# Experimental Factors Influence Diversity Metrics of the Gut Microbiome in Laboratory Mice

**Short title:** Gut Microbiome in Laboratory Mice

Rikeish R. Muralitharan[1,2], Michael E. Nakai[1], Matthew Snelson[1,3], Tenghao Zheng[1], Evany Dinakis[1], Liang Xie[1], Hamdi Jama[1], Madeleine Paterson[1], Waled Shihata[4], Flavia Wassef[5], Antony Vinh[5], Grant R Drummond[5], David M. Kaye[4,6,7], Charles R. Mackay[8,9,10],

Francine Z. Marques[1,3,4]*

[1]Hypertension Research Laboratory, School of Biological Sciences, Faculty of Science, Monash University, Melbourne, Australia; [2]Institute for Medical Research, Ministry of Health Malaysia, Kuala Lumpur, Malaysia; [3]Victorian Heart Institute, Monash University, Melbourne, Australia.; [4]Heart Failure Research Group, Baker Heart and Diabetes Institute, Melbourne, Australia; [5]Department of Microbiology, Anatomy, Physiology, La Trobe University, Melbourne, Australia; [6]Department of Cardiology, Alfred Hospital, Melbourne, Australia; [7]Central Clinical School, Faculty of Medicine Nursing and Health Sciences, Monash University, Melbourne, Australia;[8]Infection and Immunity Program, Monash Biodiscovery Institute, Monash University, Melbourne, Australia; [9]Department of Biochemistry, Monash University, Melbourne, Australia; [10]School of Pharmaceutical Sciences, Shandong Analysis and Test Center, Qilu University of Technology (Shandong Academy of Sciences), Jinan, 250014, China;

**\*Correspondence to**: A/Prof Francine Marques. Hypertension Research Laboratory, School of Biological Sciences, Monash University, 25 Rainforest Walk, Clayton, Victoria 3800, Australia. P: +61-03-99056958. E: Francine.marques@monash.edu

## Abstract

**Introduction**: Gut microbiome studies often overlook experimental factors that could influence gut microbiome diversity and could impact findings. Large-scale studies investigating these experimental factors are lacking. Thus, we aimed to determine which experimental factors influence the gut microbiome diversity in pre-clinical animal model studies.

**Methods:** We extracted DNA and sequenced the V4 region of the 16S rRNA gene of a total of 538 samples from various sections of the gastrointestinal tract of 303 young and aged male and female C57BL/6J mice of three different genotypes on five diets from three animal house facilities. As a proof-of-concept in a disease model, some mice were treated with sham or angiotensin II, a commonly studied agent used as a hypertension model. Some samples were sequenced twice as a matched-comparison group.

**Results:** Using over 17 million sequencing reads, we found that experimental factors such as animal house facility, genotype, diet, age, sex, sampling site, and technical factor (i.e., sequencing batch) affected both α- and β-diversity (weighted and unweighted UniFrac), and were associated with compositional changes in the microbiome at varying magnitude, with diet and sampling site having the largest effect. After adjustment by these factors, treatment with angiotensin II had no impact on α-diversity and was only significant in unweighted UniFrac (presence/absence of bacteria) analyses.

**Conclusion**: Our data identified several key experimental and technical factors that affect the gut microbiome in laboratory mice. Our findings support that not accounting or adjusting for these factors may lead to false-positive discoveries and non-biologically relevant findings in the gut microbiome field.

**Keywords:** microbiome, animal models, QIIME2, 16S, amplicon sequencing, short-chain fatty acids, SCFA

## Introduction

In the last decade, the gut microbiota has emerged as a novel factor contributing to many healthy and pathological states. Case-control microbiome studies are regularly conducted in experimental animal models – although these regularly show associations with a myriad of phenotypes, reproducibility of these findings from animal to human studies remain mostly challenging (Turner, 2018). One plausible explanation for the poor translation from animal to human studies may be that intrinsic (e.g., genetics) and extrinsic (e.g., environment and diet) experimental factors are often overlooked in animal models. Thus, not evaluating or adjusting for these factors can result in misleading biological conclusions.

We performed a search using a combination of keywords and identified small scale studies that have investigated some of these factors individually (Table 1). For example, the effects of diet were investigated in what is now known as conventionalised germ-free mice (Turnbaugh et al., 2008). Feeding conventionalised germ-free with western compared to complex plant polysaccharides diet resulted in significantly different gut bacterial diversity metrics (Turnbaugh et al., 2008). Another study showed that the gut microbiota of the same genotype of mice from two different vendors differed in their gut microbiome profile (Hufeldt et al., 2010). Leveraging the use of publicly available 16S rRNA gene sequencing data, a study performed a meta-analysis on the effect of aging on the gut microbiome in mice (You et al., 2022). However, a significant limitation of this study was the differences in methods between the studies and the relatively small sample size (You et al., 2022). Although important, these studies did not consider the multitude of other factors that, combined, may influence, and confound the gut microbiome in animal facilities. Thus, well-controlled large-scale studies which concomitantly evaluate several of these factors are needed.

Here, using sequencing data from 538 samples from 303 mice, we aimed to investigate the various intrinsic and extrinsic experimental and technical factors that may influence the gut microbiome, and validate the findings in a relevant disease model as a proof-of-concept. We chose the angiotensin II hypertension model, the most used model of hypertension (Jama et al., 2022), previously demonstrated in a small study to lead to changes to the gut microbiome (Cheema and Pluznick, 2019). Our study revealed that the α- and β-diversity metrics of the gut microbiome are influenced by seven factors studied, including the animal house facility, genotype, diet, age, sex, sample acquisition site, and sequencing batch, while treatment with angiotensin II (hypertension phenotype) had a minor effect that was absent once microbial composition was normalised based on abundance. Thus, we provide evidence in a large-scale multi-institute study in mice that planning and controlling for confounding experimental and technical factors is needed to ensure biological significance and meaningful data that is more likely to be replicated and translated.

## Methods

### *Animals*

Male and female mice [wild-type C57BL/6J (genotype 1), and knockout genotypes developed on the C57BL/6J background GPR41/43/109aKO (genotype 2), GPR41/43KO (genotype 3), and GPR65KO (genotype 4)], from three animal house facilities in Melbourne, Australia (Animal Facility 1, Animal Facility 2, and Animal Facility 3), were included in this study (n=302; 253 male and 49 female mice). Samples used came from independent studies where animal ethics approval was previously received. Mice were administered different diets [control 1 (AIN93G, Speciality Feeds); control 2 (Barastoc); high-fibre (SF11-025, Speciality

Feeds); no-fibre (SF09-028, Speciality Feeds); and high-amylose corn starch acetylated and butyrylated (30% HAMSAB, Ingredion, added to AIN93G)]. The only difference between the control 1, high-fibre, and no-fibre diets is the percentage of fibre, while protein and fat content is similar in the diets. Some littermate mice were randomised and underwent minipump surgery containing angiotensin II (0.5 or 0.75 mg/kg body weight/day, n=142) or 0.9% sodium chloride (sham, n=45). Gut microbiota samples were collected from young (10-12 weeks, n=264) and aged (6 months, n=38) mice. Upon euthanasia using $CO_2$ asphyxiation and cervical dislocation, contents were collected from the caecum for all samples. We also collected gastrointestinal contents for sequencing from the small intestine (duodenum, jejunum, ileum) and large intestine (caecum, colon, rectum, faeces) from some mice (n=13). The gastrointestinal contents were snap-frozen in liquid nitrogen and stored at -80$^{\circ}$C until further processing.

### *DNA extraction, Library Preparation, and Sequencing*

All samples were processed in the same laboratory. DNA was extracted using the DNeasy PowerSoil Kit (Qiagen) according to the manufacturer's protocol. The V4 region of the bacterial 16S rRNA was amplified by PCR (Veriti Thermal Cycler, ThermoFisher Scientific) using 20 ng of DNA, Platinum Hot Start PCR master mix (ThermoFisher Scientific), 515F and 806R primers (Bioneer), single-indexing and methods as described in the Earth Microbiome Project (Caporaso et al., 2012). Thus, the microbiome data reported in this study refers to amplicon sequencing data. The quantity and quality of the PCR product were determined using a Qubit (ThermoFisher Scientific). To identify contamination, non-template controls were used, and none showed amplification. PCR products of 240ng per sample were pooled and cleaned using the PureLink PCR Purification kit (ThermoFisher Scientific). Sequencing was performed across six batches in an Illumina MiSeq sequencer, generating 300 bp paired-end reads with 20% PhiX spiked in.

*Bioinformatic Analyses of Gut Microbiome*

We used the QIIME2 framework to analyse the sequence reads (Bolyen et al., 2019). The forward and reverse reads were truncated at base number 220 for forward and 200 for reverse reads. The reads were then denoised, merged, and filtered for chimera using the DADA2 plugin, resulting in an amplicon sequence variant (ASV) table at the single-nucleotide level (Callahan et al., 2016). Using fasttree (Price et al., 2010) from mafft-aligned (Katoh et al., 2002) ASVs, a phylogeny tree was created and then subsampled without replacement to 7,000 reads per sample (via q2-alignment and q2-phylogeny). From the data, we quantified two important indices commonly described in gut microbiome studies, the α- (a within-sample index) and β-diversity (a between-sample index), all generated from the rarefied samples (via q2-diversity). The α-diversity measure was determined using Shannon index, which determines both evenness and richness. α-diversity is better explained by taking into consideration both evenness and richness compared to metrics which only consider evenness or richness; thus, Shannon index was chosen. Shannon index data were regraphed in GraphPad Prism version 9 and outliers removed using robust regression and outlier removal (1%). Two-tailed Mann-Whitney U test (for 2 groups) or Kruskal-Wallis (for 3 or more groups, with adjustment for false discovery rate (FDR) was performed. β-diversity was determined using unweighted (for presence/absence) and weighted Unifrac (presence/absence adjusted by taxa abundance). Unweighted and weighted Unifrac metrics with corresponding Principal Coordinate Analysis (PCoA) distance tables using q2-diversity. ASVs were then labelled by taxonomic assignment using a naïve Bayes classifier (via q2-feature-classifier) (Bokulich et al., 2018), using the SILVA database (version 138) as a reference (Quast et al., 2013). Differences between groups were assessed using permutational multivariate analysis of variance (PERMANOVA) for the overall effect. P-values < 0.05 were considered significant.

MaAsLin2 was used to determine multivariable association between taxa counts and the various experimental factors by generalized linear regression model using trimmed mean of M-values (TMM) normalized data (Mallick et al., 2021). For the discovery of differentially abundant taxa, adjustment was made for animal house facility, genotype, diet, age, sex, sample acquisition site, and treatment with angiotensin II (hypertension phenotype). FDR-adjusted P-values of <0.05 were considered significant.

To determine the explained variance of each experimental factor on the gut microbiota composition in our study, we used VpThemAll (Forslund et al., 2021). This package enabled us to determine the full, shared, and unique variance contributed by the eight experimental factors in our study. FDR-adjusted P-values of <0.02 were considered significant. Adjusted r-squared values that were obtained were converted into percentage of variance.

## Results

### *Sequencing reads and depth*

Over 17,809,928 reads were denoised, merged, and survived chimera filtering, with read counts averaging 31,410 reads per sample (Figure 1a). At a rarefaction depth of 7,000 reads, samples showed consistent plateauing of the various diversity metrics but caused the loss of 29 samples (5.11%) (Figure 1b). We then studied eight experimental factors that could potentially affect the gut microbiome, described individually below.

### *The impact of animal house facilities*

We first studied the influence of animal house facilities by using samples from three distinct facilities in Australia. We determined that animal house facility influences bacterial diversity, both Shannon index (P<0.001, Figure 2a) and unweighted and weighted Unifrac (P=0.001, Figures 2b-c), which show distinct clustering according to the origin of the mice. Pairwise comparisons revealed significant difference between all three facilities in both unweighted and weighted Unifrac (Table S1). We identified 43 bacteria taxa that were significantly different between animal houses (Table S2). This included common gut bacteria, such as those from the genus *Bacteroides, Alistipes, Mucispirillum* and *Azospirillum*, which had high prevalence (present in >50% of samples) across the cohort.

### *The impact of the genotype*

We next examined whether the mouse genotype influenced the gut bacterial metrics, where genotype 1 is the wild-type control for genotypes 2-4. Again, we observed significant changes in both α-diversity (P=0.001, Figure 2d) as well as β-diversity based on the genotype (P=0.001, Figures 2d-f), with some genotypes (e.g., genotype 2) being the most dissimilar across groups. Pairwise comparisons for β-diversity metrics reveal all four genotypes are significantly different to each other (Table S3). We identified 68 taxa that were different across the genotypes (Table S4), particularly in mice lacking GPR41/43/109A, which lack the 3 main receptors that sense the gut microbial metabolites short-chain fatty acids (SCFA).

### *The impact of diet*

Diet is a well-known factor that impacts the microbiome (Leeming et al., 2019). We investigated the influence of diet using five different diets, four of which were manipulated in the same background, so the only difference was the amount and type of dietary fibre, a key nutrient for the microbiota. These diets had a clear impact in both α- (P<0.0001, Figure 2g) and β-diversity (P=0.001, Figure 2g-i). In the PCoAs, these diets had a clear difference in

both unweighted and weighted analyses (Figure 2g-i), with a diet lacking fibre diet being a major driver of microbiome composition, particularly in the unweighted Unifrac analysis. Pairwise comparisons for β-diversity metrics reveal all diets are mostly different to each other (Table S5). The different diets modulated 55 taxa (Table S6), with a no-fibre diet being the major driver of these changes (21 out of the 55 taxa). The no-fibre diet reduced the prevalence of bacteria known to ferment fibre and produce SCFAs, such as *Ruminococcus* and *Blautia*, while a high fibre diet increased the prevalence of *Ruminococcus* species.

### *The impact of gastrointestinal tissue sampling*

Due to differences in pH and bacterial density(O'Donnell et al., 2023), samples from the large and small intestines were expected to differ. We confirmed that the small intestine had lower α-diversity (P<0.0001, Figure 2j) and different β-diversity metrics compared to the large intestine (P=0.001, Figure 2k-l). We also observed 22 taxa that changed between these intestinal regions (Table S7), with the small intestine having lower prevalence of 15 out the 22 taxa.

### *The impact of age*

We categorised the samples into young (10-12 weeks) and aged (6 months) groups based on age. Although there was a large difference in sample size between young and aged mice, we found that aged mice had higher α-diversity (P<0.0001, Figure 2m) and distinct β-diversity metrics compared to younger mice (P=0.001, Figure 2n-o). We also identified 26 taxonomic changes between young and aged mice, with younger mice having higher prevalence of SCFA-producers such as *Ruminococcus* and *Blautia* compared to aged mice (Table S8).

### *The impact of sex*

There is growing interest in the representation of both male and female mice in animal studies, however, it is still unclear if there are sex differences in the gut microbiota (Beale et

al., 2019). Thus, we attempted to investigate whether there were differences in gut bacterial diversity. We found that male and female mice differed significantly in both α-diversity (P<0.001, Figure 2p), with female mice having higher diversity, and β-diversity metrics (Figure 2q-r). Albeit not distinct in the PCoA plots, sex reached statistical significance (P=0.001), with a small number (7) of taxonomic changes between the sexes (Table S9). Male mice had lower levels of SCFA-producing bacteria including *Ruminococcus*, *Blautia* and *Butyricicoccus* relative to female mice.

### *The impact of the phenotype*

Considering that animal house facility, diet, sample acquisition site, age, and sex influenced α- and β-diversity metrics, we applied this knowledge to a disease model which was previously associated with changes in the microbiome (Cheema and Pluznick, 2019; Kim et al., 2018). Thus, we investigated whether high blood pressure induced by angiotensin II resulted in differences in the gut microbiota using a well-matched subset of samples. Analysing all samples together, we initially identified a significant difference in α-diversity (P<0.0001) and β-diversity (P=0.001) (Figure 3a-d). We then controlled for the other factors identified above to influence the microbiome, and used samples collected from the same site from mice housed in the same animal house facility on the same diet, age, genotype, and sex (Figure 3e). We found no difference in α-diversity (P=0.219, Figure 3f), but unweighted Unifrac, a β-diversity measure that does not consider the abundance of taxa, was significantly different between the phenotypes (P=0.014, Figure 3g). There was no difference in the weighted Unifrac β-diversity measure, which considers the abundance of taxa (P=0.288, Figure 3h). We still identified 24 significantly differentially abundant taxa including *Lachnospiraceae FCS020* and *Alistipes* (Table S10).

### *The impact of sequencing batch effect*

Although our focus was experimental factors that influence the gut microbiota in the laboratory, we were able to investigate one of the many technical factors that could affect the gut microbiome – the sequencing batch. We usually sequence ~200-300 samples per MiSeq run, thus, having the samples from this study spread across two batches. Using samples collected from the large intestine that were processed in the same laboratory but sequenced twice (to create a matched comparison group, total n=218), we could determine if the sequencing batch affected α- and β-diversity metrics. We found that both α- (P=0.037, Figure 4a) and β-diversity metrics (including unweighted – P=0.001, and weighted Unifrac – P=0.006; Figure 4b-c) were significantly influenced by sequencing batch effect. This also resulted in 43 differential taxa across between the two batches (Table S11).

*Experimental factors contributed to the variation in the gut microbial composition at different degrees*

Using VPThemAll (Forslund et al., 2021), we identified that all the factors investigated in this study significantly contributed to gut microbial variations (Table 2). The full effects of all the factors combined can explain 48.85% of variations in the gut microbiome (Table 2). Shared variances explain how much variations observed in the gut microbiome are due to each experimental factor, without correction for the other variables (Table 2). Diet had the largest influence in the shared element of the data (23.3%). The unique element, which is the variation contributed by a single factor following adjustment for all the factors, was largely contributed by the compartment from which the sample was acquired from (6.8%), followed by diet (6%), and then sequencing batch (4.7%) (Table 2). Age, genotype, and animal facility each explained ~2% of variance in the gut microbial composition (Table 2). Sex and the treatment of angiotensin II had minimal effects on the gut microbial composition, each explaining less than 1% of variance (Table 2).

## Discussion

We performed a large-scale gut microbiome study in experimental samples from laboratory mice. Our study reveals that commonly studied gut microbiome indices in mice, such as the α- and β-diversity, are influenced by experimental and technical factors such as diet, sample collection site, sequencing batch effect, genotype, animal house facility, sex, and ageing. Once the microbiome of a disease model was controlled by these factors, the association between healthy (i.e., normotensive) and diseased (i.e., hypertensive) states was significantly reduced. Our findings reiterate the importance of controlling for and reporting these factors in gut microbiome studies as they drive changes in the gut microbiome that could be falsely accredited to phenotypes instead.

Guidelines on best practices in conducting gut microbiome practices (e.g., (Jarett et al., 2021; Kim et al., 2017; Marques et al., 2019; Secombe et al., 2021; Zhang et al., 2021) are available; however, the evidence supporting most of these practices in animal studies has been lacking. Similar to human studies, large-scale animal studies are essential to identify whether the effects of these experimental factors between groups. However, our study identified that some of these factors still influenced gut microbiome diversity measures. One of these factors was age. A recent meta-analysis combining several studies from different laboratories proposed that the murine aged microbiome was more diverse and different to that of younger mice, and that this was driven by a decline in SCFA-producers and overall SCFA levels (You et al., 2022). Faecal microbiota transplantation (FMT) studies from young and aged mice into germ-free mice also confirmed the differing SCFA-producing capabilities of young and aged microbiota (Lee et al., 2020). Using a larger sample size with samples processed in the same laboratory, we confirmed that α-diversity was higher in aged compared to young mice. We also identified that indeed some SCFA-producers, including those from the Lachnospiraceae family such as *Blautia*, were more abundant in younger mice. Together,

these findings suggest that deficiency of SCFA-producers and SCFAs may be associated with aging.

In comparing young and aged mice, we also discovered that young mice had significantly more abundant *Eubacterium coprostanoligenes* compared to aged mice. *Eubacterium coprostanoligenes* has been shown to be hypocholestrolemic in rabbits due to their function in reducing cholesterol to coprostanol, which is excreted in the faeces (Li et al., 1995). The microbial enzyme responsible, ismA, a cholesterol dehydrogenase, was recently discovered (Kenny et al., 2020). *Eubacterium coprostanoligenes* was also differentially abundant between male and female mice.

We identified that female mice had a higher α-diversity compared to male mice, supporting findings in human (Cuesta-Zuluaga et al., 2019). In addition, we found SCFA-producing bacteria such as *Blautia, Ruminococcus* and *Butyricicoccus*, to be more abundant in female mice compared to males. In a cohort of Chinese adults (n=551), female participants were shown to have a higher abundance of *Ruminococcus*, fitting with our findings (Gao et al., 2018). This data warrants further confirmation and investigation.

We found differences in gut bacterial diversity between the small and large intestines. This is not surprising given the differences in cell density but other factors such as pH could also be contributing to the differences in the bacterial diversity (Donaldson et al., 2016; James et al., 2020; O'Donnell et al., 2023). Future studies should determine the regional microbiome differences within the colon (e.g., proximal and distal colon) since different sources of food are fermented at different sections of the colon (So et al., 2021), leading to differences in pH (Xie et al., 2022).

As a proof-of-concept of the impact of these experimental factors in disease models, we studied the most common used hypertensive mouse model, the angiotensin II model. We

initially found an association between the microbiome and treatment with angiotensin II. However, once we considered the experimental factors we reported in this paper, the influence of angiotensin II in driving changes in the gut microbiome diversity measures was significantly reduced, with no difference in α-diversity or weighted UniFrac. We only found that the treatment of angiotensin II led to statistical differences in β-diversity, as per previous reports (Cheema and Pluznick, 2019; Kim et al., 2018), in the unweighted UniFrac analysis. Importantly, the variance explained in axes 1 (14.8%) and 2 (10.2%) was small and lower in our study than in the earlier report of axes 1 (33%) and 2 (20%) (Cheema and Pluznick, 2019). This is consistent with findings in human hypertensive cohorts, where there was no distinction between normotensive and hypertensive groups in α- or β-diversity (Nakai et al., 2021). However, we still identified some taxonomic changes that have been reported in human hypertension, such as *Parabacteroides* and *Lachnospiraceae* (Dan et al., 2019; Silveira-Nunes et al., 2020; Yan et al., 2017), suggesting that findings from our large-scale analysis may be more reliable for translational purposes. This proof-of-concept analysis in the angiotensin II hypertension model warrants further confirmatory studies using other human diseases modelled in laboratory animals. Larger and multiple cohorts or samples from multiple facilities could be the way forward to reduce the effects of experimental factors and minimise false-positive associations.

Other experimental factors not studied here may also impact the gut microbiome of experimental animals. This includes animal vendor (Rasmussen et al., 2019), which could not be addressed in our study as all animals included were bred in-house in each of the three facilities studied. Our findings suggest that gut microbiota studies should be replicated/reproduced in multiple models or at different animal house facilities using samples collected from the same gastrointestinal site in more than one cohort (Figure 5). Alternative strategies to normalise the gut microbiome include co-housing or the use of FMTs. In brief,

FMT is a method to inoculate whole gut microbiota communities to germ-free or antibiotic-treated mice to determine causality (Bokoliya et al., 2021). One study compared the effect of co-housing mice from two genotypes compared to second generation littermates from inter-crossing (Robertson et al., 2019). Co-housing mice only partially normalised the gut microbiota, but the gut microbiota of inter-crossed second-generation littermates were completely normalised (Robertson et al., 2019). Finally, while our study focused on experimental factors that influence the gut microbiota in laboratory animals, technical factors such as sequencing platform and batch effects (Allali et al., 2017) and bioinformatics analysis pipeline (Marizzoni et al., 2020) have also been shown to influence the enrichment and identification of certain gut microbiota, and need to be considered when studies are being designed. As such, we refer the reader to our full recommendations elsewhere (R Muralitharan et al., 2023).

In conclusion, we demonstrated that, in laboratory mice, gut microbiome diversity measures are influenced by common experimental factors such as diet, animal house facility, genotype, and sex, and a technical factor, sequencing batch. These factors should be accounted for and reported in gut microbiome studies. Once adjusted by these factors, treatment with angiotensin II, to induce hypertension, had a small effect only, and provides a proof-of-concept for larger studies of disease models. With this knowledge, we should advocate further for better study design and the validation of findings to improve translatability of microbiome results.

**Data availability**

Sequencing data was deposited in the NCBI SRA Database SUB12933699.

## Acknowledgements

We would like to acknowledge the Monash Animal Research Platform for support with animal work, the Monash Bioinformatics Platform for access to M3 servers. The authors wish to acknowledge the use of the services and facilities of Australian Genome Research Facility. This work was supported by a National Health & Medical Research Council (NHMRC) of Australia Project Grant (GNT1159721), and fellowship to D.M.K. F.Z.M. is supported by a Senior Medical Research Fellowship from the Sylvia and Charles Viertel Charitable Foundation, a National Heart Foundation Future Leader Fellowship (105663), and NHMRC Emerging Leader Fellowship (GNT2017382). M.S. is supported by a National Heart Foundation Postdoctoral Fellowship (106698). R.R.M. is supported by a scholarship from the Faculty of Science, Monash University. L.X. is supported by a Monash Graduate Scholarship. The Baker Heart & Diabetes Institute is supported in part by the Victorian Government's Operational Infrastructure Support Program.

## Author Contribution Statement

R.R.M. designed the study, planned, and performed most of the in vivo and in vitro experiments, data analyses, provided intellectual inputs and wrote the manuscript. M.N., T.Z., and M.S., assisted with bioinformatics analyses, E.D., L.X., H.J., M.P., W.S., F.W., A.V., D.K., and C.R.M. assisted with in-vitro experiments, maintained mice colonies, and acquired samples. F.Z.M. designed the study, secured funding, provided intellectual input, and supervised the study. All authors had full access to all data in the study, revised the manuscript critically, approved the version to be published, and had the final responsibility for the decision to submit the manuscript for publication.

# References

Allali, I., Arnold, J.W., Roach, J., Cadenas, M.B., Butz, N., Hassan, H.M., Koci, M., Ballou, A., Mendoza, M., Ali, R.*, et al.* (2017). A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. BMC Microbiology *17*, 194.

Beale, A.L., Kaye, D.M., and Marques, F.Z. (2019). The role of the gut microbiome in sex differences in arterial pressure. Biol Sex Differ *10*, 22.

Bokoliya, S.C., Dorsett, Y., Panier, H., and Zhou, Y. (2021). Procedures for Fecal Microbiota Transplantation in Murine Microbiome Studies. Frontiers in Cellular and Infection Microbiology *11*.

Bokulich, N.A., Kaehler, B.D., Rideout, J.R., Dillon, M., Bolyen, E., Knight, R., Huttley, G.A., and Gregory Caporaso, J. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome *6*, 90.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F.*, et al.* (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology *37*, 852-857.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods *13*, 581-583.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M.*, et al.* (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The ISME Journal *6*, 1621-1624.

Cheema, M.U., and Pluznick, J.L. (2019). Gut Microbiota Plays a Central Role to Modulate the Plasma and Fecal Metabolomes in Response to Angiotensin II. Hypertension *74*, 184-193.

Cuesta-Zuluaga, J.d.l., Kelley, S.T., Chen, Y., Escobar, J.S., Mueller, N.T., Ley, R.E., McDonald, D., Huang, S., Swafford, A.D., Knight, R.*, et al.* (2019). Age- and Sex-Dependent Patterns of Gut Microbial Diversity in Human Adults. mSystems *4*, e00261-00219.

Dan, X., Mushi, Z., Baili, W., Han, L., Enqi, W., Huanhu, Z., and Shuchun, L. (2019). Differential Analysis of Hypertension-Associated Intestinal Microbiota. Int J Med Sci *16*, 872-881.

Donaldson, G.P., Lee, S.M., and Mazmanian, S.K. (2016). Gut biogeography of the bacterial microbiota. Nat Rev Microbiol *14*, 20-32.

Forslund, S.K., Chakaroun, R., Zimmermann-Kogadeeva, M., Markó, L., Aron-Wisnewsky, J., Nielsen, T., Moitinho-Silva, L., Schmidt, T.S.B., Falony, G., Vieira-Silva, S.*, et al.* (2021). Combinatorial, additive and dose-dependent drug–microbiome associations. Nature *600*, 500-505.

Gao, X., Zhang, M., Xue, J., Huang, J., Zhuang, R., Zhou, X., Zhang, H., Fu, Q., and Hao, Y. (2018). Body Mass Index Differences in the Gut Microbiota Are Gender Specific. Frontiers in Microbiology *9*.

Guo, J., Song, C., Liu, Y., Wu, X., Dong, W., Zhu, H., Xiang, Z., and Qin, C. (2022). Characteristics of gut microbiota in representative mice strains: Implications for biological research. Animal Model Exp Med *5*, 337-349.

Hilbert, T., Steinhagen, F., Senzig, S., Cramer, N., Bekeredjian-Ding, I., Parcina, M., Baumgarten, G., Hoeft, A., Frede, S., Boehm, O.*, et al.* (2017). Vendor effects on murine gut microbiota influence experimental abdominal sepsis. J Surg Res *211*, 126-136.

Hildebrand, F., Nguyen, T.L.A., Brinkman, B., Yunta, R.G., Cauwe, B., Vandenabeele, P., Liston, A., and Raes, J. (2013). Inflammation-associated enterotypes, host genotype, cage and

inter-individual effects drive gut microbiota variation in common laboratory mice. Genome Biology *14*, R4.

Hufeldt, M.R., Nielsen, D.S., Vogensen, F.K., Midtvedt, T., and Hansen, A.K. (2010). Variation in the gut microbiota of laboratory mice is related to both genetic and environmental factors. Comp Med *60*, 336-347.

Jama, H.A., Muralitharan, R.R., Xu, C., O'Donnell, J.A., Bertagnolli, M., Broughton, B.R.S., Head, G.A., and Marques, F.Z. (2022). Rodent models of hypertension. Br J Pharmacol *179*, 918-937.

James, K.R., Gomes, T., Elmentaite, R., Kumar, N., Gulliver, E.L., King, H.W., Stares, M.D., Bareham, B.R., Ferdinand, J.R., Petrova, V.N.*, et al.* (2020). Distinct microbial and immune niches of the human colon. Nat Immunol *21*, 343-353.

Jarett, J.K., Kingsbury, D.D., Dahlhausen, K.E., and Ganz, H.H. (2021). Best Practices for Microbiome Study Design in Companion Animal Research. Frontiers in Veterinary Science *8*.

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res *30*, 3059-3066.

Kenny, D.J., Plichta, D.R., Shungin, D., Koppel, N., Hall, A.B., Fu, B., Vasan, R.S., Shaw, S.Y., Vlamakis, H., Balskus, E.P.*, et al.* (2020). Cholesterol Metabolism by Uncultured Human Gut Bacteria Influences Host Cholesterol Level. Cell host & microbe *28*, 245-257.e246.

Kim, D., Hofstaedter, C.E., Zhao, C., Mattei, L., Tanes, C., Clarke, E., Lauder, A., Sherrill-Mix, S., Chehoud, C., Kelsen, J.*, et al.* (2017). Optimizing methods and dodging pitfalls in microbiome research. Microbiome *5*, 52.

Kim, S., Goel, R., Kumar, A., Qi, Y., Lobaton, G., Hosaka, K., Mohammed, M., Handberg, Eileen M., Richards, Elaine M., Pepine, Carl J.*, et al.* (2018). Imbalance of gut microbiome and intestinal epithelial barrier dysfunction in patients with high blood pressure. Clinical Science *132*, 701.

Lee, J., Venna, V.R., Durgan, D.J., Shi, H., Hudobenko, J., Putluri, N., Petrosino, J., McCullough, L.D., and Bryan, R.M. (2020). Young versus aged microbiota transplants to germ-free mice: increased short-chain fatty acids and improved cognitive performance. Gut Microbes *12*, 1814107.

Leeming, E.R., Johnson, A.J., Spector, T.D., and Le Roy, C.I. (2019). Effect of Diet on the Gut Microbiota: Rethinking Intervention Duration. Nutrients *11*, 2862.

Li, L., Buhman, K.K., Hartman, P.A., and Beitz, D.C. (1995). Hypocholesterolemic effect of Eubacterium coprostanoligenes ATCC 51222 in rabbits. Letters in Applied Microbiology *20*, 137-140.

Mallick, H., Rahnavard, A., McIver, L.J., Ma, S., Zhang, Y., Nguyen, L.H., Tickle, T.L., Weingart, G., Ren, B., Schwager, E.H.*, et al.* (2021). Multivariable association discovery in population-scale meta-omics studies. PLOS Computational Biology *17*, e1009442.

Marizzoni, M., Gurry, T., Provasi, S., Greub, G., Lopizzo, N., Ribaldi, F., Festari, C., Mazzelli, M., Mombelli, E., Salvatore, M.*, et al.* (2020). Comparison of Bioinformatics Pipelines and Operating Systems for the Analyses of 16S rRNA Gene Amplicon Sequences in Human Fecal Samples. Frontiers in Microbiology *11*.

Marques, F.Z., Jama, H.A., Tsyganov, K., Gill, P.A., Rhys-Jones, D., Muralitharan, R.R., Muir, J.G., Holmes, A.J., and Mackay, C.R. (2019). Guidelines for Transparency on Gut Microbiome Studies in Essential and Experimental Hypertension. Hypertension *74*, 1279-1293.

Montonye, D.R., Ericsson, A.C., Busi, S.B., Lutz, C., Wardwell, K., and Franklin, C.L. (2018). Acclimation and Institutionalization of the Mouse Microbiota Following Transportation. Frontiers in Microbiology *9*.

Nakai, M., Ribeiro, R.V., Stevens, B.R., Gill, P., Muralitharan, R.R., Yiallourou, S., Muir, J., Carrington, M., Head, G.A., Kaye, D.M*., et al.* (2021). Essential Hypertension Is Associated With Changes in Gut Microbial Metabolic Pathways: A Multisite Analysis of Ambulatory Blood Pressure. Hypertension *78*, 804-815.

O'Donnell, J.A., Zheng, T., Meric, G., and Marques, F.Z. (2023). The gut microbiome and hypertension. Nat Rev Nephrol *19*, 153-167.

O'Donnell, J.A., Zheng, T., Meric, G., and Marques, F.Z. (2023). The gut microbiome and hypertension. Nature Reviews Nephrology *19*, 153-167.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLOS ONE *5*, e9490.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res *41*, D590-D596.

R Muralitharan, R., Snelson, M., Meric, G., Coughlan, M.T., and Marques, F.Z. (2023). Guidelines for microbiome studies in renal physiology. American journal of physiology Renal physiology.

Rasmussen, T.S., de Vries, L., Kot, W., Hansen, L.H., Castro-Mejía, J.L., Vogensen, F.K., Hansen, A.K., and Nielsen, D.S. (2019). Mouse Vendor Influence on the Bacterial and Viral Gut Composition Exceeds the Effect of Diet. Viruses *11*.

Robertson, S.J., Lemire, P., Maughan, H., Goethel, A., Turpin, W., Bedrani, L., Guttman, D.S., Croitoru, K., Girardin, S.E., and Philpott, D.J. (2019). Comparison of Co-housing and

Littermate Methods for Microbiota Standardization in Mouse Models. Cell Rep *27*, 1910-1919.e1912.

Secombe, K.R., Al-Qadami, G.H., Subramaniam, C.B., Bowen, J.M., Scott, J., Van Sebille, Y.Z.A., Snelson, M., Cowan, C., Clarke, G., Gheorghe, C.E.*, et al.* (2021). Guidelines for reporting on animal fecal transplantation (GRAFT) studies: recommendations from a systematic review of murine transplantation protocols. Gut microbes *13*, 1979878-1979878.

Silveira-Nunes, G., Durso, D.F., Jr., L.R.A.d.O., Cunha, E.H.M., Maioli, T.U., Vieira, A.T., Speziali, E., Corrêa-Oliveira, R., Martins-Filho, O.A., Teixeira-Carvalho, A.*, et al.* (2020). Hypertension Is Associated With Intestinal Microbiota Dysbiosis and Inflammation in a Brazilian Population. Frontiers in Pharmacology *11*.

Singh, G., Brass, A., Cruickshank, S.M., and Knight, C.G. (2021). Cage and maternal effects on the bacterial communities of the murine gut. Scientific Reports *11*, 9841.

So, D., Gibson, P.R., Muir, J.G., and Yao, C.K. (2021). Dietary fibres and IBS: translating functional characteristics to clinical value in the era of personalised medicine. Gut *70*, 2383-2394.

Turnbaugh, P.J., Bäckhed, F., Fulton, L., and Gordon, J.I. (2008). Diet-Induced Obesity Is Linked to Marked but Reversible Alterations in the Mouse Distal Gut Microbiome. Cell Host & Microbe *3*, 213-223.

Turner, P.V. (2018). The role of the gut microbiota on animal model reproducibility. Animal Model Exp Med *1*, 109-115.

Xie, L., Rhys-Jones, D., Muralitharan, R.R., Dinakis, E., Nakai, M., Paterson, M., Peh, A., Jama, H., Salimova, E., Anderson, D.*, et al.* (2022). Dietary fibre controls blood pressure and cardiovascular risk by lowering large intestinal pH and activating the proton-sensing receptor GPR65. bioRxiv, 2022.2011.2017.516695.

Yan, Q., Gu, Y., Li, X., Yang, W., Jia, L., Chen, C., Han, X., Huang, Y., Zhao, L., Li, P*., et al.* (2017). Alterations of the Gut Microbiome in Hypertension. Frontiers in Cellular and Infection Microbiology *7*.

You, X., Dadwal, U.C., Lenburg, M.E., Kacena, M.A., and Charles, J.F. (2022). Murine Gut Microbiome Meta-analysis Reveals Alterations in Carbohydrate Metabolism in Response to Aging. mSystems *7*, e01248-01221.

Zhang, C., Franklin, C.L., and Ericsson, A.C. (2021). Consideration of Gut Microbiome in Murine Models of Diseases. Microorganisms *9*, 1062.

# Tables

**Table 1.** Summary of previous studies that investigated the role of experimental factors that influence the gut bacterial diversity metrics in laboratory mice. Search performed on PubMed on 9th March 2023 using the following keywords: 'alpha diversity', 'beta diversity', 'experimental factors', 'laboratory', 'mice', 'animals'.

| Study | Factor(s) investigated | Total sample size | Findings |
|---|---|---|---|
| Turnbaugh et al., 2008 | Diet | 10 | Conventionalised germ-free mice fed two different diets had significant difference in β-diversity |
| Hufeldt et al., 2010 | Mice vendor | 33 | C57BL/6 mice from 2 different vendors had different caecal microbiota composition |
| | Different rooms in same breeding center | 33 | C57BL/6 mice from 1 vendor bred in two different rooms in the same breeding centre had similar microbiota composition |
| | Sex | 33 | No difference between male and female mice from same vendor housed in the same room but different cages |
| Hildebrand et al., 2013 | Genetic effects | 50 | Genetic effects influenced up to 15.65 to 18.62% of variations at lower than phylum levels among 4 genotypes |
| | Cage effects | 50 | Up to 30% of variations in the gut microbiome were explained by cage |

| | | | effects |
|---|---|---|---|
| Rausch et al., 2016 | Housing type | 94 | Difference in α- and β-diversity in mice housed in IVC compared to others |
| | Chow treatment | 104 | Difference in α- and β-diversity in autoclaved food compared to irradiated and untreated |
| | Chow provider | 104 | No difference in α-diversity between two chow provider, but different β-diversity |
| Hilbert et al., 2017 | Vendor effects | 60 | C57BL/6 from three vendors had distinct gut microbial composition |
| Ericsson et al., 2018 | Cage ventilation and bedding | 144 | Cage ventilation and bedding and the interaction, had a distinguishable effect on the caecal microbiota but not in the faecal or small intestine samples |
| Montonye et al., 2018 | Transport | 16 | Significant difference in α- and β-diversity pre- and post-arrival |
| | Animal facility | 8 | The α- and β-diversity of mice pre-shipping is significantly different to microbiota upon arrival at facilities and between facilities |
| Robertson et al., 2019 | Mice vendor | 8 | Mice from 2 vendors had no difference in α-diversity but significantly different β-diversity. Samples from pellet, colon, ileum |

| | | | |
|---|---|---|---|
| | Littermate effect and co-housing effect | 28 | Mice that were co-housed had similar gut microbiome to each other but also their non-co-housed siblings (n=12) in the colon and the ileum but not in the faecal pellet. |
| Wolff et al., 2020 | Vendor effects | 24 | Significant difference in β-diversity of genetically identical C57BL/6 from three different vendors |
| Singh et al., 2021 | Cage and maternal effects | 30 | The gut microbiome of F2 littermates of different genotypes (n=10 per group) from heterozygous parents (mdr1a-/+) clustered according to their maternal origin and cages. |
| Guo et al., 2022 | Mouse genotypes | 110 | Significant difference in α- and β-diversity of 7 background genotypes of C57BL/6 and BALB/c mice |

**Table 2.** The full, shared, and unique contribution of each experimental factor to the variations in the gut microbiota composition. FDR-adjusted p-values <0.02 were considered significant. Full represents how much all the experimental factors in our study explain the variations observed in the gut microbiome. Shared represents how much each factor explains the variation without correcting for the other variables. Unique is how much each factor explains the variations observed, after correcting for all other variables.

| | Factor | Element | *P* value | Adj.r.squared | Percentage of variance (%) | Adj. *P* value |
|---|---|---|---|---|---|---|
| 1 | full | full | NA | 0.488 | 48.800 | NA |
| | | | | | | |
| 1 | Diet | shared | 0.001 | 0.233 | 23.300 | 0.001 |
| 2 | Compartment | shared | 0.001 | 0.128 | 12.800 | 0.001 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3 | Sequencing batch | shared | 0.001 | 0.209 | 20.900 | 0.001 |
| 4 | Age | shared | 0.001 | 0.047 | 4.700 | 0.001 |
| 5 | Genotype | shared | 0.001 | 0.051 | 5.100 | 0.001 |
| 6 | Animal Facility | shared | 0.001 | 0.089 | 8.900 | 0.001 |
| 7 | Sex | shared | 0.001 | 0.014 | 1.400 | 0.001 |
| 8 | Angiotensin II treatment | shared | 0.001 | 0.028 | 2.800 | 0.001 |
| | | | | | | |
| 1 | Diet | unique | 0.001 | 0.060 | 6.000 | 0.001 |
| 2 | Compartment | unique | 0.001 | 0.068 | 6.800 | 0.001 |
| 3 | Sequencing batch | unique | 0.001 | 0.047 | 4.700 | 0.001 |
| 4 | Age | unique | 0.001 | 0.020 | 2.000 | 0.001 |
| 5 | Genotype | unique | 0.001 | 0.029 | 2.900 | 0.001 |
| 6 | Animal Facility | unique | 0.001 | 0.023 | 2.300 | 0.001 |
| 7 | Sex | unique | 0.001 | 0.004 | 0.400 | 0.001 |
| 8 | Angiotensin II treatment | unique | 0.001 | 0.004 | 0.400 | 0.001 |

**Legend**: Adj, adjusted; NA, not applicable.

## Figure Legends

**Figure 1. Study design and rarefaction curve of samples. a**, brief schematic of study design. 567 samples were sequenced from 303 animals, resulting in 17,809,928 reads. Sequencing depth of 7000 reads was used, resulting in the loss of 29 (5.1%) samples. The 538 samples were categorised according to the various experimental factors; animal house facility, diet, age, genotype, intestinal region of sample collection site, and sex. **b,** Shannon rarefaction curve of all samples showing plateauing of the sequencing depth at 7000 reads.

**Figure 2. α- and β-diversity metrics of all samples (n=538), categorised according to the various experimental factors. a-c**, animal house facility; **d-f**, mice genotype; **g-i**, diet; **j-l** intestinal region of sample collection; **m-o**, age; and, **p-r**, sex. **a, d, g, j, m, p,** Shannon index, an α-diversity metric. **b, e, h, k, n, q,** unweighted Unifrac (β-diversity metric which does not consider taxa abundance), and, **c, f, i, l, o, r,** weighted Unifrac (β-diversity which considers taxa abundance). Significance of Shannon index ($p<0.05$) determined using Mann-Whitney U (two comparison groups) or Kruskal-Wallis (more than two comparison groups) tests. Significance of unweighted and weighted Unifrac determined using PERMANOVA with adjustment for multiple comparisons ($q<0.05$).

**Figure 3. The influence of the hypertensive phenotype (using angiotensin II) on the α- and β-diversity metrics of all the samples (n=538, a-d) compared to a subset of samples acquired from the caecum of C57Bl/6 male mice of the same age from animal house facility 1 fed control diet 2 (e-h, n=89). a&e,** brief flowchart showing samples that were selected.**b&f**, Shannon index, an α-diversity metric; **c&g,** unweighted Unifrac (β-diversity metric which does not consider taxa abundance); **d&h,** weighted Unifrac (β-diversity which considers taxa abundance). Significance of Shannon index ($p<0.05$) determined using Mann-

Whitney U test. Significance of unweighted and weighted Unifrac determined using PERMANOVA with adjustment for multiple comparisons (q<0.05).

**Figure 4. α- and β-diversity metrics of a subset of matched samples (n=218). a,** Shannon index, an α-diversity metric; **b,** unweighted Unifrac (β-diversity metric which does not consider taxa abundance); **c,** weighted Unifrac (β-diversity which considers taxa abundance). Significance of Shannon index (p<0.05) determined using Mann-Whitney U test. Significance of unweighted and weighted Unifrac determined using PERMANOVA with adjustment for multiple comparisons (q<0.05).

**Figure 5. Recommendations for future gut microbiome studies.** Findings should be a) replicated in multiple cohorts in the same facility, b) validated in more than one model, c) use mice that were cohoused or F2 generation littermates, d) validate findings in more than facility, and e) consider faecal matter transplantation experiments. Figure created with BioRender.com.

**Figure 1**

**Figure 2**

**Figure 3**

**Figure 4**

**Figure 5**

a) Multiple cohorts in the same facility

b) Findings are validated in >1 models

c) Co-housing or using F2 generation littermates

d) Comparing findings across two or more facilities

e) Using faecal matter transplantation