 2015. Available from: https://qubeshub.org/resources/fastqc
59. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016 Oct 1;32(19):3047–8.
60. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018 Sep 1;34(17):i884–90.
61. SourceForge [Internet]. 2023 [cited 2024 May 16]. BBMap. Available from: https://sourceforge.net/projects/bbmap/
62. The UniVec Database [Internet]. [cited 2024 May 16]. Available from: https://www.ncbi.nlm.nih.gov/tools/vecscreen/univec/
63. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012 Apr;9(4):357–9.
64. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078–9.
65. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol. 2019 Nov 28;20(1):257.
66. Blanco-Míguez A, Beghini F, Cumbo F, McIver LJ, Thompson KN, Zolfo M, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. Nat Biotechnol. 2023 Nov;41(11):1633–44.
67. Menzel P, Ng KL, Krogh A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun. 2016 Apr 13;7(1):11257.

68. Index zone by BenLangmead [Internet]. [cited 2024 May 16]. Available from: https://benlangmead.github.io/aws-indexes/k2
69. Kaiju: Fast and sensitive taxonomic classification for metagenomics [Internet]. [cited 2024 May 16]. Available from: https://bioinformatics-centre.github.io/kaiju/downloads.html
70. Krakau S, Straub D, Gourlé H, Gabernet G, Nahnsen S. nf-core/mag: a best-practice pipeline for metagenome hybrid assembly and binning. NAR Genomics Bioinforma. 2022 Jan 13;4(1):lqac007.
71. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct *de Bruijn* graph. Bioinformatics. 2015 May 15;31(10):1674–6.
72. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ. 2019 Jul 26;7:e7359.
73. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics. 2016 Feb 15;32(4):605–7.
74. Wang Z, Huang P, You R, Sun F, Zhu S. MetaBinner: a high-performance and stand-alone ensemble binning method to recover individual genomes from complex microbial communities. Genome Biol. 2023 Jan 6;24(1):1.
75. Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome. 2018 Dec;6(1):158.
76. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015 Jul;25(7):1043–55.
77. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J. 2017 Dec;11(12):2864–8.
78. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. Borgwardt K, editor. Bioinformatics. 2022 Nov 30;38(23):5315–6.
79. Shaffer M, Borton MA, McGivern BB, Zayed AA, La Rosa SL, Solden LM, et al. DRAM for distilling microbial metabolism to automate the curation of microbiome function. Nucleic Acids Res. 2020 Sep 18;48(16):8883–900.
80. wwood/CoverM: Read coverage calculator for metagenomics [Internet]. [cited 2024 May 16]. Available from: https://github.com/wwood/CoverM
81. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010 Dec;11(1):119.
82. Bateman A. The Pfam protein families database. Nucleic Acids Res. 2004 Jan 1;32(90001):138D – 141.

83. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, et al. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. Valencia A, editor. Bioinformatics. 2020 Apr 1;36(7):2251–2.
84. The UniProt Consortium, Bateman A, Martin MJ, Orchard S, Magrane M, Ahmad S, et al. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. 2023 Jan 6;51(D1):D523–31.
85. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res. 2009 Jan 1;37(Database):D233–8.
86. Rawlings ND, Barrett AJ, Bateman A. MEROPS: the peptidase database. Nucleic Acids Res. 2010 Jan;38(suppl_1):D227–33.
87. VOGDB | Virus orthologous groups [Internet]. [cited 2024 May 16]. Available from: https://vogdb.org/
88. anttonalberdi/distillR: R package for distilling functional annotations of bacterial genomes and metagenomes [Internet]. [cited 2024 May 16]. Available from: https://github.com/anttonalberdi/distillR
89. Eisenhofer R, Odriozola I, Alberdi A. Impact of microbial genome completeness on metagenomic functional inference. ISME Commun. 2023 Dec 1;3(1):12.
90. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2022. Available from: https://www.R-project.org/
91. Pereira MB, Wallroth M, Jonsson V, Kristiansson E. Comparison of normalization methods for the analysis of metagenomic gene abundance data. BMC Genomics. 2018 Dec;19(1):274.
92. Bars-Cortina D. Alpha and beta-diversities performance comparison between different normalization methods and centered log-ratio transformation in a microbiome public dataset [Internet]. 2022 [cited 2024 May 16]. Available from: http://biorxiv.org/lookup/doi/10.1101/2022.11.07.512066
93. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014 Dec;15(12):550.
94. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. Watson M, editor. PLoS ONE. 2013 Apr 22;8(4):e61217.
95. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Schwartz R, editor. Bioinformatics. 2019 Feb 1;35(3):526–8.
96. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021 Jul 2;49(W1):W293–6.
97. Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, et al. Picante: R tools for integrating phylogenies and ecology. Bioinformatics. 2010 Jun 1;26(11):1463–4.
98. Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, et al. vegan: Community Ecology Package [Internet]. 2022. Available from: https://CRAN.R-project.org/package=vegan
99. Arbizu PM. pairwiseAdonis: Pairwise Multilevel Comparison using Adonis. 2017.

100. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016 Sep 15;32(18):2847–9.
101. kevinblighe/EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling [Internet]. [cited 2024 May 16]. Available from: https://github.com/kevinblighe/EnhancedVolcano
102. KO (KEGG ORTHOLOGY) Database [Internet]. [cited 2024 May 16]. Available from: https://www.genome.jp/kegg/ko.html
103. Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucleic Acids Res. 2021 Jul 2;49(W1):W388–96.
104. López-Ibáñez J, Pazos F, Chagoyen M. MBROLE 2.0—functional enrichment of chemical compounds. Nucleic Acids Res. 2016 Jul 8;44(W1):W201–4.
105. Kuhn M. Building Predictive Models in *R* Using the **caret** Package. J Stat Softw [Internet]. 2008 [cited 2024 May 16];28(5). Available from: http://www.jstatsoft.org/v28/i05/
106. Argelaguet R, Velten B, Arnol D, Dietrich S, Zenz T, Marioni JC, et al. Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. Mol Syst Biol. 2018 Jun;14(6):e8124.
107. Sharpton TJ. An introduction to the analysis of shotgun metagenomic data. Front Plant Sci [Internet]. 2014 Jun 16 [cited 2024 May 17];5. Available from: http://journal.frontiersin.org/article/10.3389/fpls.2014.00209/abstract
108. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol. 2017 Sep;35(9):833–44.
109. Tovo A, Menzel P, Krogh A, Cosentino Lagomarsino M, Suweis S. Taxonomic classification method for metagenomics based on core protein families with Core-Kaiju. Nucleic Acids Res. 2020 Sep 18;48(16):e93–e93.
110. Beghini F, McIver LJ, Blanco-Míguez A, Dubois L, Asnicar F, Maharjan S, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. eLife. 2021 May 4;10:e65088.
111. Almeida A, Nayfach S, Boland M, Strozzi F, Beracochea M, Shi ZJ, et al. A unified catalog of 204,938 reference genomes from the human gut microbiome. Nat Biotechnol. 2021 Jan;39(1):105–14.
112. Kieser S, Zdobnov EM, Trajkovski M. Comprehensive mouse microbiota genome catalog reveals major difference to its human counterpart. Nagarajan N, editor. PLOS Comput Biol. 2022 Mar 8;18(3):e1009947.
113. Pan H, Guo R, Zhu J, Wang Q, Ju Y, Xie Y, et al. A gene catalogue of the Sprague-Dawley rat gut metagenome. GigaScience [Internet]. 2018 May 1 [cited 2024 May 16];7(5). Available from: https://academic.oup.com/gigascience/article/doi/10.1093/gigascience/giy055/4995266

114. Lindefeldt M, Eng A, Darban H, Bjerkner A, Zetterström CK, Allander T, et al. The ketogenic diet influences taxonomic and functional composition of the gut microbiota in children with severe epilepsy. Npj Biofilms Microbiomes. 2019 Jan 23;5(1):5.
115. Li X, Wang Q, Wu D, Zhang D wen, Li S chang, Zhang S wei, et al. The effect of a novel anticonvulsant chemical Q808 on gut microbiota and hippocampus neurotransmitters in pentylenetetrazole-induced seizures in rats. BMC Neurosci. 2022 Dec;23(1):7.
116. Kim Y, Jang Y, Shin YW, Jeon D, Yoo J seok, Park DK, et al. Gut microbiome diversity in a febrile seizure mouse model. encephalitis. 2024 Jan 10;4(1):11–7.
117. Bravo JA, Forsythe P, Chew MV, Escaravage E, Savignac HM, Dinan TG, et al. Ingestion of *Lactobacillus* strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. Proc Natl Acad Sci. 2011 Sep 20;108(38):16050–5.
118. He T, Zhu YH, Yu J, Xia B, Liu X, Yang GY, et al. Lactobacillus johnsonii L531 reduces pathogen load and helps maintain short-chain fatty acid levels in the intestines of pigs challenged with Salmonella enterica Infantis. Vet Microbiol. 2019 Mar;230:187–94.
119. Thai K, Taylor MW, Fernandes T, Akinade EA, Campbell SL. Topiramate alters the gut microbiome to aid in its anti-seizure effect. Front Microbiol. 2023 Oct 10;14:1242856.
120. Ahmadian SR, Ghasemi-Kasman M, Pouramir M, Sadeghi F. Arbutin attenuates cognitive impairment and inflammatory response in pentylenetetrazol-induced kindling model of epilepsy. Neuropharmacology. 2019 Mar;146:117–27.
121. Ma J, Chen S, Li Y, Wu X, Song Z. Arbutin improves gut development and serum lipids via Lactobacillus intestinalis. Front Nutr. 2022 Sep 9;9:948573.
122. Yu Y, Nguyen DT, Jiang J. G protein-coupled receptors in acquired epilepsy: Druggability and translatability. Prog Neurobiol. 2019 Dec;183:101682.
123. Hu J, Stern M, Gimenez LE, Wanka L, Zhu L, Rossi M, et al. A G Protein-biased Designer G Protein-coupled Receptor Useful for Studying the Physiological Relevance of Gq/11-dependent Signaling Pathways. J Biol Chem. 2016 Apr;291(15):7809–20.
124. Tabouy L, Getselter D, Ziv O, Karpuj M, Tabouy T, Lukic I, et al. Dysbiosis of microbiome and probiotic treatment in a genetic model of autism spectrum disorders. Brain Behav Immun. 2018 Oct;73:310–9.
125. Liu Y, Tian X, Daniel RC, Okeugo B, Armbrister SA, Luo M, et al. Impact of probiotic Limosilactobacillus reuteri DSM 17938 on amino acid metabolism in the healthy newborn mouse. Amino Acids. 2022 Oct;54(10):1383–401.
126. Mohammadi E, Nikbakht F, Vazifekhah S, Babae JF, Jogataei MT. Evaluation the cognition-improvement effects of N-acetyl cysteine in experimental temporal lobe epilepsy in rat. Behav Brain Res. 2023 Feb;440:114263.
127. Wang H, Zhou L, Zheng Q, Song Y, Huang W, Yang L, et al. Kai-xin-san improves cognitive impairment in D-gal and Aβ25-35 induced ad rats by regulating gut microbiota and reducing neuronal damage. J Ethnopharmacol. 2024 Jul;329:118161.

128. Lynch JB, Gonzalez EL, Choy K, Faull KF, Jewell T, Arellano A, et al. Gut microbiota Turicibacter strains differentially modify bile acids and host lipids. Nat Commun. 2023 Jun 20;14(1):3669.
129. Gong X, Liu X, Chen C, Lin J, Li A, Guo K, et al. Alteration of Gut Microbiota in Patients With Epilepsy and the Potential Index as a Biomarker. Front Microbiol. 2020 Sep 18;11:517797.
130. Miljanovic N, Potschka H. The impact of Scn1a deficiency and ketogenic diet on the intestinal microbiome: A study in a genetic Dravet mouse model. Epilepsy Res. 2021 Dec;178:106826.
131. Rajakovich LJ, Fu B, Bollenbach M, Balskus EP. Elucidation of an anaerobic pathway for metabolism of L -carnitine–derived γ-butyrobetaine to trimethylamine in human gut bacteria. Proc Natl Acad Sci. 2021 Aug 10;118(32):e2101498118.
132. Pellicciari A, Posar A, Cremonini MA, Parmeggiani A. Epilepsy and trimethylaminuria: A new case report and literature review. Brain Dev. 2011 Aug;33(7):593–6.
133. Yang Y, Huang S, Liao Y, Wu X, Zhang C, Wang X, et al. Hippuric acid alleviates dextran sulfate sodium-induced colitis via suppressing inflammatory activity and modulating gut microbiota. Biochem Biophys Res Commun. 2024 May;710:149879.
134. Dalziel JE, Zobel G, Dewhurst H, Hurst C, Olson T, Rodriguez-Sanchez R, et al. A Diet Enriched with Lacticaseibacillus rhamnosus HN001 and Milk Fat Globule Membrane Alters the Gut Microbiota and Decreases Amygdala GABA a Receptor Expression in Stress-Sensitive Rats. Int J Mol Sci. 2023 Jun 21;24(13):10433.
135. Huber-Ruano I, Calvo E, Mayneris-Perxachs J, Rodríguez-Peña MM, Ceperuelo-Mallafré V, Cedó L, et al. Orally administered Odoribacter laneus improves glucose control and inflammatory profile in obese mice by depleting circulating succinate. Microbiome. 2022 Aug 25;10(1):135.
136. Zhang Y, Zhang M, Zhu W, Yu J, Wang Q, Zhang J, et al. Succinate accumulation induces mitochondrial reactive oxygen species generation and promotes status epilepticus in the kainic acid rat model. Redox Biol. 2020 Jan;28:101365.
137. Le Roy T, Moens De Hase E, Van Hul M, Paquot A, Pelicaen R, Régnier M, et al. *Dysosmobacter welbionis* is a newly isolated human commensal bacterium preventing diet-induced obesity and metabolic disorders in mice. Gut. 2022 Mar;71(3):534–43.
138. Han S, Zhuang J, Pan Y, Wu W, Ding K. Different Characteristics in Gut Microbiome between Advanced Adenoma Patients and Colorectal Cancer Patients by Metagenomic Analysis. Schmitz-Esser S, editor. Microbiol Spectr. 2022 Dec 21;10(6):e01593-22.
139. Bojović K, Ignjatović ĐDI, Soković Bajić S, Vojnović Milutinović D, Tomić M, Golić N, et al. Gut Microbiota Dysbiosis Associated With Altered Production of Short Chain Fatty Acids in Children With Neurodevelopmental Disorders. Front Cell Infect Microbiol. 2020 May 19;10:223.
140. Wei D, Chen X, Xu J, Yin Y, Peng X, Li S, et al. Identification of disordered profiles of gut microbiota and functional component in stroke and poststroke epilepsy. Brain Behav. 2023 Dec;13(12):e3318.

141. Sakamoto M, Iino T, Yuki M, Ohkuma M. Lawsonibacter asaccharolyticus gen. nov., sp. nov., a butyrate-producing bacterium isolated from human faeces. Int J Syst Evol Microbiol. 2018 Jun 1;68(6):2074–81.
142. Chaudhari DS, Jain S, Yata VK, Mishra SP, Kumar A, Fraser A, et al. Unique trans-kingdom microbiome structural and functional signatures predict cognitive decline in older adults. GeroScience. 2023 May 22;45(5):2819–34.
143. Hamamah S, Aghazarian A, Nazaryan A, Hajnal A, Covasa M. Role of Microbiota-Gut-Brain Axis in Regulating Dopaminergic Signaling. Biomedicines. 2022 Feb 13;10(2):436.
144. Dong L, Zheng Q, Cheng Y, Zhou M, Wang M, Xu J, et al. Gut Microbial Characteristics of Adult Patients With Epilepsy. Front Neurosci. 2022 Feb 16;16:803538.
145. Gallucci A, Patel DC, Thai K, Trinh J, Gude R, Shukla D, et al. Gut metabolite S-equol ameliorates hyperexcitability in entorhinal cortex neurons following Theiler murine encephalomyelitis virus-induced acute seizures. Epilepsia. 2021 Aug;62(8):1829–41.
146. Han J, Zhao X, Zhao X, Li P, Gu Q. Insight into the structure, biosynthesis, isolation method and biological function of teichoic acid in different gram-positive microorganisms: A review. Int J Biol Macromol. 2023 Dec;253:126825.
147. Crump GM, Zhou J, Mashayekh S, Grimes CL. Revisiting peptidoglycan sensing: interactions with host immunity and beyond. Chem Commun. 2020;56(87):13313–22.
148. McDonald TS, Neal ES, Borges K. Fructose 1,6-bisphosphate is anticonvulsant and improves oxidative glucose metabolism within the hippocampus and liver in the chronic pilocarpine mouse epilepsy model. Epilepsy Behav. 2021 Sep;122:108223.
149. Gruenbaum SE, Chen EC, Sandhu MRS, Deshpande K, Dhaher R, Hersey D, et al. Branched-Chain Amino Acids and Seizures: A Systematic Review of the Literature. CNS Drugs. 2019 Aug;33(8):755–70.
150. Poorshiri B, Barzegar M, Afghan M, Shiva S, Shahabi P, Golchinfar Z, et al. The effects of ketogenic diet on beta-hydroxybutyrate, arachidonic acid, and oxidative stress in pediatric epilepsy. Epilepsy Behav. 2023 Mar;140:109106.
151. Porta N, Vallée L, Boutry E, Fontaine M, Dessein AF, Joriot S, et al. Comparison of seizure reduction and serum fatty acid levels after receiving the ketogenic and modified Atkins diet. Seizure. 2009 Jun;18(5):359–64.
152. Liu Y, Ling L, Shen Y, Bi X. Metabolome-wide Mendelian randomization reveals causal effects of betaine and N-acetylornithine on impairment of renal function. Front Nutr. 2024 Apr 24;11:1371995.
153. Lan Y, Wang D, He J, Yang H, Hou Y, Di W, et al. The gut microbiome and metabolome in kidney transplant recipients with normal and moderately decreased kidney function. Ren Fail. 2023 Dec 31;45(1):2228419.
154. Wang Q, Yue J, Zhou X, Zheng M, Cao B, Li J. Ouabain regulates kidney metabolic profiling in rat offspring of intrauterine growth restriction induced by low-protein diet. Life Sci. 2020 Oct;259:118281.

155. Wang J, Wang L, Yu Q, Tang N, Mei C, Zhang H, et al. Characteristics of the Gut Microbiome and Serum Metabolome in Patients with Functional Constipation. Nutrients. 2023 Apr 6;15(7):1779.
156. Huang N, Cao B, Brietzke E, Park C, Cha D, Pan Z, et al. A pilot case-control study on the association between N-acetyl derivatives in serum and first-episode schizophrenia. Psychiatry Res. 2019 Feb;272:36–41.
157. Yang J, Yan B, Zhao B, Fan Y, He X, Yang L, et al. Assessing the Causal Effects of Human Serum Metabolites on 5 Major Psychiatric Disorders. Schizophr Bull. 2020 Jul 8;46(4):804–13.
158. McClay JL, Vunck SA, Batman AM, Crowley JJ, Vann RE, Beardsley PM, et al. Neurochemical Metabolomics Reveals Disruption to Sphingolipid Metabolism Following Chronic Haloperidol Administration. J Neuroimmune Pharmacol. 2015 Sep;10(3):425–34.

8 APPENDIX

APPENDIX 1

In this section, taxonomic annotations of the established MAG catalogue using GTDB-Tk database are presented.

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database

MAG ID	k_Kingdom; p_Phylum; c_Class; o_Order; f_Family; g_Genus; s_Species
group0_bin.89	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_14-2; s_
group1_bin.119	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_14-2; s_14-2 sp004793545
group1_bin.126	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_14-2; s_14-2 sp910589635
second_bin.8	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_14-2; s_14-2 sp910580195
group2_bin.95	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_1XD42-69; s_1XD42-69 sp003612565
second_bin.58	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_1XD42-69; s_
group1_bin.105	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_1XD8-76; s_1XD8-76 sp910573755
group2_bin.100	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_1XD8-76; s_
group1_bin.137	k_Bacteria; p_Proteobacteria; c_Alphaproteobacteria; o_RF32; f_CAG-239; g_51-20; s_51-20 sp910587405
group0_bin.139	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Acetatifactor; s_

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group1_bin.16	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Acetatifactor; s_
group1_bin.76	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Acetatifactor; s_
second_bin.109	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Acetatifactor; s_Acetatifactor sp910584865
second_bin.28	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Acetatifactor; s_
second_bin.76	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Acetatifactor; s_
second_bin.88	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Acetatifactor; s_
group0_bin.25	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Acutalibacteraceae; g_Acutalibacter; s_Acutalibacter sp910584745
group0_bin.42	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Acutalibacteraceae; g_Acutalibacter; s_
group1_bin.152	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Acutalibacteraceae; g_Acutalibacter; s_Acutalibacter muris
group1_bin.157	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Acutalibacteraceae; g_Acutalibacter; s_
group1_bin.18	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Acutalibacteraceae; g_Acutalibacter; s_Acutalibacter sp910589395
second_bin.46	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Acutalibacteraceae; g_Acutalibacter; s_Acutalibacter sp009936035

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group1_bin.51	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Agathobacter; s_Agathobacter sp910587655
group2_bin.91	k_Bacteria; p_Verrucomicrobiota; c_Verrucomicrobiae; o_Verrucomicrobiales; f_Akkermansiaceae; g_Akkermansia; s_Akkermansia muciniphila
group0_bin.30	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Rikenellaceae; g_Alistipes; s_Alistipes sp003979135
group0_bin.47	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Rikenellaceae; g_Alistipes; s_Alistipes sp910587675
group0_bin.71	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Rikenellaceae; g_Alistipes; s_Alistipes sp002358415
group1_bin.138	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Rikenellaceae; g_Alistipes; s_Alistipes sp910576555
group1_bin.156	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Rikenellaceae; g_Alistipes; s_Alistipes sp002362235
group0_bin.107	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Bacteroidaceae; g_Alloprevotella; s_Alloprevotella sp002933955
group1_bin.54	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Anaerotignaceae; g_Anaerotignum; s_Anaerotignum sp910576545
second_bin.39	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Anaerotignaceae; g_Anaerotignum; s_
group1_bin.130	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_Anaerotruncus; s_Anaerotruncus sp003612625

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

second_bin.23	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_Anaerotruncus; s_Anaerotruncus sp000403395
group2_bin.128	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_Angelakisella; s_
second_bin.103	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_Angelakisella; s_
second_bin.106	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_Angelakisella; s_Angelakisella sp910585535
group1_bin.147	k_Bacteria; p_Firmicutes_B; c_Dehalobacteriia; o_UBA4068; f_UBA5755; g_Avidehalobacter; s_
second_bin.61	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_CAG-272; g_Avispirillum; s_Avispirillum sp011957885
second_bin.62	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Avoscillospira_A; s_
group0_bin.158	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Peptostreptococcales; f_Anaerovoracaceae; g_Bacilliculturomica; s_
group1_bin.144	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Bacteroidaceae; g_Bacteroides; s_Bacteroides thetaiotaomicron
group1_bin.26	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Bacteroidaceae; g_Bacteroides; s_Bacteroides xylanisolvens
group2_bin.158	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Bacteroidaceae; g_Bacteroides; s_Bacteroides clarus
second_bin.20	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Bacteroidaceae; g_Bacteroides; s_Bacteroides uniformis

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group0_bin.72	k_Bacteria; p_Actinobacteriota; c_Actinomycetia; o_Actinomycetales; f_Bifidobacteriaceae; g_Bifidobacterium; s_Bifidobacterium globosum
second_bin.4	k_Bacteria; p_Desulfobacterota_I; c_Desulfovibrionia; o_Desulfovibrionales; f_Desulfovibrionaceae; g_Bilophila; s_Bilophila sp910585945
group2_bin.155	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Blautia_A; s_
group2_bin.30	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Blautia_A; s_
group0_bin.115	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Butyribacter; s_
group0_bin.61	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Butyribacter; s_Butyribacter sp910588655
group0_bin.68	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Butyribacter; s_Butyribacter sp003611745
group1_bin.140	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Butyribacter; s_Butyribacter sp910586555
group2_bin.47	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Butyribacter; s_Butyribacter sp009774235
second_bin.52	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Butyricicoccaceae; g_Butyricicoccus; s_
second_bin.17	k_Bacteria; p_Firmicutes; c_Bacilli; o_Erysipelotrichales; f_Erysipelotrichaceae; g_C-19; s_C-19 sp910576855

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

second_bin.63	k_Bacteria; p_Firmicutes; c_Bacilli; o_Erysipelotrichales; f_Erysipelotrichaceae; g_C-19; s_
group1_bin.20	k_Bacteria; p_Firmicutes; c_Bacilli; o_RF39; f_UBA660; g_Caccenecus; s_
group1_bin.90	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Caccovicinus; s_Caccovicinus sp910585635
group2_bin.18	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Caccovicinus; s_Caccovicinus sp910575565
group0_bin.59	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_CAG-103; s_CAG-103 sp018265835
second_bin.7	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_CAG-115; s_CAG-115 sp910587465
group0_bin.62	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_CAG-127; s_CAG-127 sp002493625
group2_bin.28	k_Bacteria; p_Cyanobacteria; c_Vampirovibrionia; o_Gastranaerophilales; f_Gastranaerophilaceae; g_CAG-196; s_CAG-196 sp002102975
group1_bin.58	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_TANB77; f_CAG-508; g_CAG-269; s_CAG-269 sp900556945
second_bin.12	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_TANB77; f_CAG-508; g_CAG-269; s_
second_bin.38	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_TANB77; f_CAG-508; g_CAG-269; s_
group0_bin.129	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_TANB77; f_CAG-508; g_CAG-273; s_CAG-273 sp000438355
group2_bin.141	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_TANB77; f_CAG-508; g_CAG-273; s_

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

second_bin.22	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_TANB77; f_CAG-508; g_CAG-273; s_
group2_bin.147	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_CAG-303; s_CAG-303 sp910584685
second_bin.74	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_CAG-317; s_CAG-317 sp000433535
group0_bin.159	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_UBA1381; f_UBA1381; g_CAG-41; s_CAG-41 sp910579415
group2_bin.118	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_UBA1381; f_UBA1381; g_CAG-41; s_
group0_bin.12	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-485; s_CAG-485 sp002493045
group0_bin.5	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-485; s_CAG-485 sp009775375
group1_bin.104	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-485; s_CAG-485 sp002361215
group1_bin.133	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-485; s_CAG-485 sp003979075
group1_bin.168	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-485; s_CAG-485 sp009775365
group1_bin.83	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-485; s_CAG-485 sp002361155
group2_bin.136	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-485; s_CAG-485 sp002491945
group2_bin.14	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-485; s_CAG-485 sp002362485
group0_bin.56	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_CAG-56; s_CAG-56 sp905200565

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group1_bin.120	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_CAG-632; s_CAG-632 sp910587045
group0_bin.103	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-873; s_CAG-873 sp011959565
group0_bin.135	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-873; s_CAG-873 sp009775195
group0_bin.151	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-873; s_CAG-873 sp009775265
group0_bin.173	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-873; s_CAG-873 sp009775225
group0_bin.76	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-873; s_CAG-873 sp001689665
group0_bin.94	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-873; s_CAG-873 sp009775535
group2_bin.111	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-873; s_CAG-873 sp910587235
group2_bin.153	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-873; s_CAG-873 sp003979155
group2_bin.20	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-873; s_CAG-873 sp910586975
group2_bin.69	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-873; s_CAG-873 sp910578095
group2_bin.70	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_CAG-873; s_CAG-873 sp002490635
group0_bin.155	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_CAG-95; s_CAG-95 sp009774465
group1_bin.181	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_CAG-95; s_CAG-95 sp910579425

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group1_bin.5	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_CAG-95; s_CAG-95 sp009911035
second_bin.21	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_CAG-95; s_
second_bin.41	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_CAG-95; s_
group2_bin.162	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_CAJFPI01; s_
second_bin.24	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_CAJJNI01; s_
second_bin.81	k_Bacteria; p_Proteobacteria; c_Alphaproteobacteria; o_RF32; f_CAG-239; g_CAZU01; s_CAZU01 sp000432275
group2_bin.110	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Choladocola; s_Choladocola sp910587765
group2_bin.90	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Choladocola; s_Choladocola sp009774155
group0_bin.28	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Choladousia; s_
group2_bin.89	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Choladousia; s_Choladousia sp002491565
group1_bin.129	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Clostridium_Q; s_Clostridium_Q sp910575795
second_bin.99	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Clostridium_Q; s_

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group0_bin.152	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_COE1; s_COE1 sp002490665
group0_bin.163	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_COE1; s_
group0_bin.53	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_COE1; s_COE1 sp009774245
group1_bin.69	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_COE1; s_COE1 sp002358575
group1_bin.75	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_COE1; s_COE1 sp009774345
group2_bin.119	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_COE1; s_
second_bin.73	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_COE1; s_COE1 sp009774375
group0_bin.106	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Copromonas; s_
group1_bin.101	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Copromonas; s_Copromonas sp910586315
group2_bin.166	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Copromonas; s_
second_bin.98	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Copromonas; s_Copromonas sp009917545
group0_bin.96	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_Borkfalkiaceae; g_Coproplasma; s_Coproplasma sp910587005
group2_bin.49	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_Borkfalkiaceae; g_Coproplasma; s_

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

second_bin.104	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_Borkfalkiaceae; g_Coproplasma; s_Coproplasma sp910587825
second_bin.5	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_Borkfalkiaceae; g_Coproplasma; s_
group0_bin.86	k_Bacteria; p_Desulfobacterota_I; c_Desulfovibrionia; o_Desulfovibrionales; f_Desulfovibrionaceae; g_Desulfovibrio; s_Desulfovibrio sp009773975
group0_bin.10	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_Duncaniella; s_Duncaniella sp910586305
group0_bin.95	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_Duncaniella; s_Duncaniella freteri
group0_bin.98	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_Duncaniella; s_Duncaniella dubosii
group1_bin.115	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_Duncaniella; s_Duncaniella sp001689575
group2_bin.24	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_Duncaniella; s_Duncaniella muris
group2_bin.48	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_Duncaniella; s_Duncaniella sp001689495
group2_bin.163	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Dysosmobacter; s_
second_bin.42	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Dysosmobacter; s_Dysosmobacter sp910577405

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

second_bin.67	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Dysosmobacter; s_
second_bin.83	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Dysosmobacter; s_Dysosmobacter sp910588005
group2_bin.25	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Eisenbergiella; s_Eisenbergiella porci
group1_bin.63	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Peptostreptococcales; f_Anaerovoracaceae; g_Emergencia; s_Emergencia sp009936045
second_bin.43	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Peptostreptococcales; f_Anaerovoracaceae; g_Emergencia; s_Emergencia sp009935805
group0_bin.87	k_Bacteria; p_Proteobacteria; c_Alphaproteobacteria; o_Rs-D84; f_Rs-D84; g_Enterousia; s_Enterousia sp013316395
group2_bin.123	k_Bacteria; p_Proteobacteria; c_Gammaproteobacteria; o_Enterobacterales; f_Enterobacteriaceae; g_Escherichia; s_Escherichia coli
group0_bin.11	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Eubacterium_F; s_Eubacterium_F sp910584065
group1_bin.111	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Eubacterium_F; s_Eubacterium_F sp910589595
group1_bin.148	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Eubacterium_F; s_Eubacterium_F sp910584875
group2_bin.46	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Eubacterium_F; s_

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group2_bin.94	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Eubacterium_G; s_Eubacterium_G sp910574895
group0_bin.113	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Eubacterium_J; s_Eubacterium_J sp910584595
group1_bin.61	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Eubacterium_J; s_
group2_bin.81	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Eubacterium_J; s_
second_bin.14	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Eubacterium_J; s_
second_bin.90	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Eubacterium_J; s_
group0_bin.52	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Acutalibacteraceae; g_Eubacterium_R; s_
group2_bin.120	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Acutalibacteraceae; g_Eubacterium_R; s_Eubacterium_R sp910577195
group0_bin.36	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Faecousia; s_
second_bin.57	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Faecousia; s_Faecousia sp000435995
group1_bin.72	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Acutalibacteraceae; g_Fimenecus; s_
second_bin.40	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_CAG-314; g_Fimimonas; s_

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group0_bin.167	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Peptostreptococcales; f_Anaerovoracaceae; g_Fimisoma; s_Fimisoma sp910588185
second_bin.19	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Frisingicoccus; s_Frisingicoccus caecimuris
group0_bin.19	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_Borkfalkiaceae; g_Gallimonas; s_Gallimonas sp910585235
group1_bin.114	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_Borkfalkiaceae; g_Gallimonas; s_Gallimonas sp910578065
group1_bin.56	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_Borkfalkiaceae; g_Gallimonas; s_
group2_bin.56	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_Borkfalkiaceae; g_Gallimonas; s_
group2_bin.83	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_Borkfalkiaceae; g_Gallimonas; s_Gallimonas sp910586985
second_bin.11	k_Bacteria; p_Firmicutes; c_Bacilli; o_CAJFEE01; f_CAJFEE01; g_Harrysmithimonas; s_Harrysmithimonas galli
group1_bin.7	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_CAG-74; g_HGM11604; s_
group0_bin.145	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_HGM13010; s_
group2_bin.66	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_HGM13010; s_

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

second_bin.30	k_Bacteria; p_Firmicutes_G; c_SHA-98; o_DTUO25; f_HGM14224; g_HGM14224; s_HGM14224 sp900761905
second_bin.32	k_Bacteria; p_Firmicutes; c_Bacilli; o_Erysipelotrichales; f_Erysipelotrichaceae; g_Holdemania; s_Holdemania sp900120005
second_bin.9	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Kineothrix; s_Kineothrix sp000403275
group0_bin.51	k_Bacteria; p_Firmicutes; c_Bacilli; o_Lactobacillales; f_Lactobacillaceae; g_Lactobacillus; s_Lactobacillus intestinalis
group2_bin.112	k_Bacteria; p_Firmicutes; c_Bacilli; o_Lactobacillales; f_Lactobacillaceae; g_Lactobacillus; s_Lactobacillus johnsonii
second_bin.82	k_Bacteria; p_Firmicutes; c_Bacilli; o_Lactobacillales; f_Lactobacillaceae; g_Lactobacillus; s_Lactobacillus taiwanensis
second_bin.94	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Lawsonibacter; s_Lawsonibacter sp900066645
group2_bin.79	k_Bacteria; p_Firmicutes; c_Bacilli; o_Lactobacillales; f_Lactobacillaceae; g_Ligilactobacillus; s_Ligilactobacillus murinus
group2_bin.19	k_Bacteria; p_Firmicutes; c_Bacilli; o_Lactobacillales; f_Lactobacillaceae; g_Limosilactobacillus; s_Limosilactobacillus reuteri_D
group1_bin.163	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Marseille-P3106; s_
group1_bin.167	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Marseille-P3106; s_

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group2_bin.106	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Marseille-P3106; s_
group2_bin.169	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Marseille-P3106; s_Marseille-P3106 sp910587345
group2_bin.175	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Marseille-P3106; s_
group2_bin.34	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Marseille-P3106; s_Marseille-P3106 sp910577415
second_bin.29	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Marseille-P3106; s_
second_bin.45	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Marseille-P3106; s_
second_bin.60	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Marseille-P3106; s_
second_bin.69	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Marseille-P3106; s_
group2_bin.113	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Mediterraneibacter; s_Mediterraneibacter torques
group0_bin.168	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Merdisoma; s_
group2_bin.31	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Merdisoma; s_
second_bin.97	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Merdisoma; s_Merdisoma sp009917555

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group1_bin.91	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_CAG-552; g_MGBC100320; s_MGBC100320 sp910577945
group2_bin.138	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_CAG-552; g_MGBC100320; s_
group2_bin.32	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_CAG-552; g_MGBC100320; s_MGBC100320 sp910588305
group2_bin.84	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_CAG-552; g_MGBC100320; s_MGBC100320 sp910577995
second_bin.55	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_CAG-552; g_MGBC113161; s_MGBC113161 sp910579245
group1_bin.35	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_CAG-274; g_MGBC114079; s_
second_bin.97	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Merdisoma; s_Merdisoma sp009917555
second_bin.105	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_MGBC116941; g_MGBC116941; s_MGBC116941 sp910585035
group0_bin.174	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_MGBC120314; s_
second_bin.49	k_Bacteria; p_Cyanobacteria; c_Vampirovibrionia; o_Gastranaerophilales; f_RUG14156; g_MGBC122484; s_MGBC122484 sp910586855

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

second_bin.6	k_Bacteria; p_Cyanobacteria; c_Vampirovibrionia; o_Gastranaerophilales; f_RUG14156; g_MGBC122484; s_MGBC122484 sp910578465
group0_bin.50	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_MGBC124762; s_
group1_bin.24	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_MGBC124762; s_
group1_bin.37	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_MGBC124762; s_MGBC124762 sp910584395
second_bin.71	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_MGBC124762; s_
group0_bin.116	k_Bacteria; p_Firmicutes; c_Bacilli; o_RF39; f_UBA660; g_MGBC162267; s_
group0_bin.111	k_Bacteria; p_Actinobacteriota; c_Coriobacteriia; o_Coriobacteriales; f_Eggerthellaceae; g_MGBC163016; s_MGBC163016 sp910588945
second_bin.16	k_Bacteria; p_Actinobacteriota; c_Coriobacteriia; o_Coriobacteriales; f_Eggerthellaceae; g_MGBC163016; s_MGBC163016 sp910588205
group0_bin.15	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Monoglobales; f_Monoglobaceae; g_Monoglobus; s_Monoglobus pectinilyticus
group2_bin.67	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Monoglobales; f_Monoglobaceae; g_Monoglobus; s_
group1_bin.66	k_Bacteria; p_Deferribacterota; c_Deferribacteres; o_Deferribacterales; f_Mucispirillaceae; g_Mucispirillum; s_Mucispirillum schaedleri

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group1_bin.1	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_Muribaculum; s_Muribaculum intestinale
group1_bin.53	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_Muribaculum; s_Muribaculum gordoncarteri
group1_bin.77	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_Muribaculum; s_Muribaculum sp910587665
group2_bin.92	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_Muribaculum; s_Muribaculum sp002473395
group0_bin.150	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Marinifilaceae; g_Odoribacter; s_Odoribacter laneus
group1_bin.23	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Marinifilaceae; g_Odoribacter; s_Odoribacter sp910578105
group1_bin.125	k_Bacteria; p_Firmicutes; c_Bacilli; o_RF39; f_UBA660; g_Onthocola_B; s_
group2_bin.76	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Tannerellaceae; g_Parabacteroides; s_Parabacteroides goldsteinii
group1_bin.67	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Paralachnospira_A; s_Paralachnospira_A sp910588405
second_bin.87	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Paralachnospira_A; s_
group0_bin.126	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_Paramuribaculum; s_Paramuribaculum sp001689535

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group2_bin.142	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_Paramuribaculum; s_Paramuribaculum sp910579675
group2_bin.41	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_Paramuribaculum; s_Paramuribaculum intestinale
second_bin.51	k_Bacteria; p_Firmicutes; c_Bacilli; o_Acholeplasmatales; f_Anaeroplasmataceae; g_Pelethenecus; s_Pelethenecus sp900540705
group0_bin.102	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Pelethomonas; s_
group0_bin.144	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Pelethomonas; s_
group2_bin.63	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Pelethomonas; s_
second_bin.10	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_Pelethomonas; s_Pelethomonas sp910587785
group2_bin.129	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Bacteroidaceae; g_Prevotella; s_Prevotella sp013166515
group1_bin.6	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Butyricicoccaceae; g_Pseudobutyricicoccus; s_Pseudobutyricicoccus sp910579575
group0_bin.125	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Peptostreptococcales; f_Anaerovoracaceae; g_QAKD01; s_
group0_bin.146	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Peptostreptococcales; f_Anaerovoracaceae; g_QAKD01; s_QAKD01 sp003343965

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

second_bin.86	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_QAMM01; s_
second_bin.75	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Christensenellales; f_Christensenellaceae; g_QANA01; s_QANA01 sp910588425
group1_bin.136	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Butyricicoccaceae; g_QXXE01; s_
group2_bin.170	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Butyricicoccaceae; g_QXXE01; s_QXXE01 sp910586575
group2_bin.61	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_RGIG4057; s_
group2_bin.98	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_TANB77; f_CAG-508; g_RGIG8482; s_
group1_bin.92	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_RGIG8607; s_RGIG8607 sp910577045
group1_bin.46	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_RGIG8773; s_RGIG8773 sp910584715
group0_bin.130	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Peptostreptococcales; f_Peptostreptococcaceae; g_Romboutsia; s_Romboutsia ilealis
group0_bin.142	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_Ruminiclostridium_E; s_
group1_bin.85	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_Ruminiclostridium_E; s_Ruminiclostridium_E sp910585505
group2_bin.149	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_Ruminiclostridium_E; s_

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group1_bin.12	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_Ruminococcus_D; s_
group1_bin.65	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Acutalibacteraceae; g_Ruminococcus_E; s_Ruminococcus_E sp002493635
group0_bin.44	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_Ruthenibacterium; s_Ruthenibacterium lactatiformans
group0_bin.82	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_Ruthenibacterium; s_
group2_bin.52	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_Ruthenibacterium; s_
group1_bin.29	k_Bacteria; p_Proteobacteria; c_Alphaproteobacteria; o_RF32; f_CAG-239; g_Scatacola_A; s_Scatacola_A faecigallinarum
group0_bin.100	k_Bacteria; p_Proteobacteria; c_Alphaproteobacteria; o_RF32; f_CAG-239; g_Scatocola; s_Scatocola sp910577205
group1_bin.182	k_Bacteria; p_Cyanobacteria; c_Vampirovibrionia; o_Gastranaerophilales; f_Gastranaerophilaceae; g_Scatousia; s_Scatousia sp000431315
group0_bin.33	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_TANB77; f_CAG-508; g_Scatovivens; s_
group0_bin.16	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Schaedlerella; s_
group1_bin.155	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Schaedlerella; s_Schaedlerella sp009911175
second_bin.47	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Schaedlerella; s_

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

second_bin.89	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Schaedlerella; s_Schaedlerella sp009774255
group1_bin.142	k_Bacteria; p_Firmicutes; c_Bacilli; o_RF39; f_UBA660; g_Scybalousia; s_
group0_bin.1	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Sporofaciens; s_Sporofaciens sp009774295
group1_bin.109	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Sporofaciens; s_Sporofaciens sp910574885
group1_bin.162	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Sporofaciens; s_Sporofaciens sp910584575
group2_bin.29	k_Bacteria; p_Cyanobacteria; c_Vampirovibrionia; o_Gastranaerophilales; f_Gastranaerophilaceae; g_Stercorousia; s_Stercorousia sp001917115
group0_bin.143	k_Bacteria; p_Firmicutes; c_Bacilli; o_Haloplasmatales_A; f_Turicibacteraceae; g_Turicibacter; s_Turicibacter sp002311155
group1_bin.31	k_Bacteria; p_Proteobacteria; c_Gammaproteobacteria; o_Burkholderiales; f_Burkholderiaceae; g_Turicimonas; s_Turicimonas muris
group0_bin.8	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA11957; s_
group0_bin.160	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_UBA1394; s_UBA1394 sp900066845
group0_bin.27	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_UBA1394; s_UBA1394 sp900538575

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group0_bin.45	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_UBA1394; s_UBA1394 sp910588145
second_bin.91	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_UBA1405; s_
group1_bin.57	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_UBA2658; s_
group2_bin.132	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA2882; s_UBA2882 sp002362385
group0_bin.117	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3282; s_UBA3282 sp002492525
group1_bin.102	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3282; s_UBA3282 sp002492725
group1_bin.146	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3282; s_UBA3282 sp002490705
group1_bin.153	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3282; s_
group1_bin.177	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3282; s_UBA3282 sp009774215
group1_bin.50	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3282; s_UBA3282 sp009774655
group2_bin.140	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3282; s_
group2_bin.161	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3282; s_
group2_bin.39	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3282; s_
group2_bin.54	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3282; s_UBA3282 sp910578715

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group2_bin.82	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3282; s_
second_bin.80	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3282; s_UBA3282 sp002491905
second_bin.95	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3282; s_
group1_bin.116	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3402; s_UBA3402 sp910588865
group1_bin.17	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3402; s_UBA3402 sp910588505
group2_bin.157	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3402; s_
second_bin.44	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3402; s_
second_bin.72	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3402; s_
second_bin.78	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA3402; s_
group0_bin.119	k_Bacteria; p_Firmicutes; c_Bacilli; o_RF39; f_UBA660; g_UBA3789; s_
group1_bin.82	k_Bacteria; p_Firmicutes; c_Bacilli; o_RF39; f_UBA660; g_UBA3789; s_
group0_bin.73	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA7050; s_UBA7050 sp002493905
group1_bin.107	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA7109; s_
group2_bin.72	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UBA7109; s_

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group0_bin.121	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_UBA7173; s_UBA7173 sp001689685
group0_bin.54	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_UBA7173; s_UBA7173 sp002491305
group1_bin.60	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_UBA7173; s_UBA7173 sp910579105
group2_bin.107	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_UBA7173; s_UBA7173 sp013316675
second_bin.102	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_UBA7173; s_UBA7173 sp001689485
second_bin.33	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_UBA7173; s_UBA7173 sp013316535
second_bin.84	k_Bacteria; p_Bacteroidota; c_Bacteroidia; o_Bacteroidales; f_Muribaculaceae; g_UBA7173; s_
group0_bin.164	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_UBA7177; s_
group1_bin.135	k_Bacteria; p_Firmicutes_B; c_Peptococcia; o_Peptococcales; f_Peptococcaceae; g_UBA7185; s_UBA7185 sp002491065
second_bin.26	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Ruminococcaceae; g_UBA866; s_UBA866 sp900543295
group0_bin.123	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Acutalibacteraceae; g_UMGS1071; s_UMGS1071 sp900541905
second_bin.64	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UMGS1370; s_UMGS1370 sp910577645
second_bin.25	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_UMGS1375; s_

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

group0_bin.108	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_TANB77; f_CAG-508; g_UMGS1994; s_
group1_bin.150	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Ventrimonas; s_Ventrimonas sp910575255
group1_bin.42	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Ventrimonas; s_Ventrimonas sp910586205
second_bin.96	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_Ventrimonas; s_Ventrimonas sp910586255
group0_bin.67	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_VSOB01; s_VSOB01 sp910585345
group1_bin.81	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_Lachnospiraceae; g_VSOB01; s_VSOB01 sp009774625
group1_bin.55	k_Bacteria; p_Cyanobacteria; c_Vampirovibrionia; o_Gastranaerophilales; f_Gastranaerophilaceae; g_Zag111; s_
group0_bin.97	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Lachnospirales; f_UBA1390; g_; s_
group2_bin.36	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_TANB77; f_CAG-508; g_; s_
group2_bin.51	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_TANB77; f_CAG-508; g_; s_
second_bin.1	k_Bacteria; p_Firmicutes; c_Bacilli; o_RF39; f_UBA660; g_; s_
second_bin.37	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_JAAYOS01; g_; s_
second_bin.56	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_; s_

Table A1.1. Taxonomic annotations of the 324 MAGs based on GTDB-Tk database (continued)

second_bin.68	k_Bacteria; p_Firmicutes; c_Bacilli; o_RF39; f_UBA660; g_; s_
second_bin.70	k_Bacteria; p_Firmicutes_A; c_Clostridia; o_Oscillospirales; f_Oscillospiraceae; g_; s_

9 CURRICULUM VITAE