Role of gut bacteria in domestication of mice

Biswa, Bhim Bahadur

Doctor of Philosophy

Department of Genetics

School of Life Science

The Graduate University for Advanced Studies, SOKENDAI

September 2024

Acknowledgments

I want to express my profound gratitude to all those who supported and guided me throughout my PhD journey. First and foremost, I extend my heartfelt thanks to my supervisor, Prof. Tsuyoshi Koide. His ongoing support and expert guidance have been instrumental in my success. Without his resilience, inspiration, encouragement, dedication, and advice, this thesis would not have been possible. His deep knowledge and guidance were always a beacon whenever I faced challenges in my research. His keen interest and support pushed me to excel. Additionally, his kindness and empathy made my time in Japan both easy and enjoyable. I am eternally grateful to him for all the support and encouragement he has provided.

I am deeply grateful to my progress committee members—Prof. Jun Kitano, Prof. Takuji Iwasato, Prof. Hironori Niki, and Prof. Hiroshi Mori—for their insightful comments, encouragement, and guidance throughout my research. Their feedback was instrumental in motivating me to continually improve my work. Additionally, I appreciate the collaboration with Prof. Ken Kurokawa and Prof. Hiroshi Mori for their valuable suggestions in designing the initial experiments and their ongoing input. I am also thankful to Prof. Atsushi Toyoda for performing the essential shotgun metagenomic sequencing that was crucial for my research.

My deepest gratitude extends to both the current and past members of the Koide lab. Their invaluable assistance, support, and guidance have been crucial to my journey, and I could not have achieved my goals without them. I am particularly thankful to Prof. Keiko Takanami, whose training in behavioural experiments and help in adjusting to life in Japan were foundational to my successes. I extend heartfelt appreciation to Mr. Hiromichi Nagayama, who taught me behavioural experiments and introduced me to R programming. Dr. Naoko Ueda's continuous encouragement and advice were vital in honing my experimental skills. Special acknowledgment goes to Dr. Lalitha Devi, whose support was indispensable throughout the ups and downs of my PhD journey. Her guidance significantly smoothed both my personal and professional experiences in Japan. Additionally, I am grateful to her for her guidance in numerous wet lab experiments.

Further, I am grateful to Dr. Kazumichi Fujiwara for his assistance with bioinformatic analyses in my thesis and for always being available to clarify my doubts. I also appreciate Mr. Abdullah Syafiq Edyanto and Ms. Bharathi Venkatachalam for contributing to a lively

lab environment and for their readiness to help whenever needed. I would like to thank Ms. Motoko Nihei for analysing all my tameness test video data and for maintaining the necessary equipment and reagents in the lab. My research greatly benefited from her diligent work. My gratitude extends to Ms. Akiko Tsuchiya for maintaining the WHS mice line and managing the breeding essential for my experiments. I am also thankful to Mr. Yuji Imai, who has supported me in every possible way, both personally and professionally. His presence has been one of the pillars of my fruitful life in Japan. I thank Mr. Daisuke Takahashi for maintaining the Animal House at NIG and providing me with mice for my experiments. I am also thankful to Mr. Kyouhei Kurosawa for conducting the tameness tests in the pyruvate injection experiment.

A special thanks to the lab secretary, Ms. Akiko Murofushi, who helped me settle in and was always welcoming and ready to assist in every small matter. Additionally, I would like to thank Mr. Shugetsu Kawachi, my "Tutor" assigned by SOKENDAI. He helped me in every possible way to settle in Japan, and without his guidance, I wouldn't have had such an easy transition to life in Japan.

I am extremely thankful to the NIG RA Fellowship, Monbukagakusho Honors Scholarship by JASSO, and SOKENDAI Special Researcher Fellowship for supporting me during my PhD study. Additionally, I owe a great debt of gratitude to both the past and current members of the Academic Services Division at the National Institute of Genetics. Their unwavering support and guidance have been crucial to my success, impacting both my personal and professional life profoundly.

I want to express my sincere thanks to my friends at NIG, especially Mr. Hassan Sibroe Abdulla Daanaa, Dr. Ramasamy Kandasamy, Dr. Piu Banerjee, Dr. Sharlyn Mae Buendia Chua, and Dr. Luis Antonio Cruz Diaz, for their unwavering support and kindness. Last but not least, special thanks to my parents, sisters, and my extended family for their enduring love, support, and motivation.

Abstract

Domestication, primarily driven by selective breeding for traits beneficial to humans, profoundly impacts animal behaviour, with tameness being a primary focus. Traditionally, studies on domestication have concentrated on genetic influences; however, my research shifts this focus towards the role of gut bacteria in shaping behaviour in mice. I analysed the gut microbiota of mice specifically bred for active tameness, examining faecal samples from 80 mice—40 selected for tameness and 40 controls—using shotgun metagenomic analysis. My findings reveal that although selective breeding for tameness does not significantly alter the taxonomic or functional diversity of the gut microbiota, there was a notable increase in the abundance of *L. reuteri* within the selected groups, accompanied by elevated pyruvate and oxytocin levels in their plasma.

To explore the functional implications of these findings, I administered pyruvate to non-selected mice, which helped ameliorate the reduction in tameness typically induced by repeated injection stress. Additionally, I isolated strains of *L. reuteri* capable of secreting extracellular pyruvate and administered one of these strains to non-tame mice via drinking water. This intervention led to a significant increase in tameness behaviour and an increase in serum oxytocin levels in these mice.

Further leveraging the metagenomic data, I compiled a comprehensive collection of 374 high-quality metagenome-assembled genomes (MAGs) of bacteria across 11 phyla, enriching our understanding of the mouse gut microbiome. This collection includes 27 novel bacterial MAGs previously unreported in the mouse gut.

My study illuminates the significant role of microbiota in the domestication process, providing new insights into the complex interplay between host genetics, behaviour, and the gut microbiota. This research opens up novel perspectives on how microbial communities within the gut can influence animal behaviour, extending our understanding of domestication beyond genetic factors to include microbial contributions to host behavioural phenotypes.

Table of Contents

List of Abbreviations

WHS	Wild-derived heterogeneous stock
QTL	Quantitative trait loci
MAG	Metagenomes Assembled Genome
SPF	Specific pathogen-free
ELISA	Enzyme-linked immunosorbent assay
DNA	Deoxyribonucleic acid
CPM	Copies per Million
FDR	False Discovery Rate
PCoA	principal coordinate analysis
ANOVA	Analysis of Variance
SCFA	short-chain fatty acids
GABA	gamma-aminobutyric acid
IP	Intraperitoneal
L. reuteri	*Limosilactobacillus reuteri*
L. helveticus	*Lactobacillus helveticus*
LRIM	*L. reuteri* Isolation Media
MRS media	De Man–Rogosa–Sharpe media
PCR	Polymerase chain reaction
JCM	Japan Collection of Microorganisms
GAM	Gifu Anaerobic Broth
OD	Optical density
CFU	Colony forming unit
PBS	Phosphate-buffered saline
qRT-PCR	Quantitative Real-time PCR
FMT	faecal microbiota transplantation
PVN	paraventricular nucleus

Chapter 1: General Introduction

For thousands of years, the domestication of animals has played a crucial role in the development of human societies, providing labour, companionship, and food resources that have significantly advanced civilization (1). The first study of animal domestication began with the pioneering work of the Belyaev group, who established the first scientific breeding program of silver foxes (*Vulpes vulpes*) (2). This initiative marked the beginning of a robust field of research aimed at deciphering the genetic and behavioural transformations involved in domestication. The process of domestication has profound effects on the morphology, physiology, and behaviour of animals, shaping their evolution in ways that enhance their utility and compatibility with human needs.

1.1 Genetic mechanism of domestication

The genetic basis of animal domestication involves complex changes at the molecular level that influence an array of behavioural, physiological, and morphological traits. Understanding these genetic mechanisms is crucial for decoding how animals have been shaped through selective breeding to meet human needs. Recent advances in genomic technology have spurred extensive research into the genetic underpinnings of domestication. Through quantitative trait loci (QTL) mapping, Matsumoto et al. have identified specific genomic regions, like the ATR1 and ATR2 loci on chromosome 11 in mice, that are associated with tameness behaviour (3). This locus contains many genes which has potential to have role in domestication process. Similarly, Kukekova et al. have pinpointed genes such as *SorCS1*, which plays a role in protein trafficking and sorting, is under selection pressure when foxes were selected for tame or aggressive behaviour (4). Moreover, research by Dou et al. has uncovered mutations, such as in the *RRM1* (Ribonucleotide Reductase Catalytic Subunit M1) gene in goats, which correlate with increased sociability and reduced anxiety, demonstrating the complex genetic orchestration behind domestication traits (5). Moreover, the integration of genomics with other biological data is expanding our understanding of how genetic changes interact with hormonal and neural pathways. For instance, the interplay between genetic predispositions and neurochemical pathways involving hormones like oxytocin can elucidate how genetic variations contribute to behavioural changes observed in domesticated animals. In conclusion, the genetic mechanisms of domestication encompass a diverse range of genomic modifications from point mutations to regulatory changes affecting gene expression. These genetic changes not only influence individual traits but also integrate into

complex biological networks, shaping the behaviour and physiology of domesticated species across generations.

1.2 Role of oxytocin in animal domestication

Oxytocin, often referred to as the "love hormone," plays a crucial role in animal domestication due to its significant effects on social bonding, trust, and reduction of fear responses towards humans. This neuropeptide is not only fundamental to maternal bonding and social interactions within species but also influences the inter-species relationships that are key to the domestication process. Oxytocin facilitates the establishment of trust and attachment between animals and humans, characteristics that are selectively enhanced in domesticated animals (6, 7). Recent studies on wild-derived heterogeneous stock (WHS) mice have shown that those WHS mice selectively bred for tameness exhibit an increased expression of oxytocin receptors in the hippocampus (8). Additionally, tameness selection leads to increases in sociability in WHS mice (9). Moreover, oxytocin's impact extends beyond behaviour to influence reproductive functions and stress responses, which are essential for the management and predictability of domesticated animals. Animals with heightened oxytocin-mediated pathways tend to have more predictable and less stress-induced behaviours, which are traits highly valued in domesticated species (10) . The calming effect of oxytocin also plays a part in facilitating handling and care activities, which are frequent in domesticated environments, thereby improving the welfare of the animals and easing the management burden on human caretakers (11). In conclusion, the role of oxytocin in animal domestication is multifaceted, influencing emotional, behavioural, and physiological traits that are essential for the successful integration of animals into human-dominated environments.

1.3 Environmental influencers of domestication process

The environment plays a crucial role in the domestication of animals, complementing genetic factors and influencing the behavioural and physiological adaptations of domesticated species. The process of domestication is not only a result of selective breeding for specific genetic traits but also involves adaptation to human-controlled environments, which includes exposure to unique sets of stressors, resources, and social structures. One of the primary environmental factors influencing domestication is the proximity and interaction with humans. Frequent and sustained interactions with humans can lead to reduced fear and increased tameness in animals. This behavioural adaptation is critical for animals living in close

quarters with humans and facilitates easier management and cooperation in agricultural or household settings. Studies have shown that animals repeatedly exposed to gentle handling by humans often exhibit calmer dispositions and reduced stress responses compared to their wild counterparts (12-14). Similarly, temperature and climate are also critical environmental factors that impact domestication. Animals bred in different climates must adapt to a variety of environmental stresses, from cold temperatures requiring changes in fur density and metabolism to hot climates necessitating adaptations in water conservation and heat tolerance. The ability of a species to adapt to these conditions can significantly influence its suitability for domestication and the specific roles it can fulfil in human society (15-17). Moreover, the type of housing and the density in which animals are kept can influence their health and social structures, affecting everything from disease transmission to hierarchical behaviours. For example, high-density housing can lead to increased stress and aggression among individuals, necessitating selective breeding for more docile and disease-resistant individuals that can thrive under such conditions (18).

In conclusion, the environmental influencers of animal domestication are diverse and impact various aspects of animal life, from behaviour and social structure to physiological adaptations. Understanding these influences is essential for managing domesticated species effectively and ethically, ensuring they are well-adapted to the conditions under human care. As domestication continues to evolve with changes in human society and technology, the environmental factors shaping this process remain a critical area of study.

1.4 Gut microbiome role in animal domestication

Numerous research findings suggest that the gut microbiota impacts host behaviour, brain development, and cognitive functions. Metabolites from these microbiotas can modify the host's metabolic processes or affect their immune response (19, 20). Various investigations have examined the link between gut microbiota and the domestication of animals, including studies on mice (21), horses (22, 23), buffalo (24), turkeys (25), chicken (26), and a variety of species together (27). These investigations have identified a range of bacterial species that potentially contribute to the domestication process. Nonetheless, such studies are often subject to multiple challenges that may affect their outcomes. Contamination of gut microbiota due to non-sterile living conditions is a frequent issue, which can obscure the natural microbial profiles of study subjects. Additionally, genetic differences between test and control groups, especially when they do not share common ancestors, can lead to variability that complicates the interpretation of results. Variations in the diets and water

sources provided to different groups of animals in a study can also introduce significant discrepancies in microbiota composition. Moreover, administration of antibiotics, which can dramatically alter microbial communities, or cross-contamination from other species further complicate these studies. To address these challenges, it is crucial to conduct studies in controlled environments where variables can be carefully managed. This approach includes maintaining a consistent and sterile environment, using standardized diets, and controlling for genetic background by using well-documented lineages. An exemplary model for such research practices is the use of WHS mice, specifically bred for tameness behaviours (3, 28) . These mice provide a stable and reliable foundation for studies, allowing researchers to more accurately discern the effects of gut microbiota on domestication without the confounding factors present in less controlled settings.

This work focuses on finding out the gut microbiome influence on animal domestication process using WHS mice as the model organism.

1.5 Organization and purpose of the dissertation

In this study, to analyse the gut microbiota of WHS mice, I performed metagenomic analysis on genomic DNA extracted from the faecal samples of both the selected and non-selected groups. Differences in the gut microbiome taxonomic composition were then analysed. I also identified the functional potential of the gut microbiome. Furthermore, I attempted to modulate tameness behaviour in the mice through two intervention strategies: firstly, by administering specific metabolites, and secondly, by introducing certain bacteria into their drinking water.

This dissertation consists of five chapters. The second focusing on investigations of tameness behaviour and gut microbiota of WHS mice. This involves analysis of taxonomic and functional diversity of gut microbiome, abundance analysis of bacteria and generation of metagenome assembled genomes (MAGs). The third chapter focus on finding out the gut metabolite which might influence tameness behaviour and then external administration of this metabolite for tameness behaviour change. The fourth chapter involve trying to influence tameness behaviour by manipulating the abundance of one bacterium or changing whole gut microbiome community by cross-fostering. Fifth and final chapter is conclusion of my study and a discussion about the possible mechanism of influence of a single bacterium on tameness behaviour.

Chapter 2: Selective breeding for tameness behaviour leads to increase in *L. reuteri* abundance

2.1 Introduction

Selective breeding has long been employed to enhance desirable traits in domestic animals, one of the most sought-after being tameness. This behavioural modification not only facilitates easier management and interaction with humans but also significantly impacts the physiological and microbiological profiles of the animals. Tameness is a behavioural phenotype that is essential for animals to effectively adapt to human interaction, and it manifests in two forms: active tameness, which is the propensity to approach humans, and passive tameness, characterized by a reduced tendency to avoid humans. To enhance active tameness, mice that showed a greater inclination to approach humans during active tameness tests—assessments specifically designed to measure this trait—were selectively bred (29). The initial breeding stock, known as wild-derived heterogeneous stock (WHS) mice, was established by crossbreeding eight different wild inbred strains from three subspecies of *Mus musculus* (3). This founder group was maintained with sixteen pairs per generation to preserve a high level of genetic diversity (30). From this original group, four separate lineages of WHS mice were developed: two were designated for tameness selection (S1 and S2) and two served as control groups (C1 and C2), as depicted in Figure 1.

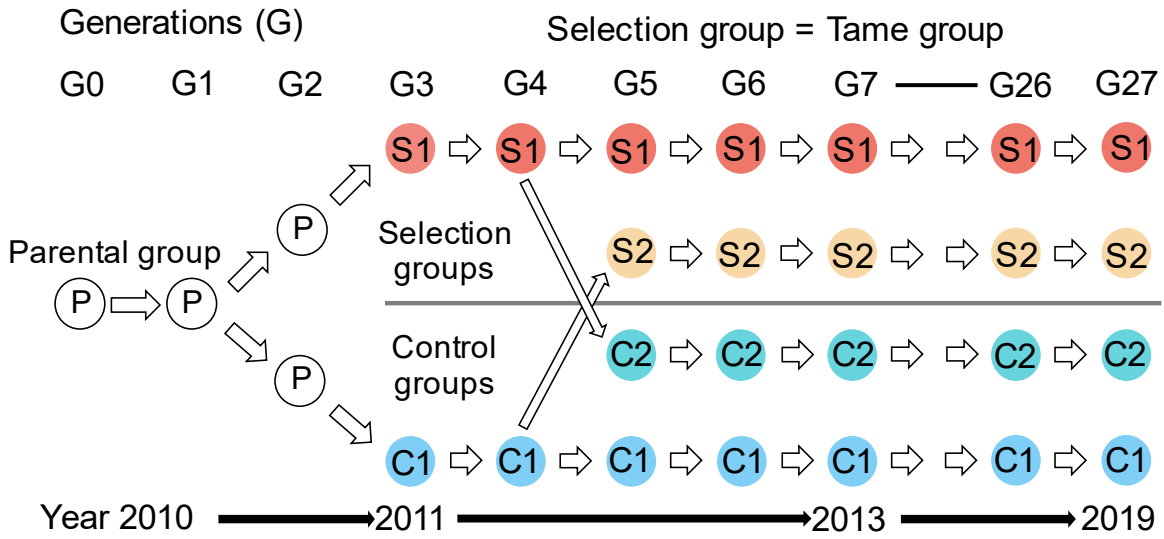


Figure 1. Scheme of wild heterogenous stock mice generation

The division of the common founder stock into S1 and C1 occurred in the third generation, with S1 subsequently bred for increased active tameness. In the fifth generation, a further separation produced another selected group, S2, from C1, and another control group,

C2, from S1. From that point, S2 was also selectively bred to enhance active tameness, while C2 continued as a non-selected group. All groups of WHS mice were housed in the same facility but were isolated from mice of other groups to prevent cross-contamination. Moreover, all mice, including those from the original wild inbred strains, were raised under specific pathogen-free (SPF) conditions with radiation-sterilized food and ultrafiltration-sterilized water, ensuring that any gut microbiota present originated solely from the common founder mice.

The objective of this study is to thoroughly examine the gut microbiome of mice selectively bred for tameness behaviour. This comprehensive analysis includes evaluating the nuances of tameness behaviour, exploring both the taxonomic and functional diversity of the gut microbiota, and conducting association analyses to determine the relationship between the abundance of specific gut bacteria and tameness behaviour. By integrating these approaches, the study aims to elucidate how selective breeding for tameness influences the microbial communities within these animals and, conversely, how these microbial variations might impact tameness behaviour.

2.2 Materials and Methods

2.2.1 Experimental Conditions

In my experiments, mice were bred and maintained under SPF conditions at the NIG animal facility. The equipment used for animal care, including standard-sized transparent polycarbonate cages, cage tops, paper bedding materials, and other necessary items, was sterilized through autoclaving. The mice were fed a standard diet (CE-2, CLEA Japan, Inc., Tokyo, Japan), which had been radiation-sterilized at 15 KGy. They had continuous access to food and to ultrafiltration sterilized water, supplemented with 3 ppm chlorine. The mice were kept in a temperature-regulated room at 23 ± 2°C with humidity maintained at 50 ± 10% and a 12-hour light-dark cycle, with lights turning on at 6:00 AM. Weaning occurred at about 3–4 weeks of age, after which the mice were housed with their same-sex littermates until the tameness test was conducted. During cage changes and behavioural tests, each mouse was gently handled using tweezers with silicone tubing on the tips to minimize stress. To maintain a sterile environment, all personnel wore polyester suits, non-woven fabric masks, and double gloves (cloth gloves beneath latex gloves). To prevent microbiome cross-contamination, cages of different mouse groups were kept apart, ensuring that each group's microbial

environment remained distinct and uncontaminated. This experiment was conducted using mice from the 27th generation of the WHS.

2.2.2 Selective breeding of WHS

The WHS mice were created by crossing eight wild strains—BFM/2Ms, PGN2/Ms, HMI/Ms, NJL/Ms, BLG2/Ms, KJR/Ms, CHD/Ms, and MSM/Ms—sourced from different countries to enhance genetic diversity. The founder stock, maintained with 16 breeding pairs, was divided into two groups in the third generation, forming S1 and C1. From that point, S1 was selectively bred for increased active tameness. In the fifth generation, another selected group, S2, was derived from C1, while another non-selected group, C2, was derived from S1. Subsequently, S2 continued to be selectively bred for higher active tameness, while C2 was maintained as the non-selected control group. For mating, mice were chosen from the five female and five male offspring of each pair based on the highest contact scores in the active tameness test. If multiple mice had the same highest contact score, the one with the higher heading score in the same test was selected.

2.2.3 Tameness test

At six weeks of age, the mice underwent the tameness test, following the protocol described in a previous study (29). The test took place during the light period, between 14:00 and 17:00, in an open-field arena constructed from grey polyvinyl chloride, measuring 40 × 40 × 40 cm (O'Hara & Co. Ltd., Tokyo, Japan). The arena floor was illuminated to approximately 100 lux at its centre (28, 31). The tameness assessment consisted of a three-minute protocol, divided into one minute each for the active and passive tameness evaluations, followed by three trials of the stay-on-hand test (Figure 2.1).

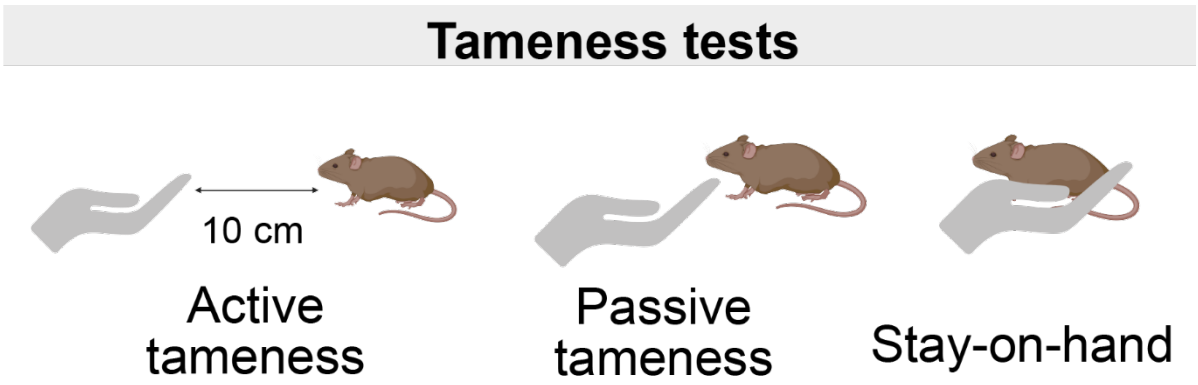


Figure 2.1. Tameness test parameters

In the active tameness test, the experimenter placed their hand 10 cm away from the test mouse for one minute, with fingers moving slightly and rhythmically to assess the mouse's

inclination to approach and interact with the human hand. In the passive tameness test, the experimenter attempted to touch the mouse with their hand, and the mouse's response was recorded for one minute to evaluate how long the mouse could tolerate being touched by a human hand. During the stay-on-hand test, each mouse was gently picked up by its tail using silicon-coated tweezers and placed on the experimenter's hand, where it was softly stroked until it attempted to leave the hand. All tests were video-recorded and analysed by another experienced researcher using the tanaMove software (zenodo.org/records/10829003) (32). The video analysis was conducted blind to the grouping of mice and scores were assigned to nine different behavioural parameters. A comprehensive description of the procedure is available in previous studies (28, 31). The behavioural data collected included nine parameters: heading, contacting, locomotion, and jumping in the active tameness test; heading, acceptance, locomotion, and jumping in the passive tameness test; and median staying time in the stay-on-hand test (Table 2.1).

Table 2.1 Explanation of Tameness Test Parameters

Active Tameness test	Heading	heading toward or approaching the hand, including standing and sniffing
	Contacting	the mouse voluntarily touches the hand excluding accidental contacts, like landing on the palm after jumping
	Locomotion	Moving around without approaching the hand
	Jumping	Jumping behaviour
Passive Tameness test	Heading	heading toward or approaching the hand, including standing and sniffing
	Accepting	mouse allows the experimenter to touch its body
	Locomotion	Moving around without approaching the hand
	Jumping	Jumping behaviour
Stay on hand test		Duration the mouse remained on the hand after being placed on the palm by its tail

2.2.3 Mice faeces collection

In all experiments, mouse faecal samples were collected immediately after the tameness test to ensure uniform timing between 17:00 and 18:00 on the same day. To maintain sterility and control environmental variables, the test mice were placed in autoclaved, sterile cages without any bedding material, allowing for natural defecation. The faeces were collected using sterile aluminium foil and placed into Eppendorf tubes, which were then promptly stored at -80°C.

2.2.4 Blood plasma collection and oxytocin ELISA

Blood plasma was collected when the mice were 11 weeks old. The collection process began by anesthetizing the mice with an intraperitoneal injection of pentobarbital (Tokyo Chemical Industry Co. Ltd., Japan) at a dose of 50 mg/kg body weight. Once the mice were anaesthetized, blood was drawn via cardiac puncture using a syringe pre-coated with heparin to prevent clotting. The blood samples were then transferred into heparinized Eppendorf tubes and left undisturbed at room temperature for 30 minutes to allow for the separation of blood components. Following this, the samples were centrifuged at 10,000 G for 10 minutes at 4°C. To measure oxytocin levels, I used an Oxytocin ELISA kit (Enzo Life Sciences, USA), adhering to the manufacturer's instructions. For baseline oxytocin measurements in WHS mice, I analysed 18 plasma samples from each group (9 males and 9 females), making a total of 72 samples. The assays were performed in duplicates (technical replicates) due to the well constraints of the ELISA plate.

2.2.5 Tameness test parameter clustering

The data analysis was conducted using R Ver. 4.3.2. The results from the tameness tests for all 80 samples underwent normalization using the "min-max" technique. Clustering was performed based on correlation distances, utilizing the Ward method via the "hclust" function. The statistical significance of the clusters was evaluated with the "pvclust" function from the R package pvclust (Ver. 2.2.0) (33), with a significance threshold set at 5% for each clustering analysis.

2.2.6 Metagenomic DNA isolation from mice faeces

In my revised approach to metagenomic DNA extraction, I adapted protocol #6 from Costea et al. (34), and used the QIAamp Fast DNA Stool Mini Kit (QIAGEN GmbH, Hilden, Germany). The procedure began with 180-220 mg of fecal samples being treated with 1 ml of lysis buffer, composed of 500 mM NaCl, 50 mM Tris-HCl (pH 8.0), 50 mM EDTA, and 4% sodium dodecyl sulfate. The mixture was homogenized for 5 minutes using a handheld homogenizer from Leda Trading Corp. Following homogenization, 10 µl of Proteinase K and sterile zirconia beads (1.0 mm, 20-30; BioSpec, Inc., USA) were added. The sample was then vortexed at maximum speed for 10 minutes using a Vortex-Genie 2 mixer (MO BIO Laboratories, Inc., USA) and incubated at 95°C for 15 minutes. After incubation, the sample was centrifuged at 16,000 x g for 5 minutes at 4°C, and the supernatant was collected into a new 2 mL tube. The remaining pellet underwent a second lysis round, resuspended in 300 µL of the lysis buffer, and processed as before. The supernatants from both lysis steps were

combined and treated with 260 µL of 10M ammonium acetate. This was followed by vortexing and a 5-minute incubation on ice. The mixture was then centrifuged at 16,000 x g for 10 minutes at 4°C, and the supernatant was equally divided into two 1.5 mL tubes, each containing 750 µL isopropanol. These tubes were incubated on ice for 30 minutes and centrifuged again at 16,000 x g for 15 minutes. The supernatant was discarded, and the pellet was washed with 0.5 mL of 70% ethanol and allowed to air-dry. The dried pellets were reconstituted in 100 µL of TE (Tris-EDTA) buffer, and the aliquots were combined. I then added 2 µL of DNase-free RNase (10 mg/mL) to the pooled DNA solution, which was incubated at 37°C for 15 minutes. After that, 10 µL of Proteinase K and 200 µL of buffer AL were added, followed by a 15-second vortexing and a 10-minute incubation at 70°C. Next, 200 µL of 96-100% ethanol was added to the lysate, which was vortexed and transferred to a QIAamp spin column for centrifugation for 1 minute. The filtrate was discarded, and the column was sequentially washed with 500 µL each of Buffer AW1 and AW2. Following a final centrifugation, the DNA was eluted with 100 µL of AE Buffer into a new 1.5 mL tube. The DNA quality was checked using a NanoVue Plus spectrophotometer, and precise quantification was performed with a Qubit 2.0 Fluorometer, following the manufacturer's guidelines. DNA integrity was verified by 0.8% agarose gel electrophoresis, and the samples were stored at -20°C for future use.

2.2.7 Shotgun metagenomic sequencing

Library preparation and sequencing were conducted at the Advanced Genomics Center of the National Institute of Genetics. In summary, genomic DNA was fragmented to an average size of 500 base pairs using a DNA shearing system (M220 focused ultrasonicator; Covaris Inc., MA, USA). A paired-end library was then constructed using the TruSeq DNA PCR-Free Library Prep kit (Illumina, CA, USA) and size-selected on an agarose gel with the Zymoclean Large Fragment DNA Recovery Kit (Zymo Research, CA, USA). The final library underwent paired-end sequencing on an Illumina NovaSeq 6000 sequencer, with a read length of 150 base pairs.

2.2.8 Quality control of metagenomic sequences

All sequencing reads were subjected to quality filtering using KneadData (Ver. 0.10.0) (https://github.com/biobakery/kneaddata), which eliminated low-quality sequences and any contamination from PhiX, the mouse genome (mouse_C57BL_6NJ.1), the human genome (hg37dec_v0.1.1), or the human transcriptome (human_hg38_refMrna.1). KneadData integrates several tools for this process: Trimmomatic (Ver. 0.33) (35) to remove adapters,

TRF(36) to eliminate repetitive sequences, and Bowtie2 (Ver. 2.4.5) (37) to filter out contaminants.

2.2.9 MAG generation and analysis

Unless otherwise stated, default settings were applied across various software. The quality-filtered sequences were assembled using metaSPAdes (k = 21, 33, 55) (Ver. 3.15.4) (38) for single assembly, or MEGAHIT (options: --k-min 21 --k-max 77) (Ver. 1.2.9) (39) for co-assembly. A range of binning strategies was employed. For the MEGAHIT assembly, binning was performed using MetaCoAG (Ver. 1.0) (40), MaxBin2 (Ver. 2.2.7) (41), and MetaBAT2 (Ver. 2.15) (42). DAStool (search_engine diamond) (Ver. 1.1.4) (43) was subsequently used to generate a nonredundant set of bins. For the metaSPAdes-generated assembly, I utilized. For the metaSPAdes-generated assembly, I utilized SemiBin (multi_easy_bin mode, Ver. 0.7.0) (44) and VAMB (multisplit mode, Ver. 3.0.8) (45). The resulting bins underwent contamination filtering with MDMcleaner (Ver. 0.8.3) (46). Bins free from chimeric sequences, as confirmed by GUNC (Ver. 1.0.5) (47), and those with MAG quality scores above 0.5 were retained for further analysis. These filtered bins were then assessed for contamination and completeness using CheckM (lineage_wf, Ver. 1.2.0) (48). To create a nonredundant bin set, the “dereplicate” command of dRep (Ver. 3.0.0) was employed with parameters (-comp 50, -con 10, -pa 0.90, -sa 0.95, -nc 0.30, -cm larger) (49). For taxonomic classification and species tree generation, I used GTDB-Tk (classify_wf, Ver. 2.1.1) with the “r207_v2” database (50). The phylogenetic tree for the MAGs was generated using FastTree (Ver. 2.1.11) (51) within GTDB-Tk using the “infer” command. The resulting tree was refined and edited using the Interactive Tree of Life (iToL) tool (52). To identify novel MAGs, I utilized the “compare” command in dRep (Ver. 3.0.0) with parameters (-pa 0.90 -sa 0.95 -nc 0.30 -cm larger) (49). This analysis was performed on a combined database, including the 1,573 mouse gut MAGs previously compiled by Kieser et al. (53) , along with the 374 MAGs generated in my study. MAGs were classified as novel if they did not cluster with any existing MAGs in this combined database. To quantify the relative abundance of each MAG, I employed CoverM (Ver. 0.6.1, https://github.com/wwood/CoverM), using minimap (54) to map reads to MAG database.

2.2.10 Taxonomic analysis

Taxonomic analysis of the gut microbiome was performed using Kraken 2 (Ver. 2.1.2) (55). The comprehensive NCBI/RefSeq database constructed by Wright et al. (NCBI RefSeq Complete V205, 1189 GB) (56) to ensure comprehensive species identification within the gut

samples. was employed to facilitate thorough species identification within the gut samples. Kraken 2 was run with default parameters, and a confidence score threshold of 0.15 was set to balance the accuracy of species identification with the proportion of sequences included in the analysis. To further refine the Kraken 2 results, Bayesian re-estimation was conducted using Braken2 (Ver. 2.6.2), which also utilized default settings (57).

2.2.11 Functional Analysis

For functional analysis in my study, I used HUMAnN3 (Ver. 3.0.1) (58) with the '--bypass-prescreen' option enabled. The filtered forward and reverse sequences from each sample were concatenated and mapped to the comprehensive ChocoPhlAn database (Ver. 201901b). Sequences that were not assigned during this step were further analysed using the UniRef90 database. To identify specific metabolic pathways, the MetaCyc database was employed (59). Copies per Million (CPM) data for each gene family and pathway were generated using the 'humann_renorm_table' command.

2.2.12 Abundance and Diversity Analysis

Diversity analysis and graph generation for this study were conducted using the MicroEco R package (60). The abundance data obtained from Braken2 were normalized by rarefying them to match the size of the smallest sample (61). To assess the statistical significance of alpha diversity, I used the Kruskal-Wallis Rank Sum test. Beta diversity was analysed using Bray-Curtis dissimilarity and supplemented by principal coordinate analysis (PCoA) on both taxonomic and functional data. For the correlation analysis, I performed a random forest classification using the 'randomForest' R package (62), incorporating mouse group data (C1, C2, S1, and S2) on the normalized Bracken2 bacterial abundance data. The p-values were adjusted using the Benjamini-Hochberg False Discovery Rate (FDR) method. Next, I normalized the tameness parameter scores using min-max normalization. These normalized tameness scores were then subjected to Pearson correlation analysis with the top 40 significant taxa identified in the random forest analysis, and the results were visualized using the 'pheatmap' R package (63) .

2.2.13 MaAsLin2 association analysis

For the association between taxonomic abundance and each mouse group, I used the Bracken2 abundance data (refer to Table S2). The data were initially normalized using Total Sum Scaling, followed by log transformation, before proceeding with the association analysis. For the functional pathway association analysis, no additional normalization or

transformation was required as HUMAnN3's output is already normalized. In both analyses, linear models were employed to compare the S1 and S2 groups with their respective controls.

2.2.14 Statistical analysis

All statistical analyses were conducted using GraphPad Prism (Ver. 10.1.2). The Shapiro-Wilk test was used to assess the normality of the dataset. For data following a normal distribution, a one-way ANOVA was applied, followed by Tukey's test for multiple comparisons. If the data did not follow a normal distribution, the Kruskal-Wallis test was used, followed by Dunn's test for multiple comparisons. A two-way ANOVA was performed to assess the impact of sex on tameness behaviour. For the statistical association analysis of plasma metabolites, taxonomic abundance, and functional abundance, MaAsLin2 (Ver. 1.12.0) (64) was used, employing a Benjamini-Hochberg q-value threshold of 0.25, in line with previous microbiome studies (65, 66). For abundance of MAGs was compared using a two tailed Mann–Whitney U test.

2.2.15 Availability of data

All 80 raw shotgun metagenomic sequencing datasets generated in this study are available in the NCBI under BioProject PRJDB15857 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJDB15857), with BioSample accession numbers from SAMD00614304 to SAMD00614383 and SRA accession numbers from DRR480456 to DRR480535. Non-redundant set of MAGs (374) including novel ones are available in zenodo database (https://doi.org/10.5281/zenodo.8289507).

2.2.16 Ethics approval

All experiments conducted in this study were performed in strict compliance with the guidelines and protocols approved by the Committee for Animal Care and Use at the National Institute of Genetics, under permit numbers R4-25 and R5-7.

2.3 Results and discussion

2.3.1 Active contacting selection pressure increases active tameness behaviour

On 27th generation of WHS mice, I assessed tameness-related behaviours in both the selected and non-selected groups of mice. I carried out evaluations for active tameness, passive tameness, and stay-on-hand tests on a total of 80 mice, comprising 10 males and 10 females from each group (Figure 2.2).

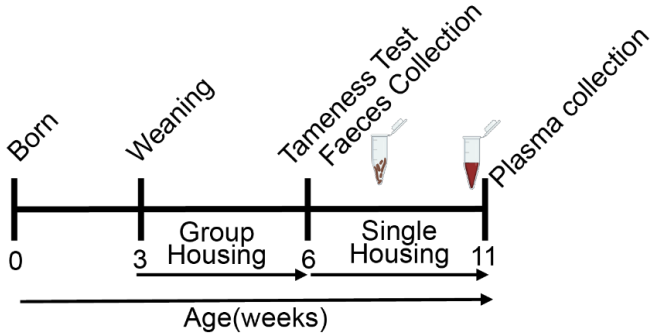


Figure 2.2. Experimental timeline

Given the genetic closeness between S1 and C2, as well as S2 and C1, I compared data between these pairs to assess the impact of selective breeding on active tameness. Since no significant sex effects were observed in the two-way analysis of variance (Two-way ANOVA; Supplementary Table S1), data from both male and female mice were pooled for all tameness analyses. For active contacting, both S1 and S2 demonstrated significantly longer contacting times compared to their respective controls ($p < 0.001$ for both), as shown in Figure 2.3B. Other parameters related to tameness, such as heading in the active tameness test for S1 ($p < 0.001$) and S2 ($p < 0.05$) depicted in Figure 2.3A, along with heading (Figure 2.3E) and accepting (Figure 2.3F) in the passive tameness test, were also higher in the selected groups compared to the non-selected groups, indicating a stronger expression of tame behaviours. In the stay-on-hand test, the median duration that mice remained on the hand was significantly longer for both S1 and S2 compared to the controls ($p < 0.001$ for both), illustrated in Figure 2.3I. Conversely, behaviours typically associated with wild-type mice, such as active jumping (Figure 2.3D) (C2, $p < 0.001$), passive locomotion (Figure 2.3G) (C2, $p < 0.001$; C1, $p < 0.05$) and jumping (Figure 2.3H) (C2, $p < 0.001$; C1, $p < 0.01$), were more prevalent in the non-selected groups, with C2 showing significantly higher levels of these behaviours compared to the selected groups. Overall, the data indicate that both S1 and S2 exhibit enhanced levels of both active and passive tameness compared to their non-selected counterparts, confirming the effect of selective breeding on enhancing tameness in these mice.

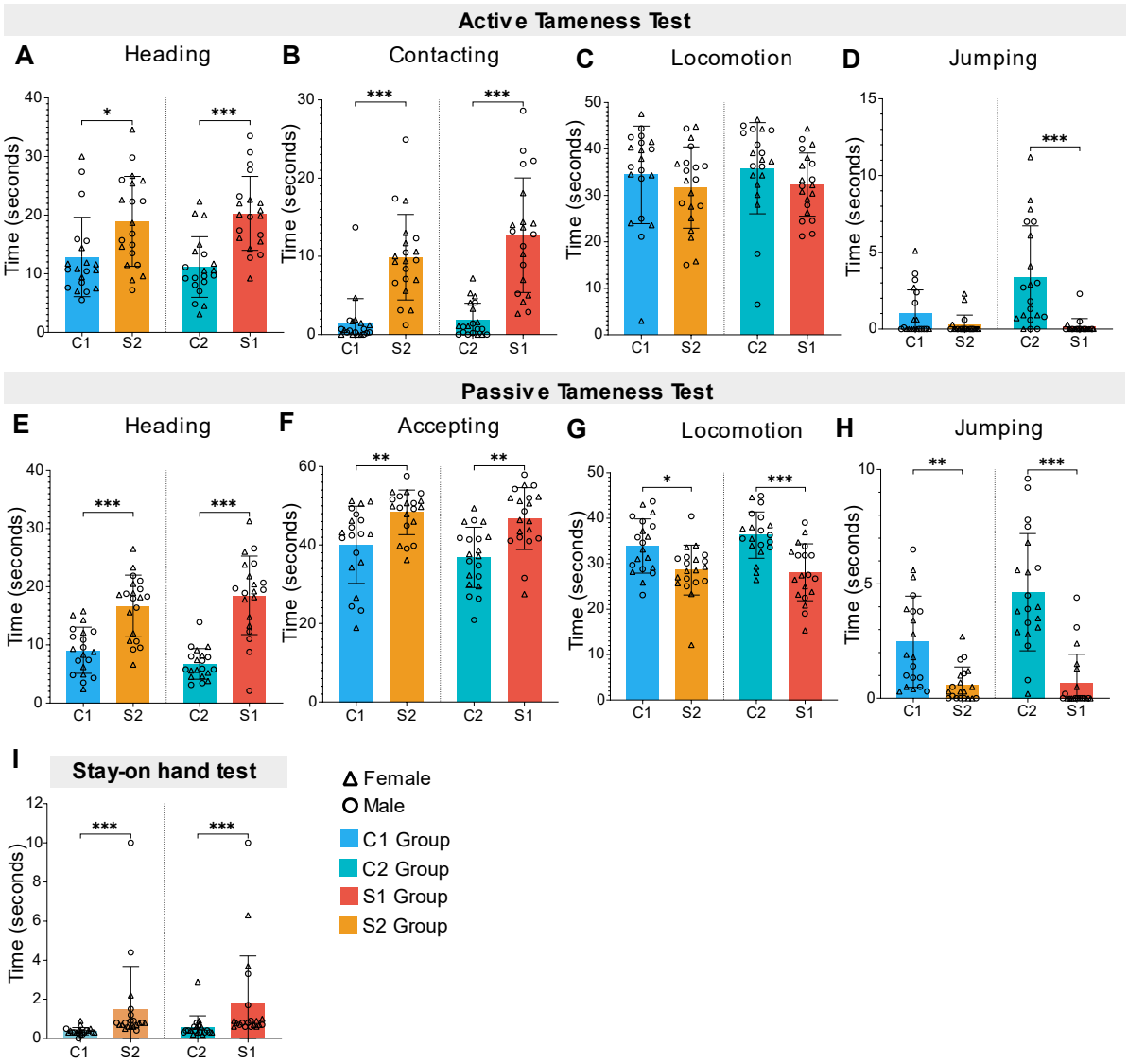


Figure 2.3. Tameness is higher in selected mice compared to control (A-D) active tameness test results, (A) heading, (B) contacting, (C) locomotion, (D) jumping, (E-H) passive tameness test results, (E) heading, (F) accepting, (G) locomotion, (H) jumping, (I) stay on hand test. N for tameness test = 80 (20 in each group with 10 male and 10 female); (*p<0.05; **p<0.01; ***p<0.001). To evaluate the normality of the dataset, we used the Shapiro–Wilk test. For normally distributed data, one-way ANOVA was performed followed by Tukey's test for multiple comparisons. For non-normally distributed data, the Kruskal–Wallis test was applied, followed by Dunn's test for multiple comparisons. Bar graphs present the mean ± SD with individual data points. For visual clarity, statistical comparison between C1 vs S2 and C2 vs S1 is only shown.

From the clustering analysis of all nine behavioural parameters, I identified two significant clusters that categorize these parameters based on their association with tameness. These clusters effectively distinguish the behaviours as either positively or negatively associated with tameness, as illustrated in Figure 2.4.

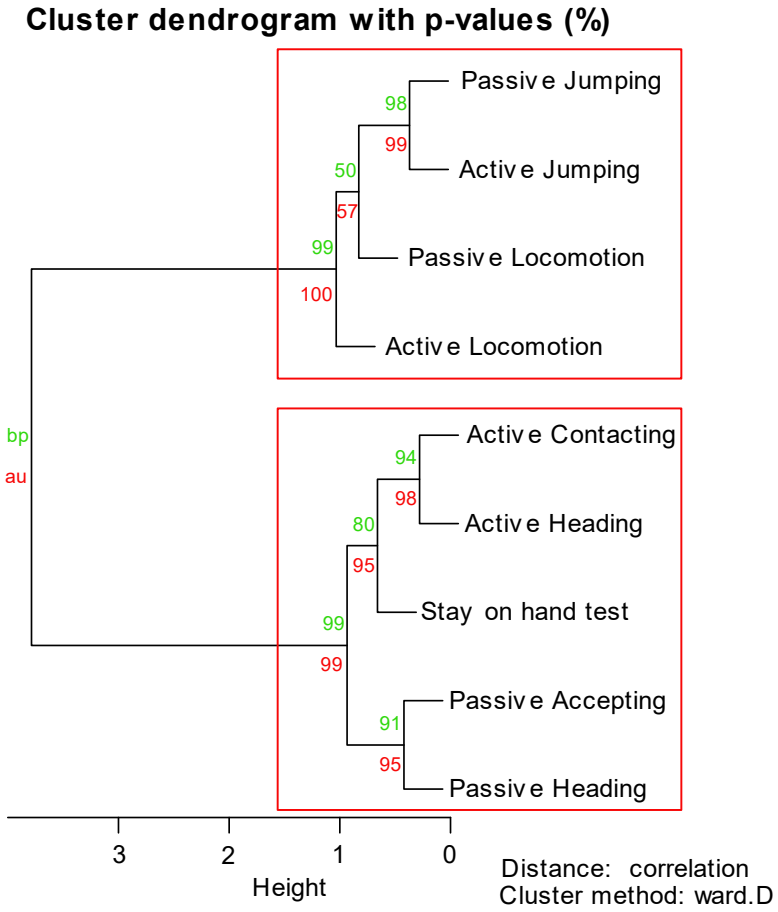


Figure 2.4. Correlation clustering of tameness test parameters. All nine parameters were normalized and clustered using the correlation distance and Ward's method. The red rectangle represents significant clusters based on the p-value calculated using pvclust. 'au' stands for (Approximately Unbiased) p-value and 'BP' for Bootstrap Probability values.

Previous studies on WHS mice showed that selected groups exhibit higher level of social behaviours as well as higher expression of hippocampal *Oxtr*, a receptor gene for oxytocin which play crucial role in social bonding (67) in these mice (3, 9). To see how selective breeding on tameness affect oxytocin pathway, I measured oxytocin levels in plasma collected from these mice. Results showed significantly higher oxytocin levels in both S2 and S1 compared to C1 (S1, $p < 0.01$; S2, $p < 0.01$). Although difference was not significant, S1 also showed higher oxytocin levels compared to the control C2 (Figure 2.5A). Furthermore, a significant mild Pearson correlation (R=0.52, *p*=2.4e-06) was observed between oxytocin levels and active contracting time (Figure 2.5B), highlighting a possible link between oxytocin and tameness behaviour.

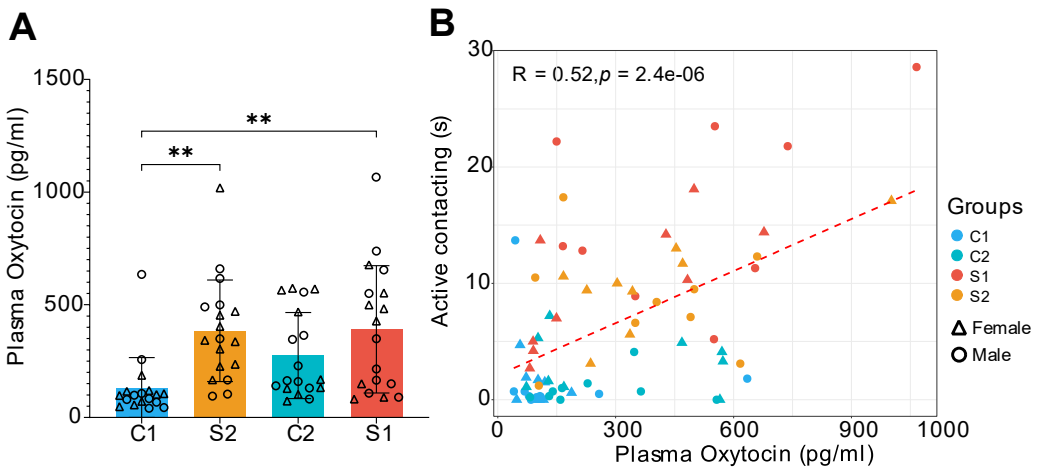


Figure 2.5. Plasma oxytocin concentration is higher in selected mice. (A) Plasma oxytocin concentration. For statistical test, a two-way ANOVA was performed followed by Tukey's test for multiple comparisons. Bar graphs show mean ± SD with individual data points. (B) Pearson correlation between oxytocin concentration and active contacting time. N = 72 (18 in each group with 9 male and 9 female) for both graphs (**p<0.01).

2.3.2 Mice gut contains many unexplored and yet to be discovered bacterial species

To investigate the gut microbiota of the selected and non-selected groups, I conducted shotgun metagenomic sequencing analysis on 80 faecal samples collected from each mouse following the tameness tests. On average, each sample yielded 31 million paired reads. MAGs were utilized to analyse the composition of the gut microbiome in WHS mice. Considering the complexity of the gut microbiota, I employed two different assemblers along with multiple binning strategies to maximize the recovery of high-quality MAGs, as depicted in Figure 2.5.

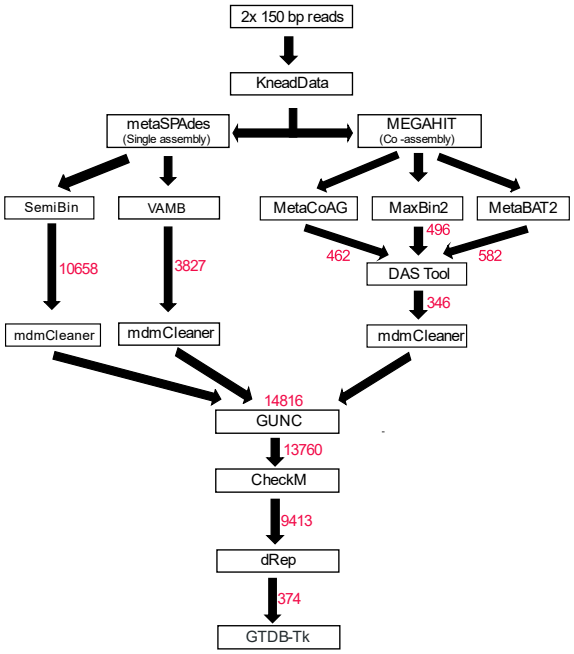


Figure 2.6. Scheme for high quality MAGs generation, count of MAGs generated were shown next to each step with red ink.

Utilizing this approach, I successfully generated 14,816 non-chimeric MAGs. Detailed information about the MAGs obtained from each binner is presented in Figure 2.7A and Figure 2.7B.

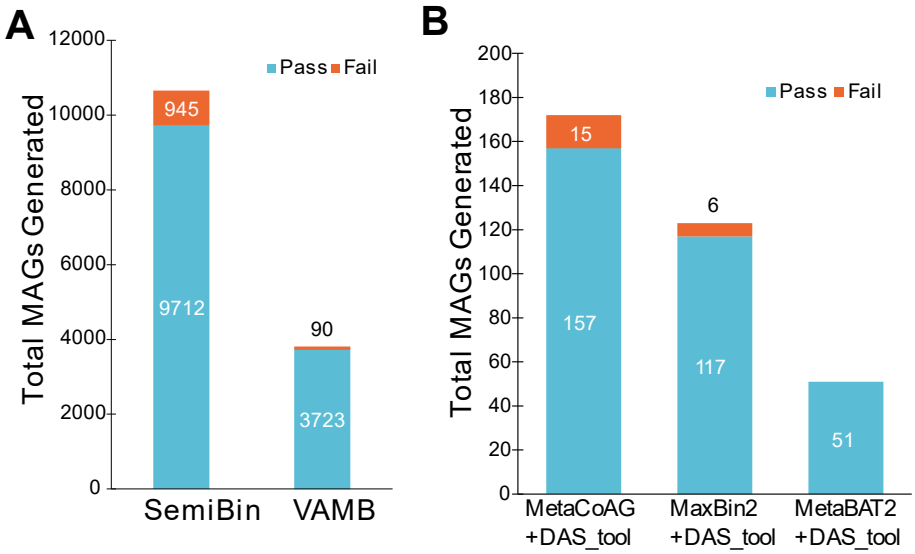


Figure 2.7. Total MAGs generated by each binner with information about number of MAGs passing GUNC chimeric test.

After eliminating the redundant MAGs, I isolated 374 bacterial MAGs that were over 50% complete and exhibited less than 10% contamination. Among these, 226 were classified as high-quality, exhibiting more than 90% completeness and less than 5% contamination, as detailed in Figure 2.8.

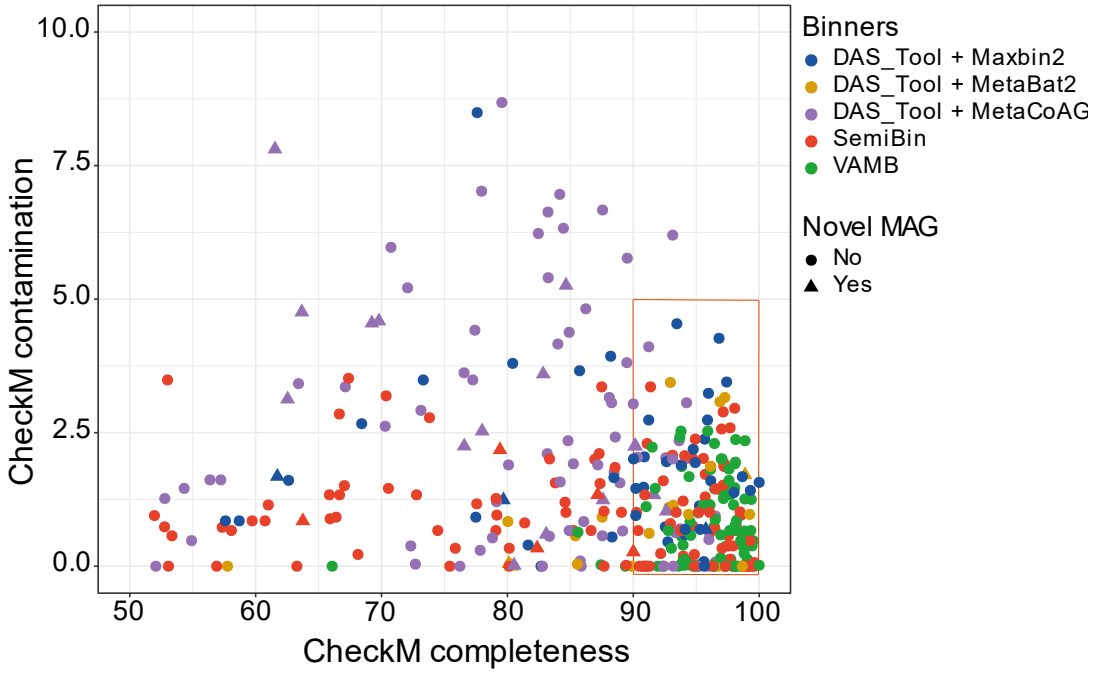


Figure 2.8. CheckM result of all 374 MAGs obtained, red rectangle encompasses MAGs which have >90% completeness and <5% contamination

The 374 MAGs I identified spanned 11 different phyla, with 281 of these MAGs attributed to the *Bacillota_A* phylum, as shown in Figure 2.9 and detailed in Supplementary Tables S2. When cross-referenced with the comprehensive mouse microbiota genome catalogue(53), it became evident that 27 of these MAGs represented species-level novelties not previously reported. Among these novel MAGs, one each was identified within the phyla *Actinomycetota*, *Bacillota B,* and *Patescibacteria;* eight were found within *Bacillota*, and 17 within *Bacillota A* (as illustrated in Figure 2.9 Supplementary Tables S3).

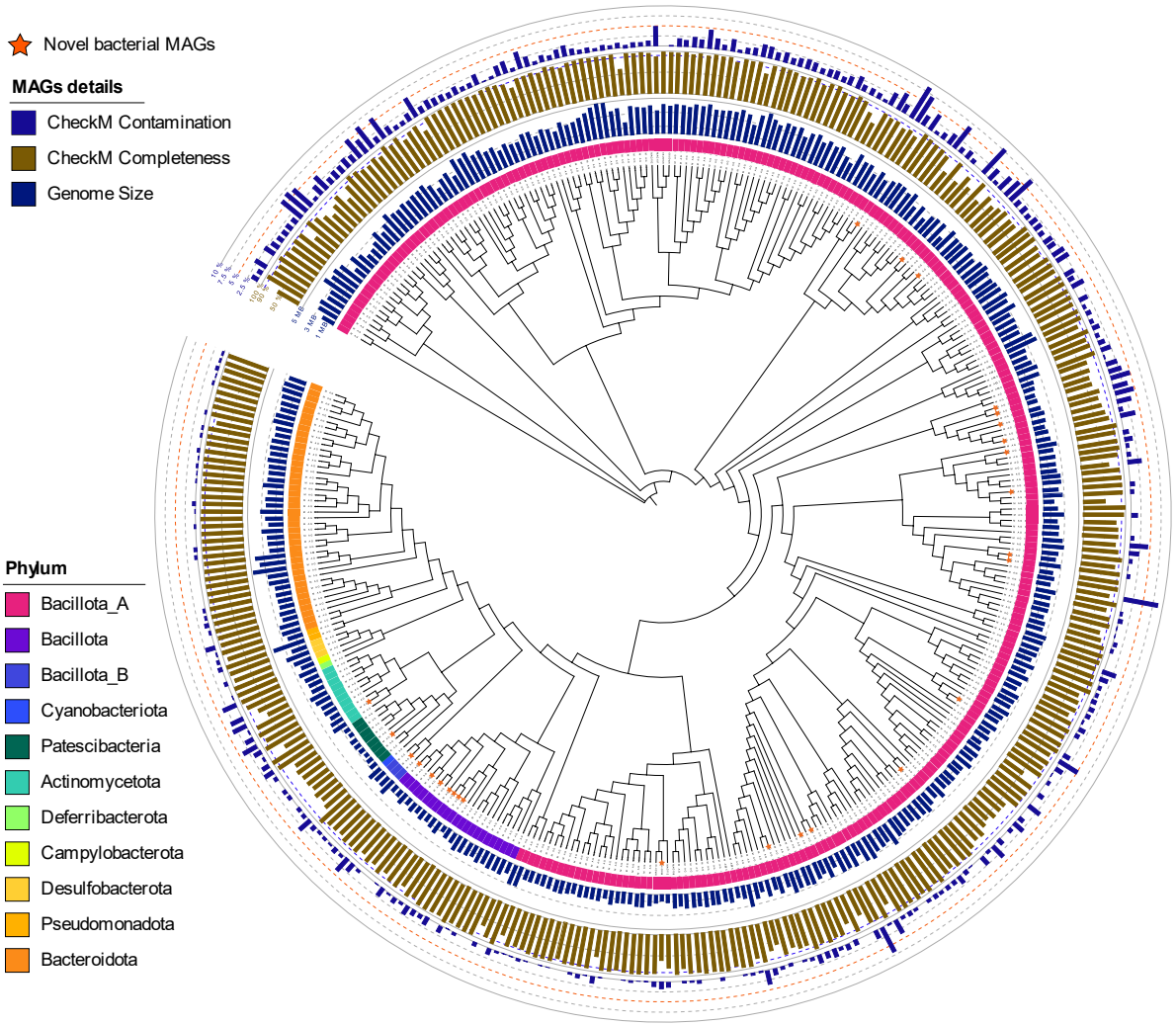


Figure 2.9. Cladogram of 374 MAGs generated. (A) All 374 MAGs were placed in this cladogram generated by GTDB output. Different phylum was represented as different colours in the innermost circle. Second circle indicates size of each MAGs; Third circle indicates CheckM completeness and fourth circle indicates CheckM contamination. All novel MAGs are indicated as orange star.

Additionally, while machine learning binners proved effective in generating a higher quantity and quality of MAGs, their limitations in binning all types of bacteria became apparent. This observation underscores the benefits of employing a combined methodology that integrates

both single assembly and co-assembly techniques, along with the use of multiple binning tools. This approach enhances the comprehensiveness and accuracy of our microbial genomic analysis, as depicted in Figure 2.10.

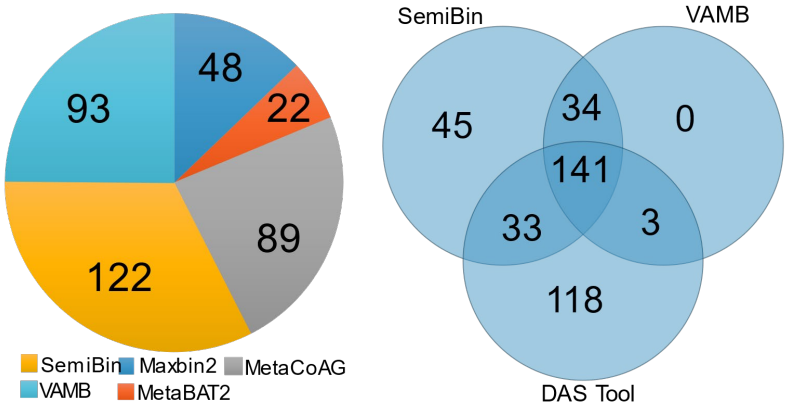


Figure 2.10 Distribution of origin of 374 MAGs generated in the current study, Ven diagram of origin of MAGs from same cluster of dRep (These data were compiled from clustering result of dRep where a MAG is considered to be of the same species if they cluster together at 95% ANI).

2.3.3 Selective breeding did not affect gut microbiota taxonomic diversity

For the taxonomic analysis aimed at capturing the full diversity of the WHS mice gut microbiome, I utilized the comprehensive NCBI/RefSeq prokaryote genome sequence database (NCBI RefSeq Complete V205) (56).

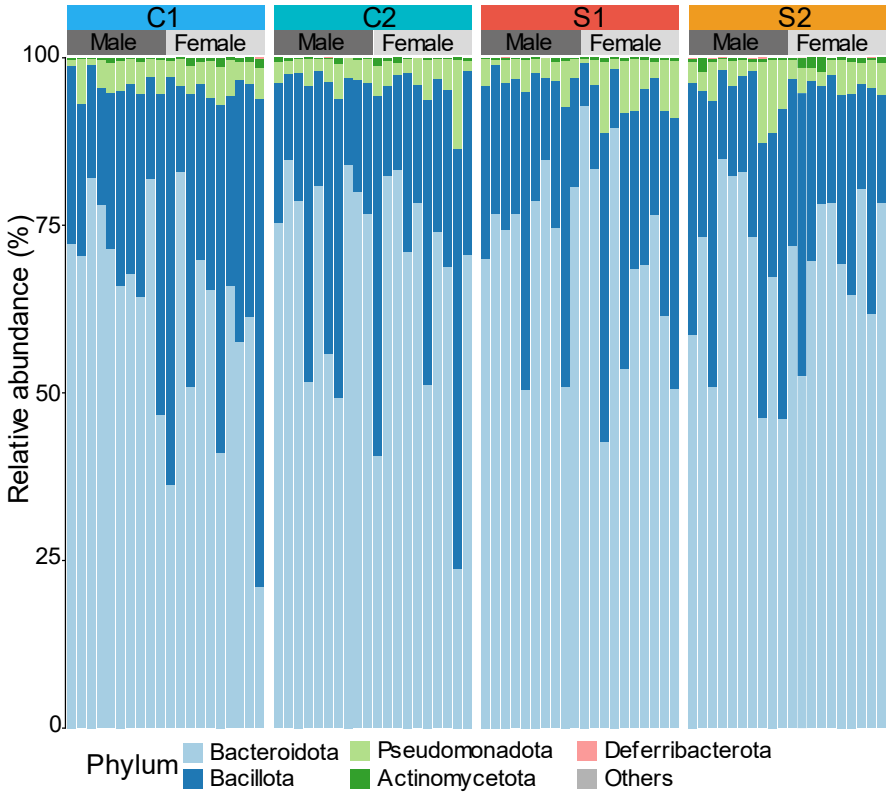


Figure 2.11. Phylum level relative abundance in all groups, five most abundant phylum is shown; full dataset includes 11 phyla, 21 classes, 45 orders, 103 families, and 339 genera. N = 80 (20 in each group with 10 male and 10 female).

This extensive database enabled the identification of 665 bacterial species, encompassing 11 phyla, 21 classes, 45 orders, 103 families, and 339 genera, providing a detailed overview of the microbial landscape present in these mice. The phylum level distribution results of this analysis are detailed in Figure 2.11.

Among the identified species, 427 were common across all four groups, accounting for 99.8% of the total microbial abundance observed in the samples (Figure 2.12A). In terms of taxonomic alpha diversity, the Chao1 richness index was significantly higher in group C1 compared to S1 ($p < 0.001$) and C2 ($p < 0.05$). Additionally, group S2 exhibited higher Chao1 richness compared to S1 ($p < 0.05$), indicating a greater species richness in these groups (Figure 2.12C). The Shannon diversity index, which measures both the abundance and evenness of species present, showed no significant differences between the groups, suggesting similar levels of overall diversity (Figure 2.12D). Moreover, beta diversity, assessed using Bray-Curtis dissimilarity to compare community composition between the groups, indicated no substantial differences, reinforcing the similarity in microbial community structure across the groups (Figure 2.12B).

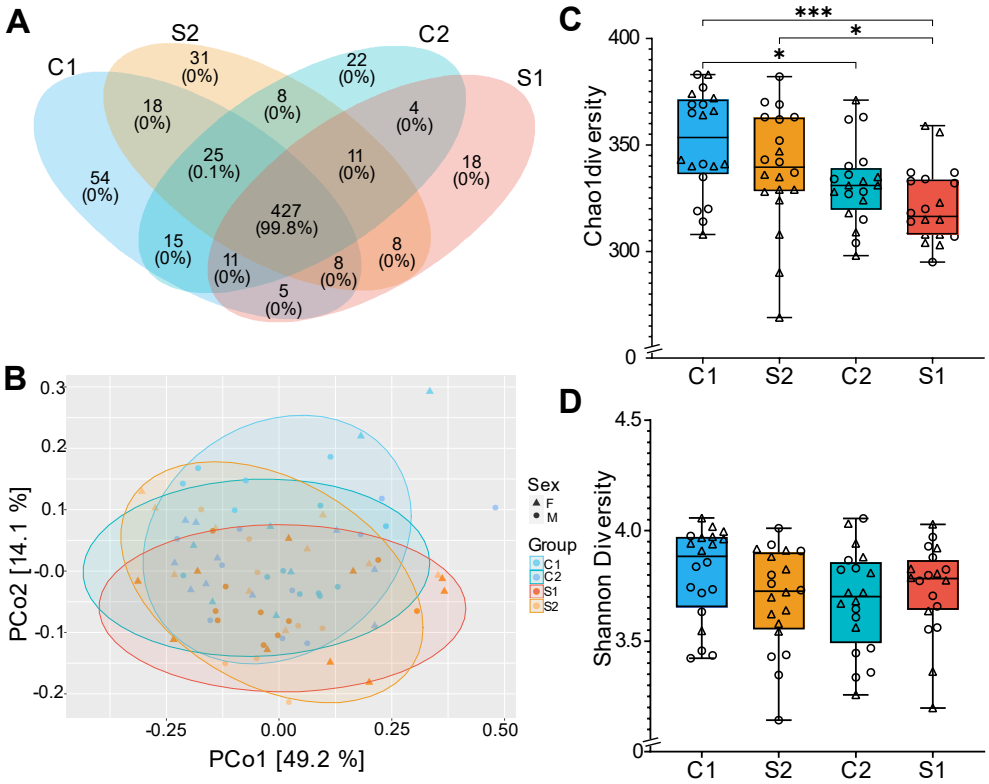


Figure 2.12. Taxonomic diversity of WHS mice (A) Ven diagram of 665 bacterial species present in WHS mice gut, (B) Beta diversity based on Bray–Curtis dissimilarity, (C) Chao1 diversity, (D) Shannon diversity. For statistical test, a two-way ANOVA was performed followed by Tukey's test for multiple comparisons. Each data point represents one sample. N = 80 (20 in each group with 10 male and 10 female). For box plot, the central line represents the median, with boxes showing the first and third quartiles. Whiskers extend to 1.5 times the interquartile range. (*p<0.05; **p<0.01; ***p<0.001).

Overall, my findings indicate that the selective breeding pressure for tameness behaviour did not consistently affect the overall taxonomic diversity of the gut microbiota in mice.

I performed Pearson correlation analysis between the tameness parameter scores and 40 significant taxa identified through random forest analysis. This analysis revealed that various *Lactobacillus* bacteria displayed a significant positive correlation with both Active and Passive Jumping behaviours, but these bacteria were significantly negatively correlated with Active Heading. Conversely, *Bacteroides* sp. CBA7301, *Ligilactobacillus animalis*, and *Prevotella bivia* showed a significant positive correlation with the specific selection pressure of active contacting, as illustrated in Figure 2.13.

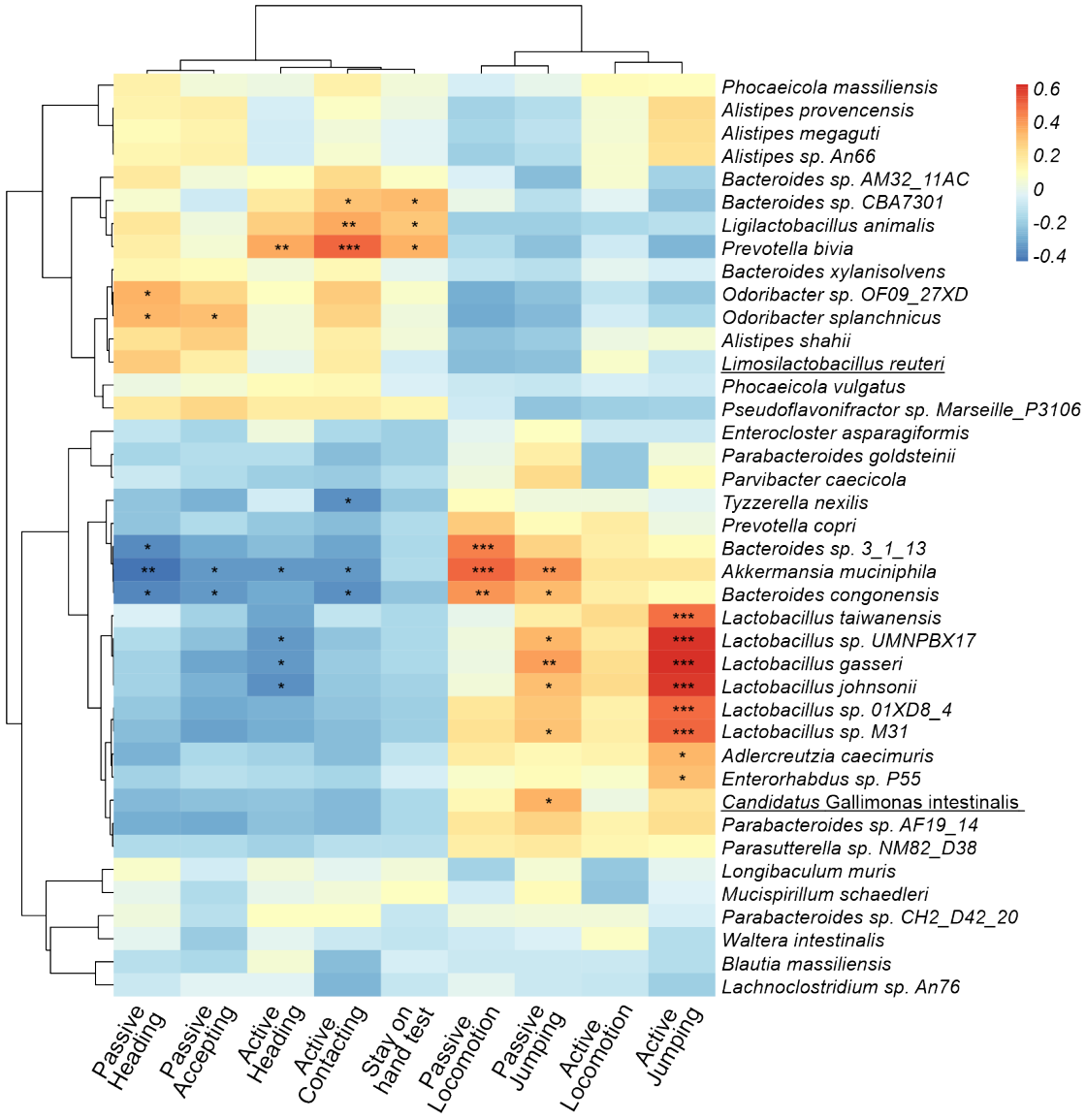


Figure 2.13. Hierarchical clustering using Pearson correlation distance between top 40 significantly different taxa (result of random forest analysis) with score of parameters of tameness test. Significant correlation is marked with (*p<0.05; **p<0.01; ***p<0.001). *P-value* was adjusted using Benjamini–Hochberg (BH) false discovery rate (FDR).

2.3.4 *L. reuteri* is enriched in both the selected groups

To ascertain the enrichment of bacterial species in tame mice and strengthen the association, I conducted a comparative analysis between each selected group and its corresponding control group. Utilizing MaAsLin2, I associated species-level gut microbiome compositions with the WHS mouse groups, as depicted in Figure 2.14 A and B. In the MaAsLin2 analysis, only species that showed significant associations in both comparisons were considered to have differences in abundance. Notably, *L. reuteri* (formerly known as *Lactobacillus reuteri*), a type of lactic acid bacteria (shown in Figure 2.14 C and D), along with *Candidatus* Gallimonas intestinalis (shown in Figure 2.14 E and F), were found to be enriched in the selected groups compared to their respective control groups.

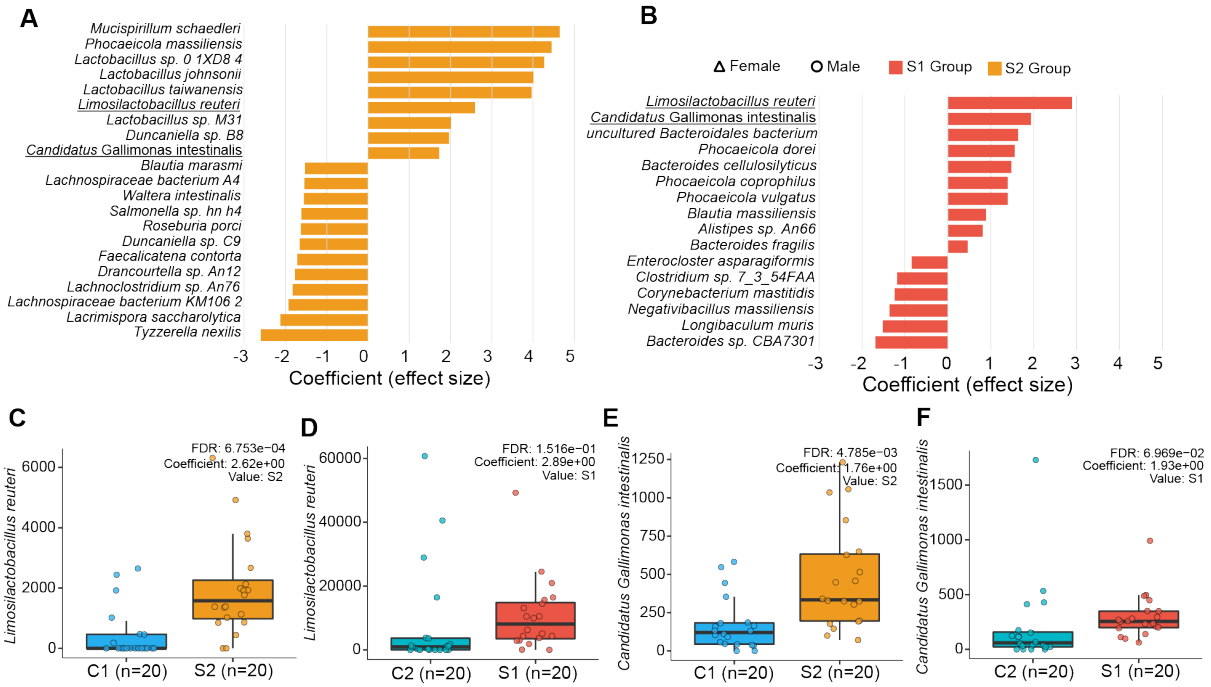


Figure 2.14. MaAsLin2 Association (A)Association analysis in MaAsLin2 between S2 group and bacterial abundance, (B) Association analysis in MaAsLin2 between S1 group and bacterial abundance, (A-B) P-value was adjusted using Benjamini–Hochberg (BH) false discovery rate (FDR). (C-D) MaAsLin2 generated graph of *L. reuteri*, (E-F) MaAsLin2 generated graph of *Candidatus* Gallimonas intestinalis. Each data point represents one sample. N = 80 (20 in each group with 10 male and 10 female). For box plot, the central line represents the median, with boxes showing the first and third quartiles. Whiskers extend to 1.5 times the interquartile range.

Among the 374 MAGs generated, I obtained one high-quality MAG from *L. reuteri* and six MAGs from the *Candidatus* Gallimonas genus. To quantify the abundance of gut bacteria, I employed CoverM to map reads to all MAGs, achieving mapping rates ranging from 75% to 93% (Figure 2.15A). Consistent with the results from Kraken2, I found that *L. reuteri* was significantly enriched in both selected groups compared to the non-selected groups (S1, $p < 0.05$; S2, $p < 0.001$) as shown in Figures 2.15B and 2.15C. However, the

relative abundance of the six *Candidatus* Gallimonas MAGs did not exhibit a significant difference between the tame and non-selected groups, as depicted in Figures 2.15D to 2.15I.

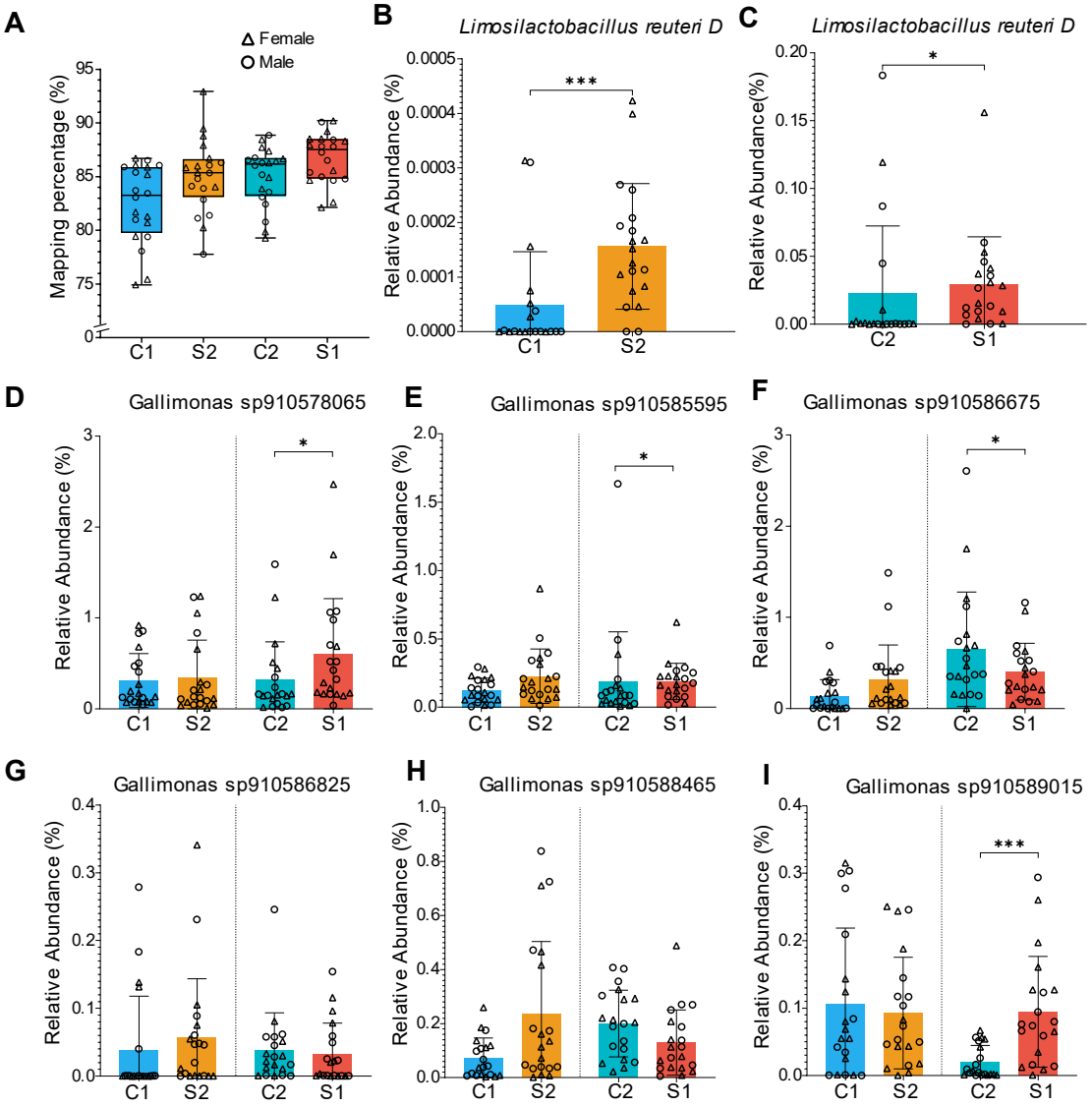


Figure 2.15: Relative abundance of MAGs (A) Mapping percentage of each sample filtered sequences against MAGs generated in this study; (B) relative abundance of *L. reuteri* MAGs in C1 & S2 (One sample was not shown as it exceeds y-axis, but utilized for statistical calculation); (C) relative abundance of *L. reuteri* MAGs in C2 & S1; (D-I) relative abundance of six *Candidatus* Gallimonas MAGs; N= 80 (20 in each group with 10 male and 10 female). (*p<0.05; **p<0.01; ***p<0.001). For box plot, the central line represents the median, with boxes showing the first and third quartiles. Whiskers extend to 1.5 times the interquartile range. (*p<0.05; **p<0.01; ***p<0.001). Bar graphs show mean ± SD with individual data points. For statistical analysis, two-tailed Mann-Whitney U test was performed.

These findings, together with previous Pearson correlation analysis, demonstrate that *L. reuteri* is the only bacterium consistently associated with tameness across all analyses conducted. The persistent link across different analytical methods underscores the potential role of *L. reuteri* in influencing tameness behaviour, highlighting its importance in the study of domestication and behavioural genetics.

2.3.5 Active contacting selection pressure did not affect gut microbiota functional diversity

The functional analysis of the metagenomic data uncovered 410,083 gene families across all samples. Among these, 240,161 gene families were common to all four groups, indicating a core set of functionalities shared across the study subjects. Additionally, each group exhibited unique gene families, highlighting the diversity within the gut microbiota's functional potential (Figure 2.16A). Next, I examined the diversity of gene families in all four groups of mice. Chao1 richness showed that the C1 group had the highest functional richness among all groups and significantly higher richness than the S1 group ($p < 0.01$) (Figure 2.16B). Unlike taxonomic diversity, the Shannon diversity of the gene family in the S1 group was significantly lower than that in the C1 group (Figure 2.16C). However, beta diversity did not differ between groups (Figure 2.16D).

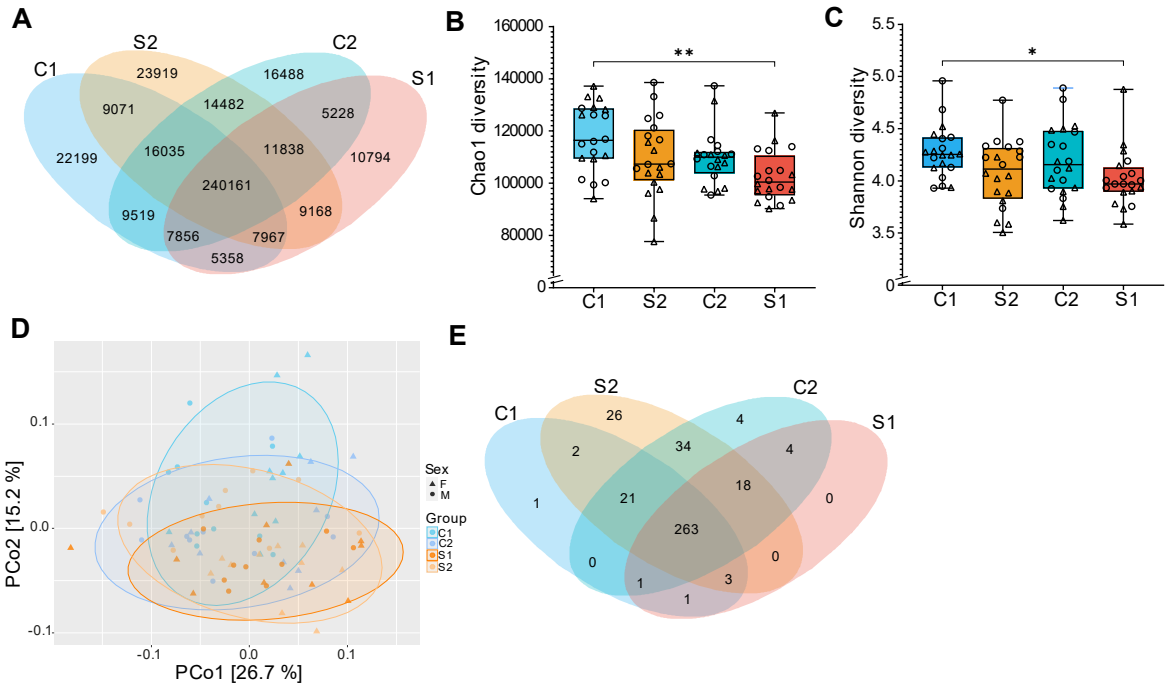


Figure 2.16: Host tameness selection pressure does not change gut microbiota functional diversity (A) Venn diagram showing all gene family identified; (B) gene family Chao1 diversity, (C) Gene family Shannon diversity; (D) Beta diversity based on Bray–Curtis dissimilarity. (E) Venn diagram of gut microbiota metabolic pathways identified. N= 80 (20 in each group with 10 male and 10 female). (*p<0.05); **p<0.01; ***p<0.001). For box plot, the central line represents the median, with boxes showing the first and third quartiles. Whiskers extend to 1.5 times the interquartile range.

Gene pathways were also identified from gene families using the MetaCyc database. Overall, I identified 378 gut microbial pathways in WHS mice of which 263 were present in all four groups (Figure 2.16E). Association analysis was performed on the MetaCyc pathways identified by functional analysis. The metabolic pathways in both S1 and S2 were compared with their respective controls using MaAsLin2. Overall, I identified 88 pathways in the S2 group and 49 pathways in the S1 group that met the q-value cutoff criteria. Among these, 35

pathways exhibited consistent trends in both the S1 and S2 groups, Supplementary Table S4. Notably, I discovered several pathways related to amino acid biosynthesis and energy generation that were enriched in the selected groups. This enrichment suggests that the selection for tameness may influence metabolic functions related to these pathways, potentially contributing to the observed behavioural traits.

Chapter 3: Tameness selection pressure increase blood pyruvate concertation

3.1 Introduction

The gut microbiome is a complex and dynamic entity that significantly influences the host beyond traditional boundaries of nutrition and immune function. One of the more intriguing aspects of this influence is the production of metabolites by gut bacteria, which can profoundly affect host behaviour. These microbial metabolites include short-chain fatty acids (SCFAs), neurotransmitters, and other biochemical compounds that interact directly with the host's nervous system, illustrating a complex pathway of gut-brain communication that is pivotal in behavioural modulation.

SCFAs, such as butyrate, acetate, and propionate, are fatty acids with fewer than six carbon atoms, produced primarily through the fermentation of dietary fibres by gut microbiota. These metabolites are not merely energy sources; they play critical roles in modulating the physiological and behavioural responses of the host. SCFAs influence behaviour by affecting brain chemistry and signalling pathways (68, 69). The gut microbiota can also produce neurotransmitters that are identical or similar to those used by the human central nervous system, such as serotonin, dopamine, and gamma-aminobutyric acid (GABA). Serotonin, predominantly found in the gastrointestinal tract, is critical for regulating mood and behaviour. The gut microbes such as certain types of *Lactobacillus* and *Bifidobacterium* are capable of producing serotonin by metabolizing tryptophan, a serotonin precursor found in the diet (70, 71). Another pathway through which gut microbiota influence behaviour is through immune system modulation. Metabolites from gut bacteria can affect the immune system's functioning, which in turn impacts neuroinflammation and brain function. By modulating immune responses, gut bacteria can potentially influence behaviours linked to mental health conditions such as autism spectrum disorders and schizophrenia (72). Additionally, other metabolic byproducts such as bile acids and bioactive peptides are converted by microbial enzymes into secondary metabolites that can serve as signalling molecules, interacting with the host's gut-brain axis. These interactions can trigger behavioural changes, including influencing satiety, stress responses, and even complex social behaviours (73). In the present study, I performed plasma metabolomic analysis of tame and control mice followed by administration of the identified metabolite to control mice.

3.2 Materials and Methods

3.2.1 Plasma metabolites analysis

To elucidate the role of gut microbiota-derived metabolites in influencing host behaviour, I conducted a comprehensive analysis of plasma metabolites. Samples were analysed by Human Metabolome Technologies, Inc. (https://humanmetabolome.com/jpn/) using Capillary Electrophoresis Time-of-Flight Mass Spectrometry (CE-TOFMS) to detect more than 1000 water-soluble metabolites. This approach, previously described (74-77), allowed us to identify and quantify metabolites in 12 mouse plasma samples, capturing both cationic and anionic metabolites.

3.2.2 Pyruvate administration

Mice from the 35th generation of the non-selected group C1 were utilized in this experiment to investigate the effects of pyruvate on tameness behaviour. To ensure objectivity and reduce experimental bias, distinct roles were assigned to three experimenters. I handled the grouping of mice and administered injections, while K. Kurosawa conducted the tameness test, and M. Nihei analysed the video recordings. To maintain impartiality, both K. Kurosawa and M. Nihei were blinded to the group assignments of the mice. The experimental protocol included intraperitoneal (IP) injections of pyruvate (Sigma-Aldrich, USA) on alternate days over a 14-day period, totalling seven injections. Control mice received equivalent volumes of saline. To minimize stress, mice were anesthetized with 5% isoflurane before each pyruvate injection. Tameness tests were conducted before and after the injection period to evaluate any changes in behaviour. To ensure consistency, both tameness tests were scheduled between 14:00 and 16:00.

3.2.3 Statistical analysis

Statistical analysis was performed using GraphPad Prism (Ver. 10.1.2) and MaAsLin2 (Ver. 1.12.0). MaAsLin2 was employed for statistical association analysis of plasma metabolites with a Benjamin-Hochberg's q-value of 0.25 (64). The effects of pyruvate treatment were evaluated using a paired two-tailed t-test.

3.2.4 Ethics approval

All experiments conducted in this study were performed in strict compliance with the guidelines and protocols approved by the Committee for Animal Care and Use at the National Institute of Genetics, under permit numbers R4-25 and R5-7.

3.3 Results and discussion

3.3.1 Plasma Pyruvate Levels in Selected Groups

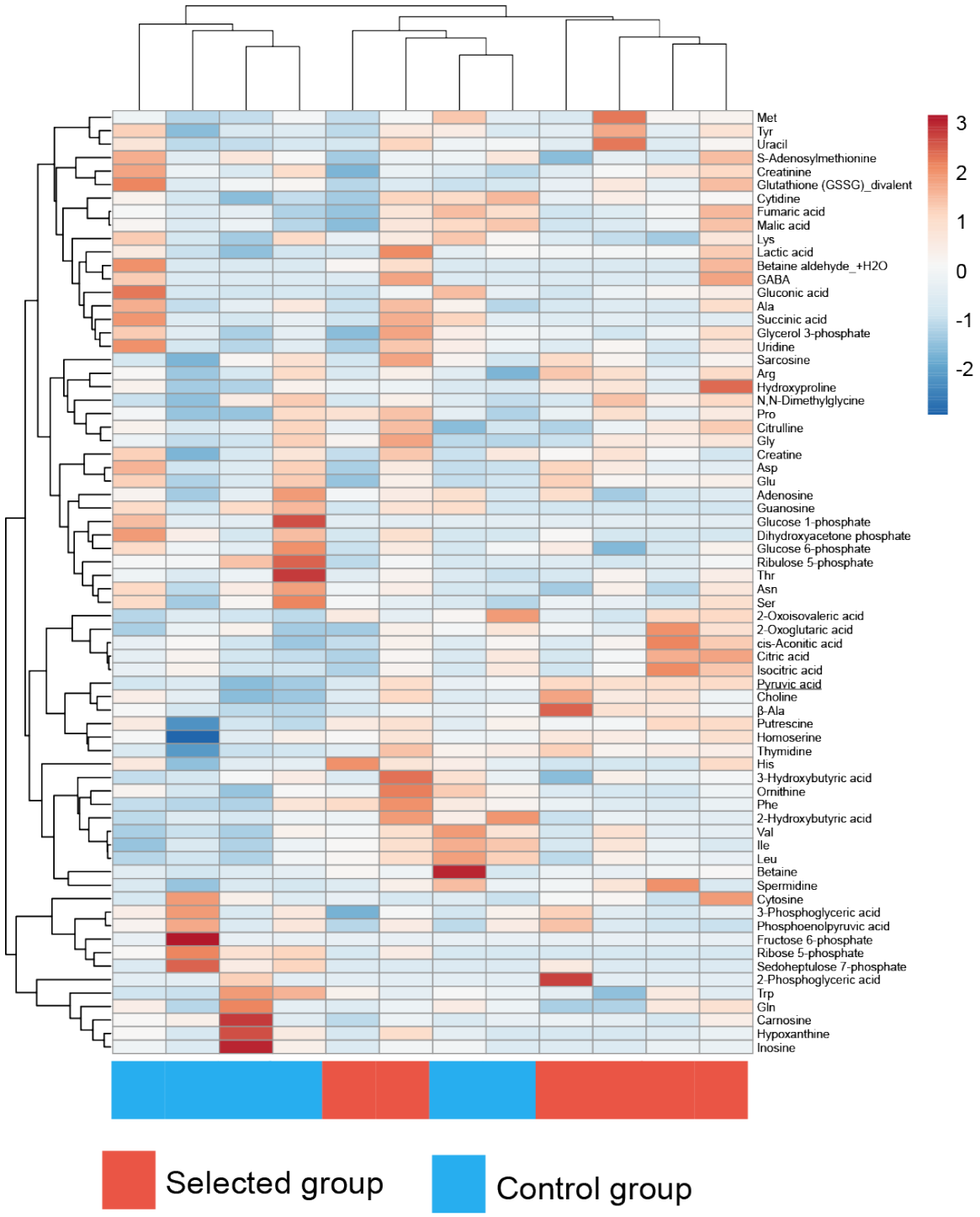


Fig. 3.1. Pyruvic acid is significantly high in selected mice. Heatmap with hierarchical clustering using Pearson correlation distance of 70 metabolites identified. *P-value* was adjusted using Benjamini–Hochberg (BH) false discovery rate (FDR).

The influence of the gut microbiota on animal behaviour is mediated through the production of metabolites, which are absorbed by the host and can lead to behavioural changes. To

determine which metabolites might be influencing changes in tameness, I performed a metabolomic analysis on plasma samples from WHS mice. Using capillary electrophoresis time-of-flight mass spectrometry (CE-TOFMS), I detected 281 peaks—179 in cation mode and 102 in anion mode—which were annotated using HMT's standard library and the known-unknown peak library. Out of these, 70 metabolites (46 in cation mode and 24 in anion mode) were detected and quantified (supplementary Table S5). However, correlation clustering of these 70 metabolites revealed no discernible differences between the control and selected groups, as shown in Figure 3.1.

Among the metabolites analysed, four were found in significantly higher concentrations in the selected mice, as shown in Figure 3.2A-D. To further validate these results, I conducted a MaAsLin2 analysis with FDR correction. After this adjustment, only pyruvic acid remained significantly associated with the selected mice, as detailed in Table S5. Additionally, I measured the plasma lactic acid concentration, which is known to maintain metabolic homeostasis with pyruvic acid. The levels of lactic acid were not significantly different between the selected and control groups, as depicted in Figure 3.2E.

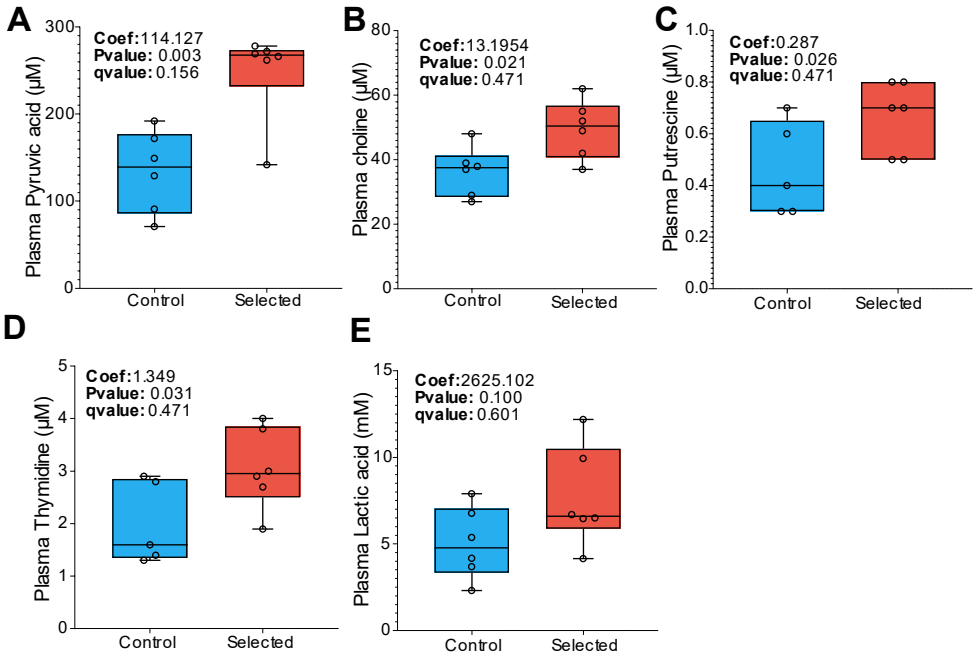


Fig. 3.2 Plasma metabolite concentration in WHS mice (A) Plasma pyruvic acid, (B-D) Significantly higher metabolites before FDR correction in MaAsLin2, (B) Plasma choline, (C) Plasma putrescine, (D) Plasma thymidine, (E) Plasma lactic acid, N = 12 (6 in each group). For box plot, the central line represents the median, with boxes showing the first and third quartiles. Whiskers extend to 1.5 times the interquartile range.

3.3.2 Pyruvate administration doesn't increase tameness behaviour

To determine any possible influence of pyruvate on mouse behaviour, I intraperitoneally (IP) injected pyruvate seven times over a period of 14 days into mice of non-selected C1 group to

observe the effects on active tameness. Active tameness test was performed before and after pyruvate injection (Figure 3.3).

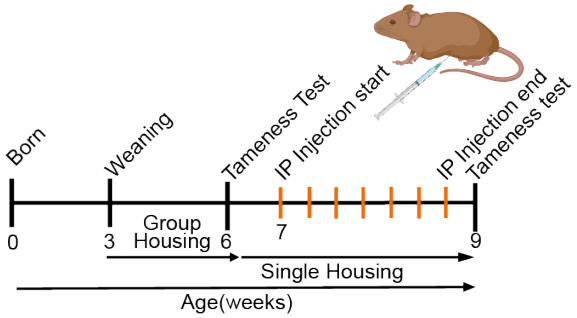


Fig. 3.3 Scheme of pyruvate administration

Active heading increased in both groups after the injection (Figure 3.4A). In the measurement of active tameness, I found that contacting, a main behavioural parameter for selective breeding, was significantly decreased ($p < 0.01$) after treatment in the saline injection group, but not in the pyruvate injection group (Figure 3.4B).

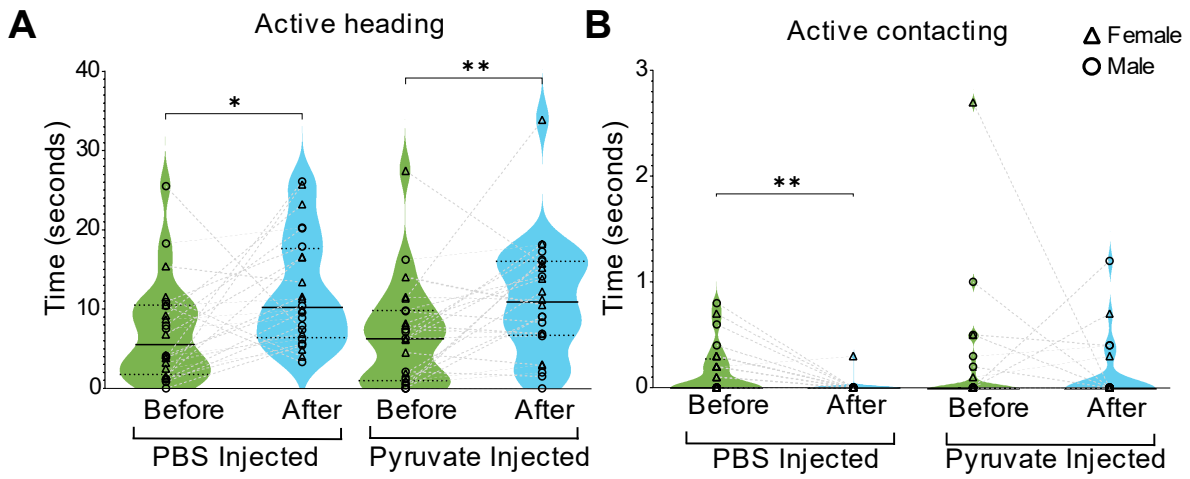


Figure 3.4. Pyruvate administration doesn't increase tameness behaviour. (A) Active tameness heading time; (B) active tameness contacting time. N= 48 (24 in each group with 12 male and 12 female). (*p<0.05; **p<0.01). Each "violin" represents the distribution and density of data points. The central line within each violin denotes the mean, with the width indicating data density. The effects of pyruvate treatment were evaluated using a paired two-tailed t-test.

The reduction in contacting in the saline treatment group may have been caused by the repeated IP injections. Therefore, the lack of reduction in contacting observed in the pyruvate injection group may indicate the effects of an increase in tameness caused by pyruvate injection. However, it was considered necessary to experiment with less stressful methods to investigate the effects of pyruvate on active tameness in mice.

Chapter 4: Tameness behaviour increased in non-selected mice by *L. reuteri* administration

4.1 Introduction

The complex ecosystem of the gut microbiota plays a pivotal role in influencing numerous aspects of host physiology and behaviour, a topic that has captured the interest of the scientific community in recent years. Research has increasingly demonstrated that the gut microbiome is not only crucial for metabolic (78-81) and immunological functions (82-84) but also exerts a significant influence on host behaviour (20, 85, 86). This relationship is observed across various species, indicating a fundamental aspect of biological interaction that transcends individual organismal boundaries. These association might be single bacterium (87, 88) or a whole community of bacteria (89-92). Recently, *L. reuteri* is getting associated with social behaviour (93-96) as well as release of oxytocin (97, 98)- an important hormone for social behaviour. My initial findings indicated that *L. reuteri* was significantly more abundant in the gut of tame mice compared to control groups. Intriguingly, these tame mice also displayed elevated levels of pyruvate and oxytocin in plasma. This led us to hypothesize a potential link between the presence of *L. reuteri*, pyruvate and oxytocin production, and possible influence on tameness behaviour. To test this hypothesis, I isolated various strains of *L. reuteri*, screened for pyruvate secretion capabilities and administered this bacterium to non-tame mice

4.2 Materials and Methods

4.2.1 *L. reuteri* isolation

Cecum and faecal samples from mice of the selected group (S1) were collected, combined, and then serially diluted. This diluted mixture was cultured on *L. reuteri* Isolation Media (LRIM) (99) agar plates, with raffinose (Tokyo Chemical Industry Co. Ltd, Japan) serving as the sole carbon source. The plates were incubated for 48 hours at 45°C under anaerobic conditions (5% CO_2). To ensure an anaerobic environment, the plates were placed in a sealed box containing an AnaeroPack sachet (Mitsubishi Gas Chemical Company, Inc., Japan). For isolating pure cultures, selected colonies were streaked twice on LRIM plates, followed by a third streaking on De Man–Rogosa–Sharpe (MRS) (Sigma-Aldrich, USA) agar plate.

4.2.2 *L. reuteri* phylogenetic tree

All isolated colonies were cultured in MRS media, and DNA was subsequently extracted using the NucleoSpin Microbial DNA kit (Takara Bio Inc., Shiga, Japan), following the manufacturer's instructions. To amplify the full 1.6 kbp region of the 16S rDNA, polymerase chain reaction (PCR) was performed using the universal primers bak4 (5'-AGGAGGTGATCCARCCGCA-3') and bak11w (5'-AGTTTGATCMTGGCTCAG-3') (100, 101). The PCR cycling conditions included an initial denaturation at 95°C for 5 minutes, followed by 35 cycles of 95°C for 15 seconds, 60°C for 30 seconds, and 72°C for 2 minutes, with a final extension at 72°C for 7 minutes. The PCR products were then purified and subjected to Sanger sequencing. The obtained sequences were taxonomically identified using NCBI BLASTn (102). For phylogenetic analysis, I included the 16S rRNA gene sequences generated in this study, along with sequences from multiple *Limosilactobacillus* species retrieved from NCBI, using *Lactobacillus helveticus* as an outgroup. All sequences were aligned using the MAFFT online server with the "FFT-NS-I" command (103, 104). The best-fit model for our data, determined using ModelFinder (105), was "TPM3u+F+I." A maximum likelihood tree was constructed using IQ-TREE 2 (106) with 1000 bootstrap replicates. The resulting tree was edited using the Interactive Tree of Life (iTOL) online server (52).

4.2.3 Pyruvate, L-Lactate, and D-Lactate secretion assay

In this experiment, *L. helveticus* (JCM1120), obtained from the Japan Collection of Microorganisms (JCM), was used as the positive control for pyruvate production. All bacterial strains, including the control, were cultured in Gifu Anaerobic Broth (GAM) medium (Nissui Pharmaceutical, Japan), supplemented with 1% glucose. The cultures were incubated for 24 hours at 37°C under anaerobic conditions. The concentrations of pyruvate, L-lactate, and D-lactate in the media were measured using biochemical assay kits from Cayman Chemicals (Michigan, USA), following the manufacturer's protocols precisely.

4.2.4 *L. reuteri* administration

All bacterial strains were cultured in MRS media for 24 hours. After cultivation, the optical density (OD) at 590 nm was measured to estimate the bacterial count in the solution, using a standard curve and colony forming unit (CFU). The bacteria were then washed with PBS, resuspended in PBS, and stored at -80°C until use. To maintain viability, a fresh batch of bacteria was cultured every week. PBS (vehicle) or bacteria (~1×10^8 CFU/cage/day) was added to the mice's drinking water daily between 17:00 and 18:00. Mice from the 37th

generation of the C1 group were used for the experiment. The administration of *L. reuteri* began immediately after weaning, around 3 weeks of age, and continued for 21 days. To minimize isolation stress, mice were housed in same-sex pairs. Faecal samples were collected from each individual both before the start and at the end of the experimental period. Additionally, daily water consumption for each cage was monitored and recorded, and the body weight of the mice was measured at the end of the experiment. Following the 21-day bacterial administration period, a tameness test was conducted, after which faecal and plasma samples were collected for further analysis.

4.2.5 qPCR quantification of *L. reuteri* and *L. helveticus*

Quantitative PCR (qPCR) was performed using a Thermal Cycler Dice Real Time System III (Takara Bio Inc., Shiga, Japan). Each 25 μL reaction contained TB Green Premix Ex Taq II (Tli RNaseH Plus) (Takara Bio Inc., Shiga, Japan) and gene-specific primers at a concentration of 1 μM. The reaction mixtures included 5 ng of metagenomic DNA, with the quantity determined using a Qubit fluorometer. The cycling conditions were as follows: an initial denaturation at 95°C for 5 minutes, followed by 40 cycles of denaturation at 95°C for 20 seconds, annealing at 62°C for 20 seconds, and extension at 72°C for 30 seconds. A melt curve analysis was then conducted over a temperature range of 60–95°C using the default settings of the system. The primers used for qPCR included universal bacterial primers EUB338(5'-ACTCCTACGGGAGGCAGCAG-3') (107), and EUB518 (5'-ATTACCGCGGCTGCTGG-3') (108); *L. reuteri* -specific 16S-23S rRNA gene primers, sg-Lreu-F (5'-GAAGATCAGTCGCAYTGGCCCAA-3'), and sg-Lreu-R (5'-TCCATTGTGGCCGATCAG-3') (109); *L. helveticus* hsp60 specific primers, F1LHelHsp (5'-CTTTGATCGCTGATGCTATGGAAAAGGTTGGTC-3'), and R1LHelHsp (5'-GATCAACAATGACTTGCCTTGTTGAACAATTTC-3') (110). . The qRT-PCR data analysis utilized the ΔCt values for the specific primers, which were normalized against the ΔCt values of the universal primers, following the ΔΔCt method.

4.2.6 Mice serum pyruvate and oxytocin concertation

To measure the concentrations of pyruvate in mice serum, I utilized biochemical assay kits from Cayman Chemicals (Michigan, USA), strictly adhering to the manufacturer's protocol. Due to the limited number of sample wells, 10 randomly selected samples (5 males and 5 females) were analysed. To measure oxytocin concentration, I used Oxytocin ELISA kit (Enzo Life Sciences, USA), following manufacturer's instructions. For looking at the effect

of *L. reuteri* treatment on blood oxytocin levels, *L. helveticus*-treated samples were excluded. Both assays were performed in duplicates (technical replication).

4.2.7 Cross-fostering

In this experiment, I used mice from the 38th generation of the C1 group and the 37th generation of the S1 group. Initially, I set up 10 breeding pairs of mice, aged between 7-8 weeks. Male mice were removed from the cages once mating was confirmed. In cases where pups from the S1 and C1 groups were born on the same day, I implemented a cross-fostering approach. To prevent infanticide and ensure acceptance by the foster mother, the pups were first mixed with the bedding from the foster mother's cage for 3-5 minutes. This step helped mask the pups' original scent. Subsequently, the pups were transferred to the home cage of their foster mother. The mice were weaned at approximately 3-4 weeks of age and were then housed with their same-sex littermates until they reached the age for the tameness test. The tameness test was conducted when the mice were 6 weeks old. Importantly, both the individuals conducting and analysing the tameness tests were blinded to whether the mice were fostered or with their birth mothers, to ensure unbiased assessments.

4.2.8 Statistical analysis

GraphPad Prism (Ver. 10.1.2) was used for all statistical analyses. To evaluate the normality of the dataset, the Shapiro–Wilk test was utilized. For data that followed a normal distribution, one-way ANOVA was applied, with subsequent Tukey's test for multiple comparisons. In cases where data did not conform to normal distribution, the Kruskal–Wallis test was employed, followed by Dunn’s test for multiple comparisons. For daily water intake, serum pyruvate and oxytocin concentration, a two-way ANOVA was performed followed by Tukey's test for multiple comparisons. For mice body weight, the Kruskal–Wallis test was employed, followed by Dunn’s test for multiple comparisons.

4.2.9 Availability of data

All 16S rDNA sequences generated in this study can be accessed from NCBI using accession IDs from LC801559 to LC801580.

4.2.10 Ethics approval

All experiments conducted in this study were performed in strict compliance with the guidelines and protocols approved by the Committee for Animal Care and Use at the National Institute of Genetics, under permit numbers R4-25 and R5-7.

4.3 Results and discussion

4.3.1 Multiple colonies of *L. reuteri* were isolated which secrete pyruvate

To investigate the effects of *L. reuteri* and pyruvate on non-tame mice, I isolated multiple strains of *L. reuteri* with varying levels of pyruvate secretion and administered these through drinking water. The bacteria were isolated from the cecum material and faeces of the selected group of WHS mice. Using a specialized medium containing raffinose, I successfully isolated 22 strains of *L. reuteri*, of which 16 were found to secrete pyruvate into GAM culture media, as demonstrated through biochemical assays (Figure 4.1). As a positive control, *L. helveticus* JCM1120, a known pyruvate-secreting species, was used (111), was used as a positive control of pyruvate secreting strain. Additionally, the secretion of D-lactate and L-lactate was assessed, given their important homeostatic relationship with pyruvate, as shown in Figure 4.1.

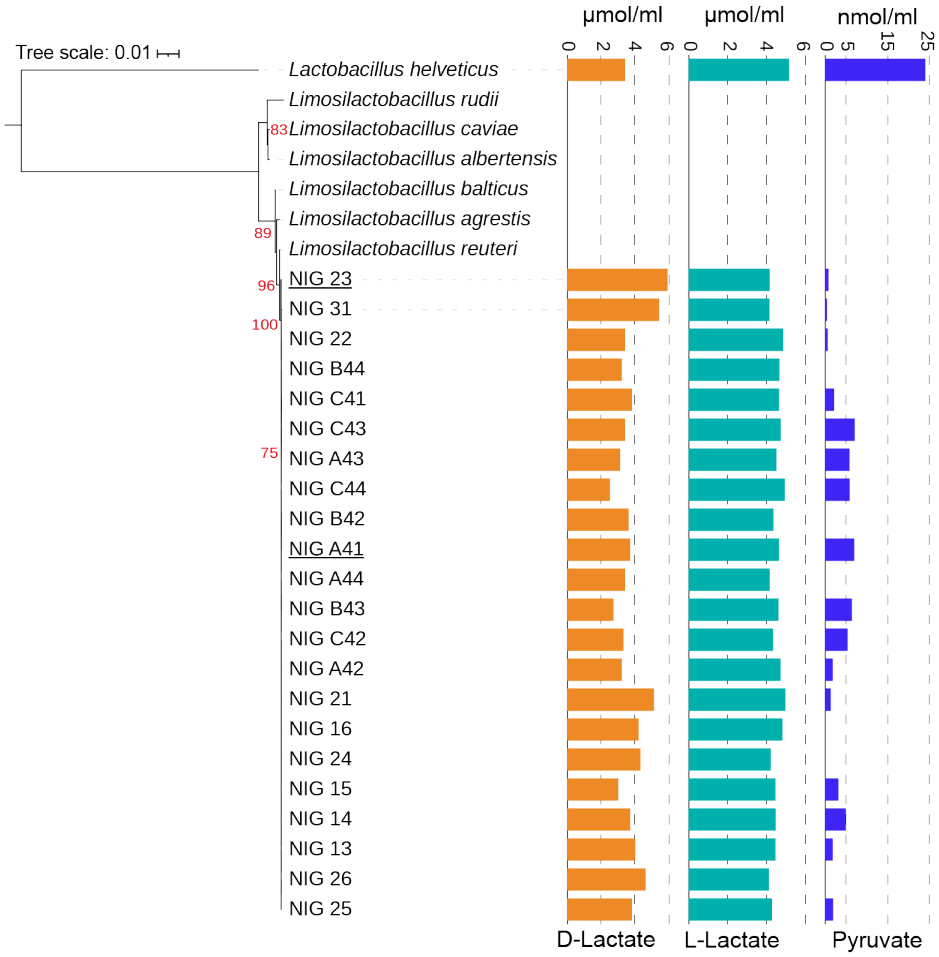


Figure 4.1. Phylogenetic tree *L. reuteri* strains isolated from WHS mice S1 group. Maximum likelihood tree of 16S rDNA region; different species of *Limosilactobacillus* genus was used for identification and *L. helveticus* was used as outgroup. A total of 1000 bootstrap replicates were performed, and bootstrap values above 70% are highlighted in red ink. Additionally, the bar graph in the figure displays the levels of pyruvate and lactate secretion in GAM media by various colonies of *L. reuteri* isolated in this study, compared to *L. helveticus*.

4.3.2 *L. reuteri* administration increases tameness behaviour.

For the experiment involving the administration of bacterial strains to the C1 group of mice, I selected NIG-A41, a high-pyruvate secreting strain, NIG-23, a low-pyruvate secreting strain, and *L. helveticus* JCM1120 was used a control. These cultured bacterial strains were administered to the non-selected C1 group mice through their drinking water over a period of 21 days. Following this period, tameness tests were conducted to evaluate the effects of the bacterial administration on the mice's behaviour (Figure 4.2).

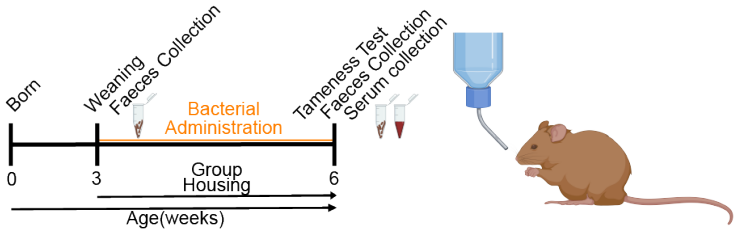


Figure 4.2. Scheme of bacterial administration

Mice treated with the pyruvate-secreting *L. reuteri* strain, NIG-A41, exhibited a significant increase in active tameness compared to the group administered with PBS ($p < 0.05$) and the group treated with *L. helveticus* JCM1120 ($p < 0.05$), as shown in Figure 4.3B. Although not statistically significant, mice treated with the lower pyruvate-secreting strain, NIG-23, also demonstrated higher active tameness compared to the PBS-treated mice. In contrast, the mice treated with *L. helveticus* did not display any increase in active tameness relative to the PBS-treated group. Additionally, the active heading scores were higher in the NIG-A41 treated group, but the difference was not significant (Figure 4.3A). The stay-on-hand test, passive heading, and passive accepting did not exhibit any significant increases in performance, indicating that the most notable effect was on active tameness behaviours (Figure 4.3).

Plasma pyruvate levels were significantly higher in the group treated with *L. helveticus* JCM1120 ($p < 0.05$), yet no significant changes in pyruvate levels were observed in mice treated with either the NIG-A41 or NIG-23 strains of *L. reuteri* compared to the PBS-treated group (Figure 4.4A). I also assessed the colonization of these bacterial strains in the host mice through quantitative PCR analysis of the bacterial genomes extracted from faeces. The results showed significantly higher levels—more than 1500 times on average—of the *L. reuteri* genomes in the faeces of mice treated with both NIG-A41 ($p < 0.01$) and NIG-23 ($p < 0.001$) compared to the PBS-treated mice (Figure 4.4B). In contrast, mice treated with *L. helveticus* JCM1120, which was originally isolated from Emmental (Swiss) cheese (112), did not exhibit a significant increase in bacterial levels in their faeces (Figure 4.4C).

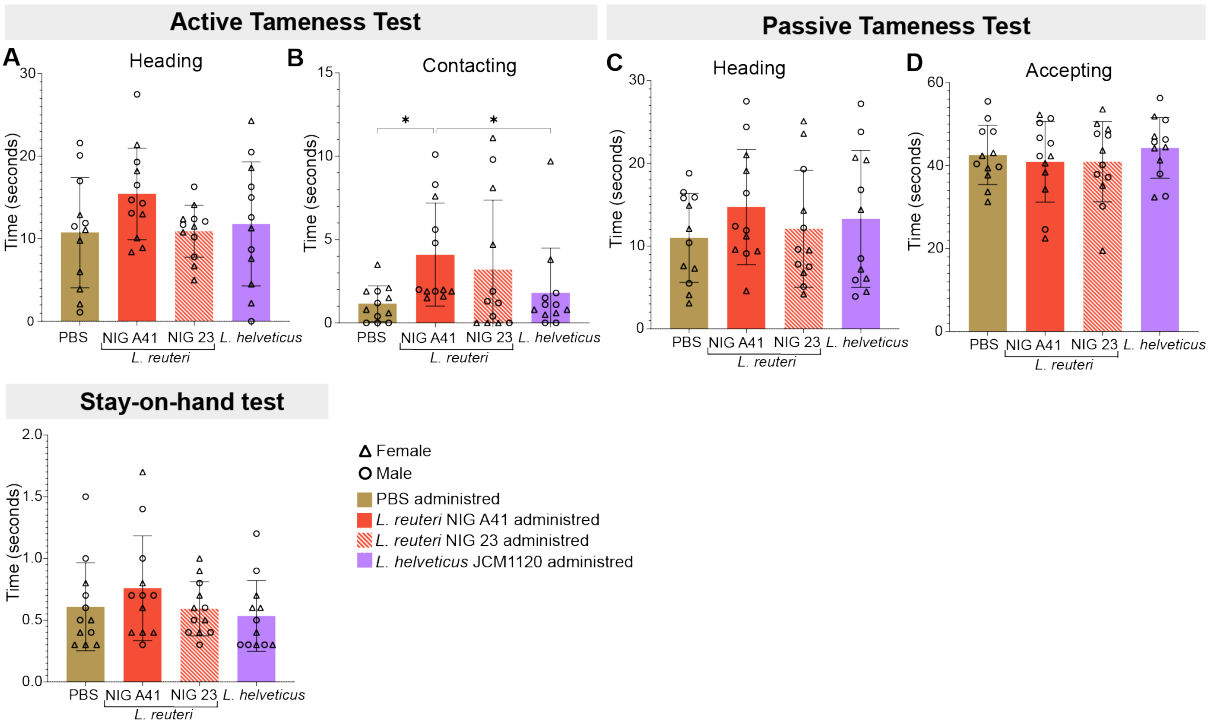


Figure 4.3. Different tameness test parameters after bacterial administration, (A) Active contacting, (B) Active heading, (C) Stay on hand test, (D) Passive heading, (E) passive accepting. N= 48 (12 in each group with 6 male and 6 female). (*p<0.05; **p<0.01; ***p<0.001). Bar graphs show mean ± SD with individual data points. For statistical test, Kruskal–Wallis test was employed, followed by Dunn's test for multiple comparisons.

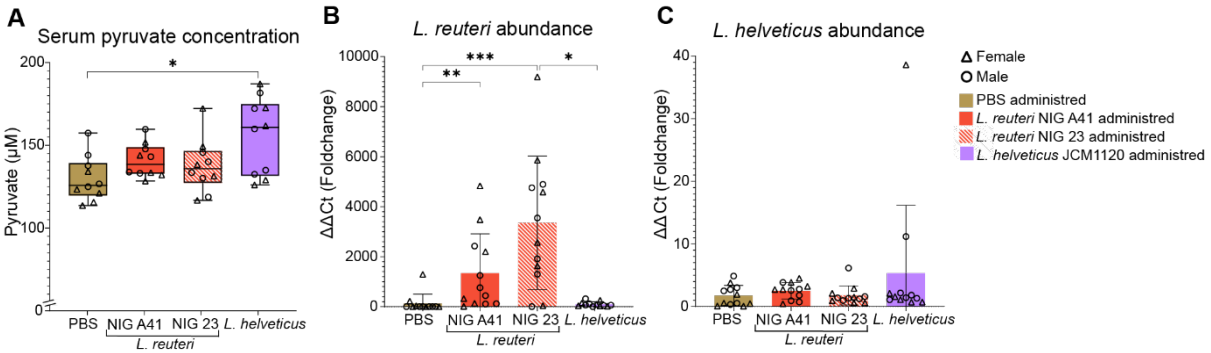


Figure 4.4. Serum pyruvate and qRT-PCR quantification. (A) Serum pyruvate level after bacterial administration, (B) qRT-PCR quantification of *L. reuteri* present in faeces; (C) qRT-PCR quantification of *L. helveticus* present in faeces. N=40 (10 in each group with 5 male and 5 female) for A. N= 48 (12 in each group with 6 male and 6 female) for B and C. (*p<0.05; **p<0.01; ***p<0.001). Bar graphs show means±SD with individual data points. For statistical test, a two-way ANOVA was performed followed by Tukey's test for multiple comparisons. For box plot, the central line represents the median, with boxes showing the first and third quartiles. Whiskers extend to 1.5 times the interquartile range.

Building on previous research that demonstrated daily administration of *L. reuteri* can elevate blood oxytocin levels, I measured oxytocin levels in the mice involved in this experiment. The results showed that serum oxytocin was significantly higher in mice treated with the *L. reuteri* strain NIG-A41 compared to the PBS-treated control group ($p < 0.01$). However, no significant difference in oxytocin levels was found in mice treated with the NIG-23 strain (Figure 4.5A). Furthermore, a mild positive Pearson correlation was observed between serum

oxytocin concentration and active contacting time (R=0.44, p=0.0075), indicating a potential link between oxytocin levels and tameness behaviour (Figure 4.5B).

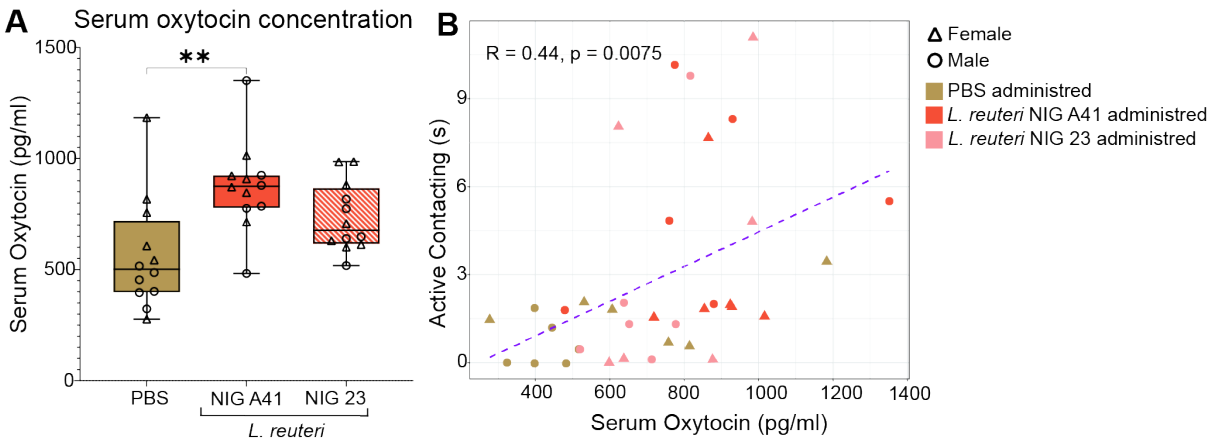


Figure 4.5. Serum oxytocin analysis (A) Serum oxytocin concentration, (B) Pearson correlation between oxytocin concentration and active contacting time. N= 36 (12 in each group with 6 male and 6 female. For box plot, the central line represents the median, with boxes showing the first and third quartiles. Whiskers extend to 1.5 times the interquartile range. For statistical test, a two-way ANOVA was performed followed by Tukey's test for multiple comparisons. (*p<0.05; **p<0.01; ***p<0.001).

During the administration period, I closely monitored changes in body weight and water intake among the different groups, but found no significant differences (Figure 4.6 A-C). These results indicate that the prolonged administration of the NIG-A41 strain of *L. reuteri* not only enhances active tameness but is also associated with elevated serum oxytocin levels and a higher count of faecal *L. reuteri.* This correlation underscores the multifaceted impact of this probiotic strain on both behavioural and physiological parameters in mice.

In conclusion, in this study I successfully isolated a specific strain of *L. reuteri* that is capable of secreting pyruvate, a key metabolic intermediate known to play various roles in cellular processes. More significantly, this strain has shown the ability to enhance active tameness behaviours in mice when administered through their drinking water. This finding not only underscores the complex interplay between microbial metabolites and host behaviour but also highlights the potential of using probiotic interventions to modulate animal behaviour in a controlled manner. The administration of this pyruvate-secreting *L. reuteri* strain has revealed promising results that could have broader implications for improving the management and welfare of domesticated animals. Furthermore, these outcomes contribute to the growing body of evidence supporting the gut-brain axis as a viable target for behavioural modification and suggest avenues for future research into the biochemical pathways involved in such interactions.

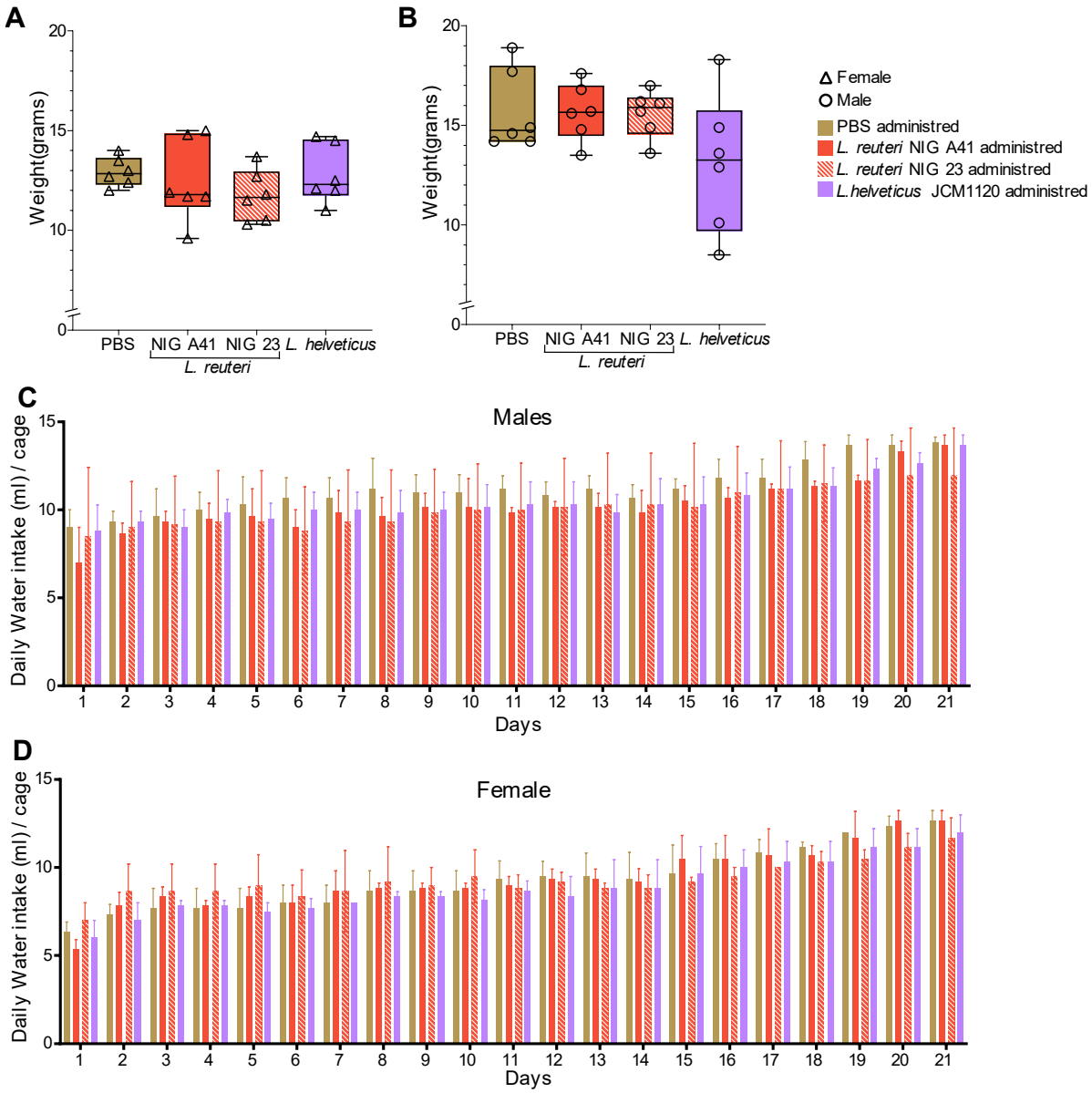
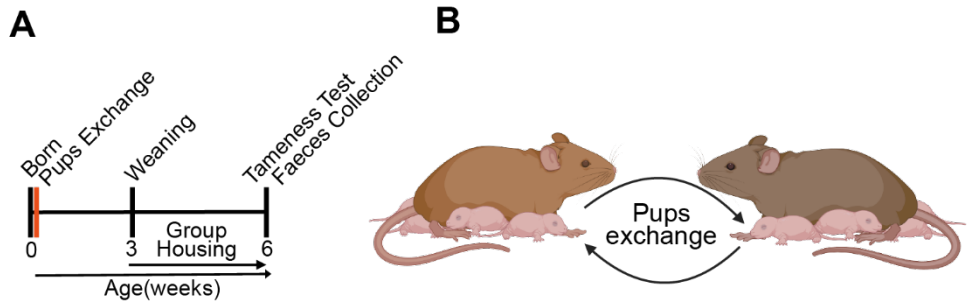


Figure 4.6. No change in body weight and water intake due to bacterial administration, (A) Body weight of females at 6 weeks of age after bacterial administration (B) Body weight of males at 6 weeks of age after bacterial administration; (C) Daily water intake in male for 21 days; (D) Daily water intake in female for 21 days. N= 48 (12 per groups). As mice were kept as same sex pairs during this experiment, so water intake data is from 6 cages per group (3 from males and 3 from females). For box plot, the central line represents the median, with boxes showing the first and third quartiles. Whiskers extend to 1.5 times the interquartile range. For statistical test, Kruskal–Wallis test was employed, followed by Dunn's test for multiple comparisons.

4.3.3 Microbiota exchange by cross-fostering doesn't change tameness behaviour

In the follow-up experiment, I expanded my scope beyond altering a single bacterial species and aimed to exchange the gut microbiota between tame and control mice. Given the potential stress of faecal microbiota transplantation (FMT) in WHS mice, I chose a cross-fostering method. It's well-documented that cross-fostered pups adopt their foster mothers' gut microbiota composition (35-37). In this experiment, I exchanged pups between S1 and C1 mice groups, aiming to study the impact on their gut microbiota (Figure 4.7A-B). After cross foster I had reasonable numbers of pubs to execute this experiment.



C

C1			
	Male	Female	Total
Non fostered	8	11	19
Fostered by S1	4	12	16

S1			
	Male	Female	Total
Non fostered	12	10	22
Fostered by C1	5	6	11

Figure 4.7. (A) Timeline scheme of cross fostering experiment. (B) Scheme of cross-fostering. (C) Number of animals obtained after cross fostering.

Results showed no significant changes in tameness behaviour in both C1 and S1 group mice after cross fostering (Figure 4.8 A-I). This suggests that early-life microbiota, acquired through cross-fostering, doesn't markedly influence tameness in these mice. Furthermore, the lack of observed behavioural changes, combined with no significant differences in taxonomic alpha and beta diversity between the S1 and C1 groups (Figure 3D-E), I decided not to conduct further experiments to verify the actual exchange of microbiome between the groups.

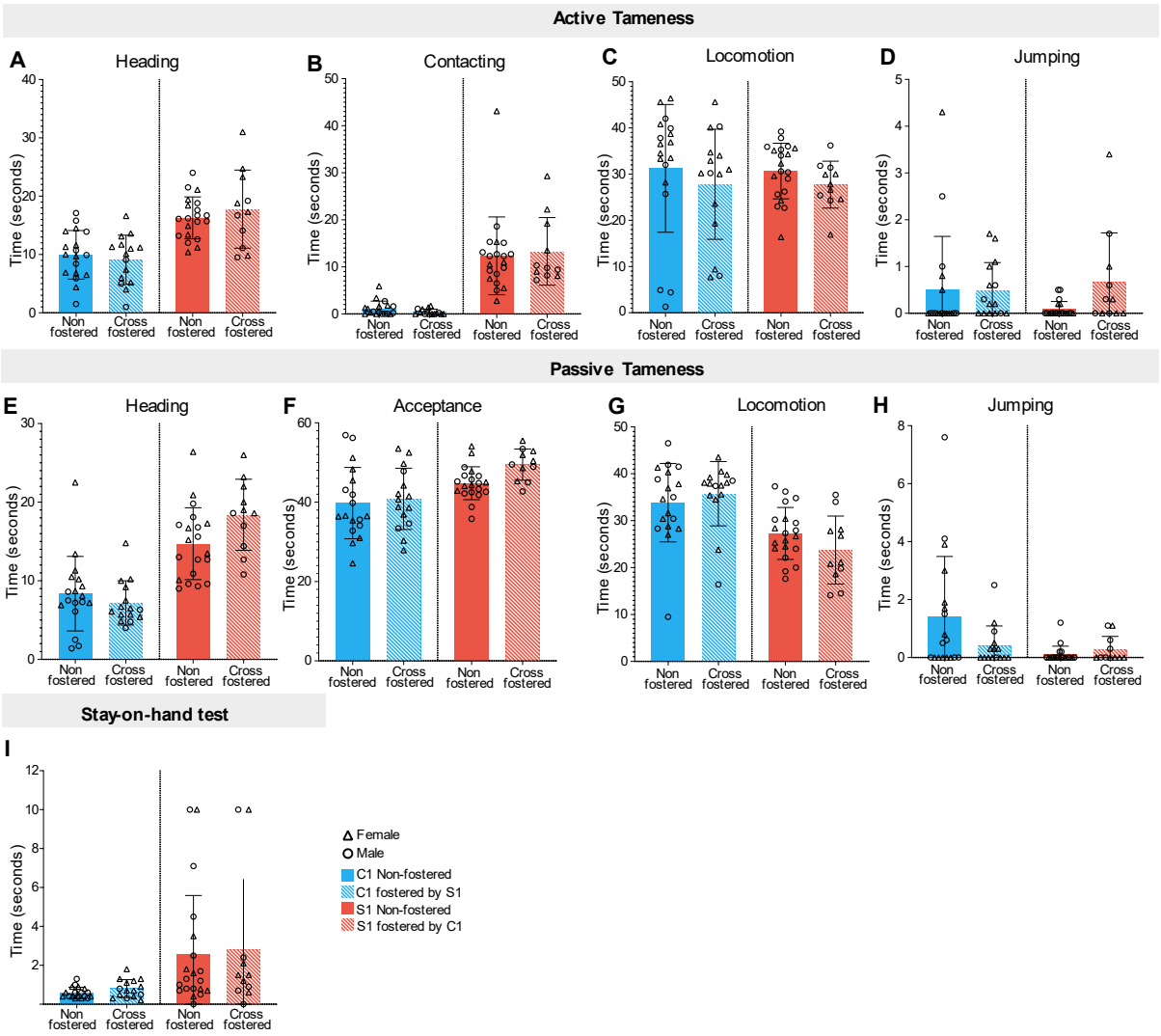


Figure 4.8. No change in tameness behaviour after cross fostering (A-D) active tameness test results, (A) heading, (B) contacting, (C) locomotion, (D) jumping, (E-H) passive tameness test results, (E) heading, (F) accepting, (G) locomotion, (H) jumping, (I) stay on hand test; Bar graphs show mean ± SD with individual data points. For statistical test, Kruskal–Wallis test was employed, followed by Dunn's test for multiple comparisons. N = 68 (C1 control = 19 (8 male, 11 female); C1 fostered by S1 = 16 (4 male, 12 female); S1 control = 22 (12 male, 10 female); S1 fostered by C1 = 11 (5 male, 6 female)).

Chapter 5: General discussion and conclusion

In this study, I investigated the complex relationship between gut microbiota and active tameness behaviour in mice, aiming to uncover potential underlying mechanisms. My research involved analysing the gut microbiota of mice selectively bred for high levels of tameness and comparing them to control groups that exhibited lower tameness. The results showed that although selective breeding for tameness did not significantly impact the taxonomic or functional diversity of the gut microbiota, it did result in a notable increase in the presence of a specific bacterial species, *L. reuteri*, within the tame mice. Additionally, a metabolic analysis of plasma revealed increased levels of pyruvate in these tame mice. Intriguingly, it was observed that chronic intraperitoneal injections of pyruvate could counteract the typical reduction in tameness caused by handling stress. Furthermore, I isolated 22 colonies of *L. reuteri* from S1 of which 16 secrete pyruvate into the culture media. And administering pyruvate secreting *L. reuteri* to control mice not only enhanced their tameness but also led to elevated serum oxytocin levels.

From the analysis of shotgun sequencing data, I successfully assembled 374 non-redundant MAGs, which included 27 novel MAGs spanning five previously unreported phyla. The mouse gut metagenome is well-documented in existing literature (53, 113, 114) and with the addition of my findings, the total number of identified MAGs in the mouse gut has now increased to 1,601. This research also demonstrated that while machine learning binners are highly effective in producing a greater quantity and quality of MAGs, they fall short in capturing all bacterial types effectively. Therefore, I recommend a combined approach that integrates traditional binning techniques with advanced machine learning methods to enhance the comprehensiveness and accuracy of future microbiome research.

The current investigation confirmed that the taxonomic and functional diversity of the gut microbiota was relatively consistent between the selected and non-selected groups of mice. This stability can likely be attributed to the controlled environmental conditions, consistent diet (115-117), and similar genetic background (118, 119). These findings suggest that the selective pressure exerted to enhance active tameness alone may not be sufficient to cause significant shifts in the diversity of the gut microbiota.

Numerous studies have linked specific bacterial genera and species with traits commonly associated with domestication in animals. For instance, in laboratory mice, the bacterial families *Akkermansiaceae, Streptococcaceae*, and *Enterobacteriaceae* are found in greater abundance compared to their wildtype counterparts (21). Similarly, domesticated

horses show a higher prevalence of archaea than feral horses (23). Among buffalo, the *Cyanobacteria* and TM7 phyla are more common in domesticated groups (24). In chickens, selective breeding for low fear has resulted in a higher presence of the bacterial orders *Clostridiales* and *Bacteroidales*, while breeding for high fear has increased the abundance of the order *Lactobacillales* (26). This finding is consistent with my observations, where multiple species of *Lactobacillus* demonstrated a significant positive correlation with increased jumping behaviour (Figure 2.13). Such studies indicate that factors beyond just host genetic selection play critical roles in shaping gut microbiota populations. Although completely isolating external environmental factors in experiments presents challenges, I managed to largely control these variables by maintaining WHS mice in consistent environmental conditions. This study also established a link between selective breeding for tameness and a higher abundance of the bacterial species *L. reuteri* (Figure 2.14).

The gut microbiota plays a crucial role in host energy and lipid metabolism by producing energy metabolites like pyruvate, fumaric acid, and citric acid, and it also affects triglyceride levels in host plasma (120). Specifically, pyruvate, which is secreted by the gut microbiome—particularly by *L. helveticus*, a close relative of *L. reuteri*—is absorbed by the host, enhancing immune system responsiveness and providing protection against *Salmonella* infection (111). Additionally, elevated levels of circulating pyruvate in the blood and brain may help protect against Alzheimer's disease (121) by mitigating age-related cognitive decline (122). In this study, I observed that mice selectively bred for high active tameness exhibited elevated plasma pyruvate levels (Figure 4A). In the repeated IP injections of pyruvate in non-selected mice, I observed no decrease in active tameness scores comparing to pre-injection stage even though such injections are highly stressful for the mice (Figure 3.4). In contrast, the repeated saline injections resulted in a significant reduction in active tameness in non-selected mice. While the exact mechanism through which pyruvate offers this protective effect against stress remains unclear, this study contributes to expand the information that underscores the importance of brain energy metabolism in various neuropsychiatric disorders (123).

I isolated twenty-two strains of *L. reuteri* by culturing colonies on selective media, and discovered that sixteen of these strains secreted pyruvate into GAM medium—a novel finding (111). This discovery is significant for understanding the potential impact of pyruvate secreted by *L. reuteri* on brain function. Previous research has established a connection between the administration of *L. reuteri* and changes in social behaviours in mice (95, 96, 124). In this study, administration of *L. reuteri* was found to enhance active tameness in mice,

though pyruvate secretion was not identified as the direct mechanism behind this behavioural change. Furthermore, bacteria isolated from the same host species demonstrated better colonization, without the sex-based differences in colonization previously reported by Donovan et al. (96). Interestingly, while enhancing *L. reuteri* abundance in mice led to increased tameness, altering the entire gut microbiota through cross-fostering did not have the same effect (Figure 4.8). This discrepancy may be attributed to the initially low native abundance of *L. reuteri* in the S1 group (Figure 2.15C). However, after administration through drinking water, the *L. reuteri* populations surged more than 1500-fold (Figure 5I), which likely contributed to the observed increase in tameness. Additionally, serum oxytocin levels were higher in mice administered *L. reuteri* compared to those given PBS.

My research also highlights the significant impact of host genetics on tameness behaviour. The observed increase in tameness in C1 mice following administration of *L. reuteri* (Figure 4.3B) was ess marked compared to that in the S1 and S2 groups (Figure 2.3B). This finding suggests that while the gut microbiota can influence behaviour, genetic factors inherent in the host due to selective breeding play a more substantial role in modifying tameness. This conclusion is further supported by the results of the cross-fostering experiment, which aimed to modify the gut microbiota of pups to match that of their foster parents (125, 126) ; this intervention did not significantly impact active tameness, as depicted in Figure 4.8. Despite these observations, the fact that long-term administration of live *L. reuteri* through drinking water significantly increased tameness behaviour, primarily by drastically enhancing its population in the gut, stands out as a significant finding. This effect underscores the potential of targeting gut microbiota to modify animal behaviour, which is a promising avenue for future research in the field of behavioural genetics and microbiome studies.

Further investigation into the role of *L. reuteri* in influencing tameness behaviour is essential to clarify the underlying mechanisms at play. Previous research by Matsumoto et al. (8) demonstrated that WHS mice selected for tameness exhibit higher expression of the gene for the oxytocin receptor in the hippocampus. This receptor is crucial for the action of oxytocin, a hormone that promotes social bonding. Moreover, the selection for tameness has been associated with increased sociability in WHS mice (9). Similarly, it has been observed that treating C57BL/6 mice with *L. reuteri* via drinking water raises plasma oxytocin levels (98). This administration also correlated with an increase in oxytocin-positive neurons in the paraventricular nucleus (PVN) of the hypothalamus (95, 124). Additionally, Sgritta et al. (93) demonstrated that *L. reuteri* can communicate with the brain and influence social behaviour

through the vagus nerve. Buffington et al. (94) found that the gut microbiota metabolite tetrahydrobiopterin, enhanced by *L. reuteri*, can alleviate social deficits in germ-free mice, a process linked to increased oxytocin release (127). Furthermore, Danhof et al. (97) discovered that oxytocin is produced and secreted by enterocytes in the intestinal epithelium, triggered by *L. reuteri*. These findings collectively suggest that oxytocin may play a central role in how *L. reuteri* influences tameness behaviour, and by extension, the domestication of animals.

Overall, my findings contribute to the expanding knowledge of the intricate relationships between behaviours, the host genome, and the gut microbiota in the context of animal domestication. The potential link between *L. reuteri* and oxytocin pathways opens new avenues for exploring the biological basis of domestication and tameness in animals. Part (https://www.biorxiv.org/content/10.1101/2024.03.11.584526v1) of the thesis is available in preprint server bioRxiv (128).

Conflict of Interest

Tsuyoshi Koide and Bhim B. Biswa applied for a patent (PCT/JP2024/18885) regarding the findings of this thesis.

References

1. Larson G, Fuller DQ. The evolution of animal domestication. Annual Review of Ecology, Evolution, and Systematics. 2014;45:115-36.
2. Belyaev DK. Domestication of animals. Science. 1969;5(1):47-52.
3. Matsumoto Y, Goto T, Nishino J, Nakaoka H, Tanave A, Takano-Shimizu T, et al. Selective breeding and selection mapping using a novel wild-derived heterogeneous stock of mice revealed two closely-linked loci for tameness. Sci Rep. 2017;7(1):4607.
4. Kukekova AV, Johnson JL, Xiang X, Feng S, Liu S, Rando HM, et al. Red fox genome assembly identifies genomic regions associated with tame and aggressive behaviours. Nat Ecol Evol. 2018;2(9):1479-91.
5. Dou M, Li M, Zheng Z, Chen Q, Wu Y, Wang J, et al. A missense mutation in RRM1 contributes to animal tameness. Sci Adv. 2023;9(25):eadf4068.
6. Knobloch HS, Grinevich V. Evolution of oxytocin pathways in the brain of vertebrates. Front Behav Neurosci. 2014;8:31.
7. Herbeck YE, Eliava M, Grinevich V, MacLean EL. Fear, love, and the origins of canid domestication: An oxytocin hypothesis. Compr Psychoneuroendocrinol. 2022;9:100100.
8. Matsumoto Y, Nagayama H, Nakaoka H, Toyoda A, Goto T, Koide T. Combined change of behavioral traits for domestication and gene-networks in mice selectively bred for active tameness. Genes Brain Behav. 2021;20(3):e12721.
9. Venkatachalam B, Biswa BB, Nagayama H, Koide T. Association of tameness and sociability but no sign of domestication syndrome in mice selectively bred for active tameness. Genes Brain Behav. 2024;23(1):e12887.
10. Olff M, Frijling JL, Kubzansky LD, Bradley B, Ellenbogen MA, Cardoso C, et al. The role of oxytocin in social bonding, stress regulation and mental health: an update on the moderating effects of context and interindividual differences. Psychoneuroendocrinology. 2013;38(9):1883-94.
11. Chen S, Sato S. Role of oxytocin in improving the welfare of farm animals - A review. Asian-Australas J Anim Sci. 2017;30(4):449-54.
12. Price EO. Behavioral development in animals undergoing domestication. Applied Animal Behaviour Science. 2002;65(3):245-71.
13. Ebinghaus A, Ivemeyer S, Knierim U. Human and farm influences on dairy cows responsiveness towards humans - a cross-sectional study. PLoS One. 2018;13(12):e0209817.
14. Harris SE, Munshi-South J, Obergfell C, O'Neill R. Signatures of rapid evolution in urban and rural transcriptomes of white-footed mice (Peromyscus leucopus) in the New York metropolitan area. PLoS One. 2013;8(8):e74938.
15. Ganeshan K, Chawla A. Warming the mouse to model human diseases. Nat Rev Endocrinol. 2017;13(8):458-65.
16. Czarnoleski M, Cooper BS, Kierat J, Angilletta MJ, Jr. Flies developed small bodies and small cells in warm and in thermally fluctuating environments. J Exp Biol. 2013;216(Pt 15):2896-901.
17. Diamond SE, Chick L, Perez A, Strickler SA, A. MR. Rapid evolution of ant thermal tolerance across an urban-rural temperature cline. Biological Journal of the Linnean Society. 2017;112(2):248-57.
18. Price EO. Animal domestication and behavior: CABI Publishing; 2002.
19. Cryan JF, O'Riordan KJ, Cowan CSM, Sandhu KV, Bastiaanssen TFS, Boehme M, et al. The Microbiota-Gut-Brain Axis. Physiol Rev. 2019;99(4):1877-2013.
20. Morais LH, Schreiber HLt, Mazmanian SK. The gut microbiota-brain axis in behaviour and brain disorders. Nat Rev Microbiol. 2021;19(4):241-55.
21. Bowerman KL, Knowles SCL, Bradley JE, Baltrūnaitė L, Lynch MDJ, Jones KM, et al. Effects of laboratory domestication on the rodent gut microbiome. ISME Communications. 2021;1(1):49.

22. Metcalf JL, Song SJ, Morton JT, Weiss S, Seguin-Orlando A, Joly F, et al. Evaluating the impact of domestication and captivity on the horse gut microbiome. Sci Rep. 2017;7(1):15497.
23. Ang L, Vinderola G, Endo A, Kantanen J, Jingfeng C, Binetti A, et al. Gut Microbiome Characteristics in feral and domesticated horses from different geographic locations. Commun Biol. 2022;5(1):172.
24. Prabhu VR, Wasimuddin, Kamalakkannan R, Arjun MS, Nagarajan M. Consequences of Domestication on Gut Microbiome: A Comparative Study Between Wild Gaur and Domestic Mithun. Front Microbiol. 2020;11:133.
25. Craft J, Eddington H, Christman ND, Pryor W, Chaston JM, Erickson DL, et al. Increased Microbial Diversity and Decreased Prevalence of Common Pathogens in the Gut Microbiomes of Wild Turkeys Compared to Domestic Turkeys. Appl Environ Microbiol. 2022;88(5):e0142321.
26. Puetz LC, Delmont TO, Aizpurua O, Guo C, Zhang G, Katajamaa R, et al. Gut Microbiota Linked with Reduced Fear of Humans in Red Junglefowl Has Implications for Early Domestication. Advanced Genetics. 2021;2(4):2100018.
27. Reese AT, Chadaideh KS, Diggins CE, Schell LD, Beckel M, Callahan P, et al. Effects of domestication on the gut microbiota parallel those of human industrialization. Elife. 2021;10.
28. Biswa BB, Koide T. A Method for Selective Breeding to Domesticate Mice. 2022. In: Behavioral Neurogenetics Neuromethods [Internet]. New York, NY: Humana; [107-22].
29. Goto T, Tanave A, Moriwaki K, Shiroishi T, Koide T. Selection for reluctance to avoid humans during the domestication of mice. Genes Brain Behav. 2013;12(8):760-70.
30. Koide T, Goto T, Takano-Shimizu T. Genomic mixing to elucidate the genetic system of complex traits. Exp Anim. 2012;61(5):503-9.
31. Nagayama H, Matsumoto Y, Tanave A, Nihei M, Goto T, Koide T. Measuring Active and Passive Tameness Separately in Mice. J Vis Exp. 2018(138).
32. Goto T, Tanave A, Koide T. tanaMove: Application for recording key inputs over time. In Selection for reluctance to avoid humans during the domestication of mice (Vol. 12, Number 8, pp. 760–70). Akira Tanave. 2013.
33. Suzuki R, Shimodaira H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics. 2006;22(12):1540-2.
34. Costea PI, Zeller G, Sunagawa S, Pelletier E, Alberti A, Levenez F, et al. Towards standards for human fecal sample processing in metagenomic studies. Nat Biotechnol. 2017;35(11):1069-76.
35. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.
36. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27(2):573-80.
37. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357-9.
38. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 2017;27(5):824-34.
39. Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, et al. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods. 2016;102:3-11.
40. Mallawaarachchi V, Lin Y. MetaCoAG: Binning Metagenomic Contigs via Composition, Coverage and Assembly Graphs. 2022. In: Research in Computational Molecular Biology RECOMB 2022 Lecture Notes in Computer Science() [Internet]. Springer; [70-85].
41. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics. 2016;32(4):605-7.
42. 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;7:e7359.

43. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, et al. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nat Microbiol. 2018;3(7):836-43.
44. Pan S, Zhu C, Zhao XM, Coelho LP. A deep siamese neural network improves metagenome-assembled genomes in microbiome datasets across different environments. Nat Commun. 2022;13(1):2326.
45. Nissen JN, Johansen J, Allesoe RL, Sonderby CK, Armenteros JJA, Gronbech CH, et al. Improved metagenome binning and assembly using deep variational autoencoders. Nat Biotechnol. 2021;39(5):555-60.
46. Vollmers J, Wiegand S, Lenk F, Kaster AK. How clear is our current view on microbial dark matter? (Re-)assessing public MAG & SAG datasets with MDMcleaner. Nucleic Acids Res. 2022;50(13):e76.
47. Orakov A, Fullam A, Coelho LP, Khedkar S, Szklarczyk D, Mende DR, et al. GUNC: detection of chimerism and contamination in prokaryotic genomes. Genome Biol. 2021;22(1):178.
48. 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;25(7):1043-55.
49. 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;11(12):2864-8.
50. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics. 2019.
51. Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS One. 2010;5(3):e9490.
52. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49(W1):W293-W6.
53. Kieser S, Zdobnov EM, Trajkovski M. Comprehensive mouse microbiota genome catalog reveals major difference to its human counterpart. PLoS Comput Biol. 2022;18(3):e1009947.
54. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34(18):3094-100.
55. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol. 2019;20(1):257.
56. Wright RJ, Comeau AM, Langille MGI. From defaults to databases: parameter and database choice dramatically impact the performance of metagenomic taxonomic classification tools. Microb Genom. 2023;9(3).
57. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in metagenomics data. PeerJ Computer Science. 2017:3:e104.
58. Beghini F, McIver LJ, Blanco-Miguez 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;10.
59. Caspi R, Altman T, Dreher K, Fulcher CA, Subhraveti P, Keseler IM, et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res. 2012;40(Database issue):D742-53.
60. Liu C, Cui Y, Li X, Yao M. microeco: an R package for data mining in microbial community ecology. FEMS Microbiol Ecol. 2021;97(2).
61. Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, et al. Vegan: community ecology package, 2.6-2. Vienna (Austria): R Foundation for Statistical Computing. 2022.
62. Liaw A, Wiener M. Classification and regression by randomForest. R news. 2002;2(3):18-22.
63. Kolde R, Kolde MR. Package 'pheatmap'. R package. 2015;1(7):790.
64. Mallick H, Rahnavard A, McIver LJ, Ma S, Zhang Y, Nguyen LH, et al. Multivariable association discovery in population-scale meta-omics studies. PLoS Comput Biol. 2021;17(11):e1009442.

65. Wang DD, Nguyen LH, Li Y, Yan Y, Ma W, Rinott E, et al. The gut microbiome modulates the protective association between a Mediterranean diet and cardiometabolic disease risk. Nat Med. 2021;27(2):333-43.
66. Strazar M, Temba GS, Vlamakis H, Kullaya VI, Lyamuya F, Mmbaga BT, et al. Gut microbiome-mediated metabolism effects on immunity in rural and urban African populations. Nat Commun. 2021;12(1):4845.
67. Froemke RC, Young LJ. Oxytocin, Neural Plasticity, and Social Behavior. Annu Rev Neurosci. 2021;44:359-81.
68. Cryan JF, Dinan TG. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. Nat Rev Neurosci. 2012;13(10):701-12.
69. Dalile B, Van Oudenhove L, Vervliet B, Verbeke K. The role of short-chain fatty acids in microbiota-gut-brain communication. Nat Rev Gastroenterol Hepatol. 2019;16(8):461-78.
70. Foster JA, McVey Neufeld KA. Gut-brain axis: how the microbiome influences anxiety and depression. Trends Neurosci. 2013;36(5):305-12.
71. Strandwitz P. Neurotransmitter modulation by the gut microbiota. Brain Res. 2018;1693(Pt B):128-33.
72. Mayer EA, Knight R, Mazmanian SK, Cryan JF, Tillisch K. Gut microbes and the brain: paradigm shift in neuroscience. J Neurosci. 2014;34(46):15490-6.
73. Sarkar A, Lehto SM, Harty S, Dinan TG, Cryan JF, Burnet PWJ. Psychobiotics and the Manipulation of Bacteria-Gut-Brain Signals. Trends Neurosci. 2016;39(11):763-81.
74. Soga T, Heiger DN. Amino acid analysis by capillary electrophoresis electrospray ionization mass spectrometry. Anal Chem. 2000;72(6):1236-41.
75. Soga T, Ueno Y, Naraoka H, Ohashi Y, Tomita M, Nishioka T. Simultaneous determination of anionic intermediates for Bacillus subtilis metabolic pathways by capillary electrophoresis electrospray ionization mass spectrometry. Anal Chem. 2002;74(10):2233-9.
76. Soga T, Ohashi Y, Ueno Y, Naraoka H, Tomita M, Nishioka T. Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. J Proteome Res. 2003;2(5):488-94.
77. Yamamoto H, Fujimori T, Sato H, Ishikawa G, Kami K, Ohashi Y. Statistical hypothesis testing of factor loading in principal component analysis and its application to metabolite set enrichment analysis. BMC Bioinformatics. 2014;15:51.
78. Agus A, Clement K, Sokol H. Gut microbiota-derived metabolites as central regulators in metabolic disorders. Gut. 2021;70(6):1174-82.
79. Brown EM, Clardy J, Xavier RJ. Gut microbiome lipid metabolism and its impact on host physiology. Cell Host Microbe. 2023;31(2):173-86.
80. Frazier K, Chang EB. Intersection of the Gut Microbiome and Circadian Rhythms in Metabolism. Trends Endocrinol Metab. 2020;31(1):25-36.
81. Wu J, Wang K, Wang X, Pang Y, Jiang C. The role of the gut microbiome and its metabolites in metabolic diseases. Protein Cell. 2021;12(5):360-73.
82. Varricchi G, Poto R, Ianiro G, Punziano A, Marone G, Gasbarrini A, et al. Gut Microbiome and Common Variable Immunodeficiency: Few Certainties and Many Outstanding Questions. Front Immunol. 2021;12:712915.
83. Yu LW, Agirman G, Hsiao EY. The Gut Microbiome as a Regulator of the Neuroimmune Landscape. Annu Rev Immunol. 2022;40:143-67.
84. Miyauchi E, Shimokawa C, Steimle A, Desai MS, Ohno H. The impact of the gut microbiome on extra-intestinal autoimmune diseases. Nat Rev Immunol. 2023;23(1):9-23.
85. Borre YE, Moloney RD, Clarke G, Dinan TG, Cryan JF. The impact of microbiota on brain and behavior: mechanisms & therapeutic potential. Adv Exp Med Biol. 2014;817:373-403.
86. Johnson KV, Foster KR. Why does the microbiome affect behaviour? Nat Rev Microbiol. 2018;16(10):647-55.

87. Kim SM, Park S, Hwang SH, Lee EY, Kim JH, Lee GS, et al. Secreted Akkermansia muciniphila threonyl-tRNA synthetase functions to monitor and modulate immune homeostasis. Cell Host Microbe. 2023;31(6):1021-37 e10.
88. Morita H, Kano C, Ishii C, Kagata N, Ishikawa T, Hirayama A, et al. Bacteroides uniformis and its preferred substrate, alpha-cyclodextrin, enhance endurance exercise performance in mice and human males. Sci Adv. 2023;9(4):eadd2120.
89. Kasahara K, Kerby RL, Zhang Q, Pradhan M, Mehrabian M, Lusis AJ, et al. Gut bacterial metabolism contributes to host global purine homeostasis. Cell Host Microbe. 2023;31(6):1038-53 e10.
90. Chevalier G, Siopi E, Guenin-Mace L, Pascal M, Laval T, Rifflet A, et al. Effect of gut microbiota on depressive-like behaviors in mice is mediated by the endocannabinoid system. Nat Commun. 2020;11(1):6363.
91. Dohnalova L, Lundgren P, Carty JRE, Goldstein N, Wenski SL, Nanudorn P, et al. A microbiome-dependent gut-brain pathway regulates motivation for exercise. Nature. 2022;612(7941):739-47.
92. Ritz NL, Brocka M, Butler MI, Cowan CSM, Barrera-Bugueno C, Turkington CJR, et al. Social anxiety disorder-associated gut microbiota increases social fear. Proc Natl Acad Sci U S A. 2024;121(1):e2308706120.
93. Sgritta M, Dooling SW, Buffington SA, Momin EN, Francis MB, Britton RA, et al. Mechanisms Underlying Microbial-Mediated Changes in Social Behavior in Mouse Models of Autism Spectrum Disorder. Neuron. 2019;101(2):246-59 e6.
94. Buffington SA, Dooling SW, Sgritta M, Noecker C, Murillo OD, Felice DF, et al. Dissecting the contribution of host genetics and the microbiome in complex behaviors. Cell. 2021;184(7):1740-56 e16.
95. Dooling SW, Sgritta M, Wang IC, Duque A, Costa-Mattioli M. The Effect of Limosilactobacillus reuteri on Social Behavior Is Independent of the Adaptive Immune System. mSystems. 2022:e0035822.
96. Donovan M, Mackey CS, Lynch MDJ, Platt GN, Brown AN, Washburn BK, et al. Limosilactobacillus reuteri administration alters the gut-brain-behavior axis in a sex-dependent manner in socially monogamous prairie voles. Front Microbiol. 2023;14:1015666.
97. Danhof HA, Lee J, Thapa A, Britton RA, Di Rienzi SC. Microbial stimulation of oxytocin release from the intestinal epithelium via secretin signaling. Gut Microbes. 2023;15(2):2256043.
98. Poutahidis T, Kearney SM, Levkovich T, Qi P, Varian BJ, Lakritz JR, et al. Microbial symbionts accelerate wound healing via the neuropeptide hormone oxytocin. PLoS One. 2013;8(10):e78898.
99. Duar RM, Frese SA, Lin XB, Fernando SC, Burkey TE, Tasseva G, et al. Experimental Evaluation of Host Adaptation of Lactobacillus reuteri to Different Vertebrate Species. Appl Environ Microbiol. 2017;83(12).
100. Greisen K, Loeffelholz M, Purohit A, Leong D. PCR primers and probes for the 16S rRNA gene of most species of pathogenic bacteria, including bacteria found in cerebrospinal fluid. J Clin Microbiol. 1994;32(2):335-51.
101. Goldenberger D, Perschil I, Ritzler M, Altwegg M. A simple "universal" DNA extraction procedure using SDS and proteinase K is compatible with direct PCR amplification. PCR Methods Appl. 1995;4(6):368-70.
102. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403-10.
103. Kuraku S, Zmasek CM, Nishimura O, Katoh K. aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. Nucleic Acids Res. 2013;41(Web Server issue):W22-8.
104. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform. 2019;20(4):1160-6.

105. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14(6):587-9.
106. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2020;37(5):1530-4.
107. Lane D. 16s/23s rRNA sequencing. In: Erko Stackebrandt MG, editor. Nucleic Acid Techniques in Bacterial Systematics. West Sussex, United Kingdom: John Wiley & Sons; 1991. p. 115.
108. Muyzer G, de Waal EC, Uitterlinden AG. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol. 1993;59(3):695-700.
109. Matsuda K, Tsuji H, Asahara T, Matsumoto K, Takada T, Nomoto K. Establishment of an analytical system for the human fecal microbiota, based on reverse transcription-quantitative PCR targeting of multicopy rRNA molecules. Appl Environ Microbiol. 2009;75(7):1961-9.
110. Herbel SR, Lauzat B, von Nickisch-Rosenegk M, Kuhn M, Murugaiyan J, Wieler LH, et al. Species-specific quantification of probiotic lactobacilli in yoghurt by quantitative real-time PCR. J Appl Microbiol. 2013;115(6):1402-10.
111. Morita N, Umemoto E, Fujita S, Hayashi A, Kikuta J, Kimura I, et al. GPR31-dependent dendrite protrusion of intestinal CX3CR1(+) cells by bacterial metabolites. Nature. 2019;566(7742):110-4.
112. Orla-Jensen S. The lactic acid bacteria. Copenhagen: Andr. Fred. Host and Son; 1919.
113. Lesker TR, Durairaj AC, Galvez EJC, Lagkouvardos I, Baines JF, Clavel T, et al. An Integrated Metagenome Catalog Reveals New Insights into the Murine Gut Microbiome. Cell Rep. 2020;30(9):2909-22 e6.
114. Xiao L, Feng Q, Liang S, Sonne SB, Xia Z, Qiu X, et al. A catalog of the mouse gut metagenome. Nat Biotechnol. 2015;33(10):1103-8.
115. Troci A, Rausch P, Waschina S, Lieb W, Franke A, Bang C. Long-Term Dietary Effects on Human Gut Microbiota Composition Employing Shotgun Metagenomics Data Analysis. Mol Nutr Food Res. 2023;67(24):e2101098.
116. Kokou F, Sasson G, Nitzan T, Doron-Faigenboim A, Harpaz S, Cnaani A, et al. Host genetic selection for cold tolerance shapes microbiome composition and modulates its response to temperature. Elife. 2018;7.
117. Vanhaecke T, Bretin O, Poirel M, Tap J. Drinking Water Source and Intake Are Associated with Distinct Gut Microbiota Signatures in US and UK Populations. J Nutr. 2022;152(1):171-82.
118. Fan P, Bian B, Teng L, Nelson CD, Driver J, Elzo MA, et al. Host genetic effects upon the early gut microbiota in a bovine model with graduated spectrum of genetic variation. ISME J. 2020;14(1):302-17.
119. Lopera-Maya EA, Kurilshikov A, van der Graaf A, Hu S, Andreu-Sanchez S, Chen L, et al. Effect of host genetics on the gut microbiome in 7,738 participants of the Dutch Microbiome Project. Nat Genet. 2022;54(2):143-51.
120. Velagapudi VR, Hezaveh R, Reigstad CS, Gopalacharyulu P, Yetukuri L, Islam S, et al. The gut microbiota modulates host energy and lipid metabolism in mice. J Lipid Res. 2010;51(5):1101-12.
121. Isopi E, Granzotto A, Corona C, Bomba M, Ciavardelli D, Curcio M, et al. Pyruvate prevents the development of age-dependent cognitive deficits in a mouse model of Alzheimer's disease without reducing amyloid and tau pathology. Neurobiol Dis. 2015;81:214-24.
122. Koivisto H, Leinonen H, Puurula M, Hafez HS, Barrera GA, Stridh MH, et al. Chronic Pyruvate Supplementation Increases Exploratory Activity and Brain Energy Reserves in Young and Middle-Aged Mice. Front Aging Neurosci. 2016;8:41.
123. Morella IM, Brambilla R, More L. Emerging roles of brain metabolism in cognitive impairment and neuropsychiatric disorders. Neurosci Biobehav Rev. 2022;142:104892.
124. Buffington SA, Di Prisco GV, Auchtung TA, Ajami NJ, Petrosino JF, Costa-Mattioli M. Microbial Reconstitution Reverses Maternal Diet-Induced Social and Synaptic Deficits in Offspring. Cell. 2016;165(7):1762-75.

125. Daft JG, Ptacek T, Kumar R, Morrow C, Lorenz RG. Cross-fostering immediately after birth induces a permanent microbiota shift that is shaped by the nursing mother. Microbiome. 2015;3:17.
126. Treichel NS, Prevorsek Z, Mrak V, Kostric M, Vestergaard G, Foesel B, et al. Effect of the Nursing Mother on the Gut Microbiome of the Offspring During Early Mouse Development. Microb Ecol. 2019;78(2):517-27.
127. Ciosek J, Guzek JW. Neurohypophysial function and pteridines: effect of (6R)-5,6,7, 8-tetrahydro-alpha-biopterin on bioassayed hypothalamo-neurohypophysial vasopressin and oxytocin in the rat. Folia Med Cracov. 1992;33(1-4):25-35.
128. Biswa BB, Mori H, Toyoda A, Kurokawa K, Koide T. Enhanced tameness by Limosilactobacillus reuteri from gut microbiota of selectively bred mice. bioRxiv. 2024:2024.03.11.584526.

Supplementary Tables

Table S1: Results of two-way ANOVA on tameness test of wild hetrogenous stock mice

Test	Trait	Effect	Df	Sum sq	Mean sq	F_value	Pr_F	Sig
Active Tameness test	Heading	Interaction	3	185.6	61.85	F (3, 72)=1.450	P=0.2354	
		Group	3	1212	404	F (3, 72)= 9.473	P<0.0001	***
		Sex	1	1.653	1.653	F (1, 72= 0.03876	P=0.8445	
		Residual	72	3071	42.65			
	Locomotion	Interaction	3	195.7	65.22	F (3, 72)= 0.7753	P=0.5116	
		Group	3	221.7	73.9	F (3, 72)= 0.8785	P=0.4564	
		Sex	1	5.778	5.778	F (1, 72)= .06869	P=0.7940	
		Residual	72	6057	84.12			
	Contacting	Interaction	3	115.4	38.47	F (3, 72)= 1.607	P=0.1952	
		Group	3	1930	643.4	F (3, 72)= 26.88	P<0.0001	***
		Sex	1	17.77	17.77	F (1, 72)= 0.7423	P=0.3918	
		Residual	72	1723	23.93			
	Jumping	Interaction	3	14.8	4.935	F (3, 72) = 1.388	P=0.2533	
		Group	3	133.8	44.6	F (3, 72) = 12.55	P<0.0001	***
		Sex	1	4.141	4.141	F (1, 72) = 1.165	P=0.2841	
		Residual	72	255.9	3.555			
Passive Tameness Test	Heading	Interaction	3	66.19	22.06	F (3, 72) = 0.9220	P=0.4347	
		Group	3	1973	657.6	F (3, 72) = 27.48	P<0.0001	***
		Sex	1	40.04	40.04	F (1, 72) = 1.673	P=0.1999	
		Residual	72	1723	23.93			
	Locomotion	Interaction	3	26.88	8.961	F (3, 72) = 0.2770	P=0.8418	
		Group	3	964.6	321.5	F (3, 72) = 9.937	P<0.0001	***
		Sex	1	121.3	121.3	F (1, 72) = 3.748	P=0.0568	
		Residual	72	2330	32.36			
	Accepting	Interaction	3	74.04	24.68	F (3, 72) = 0.3798	P=0.7678	
		Group	3	1777	592.5	F (3, 72) = 9.118	P<0.0001	***
		Sex	1	15.14	15.14	F (1, 72) = 0.2330	P=0.6308	
		Residual	72	4678	64.98			
	Jumping	Interaction	3	11.52	3.839	F (3, 72) = 1.220	P=0.3085	
		Group	3	217.1	72.38	F (3, 72) = 23.01	P<0.0001	***
		Sex	1	2.556	2.556	F (1, 72) = 0.8125	P=0.3704	
		Residual	72	226.5	3.146			
Stay on Hand test		Interaction	3	12.25	4.085	F (3, 72) = 0.7541	P=0.5236	
		Group	3	43.97	14.66	F (3, 72) = 2.706	P=0.0516	
		Sex	1	5.778	5.778	F (1, 72) = 1.067	P=0.3052	
		Residual	72	390	5.417			

Table S2: Details of all MAGs generated

Bin_Name	CheckM		Length (bp)	N50 (bp)	Binner
	completeness	contamination			
MAG01	53.01	3.49	1808608	6276	SemiBin
MAG02	92.18	0.24	2493119	14512	SemiBin
MAG03	95.56	2.02	2179988	14333	SemiBin
MAG04	85.72	3.66	1677723	7617	Maxbin2
MAG05	97.18	1.89	1827041	32067	MetaBat2
MAG06	54.9	0.48	1377323	2509	MetaCoAG
MAG07	98.87	1.71	1825185	31523	MetaBat2
MAG08	91.01	0	2001417	8696	Maxbin2
MAG09	70.36	3.19	1930864	6870	SemiBin
MAG10	95.67	2.38	2004897	40186	Maxbin2
MAG100	93.95	0.4	1740608	22932	VAMB
MAG101	93.71	0.7	2051917	21269	MetaCoAG
MAG102	90.47	0	1745187	44759	SemiBin
MAG103	96.48	0.7	1931494	122403	SemiBin
MAG104	96.48	0	1882717	38408	VAMB
MAG105	96.77	0	1844810	42407	MetaBat2
MAG106	95.97	0	2112956	49648	MetaCoAG
MAG107	93.01	0.34	1978361	66020	VAMB
MAG108	96.37	0	1996664	97010	VAMB
MAG109	81.65	0.4	1660903	45294	Maxbin2
MAG11	98	0	4216954	25825	SemiBin
MAG110	91.01	1.12	1110936	15649	VAMB
MAG111	61.73	1.68	908494	6548	Maxbin2
MAG112	79.14	1.21	1136511	8941	MetaCoAG
MAG113	94.85	0.22	1961227	86503	SemiBin
MAG114	96	0.5	1553796	7415	MetaCoAG
MAG115	86.05	0.84	2453051	35563	MetaCoAG
MAG116	93.63	2.35	2889349	10800	MetaCoAG
MAG117	88.93	1.56	2991523	15761	MetaCoAG
MAG118	98.66	1.68	2150360	25983	Maxbin2
MAG119	80.03	0.84	2356877	10681	MetaBat2
MAG12	97.96	0	3649632	36389	VAMB
MAG120	93.83	1.87	3376215	14656	VAMB
MAG121	90.2	0.95	3013243	19698	Maxbin2
MAG122	97.22	1.51	3497612	20660	VAMB
MAG123	79.16	0.96	2062495	6441	SemiBin
MAG124	96.98	2.01	3166824	14917	VAMB
MAG125	89.51	5.77	3876942	10285	MetaCoAG
MAG126	82.45	6.23	3517388	20494	MetaCoAG
MAG127	89.48	3.81	3256382	25764	MetaCoAG
MAG128	95.9	2.53	3118878	17310	VAMB

MAG129	92.53	0.74	2574159	27757	Maxbin2
MAG13	99.21	0	4658689	82373	VAMB
MAG130	88.57	2.42	3452341	20885	MetaCoAG
MAG131	94.12	0.59	2622944	24942	Maxbin2
MAG132	95.59	0.09	2650918	36113	Maxbin2
MAG133	73.12	2.92	3422968	9797	MetaCoAG
MAG134	97.01	2.57	3510608	32499	SemiBin
MAG135	86.94	2	2958168	8334	SemiBin
MAG136	96.46	1.45	3591192	30745	VAMB
MAG137	98.85	1.26	4201745	68159	VAMB
MAG138	97.22	1.92	3927368	38827	SemiBin
MAG139	99.37	1.25	4228806	48518	VAMB
MAG14	99.63	0	2798621	96069	SemiBin
MAG140	89.99	3.04	3937799	13533	MetaCoAG
MAG141	84.91	4.38	3574670	20084	MetaCoAG
MAG142	72.78	1.34	3017601	8997	SemiBin
MAG143	92.38	1.6	4004963	19605	SemiBin
MAG144	87.29	2.11	3665931	17224	SemiBin
MAG145	53.35	0.57	2414611	5789	SemiBin
MAG146	88.08	3.16	4195012	24917	MetaCoAG
MAG147	88.48	1.76	3114172	7715	MetaCoAG
MAG148	90.33	2.04	3392702	32156	MetaCoAG
MAG149	92.82	1.1	2511212	38173	MetaBat2
MAG15	99.19	0	2747811	34681	SemiBin
MAG150	95.97	1.16	4486997	25324	MetaCoAG
MAG151	92.57	2.03	2682319	20923	MetaCoAG
MAG152	96.38	0.68	4191175	27886	SemiBin
MAG153	93.35	0.63	3678367	19294	MetaCoAG
MAG154	91.22	2.74	3277890	21237	Maxbin2
MAG155	90.82	1.48	3604641	35016	Maxbin2
MAG156	97.26	3.16	3509267	24033	MetaBat2
MAG157	95.25	1.13	2839925	25884	Maxbin2
MAG158	97.13	2.89	3839519	39802	SemiBin
MAG159	87.54	0.57	3952882	24291	MetaCoAG
MAG16	97.38	0.38	2033737	33077	VAMB
MAG160	95.56	1.15	3791773	20040	VAMB
MAG161	97.21	1.87	4228429	33199	SemiBin
MAG162	97.7	1.34	3977025	60428	VAMB
MAG163	97.68	2.59	4433764	36886	SemiBin
MAG164	98.05	1.47	4020825	28733	VAMB
MAG165	96.14	1.87	4088824	22270	SemiBin
MAG166	93.69	2.42	3842099	13783	VAMB
MAG167	93.45	4.54	4083412	40179	Maxbin2
MAG168	87.18	1.9	3251919	33858	MetaCoAG

MAG169	98.06	1.92	4033739	28724	VAMB
MAG17	97.92	0	2292252	55910	SemiBin
MAG170	93.78	2.53	3299762	12592	VAMB
MAG171	83.17	2.11	3075487	14548	MetaCoAG
MAG172	56.37	1.62	1816187	9439	MetaCoAG
MAG173	96.8	4.27	3376341	24885	Maxbin2
MAG174	93.1	2.08	3992010	16783	SemiBin
MAG175	73.81	2.78	2658068	6691	SemiBin
MAG176	98.13	2.37	4801061	62257	VAMB
MAG177	94.74	2.01	3536174	19754	SemiBin
MAG178	91.09	2.3	4159983	29614	SemiBin
MAG179	95.7	1.72	4396173	26320	SemiBin
MAG18	97.23	0	2957697	52207	SemiBin
MAG180	98.05	2.96	3987512	37338	SemiBin
MAG181	96.79	1.44	4109418	25926	SemiBin
MAG182	96.13	1.81	4199262	12990	VAMB
MAG183	97.99	1.38	2695417	41885	Maxbin2
MAG184	96.31	0	2803182	19615	SemiBin
MAG185	96.87	3.08	2936983	36312	MetaBat2
MAG186	95.71	1.3	3824948	24531	SemiBin
MAG187	96.15	1.6	3618322	14038	Maxbin2
MAG188	97.59	1.82	4546894	40812	VAMB
MAG189	94.76	2.19	4098765	13288	Maxbin2
MAG19	98.18	0.75	2244941	53466	SemiBin
MAG190	96.93	0.16	3876395	37935	VAMB
MAG191	86.22	4.82	3211689	12548	MetaCoAG
MAG192	95.73	1.15	3982479	29604	VAMB
MAG193	97.99	0	4105280	31530	VAMB
MAG194	85.23	1.92	3415843	22814	MetaCoAG
MAG195	70.74	5.97	1691205	5852	MetaCoAG
MAG196	83.21	6.63	2529883	10561	MetaCoAG
MAG197	57.23	1.62	2334062	13600	MetaCoAG
MAG198	98.1	1.95	3474516	36700	VAMB
MAG199	95.97	3.24	3889252	36836	Maxbin2
MAG20	97.55	0.57	2037595	26052	SemiBin
MAG200	94.92	1.94	3744799	7492	Maxbin2
MAG201	95.89	2.74	2674111	20474	Maxbin2
MAG202	88.27	3.06	3182237	11644	MetaCoAG
MAG203	85.86	0	2539020	24359	MetaCoAG
MAG204	91.22	4.11	3221136	10744	MetaCoAG
MAG205	77.99	2.53	2815839	22464	MetaCoAG
MAG206	84.16	6.96	3644402	29838	MetaCoAG
MAG207	98.88	2.35	4042066	37673	VAMB
MAG208	93.57	0	3366242	10615	VAMB

MAG209	84	4.16	3219322	45215	MetaCoAG
MAG21	97.74	0.13	3126676	46513	VAMB
MAG210	79.56	8.68	2719292	11555	MetaCoAG
MAG211	91.5	2.23	3169148	8535	VAMB
MAG212	97.15	1.27	3297139	19277	VAMB
MAG213	94.38	0.97	3389942	29597	MetaBat2
MAG214	96.12	1.87	3267163	36228	MetaBat2
MAG215	83.8	1.56	3709574	14476	SemiBin
MAG216	88.45	1.66	3665827	16183	Maxbin2
MAG217	90.88	2.05	4421856	31522	Maxbin2
MAG218	90.2	1.46	4064037	17832	Maxbin2
MAG219	70.29	2.62	2211137	11032	MetaCoAG
MAG22	97.36	0	2159831	85176	VAMB
MAG220	57.34	0.73	1731493	4772	SemiBin
MAG221	96.51	0.88	3632000	23213	VAMB
MAG222	61.01	1.15	1735442	5697	SemiBin
MAG223	96.1	1.15	3771879	61971	VAMB
MAG224	93.94	1.22	3504177	15215	SemiBin
MAG225	97.96	1.15	4981275	45718	VAMB
MAG226	98.85	0.66	4055143	51138	VAMB
MAG227	98.08	0.77	3872523	59592	VAMB
MAG228	96.93	0.59	4883295	48918	VAMB
MAG229	98.19	0.86	4502313	44758	VAMB
MAG23	93.96	0.19	2863145	33810	SemiBin
MAG230	97.09	1.41	3903577	47684	VAMB
MAG231	80.42	3.8	3485118	35202	Maxbin2
MAG232	52.77	0.74	1568802	4282	SemiBin
MAG233	79.1	1.27	2834062	12029	SemiBin
MAG234	94.22	3.06	3318960	33991	MetaCoAG
MAG235	84.18	1.58	3300195	34575	MetaCoAG
MAG236	80.08	1.9	3279285	19545	MetaCoAG
MAG237	69.79	4.59	3155333	16622	MetaCoAG
MAG238	61.53	7.81	3173829	12439	MetaCoAG
MAG239	77.95	7.02	3050328	27333	MetaCoAG
MAG24	97.92	0	2408876	46775	SemiBin
MAG240	93.12	6.2	3250111	27656	MetaCoAG
MAG241	54.33	1.46	1694745	11132	MetaCoAG
MAG242	80.12	0	2882967	7946	SemiBin
MAG243	92.63	1.96	3075288	20482	Maxbin2
MAG244	52.78	1.27	2271824	14018	MetaCoAG
MAG245	84.45	6.33	3798084	16346	MetaCoAG
MAG246	51.95	0.95	1882830	5127	SemiBin
MAG247	95.15	1	2858457	28622	VAMB
MAG248	89.15	0.66	2349770	19173	MetaCoAG

MAG249	94.92	2.38	2957675	21567	SemiBin
MAG25	98.3	0	2457654	59104	VAMB
MAG250	97.13	1.52	3382088	72370	SemiBin
MAG251	72.07	5.21	2313511	9637	MetaCoAG
MAG252	97.59	1.61	4528278	31504	VAMB
MAG253	83.24	5.4	3513973	9475	MetaCoAG
MAG254	92.28	0	2562732	19649	VAMB
MAG255	89.37	0	2612527	10862	VAMB
MAG256	95.3	0.69	2209756	24615	Maxbin2
MAG257	93.81	1.9	2024517	10260	Maxbin2
MAG258	67.12	3.36	2214873	23992	MetaCoAG
MAG259	52.07	0	1445797	7323	MetaCoAG
MAG26	97.55	0.38	2249267	30612	SemiBin
MAG260	92.95	0.8	2689098	14689	SemiBin
MAG261	87.52	0.92	2183116	25415	MetaBat2
MAG262	92.62	1.03	2190727	18387	MetaCoAG
MAG263	91.63	1.34	2693072	18913	MetaCoAG
MAG264	97.32	0.02	1994246	37740	VAMB
MAG265	94.07	0.67	2317317	29490	VAMB
MAG266	95.19	0	2876703	19098	MetaCoAG
MAG267	94.85	0	2935115	16952	SemiBin
MAG268	67.37	3.52	1834687	7332	SemiBin
MAG269	85.23	0.67	2306157	12169	SemiBin
MAG27	94.72	0.75	2334006	27573	SemiBin
MAG270	66.67	1.34	1922095	8177	SemiBin
MAG271	87.36	1.55	2523970	14462	SemiBin
MAG272	97.05	0.34	3035251	34782	SemiBin
MAG273	94.08	2.07	2617752	17635	SemiBin
MAG274	80.15	0.34	1849911	9679	SemiBin
MAG275	91.36	3.36	3121179	22156	SemiBin
MAG276	58.07	0.67	1382363	5886	SemiBin
MAG277	100	0.02	2681410	45399	VAMB
MAG278	97.05	2.01	2953369	19355	VAMB
MAG279	98.47	1.02	3398502	34400	SemiBin
MAG28	97.34	0.09	2365222	28987	SemiBin
MAG280	99.33	0.67	3045815	29211	VAMB
MAG281	94.55	0.58	2622457	14609	VAMB
MAG282	75.42	0	1751811	8044	SemiBin
MAG283	63.28	0	1633975	6120	SemiBin
MAG284	93.61	0	2802331	15047	VAMB
MAG285	87.17	1.34	2426809	19176	SemiBin
MAG286	82.36	0.34	2322071	9655	SemiBin
MAG287	90.02	2.01	2894281	22228	Maxbin2
MAG288	87.55	6.67	2749313	8043	MetaCoAG

MAG289	86.65	0.67	2732737	19749	SemiBin
MAG29	96.42	0.75	2812843	107769	VAMB
MAG290	79.08	0.67	1904857	10962	SemiBin
MAG291	70.54	1.46	2218168	8303	SemiBin
MAG292	75.86	0.34	1894475	11697	SemiBin
MAG293	91.73	1.46	2559916	15773	VAMB
MAG294	91.3	0	2385606	15327	SemiBin
MAG295	93.39	1.01	2450669	18667	SemiBin
MAG296	84.65	1.01	2177875	17484	SemiBin
MAG297	98.15	0.67	2917592	38983	VAMB
MAG298	84.58	1.2	2887920	10996	SemiBin
MAG299	77.56	1.17	2051434	11098	SemiBin
MAG30	99.06	0.38	2435696	74684	VAMB
MAG300	97.65	0.67	2239426	62323	VAMB
MAG301	56.9	0	1818500	9687	SemiBin
MAG302	87.49	3.36	2066671	8217	SemiBin
MAG303	93.75	0.67	2205813	19439	SemiBin
MAG304	66.65	2.85	1853136	9067	SemiBin
MAG305	65.85	1.34	1687461	7098	SemiBin
MAG306	90.9	1.34	2091028	14902	SemiBin
MAG307	90.76	0	2522134	13725	VAMB
MAG308	87.66	0	2209129	17529	SemiBin
MAG309	67.05	1.51	1793601	7880	SemiBin
MAG31	97.65	0.13	2576978	31495	SemiBin
MAG310	65.92	0.89	1779410	9278	SemiBin
MAG311	53.07	0	1487640	7095	SemiBin
MAG312	88.55	1.85	2223176	16228	SemiBin
MAG313	84.65	5.26	2053768	26182	MetaCoAG
MAG314	94.43	1.68	2176154	27649	VAMB
MAG315	90.75	0	2527381	17708	SemiBin
MAG316	90.38	0.67	2741755	20835	SemiBin
MAG317	95.19	1.01	3320120	28456	SemiBin
MAG318	96.06	1.01	2441890	23800	SemiBin
MAG319	89.06	1.01	2271357	11798	SemiBin
MAG32	99.43	0.38	2867694	44237	VAMB
MAG320	92.78	0.67	2384013	16759	VAMB
MAG321	89.43	0.02	2459965	13102	SemiBin
MAG322	83.34	2.01	2127131	11417	SemiBin
MAG323	79.41	2.18	1991494	9687	SemiBin
MAG324	91.02	0	2730728	20307	SemiBin
MAG325	90.25	1.01	2386837	15949	SemiBin
MAG326	74.47	0.67	1794779	16218	SemiBin
MAG327	93.12	2.01	2545660	45246	MetaCoAG
MAG328	92.95	3.44	2535988	21060	MetaBat2

MAG329	85.57	0.04	1744250	50531	MetaBat2
MAG33	99.43	0	2813440	207094	VAMB
MAG330	77.24	3.49	2435015	17561	MetaCoAG
MAG331	97.32	0	2762404	30159	SemiBin
MAG332	95.64	0	2317612	16506	Maxbin2
MAG333	77.59	8.49	1971667	23044	Maxbin2
MAG334	94.03	0.13	2309833	13411	VAMB
MAG335	87.67	1.03	2179477	19017	SemiBin
MAG336	81.37	0.81	2254786	10057	SemiBin
MAG337	78.81	0.53	2246753	5412	MetaCoAG
MAG338	72.68	0.04	1787398	6158	MetaCoAG
MAG339	93.17	0	1840117	36231	MetaCoAG
MAG34	98.66	0	2679340	41156	SemiBin
MAG340	82.77	0	2253962	39767	MetaCoAG
MAG341	95.76	0.69	2995346	50140	Maxbin2
MAG342	80.53	0.02	2035721	21321	MetaCoAG
MAG343	94.11	0.69	2248648	9933	Maxbin2
MAG344	96.53	1.45	2961095	16093	SemiBin
MAG345	98.66	0	1959526	29302	MetaBat2
MAG346	84.81	2.35	3089816	25053	MetaCoAG
MAG347	77.41	4.42	2745337	46443	MetaCoAG
MAG348	76.23	0	1889948	6698	MetaCoAG
MAG349	84.9	0.67	2249681	36876	MetaCoAG
MAG35	99.06	0	2249274	52122	SemiBin
MAG350	92.41	0	2726334	25894	MetaCoAG
MAG351	72.32	0.38	1387982	4816	MetaCoAG
MAG352	99.29	1.42	2816891	55286	Maxbin2
MAG353	63.67	4.76	1364003	3246	MetaCoAG
MAG354	62.54	3.13	1327399	3379	MetaCoAG
MAG355	76.56	3.62	1634629	20355	MetaCoAG
MAG356	82.84	3.6	1797344	14123	MetaCoAG
MAG357	69.23	4.55	1920090	2976	MetaCoAG
MAG358	77.49	0.92	1554167	10900	Maxbin2
MAG359	73.31	3.49	1882556	6784	Maxbin2
MAG36	99.06	0.38	2444534	63111	SemiBin
MAG360	68.43	2.67	1524528	4516	Maxbin2
MAG361	68.12	0.22	1733278	6873	SemiBin
MAG362	85.6	0.64	1645997	38422	VAMB
MAG363	83.08	0.6	1537674	101761	MetaCoAG
MAG364	77.83	0.3	2321741	10973	MetaCoAG
MAG365	58.7	0.85	854266	17703	Maxbin2
MAG366	63.75	0.85	525769	6342	SemiBin
MAG367	66.07	0	611544	41631	VAMB
MAG368	63.4	3.42	1014354	19660	MetaCoAG

MAG369	60.75	0.85	915216	80603	SemiBin
MAG37	98.68	0.38	2259574	88755	VAMB
MAG370	59.72	0.85	856942	20013	SemiBin
MAG371	57.76	0	768301	11863	MetaBat2
MAG372	57.59	0.85	740849	18653	Maxbin2
MAG373	91.25	0.62	2125426	26241	MetaBat2
MAG374	99.24	0.97	4536150	62555	MetaBat2
MAG38	82.67	0	2155960	68358	Maxbin2
MAG39	99.03	0.38	2440325	37692	VAMB
MAG40	97.3	0	2347887	83629	VAMB
MAG41	98.68	0.57	2587464	29266	SemiBin
MAG42	97.92	0.38	2537971	45039	SemiBin
MAG43	99.06	0.13	2344794	131485	VAMB
MAG44	98.56	0	2092950	136651	SemiBin
MAG45	98.56	0.96	2036167	193149	SemiBin
MAG46	99.52	0.48	2151036	70928	VAMB
MAG47	99.28	0.48	1973112	203608	SemiBin
MAG48	98.53	0.48	2193604	109471	VAMB
MAG49	99.04	0	2084397	26654	SemiBin
MAG50	96.43	2.3	2717526	64325	VAMB
MAG51	98.81	0.24	2571584	86090	VAMB
MAG52	99.6	0.07	1827111	232955	SemiBin
MAG53	93.16	1.14	1936716	27045	MetaBat2
MAG54	97.41	3.45	2523649	30059	Maxbin2
MAG55	98.82	0	2490961	104783	SemiBin
MAG56	98.22	0	2113938	176229	VAMB
MAG57	92.7	0.72	3320659	17193	VAMB
MAG58	100	1.57	2116758	32185	Maxbin2
MAG59	94.24	2.02	1842255	5096	MetaCoAG
MAG60	97.17	1.42	2255027	17023	Maxbin2
MAG61	96.51	0.94	2032680	20263	MetaCoAG
MAG62	92.45	0	1652114	25292	MetaBat2
MAG63	66.39	0.92	1535590	4088	SemiBin
MAG64	98.77	0	2158116	61040	SemiBin
MAG65	99.48	0	1959357	22948	SemiBin
MAG66	96.52	0.08	1920270	14999	MetaCoAG
MAG67	96.05	0.62	2111441	15107	MetaCoAG
MAG68	85.39	0.57	1322899	11538	MetaBat2
MAG69	79.68	1.24	827146	5729	Maxbin2
MAG70	90.11	2.25	1428864	9059	MetaCoAG
MAG71	90	0.27	1689639	9884	SemiBin
MAG72	88.2	3.93	1643136	12738	Maxbin2
MAG73	87.59	1.24	1017047	7277	MetaCoAG
MAG74	83.33	0.56	969824	6847	MetaCoAG

MAG75	80.34	0	999631	7392	MetaBat2
MAG76	89.89	0	1420714	14494	MetaBat2
MAG77	90.17	2.25	1451288	6785	MetaCoAG
MAG78	93.82	0.56	1474276	15100	VAMB
MAG79	76.57	2.25	1031096	2435	MetaCoAG
MAG80	80.11	0.05	1072512	6451	MetaBat2
MAG81	95.16	0	1623870	93170	VAMB
MAG82	93.88	0.03	1543441	49271	VAMB
MAG83	96.37	0.03	1780096	105034	VAMB
MAG84	95.16	0	1621441	59880	VAMB
MAG85	87.43	0.03	1400773	29769	VAMB
MAG86	94.35	0.08	1593014	60995	VAMB
MAG87	94.76	0.03	1518716	73870	VAMB
MAG88	95.56	0.03	1941555	126392	VAMB
MAG89	94.49	0.83	1617760	54874	VAMB
MAG90	94.09	0.03	1666257	51338	VAMB
MAG91	88.31	0.55	1557801	46958	Maxbin2
MAG92	92.74	0.46	1463185	30290	Maxbin2
MAG93	85.79	0.1	1384544	32547	MetaCoAG
MAG94	62.62	1.61	1120420	3818	Maxbin2
MAG95	95.16	0.03	1959627	110160	VAMB
MAG96	94.76	0	1565772	58298	VAMB
MAG97	94.35	0.03	1529896	27520	VAMB
MAG98	94.09	0.03	1545399	32818	VAMB
MAG99	91.94	0	1613422	54538	VAMB

Table S3: Taxonomic diversity of all MAGs generated

Bin Name	Phylum	Class	Order	Family	Genus	species	Novel
MAG01	Actinomycetota	Coriobacteriia	Coriobacteriales	Eggerthellaceae	Adlercreutzia	Adlercreutzia caecimuris	No
MAG02	Actinomycetota	Coriobacteriia	Coriobacteriales	Eggerthellaceae	Adlercreutzia	Adlercreutzia muris	No
MAG03	Actinomycetota	Coriobacteriia	Coriobacteriales	Eggerthellaceae	D16-63		No
MAG04	Actinomycetota	Coriobacteriia	Coriobacteriales	Eggerthellaceae	D16-63	D16-63 sp910588095	No
MAG05	Actinomycetota	Coriobacteriia	Coriobacteriales	Eggerthellaceae	D16-63		No
MAG06	Actinomycetota	Coriobacteriia	Coriobacteriales	Eggerthellaceae	MGBC163016	MGBC163016 sp910588945	No
MAG07	Actinomycetota	Coriobacteriia	Coriobacteriales	Eggerthellaceae	MGBC163016		Yes
MAG08	Actinomycetota	Coriobacteriia	Coriobacteriales	Eggerthellaceae	Parvibacter	Parvibacter caecicola	No
MAG09	Actinomycetota	Coriobacteriia	Coriobacteriales	Eggerthellaceae	Enterorhabdus	Enterorhabdus sp009911695	No
MAG10	Actinomycetota	Coriobacteriia	Coriobacteriales	Eggerthellaceae	ZJ304	ZJ304 sp910584145	No
MAG100	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Gallimonas	Gallimonas sp910589015	No
MAG101	Bacillota_A	Clostridia	Christensenellales	Pumilibacteraceae	HGM11386		No
MAG102	Bacillota_A	Clostridia	Christensenellales	Pumilibacteraceae	Anaerocaecibacter	Anaerocaecibacter sp910577995	No
MAG103	Bacillota_A	Clostridia	Christensenellales	Pumilibacteraceae	Anaerocaecibacter	Anaerocaecibacter muris	No
MAG104	Bacillota_A	Clostridia	Christensenellales	Pumilibacteraceae	Anaerocaecibacter		No
MAG105	Bacillota_A	Clostridia	Christensenellales	Pumilibacteraceae	MGBC113161	Pumilibacter sp910579245	No
MAG106	Bacillota_A	Clostridia	Christensenellales	Pumilibacteraceae	MGBC113161	Pumilibacter muris	No
MAG107	Bacillota_A	Clostridia	Christensenellales	Pumilibacteraceae	MGBC157735	MGBC157735 sp910587445	No
MAG108	Bacillota_A	Clostridia	Christensenellales	Pumilibacteraceae	MGBC157735	MGBC157735 sp910588155	No
MAG109	Bacillota_A	Clostridia	Christensenellales	Pumilibacteraceae	MGBC157735	MGBC157735 sp022793815	No
MAG11	Bacteroidota	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	Bacteroides acidifaciens	No
MAG110	Bacillota_A	Clostridia	Christensenellales	UBA1242	Caccovivens	Caccovivens sp910584295	No
MAG111	Bacillota_A	Clostridia	Christensenellales	UBA1242	Caccovivens		Yes
MAG112	Bacillota_A	Clostridia	Christensenellales	UBA1242			No
MAG113	Bacillota_A	Clostridia	Christensenellales	UBA3700	RACS-045	RACS-045 sp910579785	No
MAG114	Bacillota_A	Clostridia	Clostridiales	Clostridiaceae	Dwaynesavagella	Dwaynesavagella sp000270205	No
MAG115	Bacillota_A	Clostridia	Lachnospirales	Anaerotignaceae	Anaerotignum	Anaerotignum sp910576545	No
MAG116	Bacillota_A	Clostridia	Lachnospirales	Anaerotignaceae	Coprocola	Coprocola sp910575055	No
MAG117	Bacillota_A	Clostridia	Lachnospirales	Anaerotignaceae	Coprocola	Coprocola sp910587735	No
MAG118	Bacillota_A	Clostridia	Lachnospirales	Anaerotignaceae	JANFXS01	JANFXS01 sp022791155	No
MAG119	Bacillota_A	Clostridia	Lachnospirales	CAG-274	MGBC114079	MGBC114079 sp910579745	No
MAG12	Bacteroidota	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	Bacteroides sp002491635	No
MAG120	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	14-2	14-2 sp000403315	No
MAG121	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	14-2	14-2 sp009774455	No
MAG122	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	14-2	14-2 sp910575005	No
MAG123	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	14-2	14-2 sp910577785	No
MAG124	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	14-2	14-2 sp910585625	No
MAG125	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	14-2	14-2 sp910585895	No
MAG126	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	14-2	14-2 sp910588285	No
MAG127	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	14-2	14-2 sp910588885	No
MAG128	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	1XD42-69	1XD42-69 sp003612565	No
MAG129	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	1XD42-69	1XD42-69 sp009911505	No

MAG13	Bacteroidota	Bacteroidia	Bacteroidales	Bacteroidaceae	Phocaeicola	Phocaeicola vulgatus	No
MAG130	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	1XD42-69	1XD42-69 sp910588565	No
MAG131	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	1XD42-69	1XD42-69 sp910589105	No
MAG132	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	1XD42-69		No
MAG133	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	1XD42-69	1XD42-69 sp022794015	No
MAG134	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	1XD8-76	1XD8-76 sp910573755	No
MAG135	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	1XD8-76	1XD8-76 sp022799355	No
MAG136	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Acetatifactor	Acetatifactor sp003612485	No
MAG137	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Acetatifactor	Acetatifactor sp011959105	No
MAG138	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Acetatifactor	Acetatifactor sp910578215	No
MAG139	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Acetatifactor	Acetatifactor sp910579755	No
MAG14	Bacteroidota	Bacteroidia	Bacteroidales	Bacteroidaceae	Prevotella	Prevotella sp002933775	No
MAG140	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Acetatifactor	Acetatifactor sp910584435	No
MAG141	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Acetatifactor	Acetatifactor sp910584865	No
MAG142	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Acetatifactor	Acetatifactor sp910585425	No
MAG143	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Acetatifactor	Acetatifactor sp910585665	No
MAG144	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Acetatifactor	Acetatifactor sp910585805	No
MAG145	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Acetatifactor	Acetatifactor sp910588225	No
MAG146	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Acetatifactor	Acetatifactor sp910589035	No
MAG147	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Acetatifactor	Acetatifactor sp910589655	No
MAG148	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Acetatifactor	Acetatifactor sp022793065	No
MAG149	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Agathobacter		No
MAG15	Bacteroidota	Bacteroidia	Bacteroidales	Marinifilaceae	Odoribacter	Odoribacter sp910589025	No
MAG150	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Butyribacter	Butyribacter sp009774235	No
MAG151	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Butyribacter	Butyribacter sp910575435	No
MAG152	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Butyribacter	Butyribacter sp910579085	No
MAG153	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Caccovicinus	Caccovicinus sp009774615	No
MAG154	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Caccovicinus	Caccovicinus sp910575565	No
MAG155	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Caccovicinus	Caccovicinus sp910588895	No
MAG156	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	CAG-303	CAG-303 sp910585215	No
MAG157	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	CAG-56	CAG-56 sp004793585	No
MAG158	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	CAG-95	CAG-95 sp000403495	No
MAG159	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	CAG-95	CAG-95 sp003612395	No
MAG16	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	CAG-485	CAG-485 sp002361155	No
MAG160	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	CAG-95	CAG-95 sp009911035	No
MAG161	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	CAG-95	CAG-95 sp009917455	No
MAG162	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	CAG-95	CAG-95 sp011959285	No
MAG163	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	CAG-95	CAG-95 sp910579425	No
MAG164	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	CAG-95	CAG-95 sp910579685	No
MAG165	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	CAG-95	CAG-95 sp910586025	No
MAG166	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	CAG-95	CAG-95 sp910587295	No
MAG167	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	CAG-95		No
MAG168	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Choladocola	Choladocola sp009774135	No
MAG169	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Choladocola	Choladocola sp009774145	No

MAG17	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	CAG-485	CAG-485 sp002361215	No
MAG170	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Choladocola	Choladocola sp910574965	No
MAG171	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Choladocola	Choladocola sp910575445	No
MAG172	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Choladousia	Choladousia sp003612585	No
MAG173	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Clostridium_Q	Clostridium_Q sp009911305	No
MAG174	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	COE1	COE1 sp000403215	No
MAG175	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	COE1	COE1 sp000403335	No
MAG176	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	COE1	COE1 sp002490665	No
MAG177	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	COE1	COE1 sp003513705	No
MAG178	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	COE1	COE1 sp009774345	No
MAG179	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	COE1	COE1 sp009774395	No
MAG18	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	CAG-485	CAG-485 sp002362485	No
MAG180	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	COE1	COE1 sp910577115	No
MAG181	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	COE1	COE1 sp910585765	No
MAG182	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	COE1	COE1 sp910586585	No
MAG183	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Eubacterium_F	Eubacterium_F sp910574015	No
MAG184	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Eubacterium_F	Eubacterium_F sp910575665	No
MAG185	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Eubacterium_F	Eubacterium_F sp910584245	No
MAG186	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Eubacterium_J	Eubacterium_J sp009774535	No
MAG187	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Eubacterium_J	Eubacterium_J sp910574915	No
MAG188	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Eubacterium_J	Eubacterium_J sp910579075	No
MAG189	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Eubacterium_J	Eubacterium_J sp910586445	No
MAG19	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	CAG-485	CAG-485 sp003979075	No
MAG190	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Kineothrix	Kineothrix sp000403275	No
MAG191	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Kineothrix	Kineothrix sp011958945	No
MAG192	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Kineothrix	Kineothrix sp910578405	No
MAG193	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Kineothrix	Kineothrix sp910588855	No
MAG194	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Lachnospira	Lachnospira sp910589305	No
MAG195	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	MD308	MD308 sp910575105	No
MAG196	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	MD308	MD308 sp910578455	No
MAG197	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Merdisoma	Merdisoma sp910574255	No
MAG198	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	MGBC114844	MGBC114844 sp910579865	No
MAG199	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	MGBC131033	MGBC131033 sp910585655	No
MAG20	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	CAG-485	CAG-485 sp910579305	No
MAG200	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	MGBC136627	MGBC136627 sp910585935	No
MAG201	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	MGBC164599	MGBC164599 sp910575625	No
MAG202	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	MGBC164599	MGBC164599 sp910588825	No
MAG203	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	MGBC164599		No
MAG204	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	MGBC165282	MGBC165282 sp910589265	No
MAG205	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Paralachnospira_A		Yes
MAG206	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Paralachnospira_A		No
MAG207	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	RGIG4057	RGIG4057 sp910577675	No
MAG208	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	RGIG4057		No
MAG209	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	RGIG4284	RGIG4284 sp910576205	No

MAG21	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	CAG-873	CAG-873 sp001689665	No
MAG210	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	RGIG7193	RGIG7193 sp022791755	No
MAG211	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	RGIG8607	RGIG8607 sp910577045	No
MAG212	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	RGIG8607	RGIG8607 sp910588535	No
MAG213	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Roseburia	Roseburia sp910587865	No
MAG214	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	RUG115	RUG115 sp910578885	No
MAG215	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Schaedlerella	Schaedlerella sp000403295	No
MAG216	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Schaedlerella	Schaedlerella sp910574865	No
MAG217	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Schaedlerella	Schaedlerella sp910589385	No
MAG218	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Schaedlerella		No
MAG219	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Sporofaciens	Sporofaciens musculi	No
MAG22	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	CAG-873	CAG-873 sp002490635	No
MAG220	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Sporofaciens	Sporofaciens sp910577685	No
MAG221	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3282	UBA3282 sp003611805	No
MAG222	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3282	UBA3282 sp009774175	No
MAG223	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3282	UBA3282 sp009774215	No
MAG224	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3282	UBA3282 sp009774655	No
MAG225	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3282	UBA3282 sp910575245	No
MAG226	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3282	UBA3282 sp910575775	No
MAG227	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3282	UBA3282 sp910577735	No
MAG228	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3282	UBA3282 sp910578115	No
MAG229	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3282	UBA3282 sp910579395	No
MAG23	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	CAG-873	CAG-873 sp009775265	No
MAG230	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3282	UBA3282 sp910584965	No
MAG231	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3402	UBA3402 sp002358555	No
MAG232	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3402	UBA3402 sp910575585	No
MAG233	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3402	UBA3402 sp910577935	No
MAG234	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3402	UBA3402 sp910579515	No
MAG235	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3402	UBA3402 sp910586845	No
MAG236	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3402	UBA3402 sp910588865	No
MAG237	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3402	UBA3402 sp022796015	Yes
MAG238	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3402		Yes
MAG239	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA3402	UBA3402 sp022792815	No
MAG24	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	CAG-873	CAG-873 sp009775535	No
MAG240	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA7109	UBA7109 sp910575605	No
MAG241	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	UBA7109	UBA7109 sp910576335	No
MAG242	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Ventrimonas	Ventrimonas sp003611875	No
MAG243	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Ventrimonas	Ventrimonas sp009911065	No
MAG244	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Ventrimonas	Ventrimonas sp910577655	No
MAG245	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Ventrimonas	Ventrimonas sp910577765	No
MAG246	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	Ventrimonas	Ventrimonas sp910586205	No
MAG247	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	VSOB01	VSOB01 sp009774625	No
MAG248	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	VSOB01	VSOB01 sp910585205	No
MAG249	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	VSOB01	VSOB01 sp910587635	No

MAG25	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	CAG-873	CAG-873 sp011959565	No
MAG250	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	VSOB01	VSOB01 sp910589075	No
MAG251	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	W0P33-017		No
MAG252	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae			No
MAG253	Bacillota_A	Clostridia	Lachnospirales	Lachnospiraceae	JAAVZI01	JAAVZI01 sp022791575	No
MAG254	Bacillota_A	Clostridia	Oscillospirales	Acutalibacteraceae	Acutalibacter	Acutalibacter muris	No
MAG255	Bacillota_A	Clostridia	Oscillospirales	Acutalibacteraceae	Acutalibacter	Acutalibacter sp009917525	No
MAG256	Bacillota_A	Clostridia	Oscillospirales	Acutalibacteraceae	Acutalibacter	Acutalibacter sp009936035	No
MAG257	Bacillota_A	Clostridia	Oscillospirales	Acutalibacteraceae	Acutalibacter	Acutalibacter sp009936055	No
MAG258	Bacillota_A	Clostridia	Oscillospirales	Acutalibacteraceae	Acutalibacter	Acutalibacter sp910574935	No
MAG259	Bacillota_A	Clostridia	Oscillospirales	Acutalibacteraceae	Acutalibacter	Acutalibacter sp910580045	No
MAG26	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	CAG-873	CAG-873 sp910578365	No
MAG260	Bacillota_A	Clostridia	Oscillospirales	Acutalibacteraceae	Acutalibacter	Acutalibacter sp910587835	No
MAG261	Bacillota_A	Clostridia	Oscillospirales	Acutalibacteraceae	Acutalibacter	Acutalibacter sp910587995	No
MAG262	Bacillota_A	Clostridia	Oscillospirales	Acutalibacteraceae	Acutalibacter		Yes
MAG263	Bacillota_A	Clostridia	Oscillospirales	Acutalibacteraceae	Acutalibacter		Yes
MAG264	Bacillota_A	Clostridia	Oscillospirales	Acutalibacteraceae	Eubacterium_R	Eubacterium_R sp011958665	No
MAG265	Bacillota_A	Clostridia	Oscillospirales	Butyricicoccaceae	Pseudobutyricicoccus	Pseudobutyricicoccus sp910579575	No
MAG266	Bacillota_A	Clostridia	Oscillospirales	Butyricicoccaceae	QXXE01	QXXE01 sp009911355	No
MAG267	Bacillota_A	Clostridia	Oscillospirales	Butyricicoccaceae	QXXE01	QXXE01 sp910589115	No
MAG268	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp000403435	No
MAG269	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp009774015	No
MAG27	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	CAG-873	CAG-873 sp910587235	No
MAG270	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp910575505	No
MAG271	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp910575815	No
MAG272	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp910577405	No
MAG273	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp910579365	No
MAG274	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp910579435	No
MAG275	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp910583695	No
MAG276	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp910583785	No
MAG277	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp910584995	No
MAG278	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp910585245	No
MAG279	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp910585385	No
MAG28	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	CAG-873	CAG-873 sp910587545	No
MAG280	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp910587525	No
MAG281	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp910588005	No
MAG282	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp910588175	No
MAG283	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp910589315	No
MAG284	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter		Yes
MAG285	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp022797935	Yes
MAG286	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter		Yes
MAG287	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp022795445	No
MAG288	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp022791875	No
MAG289	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp022791135	No

MAG29	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	Duncaniella	Duncaniella dubosii	No
MAG290	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter	Dysosmobacter sp022801815	No
MAG291	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Dysosmobacter		No
MAG292	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Enterenecus	Enterenecus sp910575695	No
MAG293	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Enterenecus	Enterenecus sp910585155	No
MAG294	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Enterenecus	Enterenecus sp910585265	No
MAG295	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Enterenecus	Enterenecus sp910587255	No
MAG296	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Enterenecus	Enterenecus sp910588495	No
MAG297	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Enterenecus		No
MAG298	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Enterenecus		No
MAG299	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Enterenecus		No
MAG30	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	Duncaniella	Duncaniella freteri	No
MAG300	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	HGM13010		No
MAG301	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Lawsonibacter	Lawsonibacter sp003612335	No
MAG302	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Lawsonibacter	Lawsonibacter sp910575265	No
MAG303	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Lawsonibacter	Lawsonibacter sp910577805	No
MAG304	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Lawsonibacter	Lawsonibacter sp910578735	No
MAG305	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Lawsonibacter	Lawsonibacter sp910589085	No
MAG306	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Lawsonibacter	Lawsonibacter sp022801435	No
MAG307	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Lawsonibacter	Lawsonibacter sp022796195	No
MAG308	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Lawsonibacter	Lawsonibacter sp022793355	No
MAG309	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Lawsonibacter	Lawsonibacter sp022798855	No
MAG31	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	Duncaniella	Duncaniella muris	No
MAG310	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Lawsonibacter	Lawsonibacter sp022794495	No
MAG311	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Lawsonibacter		No
MAG312	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Marseille-P3106	Marseille-P3106 sp910587345	No
MAG313	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Marseille-P3106	Marseille-P3106 sp022800215	Yes
MAG314	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Marseille-P3106		No
MAG315	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	NSJ-51	NSJ-51 sp910585415	No
MAG316	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Pelethomonas	Pelethomonas sp009774055	No
MAG317	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Pelethomonas	Pelethomonas sp014237695	No
MAG318	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Pelethomonas	Pelethomonas sp910577915	No
MAG319	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Pelethomonas	Pelethomonas sp910579965	No
MAG32	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	Duncaniella	Duncaniella sp001689575	No
MAG320	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Pelethomonas	Pelethomonas sp910587645	No
MAG321	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Pelethomonas	Pelethomonas sp910587785	No
MAG322	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Pelethomonas	Pelethomonas sp910588365	No
MAG323	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Pelethomonas	Pelethomonas sp022804995	Yes
MAG324	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Pelethomonas	Pelethomonas sp022800665	No
MAG325	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Pelethomonas	Pelethomonas sp022797115	No
MAG326	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Pelethomonas	Pelethomonas sp022795905	No
MAG327	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	JAMPGF01		No
MAG328	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	Avoscillospira_A	Avoscillospira_A sp022792945	No
MAG329	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	CAKUEI01	CAKUEI01 sp022793685	No

MAG33	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	Muribaculum	Muribaculum gordoncarteri	No
MAG330	Bacillota_A	Clostridia	Oscillospirales	Oscillospiraceae	CAJFPI01		No
MAG331	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	Anaerotruncus	Anaerotruncus sp000403395	No
MAG332	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	Anaerotruncus	Anaerotruncus sp003612625	No
MAG333	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	Anaerotruncus	Anaerotruncus sp910586875	No
MAG334	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	Angelakisella	Angelakisella sp013316495	No
MAG335	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	Angelakisella	Angelakisella sp910585525	No
MAG336	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	Angelakisella	Angelakisella sp910585535	No
MAG337	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	Angelakisella	Angelakisella sp022795575	No
MAG338	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	Angelakisella	Angelakisella sp022794295	No
MAG339	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	Avimicrobium		No
MAG34	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	Muribaculum	Muribaculum intestinale	No
MAG340	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	Faecivivens	Faecivivens sp022793675	No
MAG341	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	Harryflintia		Yes
MAG342	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	Harryflintia	Harryflintia sp022793895	Yes
MAG343	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	Negativibacillus	Negativibacillus sp910578755	No
MAG344	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	Ruminiclostridium_E	Ruminiclostridium_E sp910588335	No
MAG345	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	UBA1394	UBA1394 sp910588145	No
MAG346	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	UBA1405	UBA1405 sp910589145	No
MAG347	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	UBA1405	UBA1405 sp022800935	No
MAG348	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	UBA866	UBA866 sp022791945	No
MAG349	Bacillota_A	Clostridia	Oscillospirales	Ruminococcaceae	JAAVYW01	JAAVYW01 sp022791315	No
MAG35	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	Muribaculum	Muribaculum sp002473395	No
MAG350	Bacillota_A	Clostridia	Peptostreptococcales	Anaerovoracaceae	Emergencia	Emergencia sp009936045	No
MAG351	Bacillota_A	Clostridia	Peptostreptococcales	Anaerovoracaceae	JAAVYD01	JAAVYD01 sp022790955	No
MAG352	Bacillota_A	Clostridia	Peptostreptococcales	Anaerovoracaceae	RGIG446		No
MAG353	Bacillota_A	Clostridia	TANB77	CAG-508	CAG-269		Yes
MAG354	Bacillota_A	Clostridia	TANB77	CAG-508	CAG-269	CAG-269 sp022794055	Yes
MAG355	Bacillota_A	Clostridia	TANB77	CAG-508	CAG-273		No
MAG356	Bacillota_A	Clostridia	TANB77	CAG-508	Scatovivens		Yes
MAG357	Bacillota_A	Clostridia	TANB77	CAG-508	Scatovivens	Scatovivens sp022790975	Yes
MAG358	Bacillota_A	Clostridia	TANB77	CAG-508	Scatovivens		No
MAG359	Bacillota_A	Clostridia	TANB77	CAG-508	Scatovivens		No
MAG36	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	Paramuribaculum	Paramuribaculum sp001689565	No
MAG360	Bacillota_A	Clostridia	TANB77	CAG-508	CAKPSH01	CAKPSH01 sp022795675	No
MAG361	Bacillota_A	Clostridia	TANB77	CAG-508			No
MAG362	Bacillota_B	Dehalobacteriia	UBA4068	UBA5755	Avidehalobacter	Avidehalobacter sp022797295	No
MAG363	Bacillota_B	Dehalobacteriia	UBA4068	JAAVYR01	JAAVYR01		Yes
MAG364	Bacillota_B	Peptococcia	Peptococcales	Peptococcaceae			No
MAG365	Patescibacteria	Saccharimonadia	Saccharimonadales	Nanosyncoccaceae	Nanosyncoccus	Nanosyncoccus sp003979185	No
MAG366	Patescibacteria	Saccharimonadia	Saccharimonadales	Nanosyncoccaceae	Nanosyncoccus		Yes
MAG367	Patescibacteria	Saccharimonadia	Saccharimonadales	Nanosyncoccaceae	Nanosyncoccus		No
MAG368	Patescibacteria	Saccharimonadia	Saccharimonadales	Nanosyncoccaceae	Nanosyncoccus		No
MAG369	Patescibacteria	Saccharimonadia	Saccharimonadales	Nanosyncoccaceae	Nanosyncoccus		No

MAG37	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	UBA3263	UBA3263 sp001689615	No
MAG370	Patescibacteria	Saccharimonadia	Saccharimonadales	Nanosyncoccaceae	Nanosyncoccus		No
MAG371	Patescibacteria	Saccharimonadia	Saccharimonadales	Nanosyncoccaceae	Nanosyncoccus		No
MAG372	Patescibacteria	Saccharimonadia	Saccharimonadales	Nanosyncoccaceae	Nanosyncoccus		No
MAG373	Pseudomonadota	Gammaproteobacteria	Burkholderiales	Burkholderiaceae	Parasutterella	Parasutterella sp009767915	No
MAG374	Pseudomonadota	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae	Escherichia	Escherichia coli	No
MAG38	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	UBA3263		No
MAG39	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	UBA7173	UBA7173 sp001689485	No
MAG40	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	UBA7173	UBA7173 sp001689685	No
MAG41	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	UBA7173	UBA7173 sp002491305	No
MAG42	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	UBA7173	UBA7173 sp013316675	No
MAG43	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	JAFLTL01	JAFLTL01 sp910578395	No
MAG44	Bacteroidota	Bacteroidia	Bacteroidales	Rikenellaceae	Alistipes	Alistipes sp002428825	No
MAG45	Bacteroidota	Bacteroidia	Bacteroidales	Rikenellaceae	Alistipes	Alistipes sp009774895	No
MAG46	Bacteroidota	Bacteroidia	Bacteroidales	Rikenellaceae	Alistipes	Alistipes sp910575495	No
MAG47	Bacteroidota	Bacteroidia	Bacteroidales	Rikenellaceae	Alistipes	Alistipes sp910580175	No
MAG48	Bacteroidota	Bacteroidia	Bacteroidales	Rikenellaceae	Alistipes	Alistipes sp910587675	No
MAG49	Bacteroidota	Bacteroidia	Bacteroidales	Rikenellaceae	Avirikenella	Avirikenella sp910588925	No
MAG50	Bacteroidota	Bacteroidia	Bacteroidales	UBA932	Cryptobacteroides	Cryptobacteroides sp002298075	No
MAG51	Bacteroidota	Bacteroidia	Bacteroidales	UBA932	Cryptobacteroides	Cryptobacteroides sp009774765	No
MAG52	Campylo-bacterota	Campylobacteria	Campylobacterales	Helicobacteraceae	Helicobacter_D	Helicobacter_D rodentium	No
MAG53	Cyanobacteriota	Vampirovibrionia	Gastranaerophilales	RUG14156	MGBC122484	MGBC122484 sp910578465	No
MAG54	Deferribacterota	Deferribacteres	Deferribacterales	Mucispirillaceae	Mucispirillum	Mucispirillum sp910586745	No
MAG55	Desulfobacterota	Desulfovibrionia	Desulfovibrionales	Desulfovibrionacea e	Desulfovibrio	Desulfovibrio sp910575755	No
MAG56	Desulfobacterota	Desulfovibrionia	Desulfovibrionales	Desulfovibrionacea e	Mailhella	Mailhella sp003512875	No
MAG57	Desulfobacterota	Desulfovibrionia	Desulfovibrionales	Desulfovibrionacea e	Mailhella	Mailhella sp910579055	No
MAG58	Bacillota	Bacilli	Erysipelotrichales	Coprobacillaceae	Fimiplasma	Fimiplasma sp910588165	No
MAG59	Bacillota	Bacilli	Erysipelotrichales	Erysipelotrichaceae	C-19	C-19 sp910576345	No
MAG60	Bacillota	Bacilli	Erysipelotrichales	Erysipelotrichaceae	C-19	C-19 sp910576855	No
MAG61	Bacillota	Bacilli	Erysipelotrichales	Erysipelotrichaceae	C-19	C-19 sp910586285	No
MAG62	Bacillota	Bacilli	Erysipelotrichales	Erysipelotrichaceae	CAJTFG01	CAJTFG01 sp910575705	No
MAG63	Bacillota	Bacilli	Haloplasmatales_A	Turicibacteraceae	Turicibacter	Turicibacter sp002311155	No
MAG64	Bacillota	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	Lactobacillus taiwanensis	No
MAG65	Bacillota	Bacilli	Lactobacillales	Lactobacillaceae	Ligilactobacillus	Ligilactobacillus murinus	No
MAG66	Bacillota	Bacilli	Lactobacillales	Lactobacillaceae	Limosilactobacillus	Limosilactobacillus reuteri_D	No
MAG67	Bacillota	Bacilli	Lactobacillales	Lactobacillaceae	Limosilactobacillus	Limosilactobacillus sp009663775	No
MAG68	Bacillota	Bacilli	RF39	UBA660	Caccenecus	Caccenecus sp910585045	No
MAG69	Bacillota	Bacilli	RF39	UBA660	DUWA01		Yes
MAG70	Bacillota	Bacilli	RF39	UBA660	MGBC102946		Yes
MAG71	Bacillota	Bacilli	RF39	UBA660	MGBC102946		Yes
MAG72	Bacillota	Bacilli	RF39	UBA660	MGBC108787	MGBC108787 sp910583945	No
MAG73	Bacillota	Bacilli	RF39	UBA660	MGBC108787		Yes
MAG74	Bacillota	Bacilli	RF39	UBA660	MGBC108787		No
MAG75	Bacillota	Bacilli	RF39	UBA660	RUG11867		Yes

MAG76	Bacillota	Bacilli	RF39	UBA660	RUG11867		No
MAG77	Bacillota	Bacilli	RF39	UBA660	UBA3789	UBA3789 sp910589335	No
MAG78	Bacillota	Bacilli	RF39	UBA660	UBA5026		No
MAG79	Bacillota	Bacilli	RF39	UBA660	Ventrenecus		Yes
MAG80	Bacillota	Bacilli	RF39	UBA660	CAKPCJ01		Yes
MAG81	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Coproplasma	Coproplasma sp910577555	No
MAG82	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Coproplasma	Coproplasma sp910578195	No
MAG83	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Coproplasma	Coproplasma sp910578565	No
MAG84	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Coproplasma	Coproplasma sp910578765	No
MAG85	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Coproplasma	Coproplasma sp910579525	No
MAG86	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Coproplasma	Coproplasma sp910579705	No
MAG87	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Coproplasma	Coproplasma sp910585705	No
MAG88	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Coproplasma	Coproplasma sp910586035	No
MAG89	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Coproplasma	Coproplasma sp910586065	No
MAG90	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Coproplasma	Coproplasma sp910587005	No
MAG91	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Coproplasma	Coproplasma sp910587365	No
MAG92	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Coproplasma	Coproplasma sp910588645	No
MAG93	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Coproplasma		No
MAG94	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Coproplasma		No
MAG95	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Gallimonas	Gallimonas sp910578065	No
MAG96	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Gallimonas	Gallimonas sp910585595	No
MAG97	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Gallimonas	Gallimonas sp910586675	No
MAG98	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Gallimonas	Gallimonas sp910586825	No
MAG99	Bacillota_A	Clostridia	Christensenellales	Borkfalkiaceae	Gallimonas	Gallimonas sp910588465	No

Table S4: MaAsLin2 association analysis on the MetaCyc pathways

feature	value	coef	stderr	N	N.not.0	pval	qval
PWY 5981: CDP diacylglycerol biosynthesis III	S2	2.93	0.49	40	24	0.0000	0.0001
PWY 241: C4 photosynthetic carbon assimilation cycle NADP ME type	S2	2.73	0.50	40	21	0.0000	0.0002
PWY 7117: C4 photosynthetic carbon assimilation cycle PEPCK type	S2	2.66	0.48	40	21	0.0000	0.0002
PWY 6936: seleno amino acid biosynthesis plants	S2	2.61	0.51	40	26	0.0000	0.0004
PWY 702: L methionine biosynthesis II	S2	2.37	0.49	40	26	0.0000	0.0007
PWY 6549: L glutamine biosynthesis III	S2	2.00	0.40	40	24	0.0000	0.0004
P124 PWY: Bifidobacterium shunt	S2	1.19	0.31	40	21	0.0004	0.0078
PWY 6906: chitin derivatives degradation	S2	0.96	0.41	40	32	0.0255	0.1016
PWY 5497: purine nucleobases degradation II anaerobic	S2	0.84	0.34	40	36	0.0192	0.0845
PWY 7234: inosine 5 phosphate biosynthesis III	S2	0.82	0.28	40	39	0.0068	0.0457
SALVADEHYPOX PWY: adenosine nucleotides degradation II	S2	0.80	0.27	40	40	0.0045	0.0375
UDPNAGSYN PWY: UDP N acetyl D glucosamine biosynthesis I	S2	0.76	0.23	40	40	0.0021	0.0246
HISDEG PWY: L histidine degradation I	S2	0.74	0.15	40	40	0.0000	0.0004
PWY66 399: gluconeogenesis III	S2	0.69	0.38	40	39	0.0778	0.1975
PWY 1269: CMP 3 deoxy D manno octulosonate biosynthesis	S2	0.61	0.23	40	40	0.0118	0.0605
PWY 5659: GDP mannose biosynthesis	S2	0.61	0.20	40	40	0.0046	0.0375
PWY 7977: L methionine biosynthesis IV	S2	0.56	0.19	40	40	0.0061	0.0457
OANTIGEN PWY: O antigen building blocks biosynthesis E coli	S2	0.53	0.19	40	40	0.0078	0.0487
PWY 5030: L histidine degradation III	S2	0.52	0.21	40	40	0.0204	0.0852
PWY 6124: inosine 5 phosphate biosynthesis II	S2	0.50	0.19	40	40	0.0117	0.0605
X1CMET2 PWY: folate transformations III E coli	S2	0.47	0.17	40	40	0.0082	0.0487
PPGPPMET PWY: ppGpp metabolism	S2	0.45	0.25	40	40	0.0729	0.1923
PWY 5695: inosine 5 phosphate degradation	S2	0.43	0.15	40	40	0.0076	0.0487
PWY 6608: guanosine nucleotides degradation III	S2	0.43	0.19	40	40	0.0302	0.1058
PWY 6123: inosine 5 phosphate biosynthesis I	S2	0.41	0.16	40	40	0.0157	0.0727
SER GLYSYN PWY: superpathway of L serine and glycine biosynthesis I	S2	0.41	0.20	40	40	0.0440	0.1334
PWY 7663: gondoate biosynthesis anaerobic	S2	0.40	0.18	40	40	0.0329	0.1113
PWY0 162: superpathway of pyrimidine ribonucleotides de novo biosynthesis	S2	0.39	0.15	40	40	0.0122	0.0605
PWY 5973: cis vaccenate biosynthesis	S2	0.37	0.16	40	40	0.0297	0.1058
PWY 6703: preQ0 biosynthesis	S2	0.36	0.16	40	40	0.0302	0.1058
PWY 6353: purine nucleotides degradation II aerobic	S2	0.34	0.16	40	40	0.0429	0.1334
PWY 7197: pyrimidine deoxyribonucleotide phosphorylation	S2	0.33	0.15	40	40	0.0354	0.1159
RIBOSYN2 PWY: flavin biosynthesis I bacteria and plants	S2	0.32	0.13	40	40	0.0136	0.0657
PANTO PWY: phosphopantothenate biosynthesis I	S2	0.32	0.10	40	40	0.0018	0.0243
PWY 5686: UMP biosynthesis I	S2	0.32	0.10	40	40	0.0033	0.0305
PWY 7790: UMP biosynthesis II	S2	0.32	0.10	40	40	0.0033	0.0305
PWY 7791: UMP biosynthesis III	S2	0.32	0.10	40	40	0.0033	0.0305
PWY 6125: superpathway of guanosine nucleotides de novo biosynthesis II	S2	0.31	0.19	40	40	0.1002	0.2338
PWY 7228: superpathway of guanosine nucleotides de novo biosynthesis I	S2	0.31	0.17	40	40	0.0847	0.2072
PWY 6126: superpathway of adenosine nucleotides de novo biosynthesis II	S2	0.31	0.17	40	40	0.0796	0.1994

PWY 6606: guanosine nucleotides degradation II	S2	0.30	0.18	40	40	0.0992	0.2338
PWY 7229 superpathway of adenosine nucleotides de novo biosynthesis I	S2	0.30	0.14	40	40	0.0447	0.1334
PWY 3841: folate transformations II plants	S2	0.29	0.10	40	40	0.0059	0.0457
ASPASN PWY: superpathway of L aspartate and L asparagine biosynthesis	S2	0.29	0.14	40	40	0.0403	0.1298
PWY 841: superpathway of purine nucleotides de novo biosynthesis I	S2	0.28	0.13	40	40	0.0340	0.1130
TRNA CHARGING PWY: tRNA charging	S2	0.27	0.08	40	40	0.0023	0.0265
PYRIDNUCSYN PWY: NAD de novo biosynthesis I from aspartate	S2	0.27	0.13	40	40	0.0425	0.1334
PWY 3001: superpathway of L isoleucine biosynthesis I	S2	0.26	0.15	40	40	0.0866	0.2093
PWY 7221: guanosine ribonucleotides de novo biosynthesis	S2	0.26	0.09	40	40	0.0087	0.0503
PWY 1042: glycolysis IV	S2	0.24	0.10	40	40	0.0218	0.0884
PWY 6609: adenine and adenosine salvage III	S2	0.24	0.11	40	40	0.0294	0.1058
PEPTIDOGLYCANSYN PWY: peptidoglycan biosynthesis I meso diaminopimelate containing	S2	0.23	0.10	40	40	0.0296	0.1058
PWY 8178: pentose phosphate pathway non oxidative branch II	S2	0.22	0.11	40	40	0.0614	0.1686
PWY 6387: UDP N acetylmuramoyl pentapeptide biosynthesis I meso diaminopimelate containing	S2	0.21	0.11	40	40	0.0540	0.1543
PWY 6121: 5 aminoimidazole ribonucleotide biosynthesis I	S2	0.19	0.08	40	40	0.0261	0.1017
PWY 6122: 5 aminoimidazole ribonucleotide biosynthesis II	S2	0.19	0.12	40	40	0.1101	0.2483
PWY 6277: superpathway of 5 aminoimidazole ribonucleotide biosynthesis	S2	0.19	0.12	40	40	0.1101	0.2483
CALVIN PWY: Calvin Benson Bassham cycle	S2	0.18	0.10	40	40	0.0844	0.2072
PWY 6163: chorismate biosynthesis from 3 dehydroquinate	S2	0.17	0.06	40	40	0.0063	0.0457
ARO PWY: chorismate biosynthesis I	S2	0.15	0.06	40	40	0.0206	0.0852
NONMEVIPP PWY: methylerythritol phosphate pathway I	S2	0.15	0.06	40	40	0.0107	0.0605
DTDPRHAMSYN PWY: dTDP beta L rhamnose biosynthesis	S2	-0.07	0.04	40	40	0.0645	0.1722
GLYCOGENSYNTH PWY: glycogen biosynthesis I from ADP D Glucose	S2	-0.22	0.11	40	40	0.0525	0.1522
PWY 7400: L arginine biosynthesis IV archaebacteria	S2	-0.25	0.13	40	40	0.0597	0.1683
RHAMCAT PWY: L rhamnose degradation I	S2	-0.31	0.12	40	40	0.0122	0.0605
COBALSYN PWY: superpathway of adenosylcobalamin salvage from cobinamide I	S2	-0.33	0.19	40	40	0.0941	0.2248
PWY 7242: D fructuronate degradation	S2	-0.36	0.20	40	40	0.0749	0.1951
PWY0 1296: purine ribonucleosides degradation	S2	-0.41	0.14	40	40	0.0066	0.0457
GLUCUROCAT PWY: superpathway of beta D glucuronosides degradation	S2	-0.42	0.13	40	40	0.0033	0.0305
GLCMANNANAUT PWY: superpathway of N acetylglucosamine N acetylmannosamine and N acetylneuraminate degradation	S2	-0.43	0.10	40	40	0.0001	0.0012
PWY 6731: starch degradation III	S2	-0.47	0.18	40	40	0.0117	0.0605
PWY 7356: thiamine diphosphate salvage IV yeast	S2	-0.50	0.21	40	39	0.0196	0.0846
FERMENTATION PWY: mixed acid fermentation	S2	-0.52	0.29	40	40	0.0771	0.1975
PWY 5505: L glutamate and L glutamine biosynthesis	S2	-0.52	0.21	40	40	0.0158	0.0727
GALACT GLUCUROCAT PWY: superpathway of hexuronide and hexuronate degradation	S2	-0.58	0.21	40	40	0.0080	0.0487
PWY 6897: thiamine diphosphate salvage II	S2	-0.64	0.32	40	39	0.0501	0.1474
PWY 6507: 4 deoxy L threo hex 4 enopyranuronate degradation	S2	-0.65	0.26	40	40	0.0162	0.0729
PWY 7357: thiamine phosphate formation from pyrithiamine and oxythiamine yeast	S2	-0.65	0.31	40	39	0.0446	0.1334
PWY 7315: dTDP N acetylthomosamine biosynthesis	S2	-0.68	0.15	40	40	0.0000	0.0011
THISYNARA PWY: superpathway of thiamine diphosphate biosynthesis III eukaryotes	S2	-0.71	0.32	40	39	0.0318	0.1093
PWY 7345: superpathway of anaerobic sucrose degradation	S2	-0.97	0.29	40	36	0.0019	0.0246
ALLANTOINDEG PWY: superpathway of allantoin degradation in yeast	S2	-1.00	0.61	40	32	0.1063	0.2451

PWY 6292: superpathway of L cysteine biosynthesis mammalian	S2	-1.03	0.53	40	24	0.0615	0.1686
PWY 7198: pyrimidine deoxyribonucleotides de novo biosynthesis IV	S2	-1.04	0.35	40	40	0.0046	0.0375
PWY 621: sucrose degradation III sucrose invertase	S2	-1.07	0.32	40	36	0.0018	0.0243
PWY 6749: CMP legionaminate biosynthesis I	S2	-1.23	0.36	40	38	0.0014	0.0223
PWY 6143: CMP pseudaminate biosynthesis	S2	-1.36	0.36	40	37	0.0006	0.0105
METH ACETATE PWY: methanogenesis from acetate	S2	-1.51	0.33	40	25	0.0000	0.0011
PWY 1269: CMP 3 deoxy D manno octulosonate biosynthesis	S1	0.57	0.16	40	40	0.0011	0.0729
PWY 7234: inosine 5 phosphate biosynthesis III	S1	0.53	0.19	40	40	0.0087	0.0825
PWY 5030: L histidine degradation III	S1	0.52	0.18	40	40	0.0072	0.0729
PWY 7977: L methionine biosynthesis IV	S1	0.40	0.16	40	40	0.0184	0.1492
PWY 5659: GDP mannose biosynthesis	S1	0.40	0.15	40	40	0.0106	0.0920
X1CMET2 PWY: folate transformations III E coli	S1	0.38	0.14	40	40	0.0100	0.0909
PWY 6124: inosine 5 phosphate biosynthesis II	S1	0.37	0.12	40	40	0.0050	0.0729
GALACTUROCAT PWY: D galacturonate degradation I	S1	0.36	0.18	40	40	0.0521	0.2189
HISDEG PWY: L histidine degradation I	S1	0.36	0.16	40	40	0.0283	0.1587
PWY 7199: pyrimidine deoxyribonucleosides salvage	S1	0.36	0.16	40	40	0.0294	0.1590
PWY 6123: inosine 5 phosphate biosynthesis I	S1	0.33	0.11	40	40	0.0040	0.0729
PWY 7663: gondoate biosynthesis anaerobic	S1	0.28	0.12	40	40	0.0259	0.1587
SER GLYSYN PWY: superpathway of L serine and glycine biosynthesis I	S1	0.27	0.12	40	40	0.0377	0.1840
PWY 5973: cis vaccenate biosynthesis	S1	0.26	0.11	40	40	0.0216	0.1492
PWY 6122: 5 aminoimidazole ribonucleotide biosynthesis II	S1	0.24	0.08	40	40	0.0073	0.0729
PWY 6277: superpathway of 5 aminoimidazole ribonucleotide biosynthesis	S1	0.24	0.08	40	40	0.0073	0.0729
PWY 3841: folate transformations II plants	S1	0.23	0.07	40	40	0.0026	0.0729
PWY 5695: inosine 5 phosphate degradation	S1	0.23	0.10	40	40	0.0273	0.1587
PWY 5686: UMP biosynthesis I	S1	0.21	0.07	40	40	0.0056	0.0729
PWY 7790: UMP biosynthesis II	S1	0.21	0.07	40	40	0.0056	0.0729
PWY 7791: UMP biosynthesis III	S1	0.21	0.07	40	40	0.0056	0.0729
PEPTIDOGLYCANSYN PWY: peptidoglycan biosynthesis I meso diaminopimelate containing	S1	0.21	0.09	40	40	0.0213	0.1492
PYRIDNUCSYN PWY: NAD de novo biosynthesis I from aspartate	S1	0.21	0.09	40	40	0.0253	0.1587
PWY 6387: UDP N acetylmuramoyl pentapeptide biosynthesis I meso diaminopimelate containing	S1	0.21	0.09	40	40	0.0284	0.1587
PWY 1042: glycolysis IV	S1	0.20	0.06	40	40	0.0029	0.0729
PWY 6609: adenine and adenosine salvage III	S1	0.19	0.09	40	40	0.0496	0.2176
PWY 7221: guanosine ribonucleotides de novo biosynthesis	S1	0.18	0.08	40	40	0.0212	0.1492
PWY 6385: peptidoglycan biosynthesis III mycobacteria	S1	0.17	0.09	40	40	0.0525	0.2189
PANTO PWY: phosphopantothenate biosynthesis I	S1	0.17	0.08	40	40	0.0364	0.1818
PWY66 429: fatty acid biosynthesis initiation mitochondria	S1	0.17	0.08	40	40	0.0500	0.2176
TRNA CHARGING PWY: tRNA charging	S1	0.17	0.07	40	40	0.0195	0.1492
PWY 6121: 5 aminoimidazole ribonucleotide biosynthesis I	S1	0.14	0.06	40	40	0.0309	0.1629
DTDPRHAMSYN PWY: dTDP beta L rhamnose biosynthesis	S1	-0.17	0.06	40	40	0.0063	0.0729
PWY 6151: S adenosyl L methionine salvage I	S1	-0.24	0.13	40	40	0.0603	0.2459
PWY 6901: superpathway of glucose and xylose degradation	S1	-0.26	0.11	40	40	0.0244	0.1587
GLCMANNANAUT PWY: superpathway of N acetylglucosamine N acetylmannosamine and N acetylneuraminate degradation	S1	-0.27	0.12	40	40	0.0351	0.1800

PWY 7315: dTDP N acetylthomosamine biosynthesis	S1	-0.47	0.15	40	40	0.0029	0.0729
GLUCUROCAT PWY: superpathway of beta D glucuronosides degradation	S1	-0.56	0.18	40	40	0.0041	0.0729
PWY 5384: sucrose degradation IV sucrose phosphorylase	S1	-0.62	0.29	40	39	0.0393	0.1871
PWY 6588: pyruvate fermentation to acetone	S1	-0.63	0.26	40	38	0.0213	0.1492
PWY0 1298: superpathway of pyrimidine deoxyribonucleosides degradation	S1	-0.66	0.32	40	35	0.0464	0.2157
GALACT GLUCUROCAT PWY: superpathway of hexuronide and hexuronate degradation	S1	-0.77	0.38	40	38	0.0497	0.2176
PWY 6507 4 deoxy L threo hex 4 enopyranuronate degradation	S1	-0.79	0.35	40	39	0.0286	0.1587
PWY4FS: 7 phosphatidylglycerol biosynthesis I plastidic	S1	-0.91	0.27	40	40	0.0019	0.0729
PWY4FS: 8 phosphatidylglycerol biosynthesis II non plastidic	S1	-0.91	0.27	40	40	0.0019	0.0729
PHOSLIPSYN PWY: superpathway of phospholipid biosynthesis I bacteria	S1	-1.02	0.25	40	40	0.0002	0.0380
PWY 4041: gamma glutamyl cycle	S1	-1.10	0.35	40	38	0.0035	0.0729
PWY 6143: CMP pseudaminate biosynthesis	S1	-1.23	0.43	40	37	0.0069	0.0729
PYRIDOXSYN PWY: pyridoxal 5 phosphate biosynthesis I	S1	-1.60	0.52	40	21	0.0042	0.0729

Table S5: Quantitative Estimation of Target Metabolites

	Metabolite	Concentration (μM)												MaAsLin2 Comparative Analysis			
		Control						Selected						Selected vs Control			
		01	02	03	04	05	06	07	08	09	10	11	12	coef	stderr	pval	qval
1	2-Hydroxybutyric acid	13	19	8.4	15	27	45	19	17	19	10	44	16	-0.09	7.19	0.99	0.99
2	2-Oxoglutaric acid	57	24	26	47	52	67	42	92	24	52	59	68	10.75	11.80	0.38	0.73
3	2-Oxoisovaleric acid	6.3	5.2	4.7	8.5	13	23	6.6	19	15	9.5	9.7	18	2.83	3.52	0.44	0.75
4	2-Phosphoglyceric acid	1.2	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	2.2	N.D.	N.D.	N.A.	N.A.	N.A.	N.A.
5	3-Hydroxybutyric acid	272	320	152	178	363	242	310	238	202	89	525	269	17.96	67.84	0.80	0.99
6	3-Phosphoglyceric acid	3.4	7.3	6.9	11	2.8	6.4	3.9	3.2	N.D.	9.1	5.7	4.3	-1.95	1.74	0.29	0.73
7	Adenosine	0.10	0.3	0.2	N.D.	0.2	0.07	N.D.	0.06	0.14	0.2	0.2	0.08	-0.03	0.06	0.61	0.88
8	Ala	262	348	434	211	322	197	247	241	215	233	422	378	-6.17	51.30	0.91	0.99
9	Arg	72	104	88	60	77	54	102	80	73	110	92	104	17.79	9.50	0.09	0.52
10	Asn	40	52	46	27	30	30	39	27	35	25	40	42	-2.66	5.09	0.61	0.88
11	Asp	4.2	8.2	9.0	3.2	3.0	3.2	6.6	4.6	2.4	8.0	6.7	4.4	0.31	1.38	0.83	0.99
12	Betaine	52	63	57	34	264	41	68	81	97	76	81	92	-2.56	36.32	0.95	0.99
13	Betaine aldehyde_+H_2O	N.D.	N.D.	0.4	N.D.	N.D.	N.D.	N.D.	N.D.	0.13	N.D.	0.3	0.3	N.A.	N.A.	N.A.	N.A.
14	Carnosine	1.2	0.3	0.5	0.5	0.3	0.4	0.4	0.2	0.2	0.4	0.4	0.5	-0.18	0.15	0.26	0.73
15	Choline	27	29	48	39	37	38	49	52	37	62	55	42	13.20	4.81	0.02	0.50
16	*cis*-Aconitic acid	9.7	8.0	12	13	11	14	14	21	9.3	11	14	18	3.14	1.95	0.14	0.59
17	Citric acid	249	236	277	319	257	339	301	408	222	250	313	416	38.96	36.51	0.31	0.73
18	Citrulline	59	77	67	56	46	54	61	71	60	50	80	78	6.94	6.56	0.31	0.73
19	Creatine	99	157	179	60	88	153	160	130	86	134	183	94	8.49	24.36	0.73	0.99
20	Creatinine	10	15	17	12	10	9.1	12	14	7.3	11	11	15	-0.42	1.74	0.81	0.99
21	Cytidine	1.2	2.6	6.0	3.1	7.8	9.0	6.3	5.0	1.8	4.5	8.2	5.6	0.24	1.54	0.88	0.99
22	Cytosine	0.5	0.14	0.08	1.0	0.2	N.D.	0.4	N.D.	0.2	0.4	0.09	1.0	0.05	0.22	0.84	0.99
23	Dihydroxyacetone phosphate	N.D.	2.7	3.2	1.5	N.D.	N.D.	N.D.	N.D.	N.D.	1.1	2.0	N.D.	N.A.	N.A.	N.A.	N.A.
24	Fructose 6-phosphate	N.D.	N.D.	N.D.	0.6	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.A.	N.A.	N.A.	N.A.
25	Fumaric acid	13	8.6	15	12	22	20	12	15	7.1	10	18	22	-0.86	3.09	0.79	0.99
26	GABA	N.D.	N.D.	0.4	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	0.5	0.5	N.A.	N.A.	N.A.	N.A.
27	Gln	947	648	743	546	747	665	541	768	620	536	670	797	-60.83	71.36	0.41	0.75
28	Glu	31	63	62	23	26	28	45	46	19	61	48	48	6.00	9.47	0.54	0.83
29	Gluconic acid	12	14	76	15	59	23	26	33	9.4	13	28	38	-8.68	12.04	0.49	0.80
30	Glucose 1-phosphate	N.D.	1.2	0.8	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.A.	N.A.	N.A.	N.A.
31	Glucose 6-phosphate	2.4	7.2	5.3	3.0	1.7	3.3	N.D.	1.7	1.4	4.2	4.4	3.8	-1.24	1.12	0.29	0.73
32	Glutathione (GSSG)_divalent	5.6	6.3	14	3.6	2.0	2.8	8.0	4.4	1.5	4.6	3.5	11	-0.16	2.23	0.95	0.99
33	Gly	183	329	235	175	170	168	287	277	262	181	366	290	67.58	35.51	0.09	0.52
34	Glycerol 3-phosphate	6.1	11	21	11	16	13	8.7	13	4.4	13	24	19	0.57	3.55	0.88	0.99
35	Guanosine	0.10	0.12	0.10	N.D.	0.10	N.D.	N.D.	N.D.	N.D.	N.D.	0.09	N.D.	N.A.	N.A.	N.A.	N.A.
36	His	64	67	84	45	80	66	59	58	109	56	86	92	9.05	10.66	0.42	0.75
37	Homoserine	0.4	0.5	0.5	N.D.	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.16	0.08	0.08	0.52

38	Hydroxyproline	15	31	31	13	21	21	35	23	26	34	26	53	10.83	5.54	0.08	0.52
39	Hypoxanthine	6.0	2.2	1.4	N.D.	0.6	N.D.	N.D.	0.3	N.D.	N.D.	3.0	0.8	-1.04	1.05	0.35	0.73
40	Ile	67	98	58	82	155	145	126	95	108	82	131	82	3.28	18.68	0.86	0.99
41	Inosine	7.5	2.1	1.0	N.D.	1.2	N.D.	N.D.	0.8	N.D.	0.2	1.2	0.7	-1.50	1.17	0.23	0.73
42	Isocitric acid	20	20	24	27	21	29	26	38	18	20	27	34	3.47	3.45	0.34	0.73
43	Lactic acid	2,300	4,163	7,888	3,684	5,379	6,765	6,467	6,691	4,153	6,499	12,183	9,936	2625.10	1448.77	0.10	0.53
44	Leu	102	158	106	117	246	217	176	147	168	109	205	146	0.70	28.29	0.98	0.99
45	Lys	152	282	296	182	299	235	169	150	207	190	254	262	-35.51	31.47	0.29	0.73
46	Malic acid	84	55	96	71	123	130	70	84	42	62	111	132	-9.61	18.03	0.61	0.88
47	Met	17	37	35	16	62	29	79	40	18	24	38	41	7.53	11.04	0.51	0.80
48	*N*,*N*-Dimethylglycine	15	19	7.4	4.6	9.7	7.3	20	15	8.6	8.8	14	18	3.56	2.93	0.25	0.73
49	Ornithine	33	70	75	49	102	77	50	48	66	53	124	66	0.16	15.22	0.99	0.99
50	Phe	43	75	46	44	72	67	54	52	80	59	95	59	8.90	9.33	0.36	0.73
51	Phosphoenol pyruvic acid	2.9	6.1	5.2	8.8	1.7	5.7	2.9	2.2	1.7	8.0	5.7	2.0	-1.32	1.45	0.38	0.73
52	Pro	56	109	87	55	79	60	107	79	106	82	118	99	23.99	10.78	0.05	0.52
53	Putrescine	0.3	0.3	0.7	N.D.	0.4	0.6	0.5	0.8	0.7	0.5	0.7	0.8	0.29	0.11	0.03	0.50
54	Pyruvic acid	71	91	129	149	172	192	262	266	142	278	269	272	114.13	28.67	0.00	0.16
55	Ribose 5-phosphate	1.1	1.3	0.6	1.9	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	0.8	N.D.	N.A.	N.A.	N.A.	N.A.
56	Ribulose 5-phosphate	11	16	5.2	5.0	1.7	2.3	1.4	2.8	N.D.	2.4	4.9	5.0	-4.17	2.43	0.12	0.57
57	*S*-Adenosylmethionine	0.2	0.2	0.3	0.2	0.2	0.2	0.2	0.14	0.09	0.08	0.2	0.3	-0.06	0.04	0.14	0.59
58	Sarcosine	14	17	9.5	5.3	14	9.3	14	9.3	9.1	18	21	14	2.69	2.54	0.31	0.73
59	Sedoheptulose 7-phosphate	0.5	0.7	N.D.	1.2	N.D.	N.D.	N.D.	N.D.	N.D.	0.5	N.D.	N.D.	N.A.	N.A.	N.A.	N.A.
60	Ser	130	200	160	75	99	82	112	94	124	115	110	157	-5.64	21.75	0.80	0.99
61	Spermidine	0.2	0.14	0.2	N.D.	0.6	0.3	0.5	0.7	0.2	0.3	0.4	0.2	0.14	0.12	0.26	0.73
62	Succinic acid	18	44	158	30	117	39	36	26	12	31	143	34	-20.80	30.22	0.51	0.80
63	Thr	159	333	161	134	135	111	184	138	142	125	183	194	-11.23	35.04	0.76	0.99
64	Thymidine	1.3	1.4	1.6	N.D.	2.8	2.9	2.7	3.0	1.9	3.8	4.0	2.9	1.35	0.54	0.03	0.50
65	Trp	94	89	43	42	61	56	32	70	65	53	54	51	-9.97	10.51	0.37	0.73
66	Tyr	64	68	109	43	92	65	120	70	55	71	95	99	11.28	13.64	0.43	0.75
67	Uracil	1.9	2.9	8.5	1.7	5.2	5.6	14	3.6	2.8	4.0	10	6.3	2.53	2.11	0.26	0.73
68	Uridine	11	20	53	16	32	24	29	17	9.6	19	44	35	-0.28	8.10	0.97	0.99
69	Val	165	257	164	183	351	287	289	214	242	193	293	220	7.60	35.47	0.83	0.99
70	β-Ala	0.8	0.8	2.2	1.2	1.7	2.2	3.1	2.9	1.2	5.0	1.9	2.1	1.23	0.60	0.07	0.52