Animal Microbiome

RESEARCH Open Access

Microbiome signature suggestive of lactose-intolerance in rhesus macaques (*Macaca mulatta*) with intermittent chronic diarrhea

Annemiek Maaskant[1,2*†], Bas Voermans[3,4*†], Evgeni Levin[3], Marcus C. de Goffau[3,5], Nicole Plomp[6], Frank Schuren[6], Edmond J. Remarque[1], Antoine Smits[1], Jan A. M. Langermans[1,2], Jaco Bakker[1†] and Roy Montijn[6†]

Abstract

Background Chronic diarrhea is a common cause of mortality and morbidity in captive rhesus macaques (*Macaca mulatta*). The exact etiology of chronic diarrhea in macaques remains unidentified. The occurrence of diarrhea is frequently linked to dysbiosis within the gut microbiome. Research into microbiome signatures correlated with diarrhea in macaques have predominantly been conducted with single sample collections. Our analysis was based on the metagenomic composition of longitudinally acquired fecal samples from rhesus macaques with chronic diarrhea and clinically healthy rhesus macaques that were obtained over the course of two years. We aimed to investigate potential relationships between the macaque gut microbiome, the presence of diarrhea and diet interventions with a selection of commercially available monkey diets.

Results The microbiome signature of macaques with intermittent chronic diarrhea showed a significant increase in lactate producing bacteria e.g. *lactobacilli*, and an increase in fermenters of lactate and succinate. Strikingly, two lactose free diets were associated with a lower incidence of diarrhea.

Conclusion A lactose intolerance mechanism is suggested in these animals by the bloom of *Lactobacillus* in the presence of lactose resulting in an overproduction of intermediate fermentation products likely led to osmotically induced diarrhea. This study provides new insights into suspected microbiome-lactose intolerance relationship in rhesus macaques with intermittent chronic diarrhea. The integration of machine learning with metagenomic data analysis holds potential for developing targeted dietary interventions and therapeutic strategies and therefore ensuring a healthier and more resilient primate population.

†Annemiek Maaskant and Bas Voermans have contributed equally to this work to be shared first authors.

†Jaco Bakker and Roy Montijn have contributed equally to this work to be shared last authors.

*Correspondence:
Annemiek Maaskant
maaskant@bprc.nl
Bas Voermans
b.voermans@amsterdamumc.nl
Full list of author information is available at the end of the article

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Keywords Machine learning, Monkey, Gut microbiota, Lactate, Succinate, Intestinal health, Food hypersensitivity, SCFA, Diet

Background

Chronic diarrhea is a common cause of mortality and morbidity in captive rhesus macaques (*Macaca mulatta*). Diarrhea may result in dehydration, weight-loss and electrolyte disbalance [1–3]. Both juvenile and adult rhesus macaques can be affected, resulting in loss of animals and breeding potential and, more importantly, a decreased welfare of these animals. The etiology of diarrhea in rhesus macaques is likely to be multi-factorial and diverse as many risk factors can lead to the occurrence of diarrhea [2, 4, 5]. Frequently reported bacterial pathogens causing diarrhea in macaques are Giardia, Shigella, Campylobacter, Yersinia, Salmonella, and Escherichia coli [3, 6–9]. Viruses, such as adeno and rotaviruses, can induce diarrhea, but in general, viruses are not strongly associated with idiopathic diarrhea [10–12]. Furthermore, dietary factors have also been associated with diarrhea. Consumption of high-fat diets can result in malabsorption and protein losing enteropathy, gluten sensitivity and lactose intolerance have been reported as possible causes of diarrhea [13–17].

Lactose intolerance is currently not considered as a major contributor to diarrhea in rhesus macaques. Although literature in this area is sparse in macaques, lactose intolerance is well described in humans [14, 17]. Lactose intolerance is the inability to digest lactose in the small intestine due to decline in lactase expression, naturally occurring in mammals after weaning [18, 19]. The undigested lactose enters the large intestine and leads to osmotic trapping of water and an increased osmotic load by fermentation of lactose to short chain fatty acids (SCFA) [20–23]. As a result, symptoms such as bloating, abdominal pain, and diarrhea after ingestion of lactose can occur [21, 22, 24]. Different tests are available to diagnose lactose intolerance, each of which investigates a different aspect of the process and presents with its own limitations [22]. Most of them are designed for human purposes e.g., a breath test or a genetic test. A less human specific test for lactose intolerance involves avoiding lactose combined with subsequent cessation of symptoms [14, 22].

Various studies have investigated the influence of diet on the microbiome in nonhuman primate (NHP) models for human diseases [25–27]. Strong relationship between gut microbiota and diseases e.g., diabetes, neurodegenerative and cardiovascular diseases have been shown [28–33]. However, very few studies have primarily focused on the influence of diet for the welfare and gastrointestinal health of macaques. As diarrhea is a common health problem in captive macaques, the relationship between their microbiome and diarrhea has already been investigated. It has been observed that rhesus macaques with diarrhea show in general an increased abundance of *Campylobacter*, *Faecalibacterium*, *Roseburia*, in a less frequent degree an abundance of *Ruminococcus*, *Megasphaera*, *Eubacterium*, *Dialister*, *Clostridium* and *Bacteroides* compared to non-symptomatic controls [34–37]. Compared to our study, only a limited number of fecal samples per animal were collected, no longitudinal follow up was included, and conventional microbiome analyses were performed.

Innovations in machine learning (ML) made it possible to use this modality to uncover complex relations between components within the microbiome [38–40]. In humans, ML has opened up the possibility of correlating the enormous wealth of measurements obtained through microbiome analysis and other omics approaches. ML can provide interesting leads that could also be of relevance with relatively small cohorts. For example, a relatively small cohort of 50–150 samples can already measure more than 500 different microbial strains. This makes fecal metagenomics data high dimensional with low sample size. Wastyk et al. identified 2204 unique amplicon sequence variants in 36 participants, and Qin et al. [41–43] identified 2580 individual taxa in 131 individual samples. This renders metagenomic data largely unsuitable to traditional statistical models. Furthermore, ML approaches often employ non-linear multivariate predictive models, which can take interactions between microbes into account and reduce the high-dimensional space to a single outcome metric (e.g. Receiver Operating Characteristic-Aera Under the Curve (ROC-AUC) for binary classification or root-mean-squared-error for regression). Additionally, when combined with (traditional) metagenomic analysis and biological interpretation, concordance of the outcomes increases confidence and robustness of the results. To the authors knowledge these advanced ML techniques have not yet been applied in NHP studies.

We aimed to investigate potential relationships between the macaque gut microbiome, the presence of diarrhea and diet interventions with commercially available monkey diets. We employed ML for the first time to differentiate between intestinal microbiome

signatures associated with animals experiencing idiopathic mild to profuse diarrhea and those with normal feces over a two-year period. Our analysis was based on the metagenomic composition of longitudinally acquired fecal samples from rhesus macaques with chronic intermittent diarrhea and clinically healthy rhesus macaques.

Methods

Animals, husbandry and housing

Between 2020 and 2022, fourteen captive bred rhesus macaques (*Macaca mulatta*) were included in this study (Table 1). Ten animals had a history of chronic diarrhea, which was unresponsive to conventional veterinary treatment. Four macaques, without a history of gastrointestinal symptoms, were included as controls. All animals were of Indian origin and bred and raised at the Biomedical Primate Research Centre (BPRC, Rijswijk, The Netherlands) in naturalistic multi-generational family groups with no forced weaning.

The animals with chronic diarrhea had frequently been treated with ivermectin, oxytetracycline and enrofloxacin. In addition, no probiotics were used to support these treatments. Both the control and diarrhea group had not received antimicrobial treatment at least 30 days prior to the start of the study. During the study, the animals did not receive any antimicrobial or probiotic treatment. The humane endpoint of this study was established when diarrhea symptoms necessitated veterinary intervention. One animal with diarrhea met the humane endpoint early in the study. With only one collected sample, this animal was excluded from further analysis. The remaining macaques with diarrhea (n = 9) were housed in two pairs and one group of three animals; the controls were housed in two pairs. The animal enclosures were divided into an indoor and outdoor compartment; the indoor floor was provided with wood fiber bedding (Lignocel3-4, JRS, Rosenberg, Germany) and the outdoor compartment consisted of sand bedding.

The wood fiber bedding of the indoor compartments was replaced weekly; one week without additional cleaning procedures and one week with high-pressure water cleaning including disinfection (Anistel Surface disinfectant, Tristel Solutions Limited, Cambridgeshire, United Kingdom) as described elsewhere [44].

Standard environmental enrichment in these enclosures consisted of several climbing structures, beams, fire hoses, and sitting platforms. The indoor temperature was set at 18 °C, with a 12 h light:dark cycle. The animals were fed commercially available monkey pellets supplemented with limited amounts of fruit, vegetables, or grain mixtures. Water was available ad libitum, provided by automatic water dispensers. Animal caretakers observed all animals at least twice daily for injuries and illness; abnormalities were reported to the veterinarians. Individual electronic health records were kept for each animal.

Table 1 Individual animal characteristics, including subject' social grouping (enclosure), age in years (yrs), bodyweight in kilograms (kg) and gender

Enclosure	Animal ID	Group	Age (yrs)	Bodyweight (kg)	Gender
1	RM1	Diarrhea	4	8.1	Male
	RM2	Diarrhea	4	8.3	Male
2	RM3	Diarrhea	5	7.9	Female
	RM4	Diarrhea	7	8.5	Female
3	RM5	Diarrhea	5	10.8	Male
	RM6	Diarrhea	5	10.0	Male
	RM7	Diarrhea	5	9.5	Male
4	RM8	Control	3	7.4	Male
	RM9	Control	3	6.0	Male
5	RM10	Control	5	10.1	Male
	RM11	Control	5	10.7	Male
6	RM12	Diarrhea	7	11.1	Male
	RM13	Diarrhea	7	12.4	Male

Diets

Five commercially available diets were selected for this study. Each diet differed in ingredients, content or preparation. Full product names as provided by the manufacturers are shown in supplementary Table S1. The main nutrients and differences in dietary composition are shown in Table 2. Diet B and C were extruded; during this process the raw ingredients are exposed to high pressure and high temperatures (> 100 °C). Diet B was manufactured during a high pressure and high temperature extrusion process and Diet C with a conventional extrusion process. All animals were assigned at random to a specific diet. After each dietary cycle

Table 2 percentage of crude nutrients (protein, fat, fiber) of the different diets

Diet	Protein (%)	Fat (%)	Fiber (%)	Gluten	Lactose
A	25.2	4.3	4.2	+	+
B	22.2	4.6	4.5	+	+
C	22.2	4.6	4.5	+	+
D	22.0	5.0	3.5	+	–
E	17.0	3.5	14.0	+	–

Presence of gluten and lactose are indicated as either positive (+) or negative (–)

with a duration of approximately three months, a washout period with a similar duration was implemented. Diet A was used for washing out because this was the standard diet at the BPRC.

Fecal scoring and collection

At the end of each dietary cycle, on the last standard cleaning day at approximately 3 pm the animals were fed a food coloring product to identify individual feces the next morning. The next morning, between 9 and 11 am, the feces was collected and divided into three cryovials, each containing 1 g, stored at −80 °C until further processing. On two occasions, there was no sample obtained as only indoor samples were collected.

One animal per pair or group of three was randomly selected to receive blue food coloring, and for the group of three, an additional animal was selected to receive indigestible, non-toxic glitter. All animals received the food coloring or glitters in alternating sequence, to exclude the possible influence of these identification methods. Animals receiving food coloring or glitter drank 20 mL of diluted syrup with 0.3 mL blue food coloring (Wilton Industries, Icing color, Royal blue, Naperville, USA) or 0.35 g of glitter (The sparkle range, rainbowdust.co.uk, Cuerden Greenmill, UK). The individual not receiving coloring or glitter was offered diluted syrup only. In addition, this coloring procedure was performed every two weeks, after which the Waltham Faeces Scoring System was applied by two neutral observers [45, 46]. The Waltham scale utilizes a scale of 1–5 with half numerical increments, covering a range of very hard (score 1) and dry to entirely liquid feces (score 5) [45, 46]. The mean Waltham score was calculated over a period of three months prior to each fecal collection to determine the overall stool consistency. The cut-off score for diarrhea was set at a mean score of 3.6. This also entails that control animals whose mean Waltham score exceeded this value were labeled as diarrhea-positive and vice versa.

Fecal metagenomics

DNA isolation and metagenomic sequencing

Total DNA was extracted from the samples using an Agowa/PurePrep protocol. To each 150 l sample, 500 l zirconium beads (0.1 mm) and 800 l CD1 solution (DNeasy 96 Powersoil Pro QIAcube HT kit) were added. Cells were disrupted by bead beating twice for 2 min, with cooling on ice in between and afterwards. After centrifugation for 6 min at 3000 RPM, 350 l supernatant was mixed with 300 l Agowa binding buffer and 10 l Agowa magnetic beads. Samples were further purified using the PurePrep 96 system (Molgen, The Netherlands) with two wash steps and a final elution step in 65 l. Libraries for whole-genome sequencing were prepared using the Illumina DNA prep protocol according to the instructions of Illumina (Illumina DNA Prep ReferenceGuide). Blank controls were included for DNA isolation and sequencing, consisting of all components used for sample DNA isolation except for sample material. These controls contained very low levels of DNA, and after sequencing, they did not result in sequence data.

DNA concentrations were standardized across samples. After tagmentation and clean-up steps, PCR-mediated standard indexed i5 and i7 adapters were added and the library was amplified. Next, the libraries were cleaned-up and pooled. Whole genome sequencing was performed using the Illumina MiSeq sequencer applying MiSeq V3 chemistry.

Metagenomic profiling

Fastp was applied to preprocess the reads to filter out low-quality and too short reads [47]. In addition, fastp trims out all reads from the front and the tail and cuts potentially present adapter sequences. Profiling of trimmed reads was performed with a kraken2 pipeline with a custom database built through kraken2's proprietary method [48]. A database consisting of archaea, bacteria, plasmid, virus, human, fungi and UniVec Core nucleotide/protein sequences was used. The latter is a subset of the NCBI-supplied database of vector, adapter, linker and primer sequences that may be contamination factors. The GRCh38 human genome assembly, as available on the NCBI database, was used to filter out potential human reads that may have ended up in the samples due to contamination during or before sample collection [49]. Parameter choices for kmer length were unchanged from the default settings of kraken2. After profiling, a Bayesian re-estimation of abundance was performed using the Bracken tool [50]. All profiling tools were installed from bioconda channels [51].

Machine learning

Software packages

All analyses were performed in Python 3.9 and for statistical test, scipy (version 1.12.0) was used. A stacked L2-model comprising three separately trained predictors: CatBoost, XGBoost and Extratrees was applied [52, 53]. The parameters of the XGBoost (version 2.0.3) and Extratrees models were selected using stratified three-fold cross-validation. The CatBoost (version 1.2.5) model was trained using its default parameters as this has been shown to yield to high performance [54]. The Extratrees model was used as is made available in the scikit-learn package (version 1.4.1). Plots were made using a combination of seaborn (version 0.11.2) and matplotlib (version 3.8.3). Figure annotation was performed using the

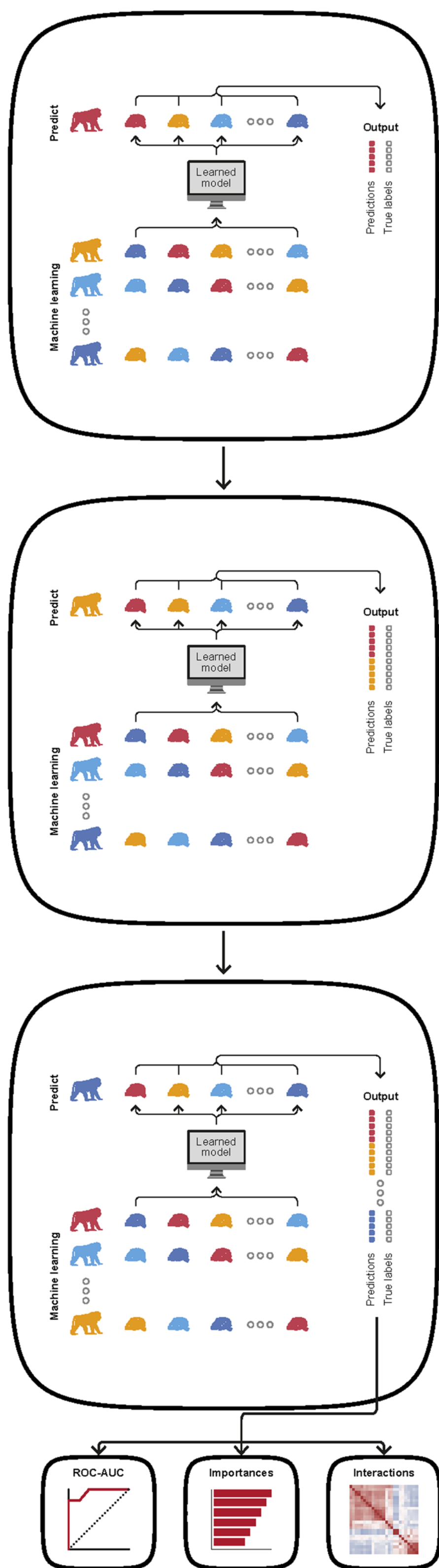

Fig. 1 Schematic representation of the Leave-One-Monkey-Out approach used for the machine-learning algorithm for biomarker discovery. Each monkey is used as the independent test data in an iterative manner and is thus "left out" of the samples on which the machine-learning model is developed. One stability run is defined as a single loop where each monkey is left-out once

statannotation package (version 0.6.0). Microbial diversity measures were calculated using scikit-bio (version 0.6.0).

Stability runs

The samples of all macaques were processed as one cohort. As such, there was no a priori appointed external validation dataset. However, as arbitrary sample selection might result in biases, our analysis was therefore repeated 50 times where each repetition is named a stability run. This ensures the correct assessment of stochastic effects present in the results of our method.

Leave-one-monkey-out

Samples for predictive models should be independent and identically distributed between the training and testing sets. In this study, we used longitudinally collected samples from each animal. Therefore, samples cannot be selected without taking the individual animal into account as having samples from the same animal in both the training and testing sets breaks the assumption of an independent distribution. As such, a so-called Leave-One-Monkey-Out approach was applied (Fig. 1). To test the performance of our model on novel, independent data, iterative approach was applied. All samples from one left-out monkey were used as test data, only to be used for evaluation of the model. Subsequently, all other samples were used to train our model. This model was then applied to predict the diarrhea phenotype in all samples of the left-out monkey according to the microbial signature of its samples. Model predictions on all left-out monkeys are then concatenated and a single Receiver Operator Characteristic (ROC) curve was created across all predictions and true labels of the left-out samples. The Area Under the Curve (AUC) of this combined ROC curve was used for model evaluation. A full loop where each monkey is used as the left-out monkey is the definition of a single stability run.

Feature selection

Prior to passing data to our model, two methods of dimensionality reduction were performed. Unsupervised sparsity selection was performed on the non-zero relative abundance of each microbial species, to remove species that are sparser and are thus more susceptible to stochastic effects. A threshold was set, such that species that had

a relative abundance of zero in more than 80% of samples were removed from the dataset. Subsequently, supervised multivariate recursive feature elimination was performed as available within the scikit-learn package [55]. As an evaluator for feature selection, a step size of 1 and scikit-learn Extra-Trees classifier with default parameters was used. Ten repetitions of threefold cross-validation were performed to obtain the average validation ROC-AUC for a range between 1–150 of kept features.

Feature importance

A permutation feature importance algorithm was performed on all models to verify its predictions. Each feature was presented to the model ten times, but the values were shuffled every time. The feature importance of each microbial species was defined as the mean absolute change in model performance when its relative abundance values were shuffled.

Permutation test

In typical ML settings, the number of samples should outweigh the number of measurements (in this case microbial species) by at least a factor 10 as a rule of thumb [56]. In the current setting, the measurements outweigh the samples by an approximate factor of 10. For this reason, we implemented a feature selection method based on all samples. Because this has implications for risk of overfitting of the trained models, a permutation test was performed to quantify the overfitting of our models. This test measures the performance of a ML algorithm compared to a randomly permuted reference outcome. If there is no significant difference between a randomly permuted outcome and the true labels, it is possible that the performance of the measured might be caused by spurious effects, programming error, or biases induced by selecting features on the full set of samples.

We ran a permutation test consisting of 50 additional stability runs, the same as for the normal runs, where the exact same method as for the actual labels was used with randomly permuted labels. The same feature selection algorithm, parameters and models were used. In addition, the same number of features were selected to be used for the model, but on the permuted labels. This permutation test will show the difference in performance between the true labels and completely random labels. Differences in performance were assessed using an unpaired two-sided independent t-test.

Statistical analysis

Statistical testing to assess differences in group distribution were performed by either an unpaired t-test, a Kruskal Wallis test or a Mann–Whitney U rank test. Multiple testing correction was performed using Benjamin-Hochberg correction when applicable. A one-way ANOVA test was used to analyze the differences in alpha-diversity (Shannon-index) between the diets. The specific test used to calculate p-values is indicated for each specific case. All tests and corrections were implemented using the SciPy package [57]. The effect of a lactose containing diet was evaluated using a mixed logistic model with diarrhea as outcome variable, diet containing lactose as fixed and animal as random variable. Reported as odd ratio (OR) with 95% confidence intervals CI (lower limit-upper limit). This specific test was performed in R studio v4.1.3. *P*-values of <0.05 were considered statistically significant.

A power calculation was conducted with pwrEWAS to determine the sample size necessary to reliably predict the difference among the groups [58]. Based on our computations adapted for whole-genome shotgun data, a sample size of 110 microbial measurements was estimated to be sufficient.

Results

We analyzed a total of 114 samples from 13 macaques: 35 samples of four control animals and 79 samples of the nine diarrhea animals. Based on the Waltham score threshold of 3.6, we labeled 78 samples as non-diarrhea samples and 36 samples as diarrhea. Genomic transcripts of diarrheal pathogens commonly identified in rhesus macaques *Campylobacter coli*, *Campylobacter jejuni*, *Shigella flexneri*, and *Yersinia enterocolitica* represented an insignificant percentage of reads (max. relative abundance $<0.07\%$) ruling them out as causative agents. The top five most abundant phyla were *Firmicutes* ($91.8\% \pm 7.38\%$), *Bacteroidetes* ($1.84\% \pm 5.41\%$), *Actinobacteria* ($1.67\% \pm 1.48\%$), *Proteobacteria* ($0.37\% \pm 0.53\%$), and *Euryarchaeota* ($0.25\% \pm 1.28\%$). An unpaired t-test gave the following *P*-values for *Firmicutes* ($P=0.64$), *Bacteroidetes* ($P=0.37$), *Actinobacteria* ($P=0.18$), *Proteobacteria* ($P=0.75$) and *Euryarchaeota* ($P=0.30$), showing no significant differences on a phylum level between diarrhea and non-diarrhea animals. In addition, we observed no significant difference in α-diversity (using Shannon-index) on species level between the healthy group and the diarrhea group ($P=0.097$). Boxplots of α-diversity of normal fecal samples vs. diarrhea are shown in supplementary Fig. S1. Scatter plots of the first two principal coordinates of the β-diversity PCoA based on Bray–Curtis dissimilarity based on the relative abundances at the genus level are shown in Fig. S2. A bar plot of explained variance for each principal coordinate is shown in Fig. S3. Significant differences in the first four principal coordinates were calculated using a Mann–Whitney U test for binary values and a Kruskal Wallis test for multiclass values. All corrected *P*-values are summarized in

table S2. Sample name i.e., individual macaques, showed significant differences in the first three principal coordinates ($P_{PCo1}=0.0441$, $P_{PCo2}=0.0483$ and $P_{PCo3}=0.0441$) Fig. S4. This shows that there are differences in microbiome between the animals on a large-scale. However, none of these differences are associated with fecal consistency Fig. S5. The timepoints at which a sample was collected shows only a significant difference at $P_{PCo3}=0.0441$) fig. S6.

Model performance and permutation test

The average ROC curves of each of the models is shown in Fig. 2. With an average AUC of 0.825 ± 0.009, our models were able to accurately predict diarrhea in the rhesus macaques using the biomarkers i.e., microbial species, that our feature-selection algorithm has identified. The model yields optimal performance in discriminating between normal feces and diarrhea, with twenty-five microbial species identified as the most critical biomarkers. The average AUC of the permuted responder column was equal to 0.694 ± 0.077. The fact that the AUC is higher than a true coin toss result, i.e. an average AUC of 0.5 suggests that our feature selection algorithm is inducing some bias in our results. An independent t-test ($P=1.3E{-}35$) showed that the true labels result in significantly higher performance, indicating that the true labels are required to reach the highest AUC value. The comparative performances of the true and permuted responder labels are visualized in the boxplot shown in Fig. 3.

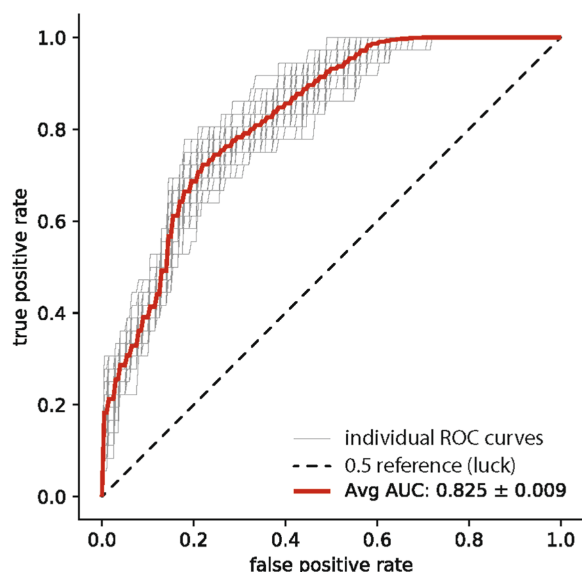

Fig. 2 Area Under the Receiver Operating Characteristic curve (AUC) shows the performance of the machine learning method using an L2-model, mean AUC ± SD

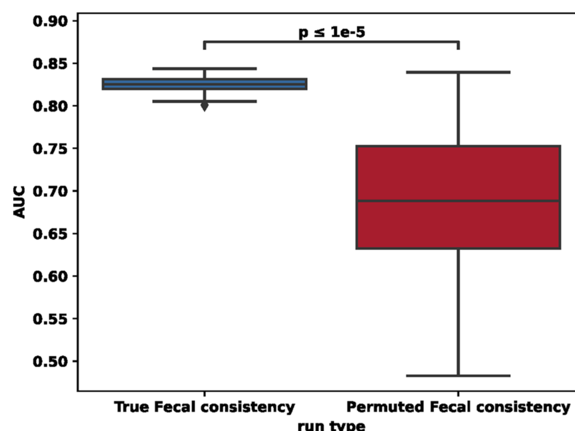

Fig. 3 Boxplots illustrating the comparative performances of the model using true and randomly permuted fecal consistency labels. The true labels show a significantly higher AUC ($P=1.3E{-}35$)

Signature gut microbiome

The twenty-five most discriminative microbial species identified in regard to diarrheal status and their relative feature importance as derived by ML models, are shown in Fig. 4. Eighteen out of the 25 (72%) identified species were more abundant in the diarrhea group, compared to the animals with normal feces. The species comprising the identified as being most discriminative microbiome were mainly derived from three phyla, of which *Firmicutes* was the most important, followed by *Actinobacteria* and *Bacteroidetes*. Members of the *Lactobacillaceae* family were most (40%) represented in animals with diarrhea, followed by *Lachnospiraceae* (20%), *Veillonellaceae* (16%), *Atopobiaceae* (8%). Full taxonomic classifications of these 25 species are provided in Table S3.

Clustering

To biologically interpret the features identified with ML, we generated hierarchically clustered heatmaps (Figs. 5 and 6). Figure 5 visualizes the correlations between the 25 biomarkers identified with the use of ML. Most of these biomarkers can be divided in two main clusters. One cluster consisting of species belonging to the *Lactobacillaceae* cluster, and one cluster (ML cluster) consisting of multiple bacteria families and species including *Ligilactobacillus ruminis*, *Megasphaera elsdenii*, *Dialister massiliensis*, *Dialister hominis*, *Anaerobutyricum hallii*, *Dorea longicatena*, *Blautia sp. sc05b48*, *Parolsenella catena*, and *Olsenella timonensis*. Figure 6 shows the 250 most abundant organisms, resulting in various visually apparent clusters. The normal commensal species were represented by the large central block (highlighted in green), whilst diarrhea and/or dysbiosis associated bacteria are

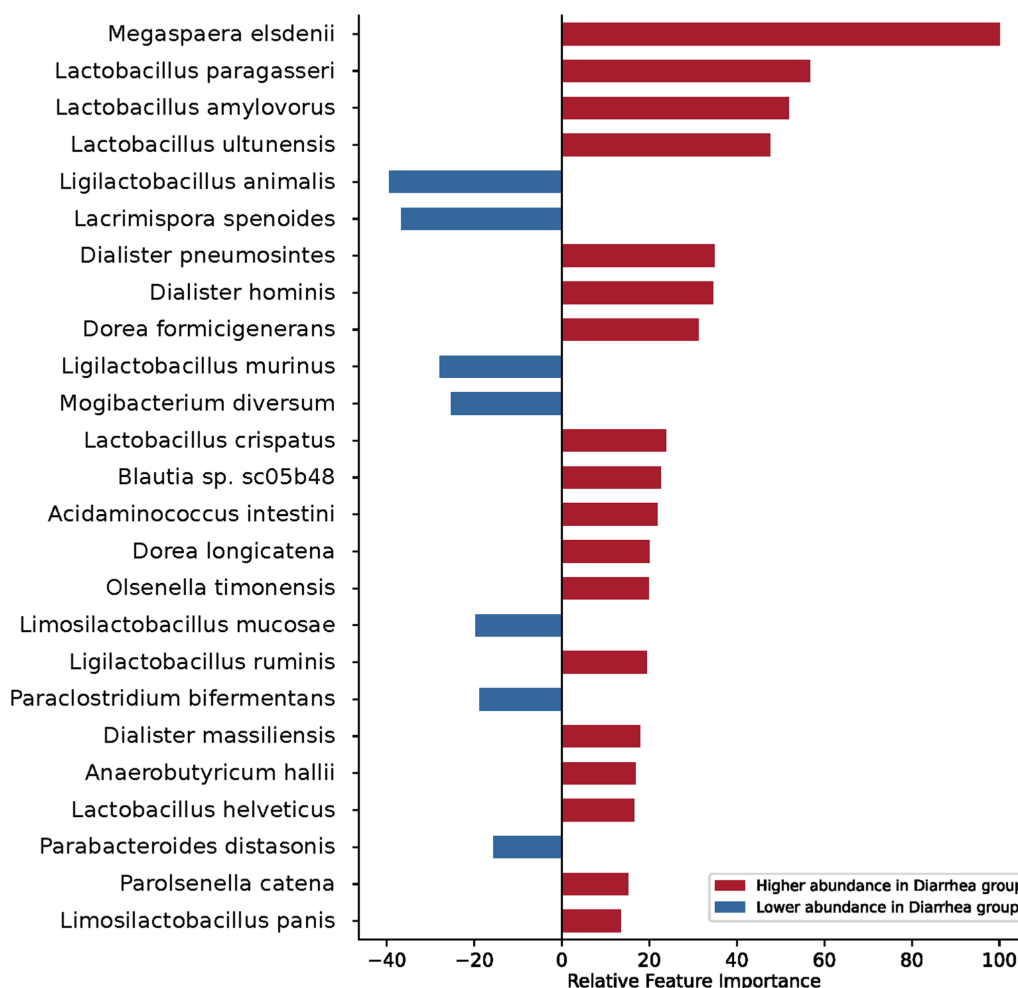

Fig. 4 Mean relative feature importances of the biomarkers, i.e., microbial species, in the models. The direction and color of the bar indicates in which group the mean of each bacterial species was higher

located in blocks at the top and bottom of the heatmap. Notably, Fig. 6 also illustrates the positioning of the *Lactobacillaceae* cluster and the ML-cluster shown in Fig. 5. The summed-up abundance of species of the ML-cluster and *Lactobacillus* genus are both higher in the diarrhea group ($P = 0.0139$ and 0.0139, respectively, Fig. 7 and 8, Mann–Whitney U rank). Many of the species shown in Fig. 5 are lactate and succinate producers (including all *Lactobacillales* species). The species that cluster within the ML-cluster are fermenters of intermediate fermentation products, such as lactate and succinate.

Diets and waltham score

A one-way ANOVA test revealed no significant differences ($P = 0.38$) in α-diversity of the microbiome between the diets (Fig. S7). Yet, less animals were observed with a Waltham score of > 3.6 in the diarrhea group when fed with Diet D and E (Fig. 9). Two samples of the diarrhea group were unavailable because no sample could be collected indoors. Strikingly, in the diarrhea group none of the animals had diarrhea when fed with Diet E. In addition, within the diarrhea group six out of nine animals had normal feces when fed Diet D compared to three animals out of nine for Diet A. The odds ratio of having diarrhea when fed a lactose containing diet was 8.79 (CI 1.7–45.4, P = 0.009) compared to having no diarrhea 0.11 (CI 0.02–0.59, P = 0.009).

As for β-diversity, the differences between each individual diet were significant in the third and fourth principal coordinates ($P_{PCo3} = 0.0136$, $P_{PCo4} = 0.000635$). Diets D and E show the biggest differences in distribution in these principal coordinates (Fig. 10A, B). Therefore, we grouped the diets based on whether they are lactose free or not (Fig. 10C, D). Based on lactose content, we observe significant differences in the third and fourth principal coordinate ($P_{PCo3} = 0.0136$

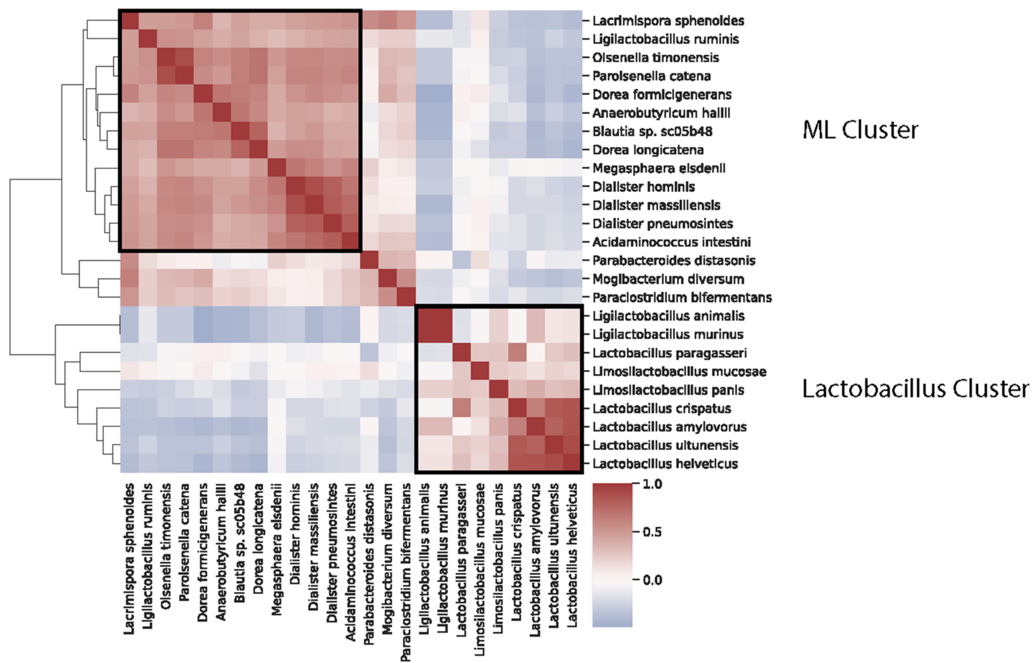

Fig. 5 Clustered heatmap of the correlations using Spearmann-rho coefficients of the 25 bacterial species identified with the use of machine learning (ML). The observed Lactobacillaceae cluster and the ML cluster are highlighted separately

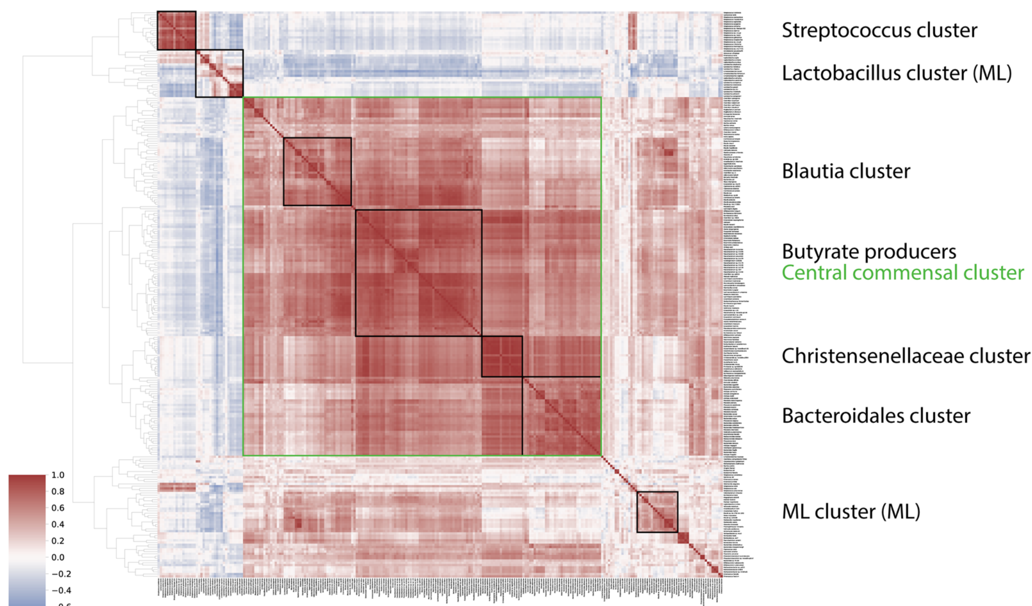

Fig. 6 Clustered heatmap of the correlations using spearman-rho coefficients of the top 250 most abundant microbes on species level. Characterizable clusters of species are highlighted separately in the clustered heatmap. The two clusters that are identified by machine learning are indicated with (ML)

and $P_{PCo4} = 0.000505$). Eigenvectors obtained in the β-diversity analysis with respect to the third and fourth principal coordinates are shown in a biplot (Fig. S8). We chose these principal coordinates as they show the most interesting differences with respect to individual diets and lactose content. The top-5 most important genera in this plane are *Prevotella, Streptococcus, Ligilactobacillus, Limosilactobacillus* and *Vescimonas*,

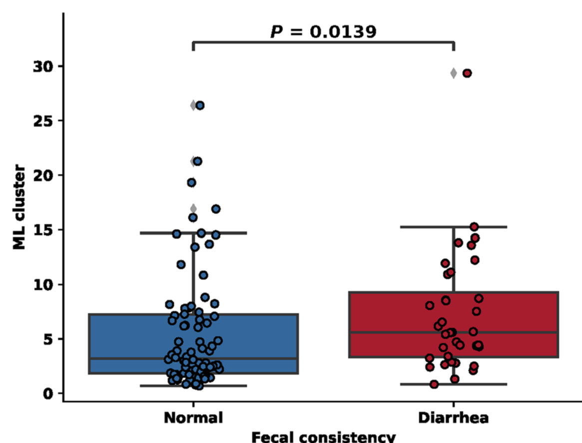

Fig. 7 Boxplot of the summed relative abundances of the machine-learning (ML) cluster in Fig. 6. The macaques with normal feces show a significantly lower abundance of these species compared to the macaques with diarrhea ($P=0.0139$, BH-corrected Mann–Whitney U rank)

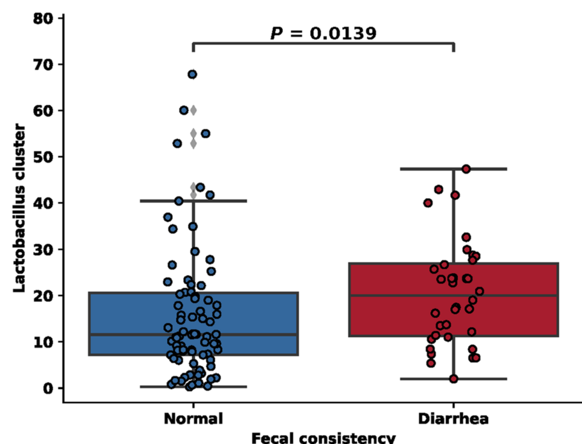

Fig. 8 Boxplot of the summed relative abundances of the species belonging to the Lactobacillus genus. The macaques with normal feces show a significantly lower abundance of these Lactobacillus species compared to the macaques with diarrhea ($P=0.0139$, BH- corrected Mann–Whitney U rank)

where *Limosilactobacillus* is the most correlated with the difference between lactose-free and lactose containing diets (lower-right vs. top-left). The overall less abundant genus, *Faecalibacillus,* showed a significantly increased abundance between lactose free (Diet D and E) and lactose containing diets (Diet A, B and C) ($P=0.036$, Mann–Whitney U test). We observed no significant differences with other diet combinations.

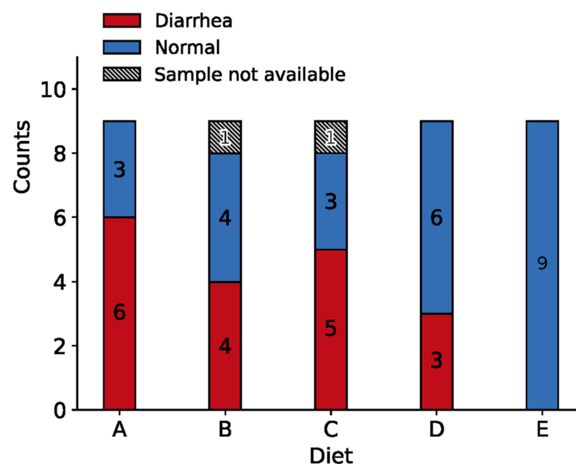

Fig. 9 The number of animals classified as normal (≥ 3.5) versus diarrhea (≤ 3.6) for each diet based on mean Waltham score

Discussion

The current study shows a distinct microbiome signature in rhesus macaques with chronic intermittent diarrhea hints at a potential mechanism of action. The classification performance of our ML model indicates a strong ability to distinguish between diarrhea and non-diarrhea macaques, which is an important prerequisite for the interpretation of identified biomarkers. The most clinically relevant result within this signature was an overabundance of *Lactobacilli* and a cluster of various species who have in common that they are fermenters of intermediate fermentation products such as lactate and succinate. The present study was designed to determine the effect of diet interventions on the gut microbiome and diarrhea status of rhesus macaques. Diet D and E were associated with a reduced diarrhea incidence and and β-diversity analyses show them to have a significantly different microbiome composition compared to the other diets.

In line with Westreich et al. [36] we observed a higher abundance of *Lactobacillus* and *Megasphaera* in the diarrhea group. However, this outcome contrasts Yang et al. [37] who described a depletion of *Lactobacilli* in rhesus macaques with diarrhea. In comparison to our study the age distribution in Yang et al. [59] displayed a broad range. Their overall age range consisted of rhesus macaques between 2 and 19 years (median 8 years) compared to a relatively small range of 3–7 years (median 5 years) in our study. Yet, it has been described that the abundance of *lactobacilli* in young adult, adult and old cynomolgus macaques was relatively stable and therefore an unlikely explanation. More likely, the discrepancies between these studies could be caused by differences in origin, geography, husbandry and housing. Moreover,

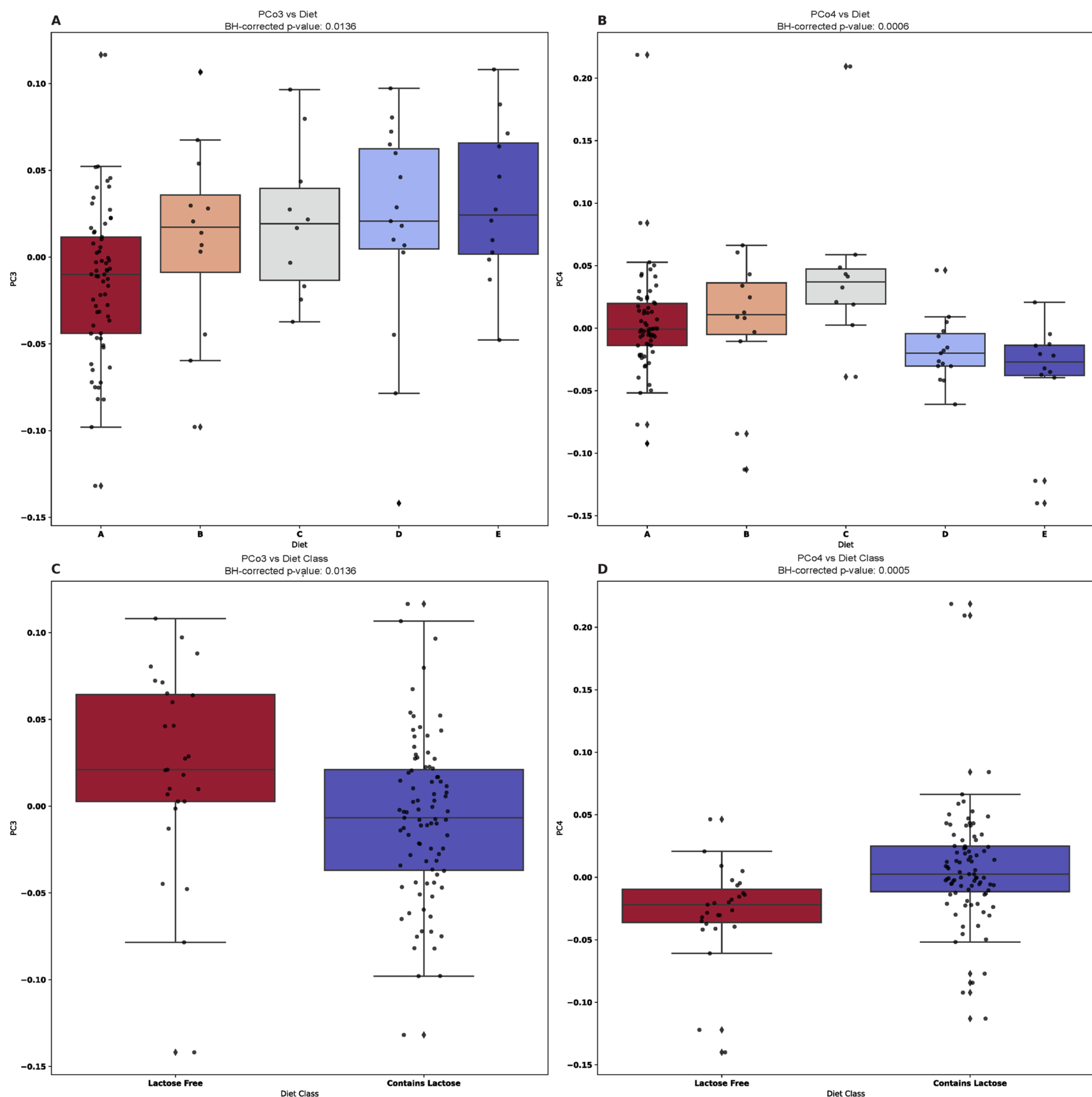

Fig. 10 The differences between each individual diet were significant in the third **A** and fourth **B** principal coordinates ($P_{\text{PCo3}} = 0.0136$, $P_{\text{PCo4}} = 0.000635$). Diets D and E show the most differences in distribution in these principal coordinates (A–B). Grouped in lactose free vs lactose containing diets show a significant difference in the third **C** and fourth **D** principal coordinate ($P_{\text{PCo3}} = 0.0136$ and $P_{\text{PCo4}} = 0.000505$)

difference in sampling frequency, one single sample per animal compared to the longitudinal follow up in our study could also have attributed. In addition, although the total number of animals included in our study was lower, the total number of analyzed samples (n = 114) was much higher.

Shifts in the gut microbiome can be induced by dietary interventions. It is known that high-fat diets, mediterranean versus westernized diets, wild versus provisioned or captivity diets induce different microbiome compositions in NHP [26, 59–61]. Yet, in our study differences in microbiome between the individual diet interventions were nonsignificant in the first and second principal coordinates. This is most likely due to the effect of the observed inter-individual variance. Other reasons are (1) the limited number of animals, (2) that

the composition and ingredients of the diets did not differ enough from each other, as they were all commercial monkey diets and lastly (3) the inherent instability of the gut microbiome of these idiopathic diarrheal animals [62, 63]. However, the third and fourth principal coordinates show indeed a significant difference. Where the first and second principal coordinates probably account for more inter-individual variability, the third and fourth principal coordinates are likely more aligned with the effect of diet.

We observed on genus level a significantly more abundant *Faecalibacillus* in the lactose free diets compared to the lactose containing diets. Interestingly, Huang et al. [64] observed potential probiotic effects of *Faecalibacillus intestinalis*. Although the exact meaning of the *Faecalibacillus* difference remains unclear, there could be a possible positive relationship between the lower diarrhea incidence, lactose free diets and the abundance of *Faecalibacillus* in our macaques.

Firrman et al. [65] examined the impact of lactose exposure on the gut microbiome in vitro, using fecal samples originating from healthy human donors. Lactose treatment decreased the relative abundance of *Bacteroidaceae* and increased lactic acid bacteria such as *Lactobacillacea*, *Enterococceae*, and *Streptococcaceae.* This corresponded with an increased abundance of lactate utilizers such as *Veillonellaceae.* In line with these findings, we observed an increased abundance of *Lactobacillacea* and *Veillonellaceae* in the diarrhea microbiome signature. In contrast, for the lactose free we observed a decrease in the *Bacteroidales* cluster. In addition, we did not observe a significant difference in abundance in *Enterococceae* and *Streptococcaceae*, this could be explained by the fact that the most abundant *Streptococcaceae* in our dataset were of the species *equini* and *lutetiensis,* both are unable to ferment lactose [66, 67]. In addition, the abundance of *Enterococceae* was overall very low (0.03%).

We observed a reduction of diarrhea incidence in the diarrhea group when fed Diet D and E. In retrospect, the most striking commonality between these diets was the absence of lactose. In contrast, the other diets consisted of lactose rich ingredients such as whey powder or milk and cream powder. Diets D and E showed also the biggest differences in distribution in the principal coordinates compared to the other lactose containing diets A, B and C. Grouped together in lactose free and lactose containing diets, we observed a significant difference in the third and fourth principal coordinate. The reduction of the diarrhea incidence combined with the differences in β-diversity when the animals where fed Diet D and E, suggests that lactose contributes to diarrhea symptoms.

In addition, we observed an overabundance of *Lactobacilli*, which produce lactate and succinate from lactose [68, 69], and a cluster of species subsequently capable of fermenting intermediate fermentation products such as lactate and succinate (ML-cluster).

The overabundance of lactate and succinate consumers (ML-cluster) is indicative of an overabundance of available lactate and succinate in the large intestine, as has indeed been shown by others [70, 71]. Figure 11 illustrates our hypothesis regarding the suggested lactose

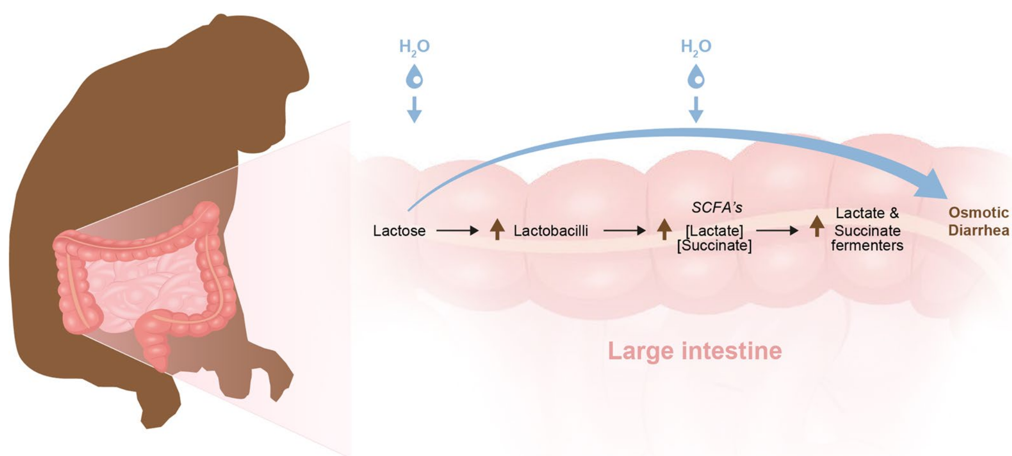

Fig. 11 Illustration of the hypothesized lactose intolerance relationship and the observed microbiome signature in macaques with diarrhea. Due to malabsorption of lactose in the small intestine water is trapped, subsequently the lactose enters the large intestine to primarily be fermented by lactobacilli. Consequently, the abundance of lactobacilli increases by the excess of substrate. The lactobacilli ferment the lactose into short chain fatty acids (SCFA) resulting in elevated concentrations of lactate, succinate in the lumen. The increased concentration of lactate and succinate leads to proliferations of lactate- and succinate-consuming bacteria. Water is attracted by the lumen of the intestine due to the osmotic effect of the increased SCFAs concentration, together with the earlier trapped water in the intestine resulting in osmotic diarrhea

intolerance relationship and diarrhea in our macaques. An overabundance of available lactate and succinate, produced by lactose consumers such as *Lactobacillus*, might be due to an overabundance of lactose which is not taken up properly in lactose intolerant individuals i.e., lactose malabsorption. It is known that inefficient digestion of lactose due to reduced or impaired lactase activity in the small intestine results in a higher concentration of lactose entering the caecum and large intestine [21, 24] and it is furthermore shown that lactose intake in lactose intolerant individuals/humans can induce an increase in *Lactobacilli* [20, 72]. Normally, lactate and succinate are a by-product of microbial anaerobic fermentation of dietary fiber that is rapidly converted into propionate in the case of succinate and in the case of lactate into acetate, propionate and butyrate [20, 73]. These are subsequently absorbed by colonocytes. Therefore, the succinate concentration in the lumen (or feces) is usually low (1–3 mmol/kg). In pathophysiological situations the concentrations of intermediate fermentation products can increase in the lumen of the large intestine. As a result of the increased concentration of intermediate short-chain fatty acids (SCFAs) such as succinate, lactate and formate secretory diarrhea can occur [69, 72, 74, 75]. It has similarly been shown in rats by Xue et al. [72] that lactose intolerance results in osmotically induced diarrhea. Therefore, our data suggests a possible etiology of lactose intolerance in our initially idiopathic diarrhea group. There is a relatively small body of literature describing lactose intolerance in NHP. Hart et al. [14] reported lactose intolerance in two sibling rhesus macaques. Both infants were hand raised with milk formula, both suffered from diarrhea and slow weight gain. These symptoms dissolved when the animals were fed a lactose-free formula. Wen et al. [17] observed a 100% incidence of lactose intolerance in seven adult wild caught rhesus macaques when challenged with lactose rich diets. However, adaptation to long term lactose feeding did occur. Our study group comprised individuals of the fourth generation or beyond at the BPRC and the same lactose-containing Diet A, was fed for decades. However, our diets had relatively low lactose levels at 3.4% and 3.5% compared to 20% lactose containing diets fed in the study by Wen et al. [17]. Adaptation to lactose levels in macaques at the BPRC is expected. Yet, our study indicates that some macaques with diarrhea remain lactose intolerant despite multigenerational captivity and lifelong exposure to dietary lactose.

In humans, it has been described that lactose intolerant individuals usually can tolerate a dose of 12–24 g lactose per day [76, 77]). In macaques, a corresponding estimated tolerated dose would be 1.6–3.2 g of lactose, assuming a standard human bodyweight of 70 kg and adjusted for the mean bodyweight of the macaques in our study. The lactose containing diets contained of 3.4 and 3.5% lactose which results in an intake of approximately 5 g lactose per day. This indicates that the daily extrapolated macaque lactose toleration threshold was exceeded with 313–156% when fed Diet A, B and C. In combination with our finding that lactose-free Diets D and E resulted in significant lower incidence of diarrhea compared to lactose-containing diets this suggests that lactose could have contributed to the diarrhea.

In addition to being free of lactose, Diet E had higher dietary fiber levels. This might explain why none of the animals experienced diarrhea. For example, the two increased *Atopobiaceae* species of the genus *Olsenella* and *Parolsenella* tend to be susceptible for fiber supplementation in dogs in relation to diarrhea. Levels of *Atopobiaceae* genus *Atopobium* were at comparable levels between diarrhea and control dogs, but fiber intervention resulted in a decrease in diarrhea dogs and an increase in control dogs [78]. In addition, in marmosets, an increase in *Parolsenella catena* was observed upon yoghurt supplementation, which limited Multiple-sclerosis pathology [79]. Furthermore, in Asian colobines, a leaf-eating old world monkey, the *Atopobium* genus was relatively more abundant in wild colobines compared to their captive counterparts, possibly highlighting its association with fiber-rich wild diets [80]. Yang et al. [37] observed no significant changes in *Olsenella* species between the diarrhea group and the control group. Although, in humans, a relation between members of the *Atopobiaceae* family and gastrointestinal health has been proposed [81, 82], these inconclusive reports implicate that its exact role is not clear.

In our study, a predictive modeling ML algorithm was employed that demonstrated robust performance, as evidenced by an average Receiver Operating Characteristic Area Under the Curve (ROC-AUC) of 0.825. This ROC-AUC score is indicative of the algorithm's effective discrimination between diarrhea and normal Waltham scores in rhesus macaques, underscoring its utility in our research. This result showed that such a non-linear multivariate predictive modeling approach can provide additional insights into fecal metagenomics data when sample sizes are limited. The permutation feature importances calculated allowed for identification and quantification of the relevant patterns in the fecal metagenome rhesus macaques associated with lactate and succinate consumption and production. However, it is important to note that our models are primarily developed for biomarker discovery, and as such, feature selection was conducted on the entire sample pool. This approach, while beneficial for biomarker identification, may limit the generalizability of the models for future applications, as the

selected features may be over-fitted to the current dataset. Despite this, rigorous steps have been taken to ensure the validity of our findings, i.e., the comparison with the results of a randomly permuted version of itself yielded a significant lowering of performance. The biomarkers identified through ML were further validated using additional analyses and corroborated with existing literature. This multi-faceted approach reinforced the validity of our results.

The limitation of this study was the relatively low number of animals. Compared to similar monkey studies with a range of 24–96 cross sectional samples, our study included 114 samples from 13 individual animals [26, 36, 37, 61]. With our longitudinal follow up we complied with the suggestion that longitudinal and intervention studies combined with methodological approaches are of importance to advance research on microbiome association and modulation [83]. However, for ML and statistics in general more samples and more animals would have resulted in more robust data. On the other hand, the results of both ML and biological analysis were consistent, thereby increasing confidence and robustness of the outcome.

Overall, collecting large sample sizes is a common challenge in animal research. Availability of patients and matching controls can be limited. Since we also studied patients, this limited the inclusion of different genders and the total number of animals. Date of last treatment, and the time-frame in which patients had arrived from the breeding colony also influenced the selection of the animals. Although the fecal sample collections were conducted within consistent time ranges, variability in defecation time and sample collection could not be controlled. Meanwhile, the macaques used in this study were housed and cared for identically thus decreasing inter-animal variability. Controlling variability, i.e. noise, is a known strategy in preclinical animal research to reduce sample size. In addition, in our study biases were also reduced by the repeated measures design: feces of each animal were scored and recorded every two weeks during the entire study. Furthermore, the random assignment of the diets reduced the impact of unknown confounders.

Conclusion

This study suggests microbiome-lactose intolerance relationship in rhesus macaques with intermittent chronic diarrhea through the application of ML data analysis techniques and biological interpretation. We combined metagenomic profiling with ML algorithms, offering an additional avenue to investigate the complex interactions between diet, gut microbiota, and health outcomes in rhesus macaques. In particular, our approach has not only provided insights into the specific microbial signatures possibly associated with lactose intolerance but is also bridging the gap between computational biology and animal health, offering a potential blueprint for similar research in other species and contexts.

The integration of ML with metagenomic data analysis holds potential for developing targeted dietary interventions and therapeutic strategies. As technology continuous to evolve, our ability to understand and possibly manipulate the microbiome through dietary interventions will improve, leading to better health outcomes of rhesus macaques. Although additional research is warranted, in the future this understanding will not only enhance the welfare of macaques in zoos and research facilities but also pave the way for improved management practices, ensuring healthier and more resilient primate populations.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s42523-024-00338-z.

Additional file1
Additional file2
Additional file3
Additional file4
Additional file5
Additional file6
Additional file7
Additional file8
Additional file9
Additional file10
Additional file11

Acknowledgements
The authors would like to thank the Animal Science department of the BPRC, especially the animal caretakers for excellent care of the animals and for assisting with sample collection, Francisca van Hassel for her assistance with figure design, Thea de Koning for reading and editing. We also would like to thank Donghyeok Lee, data scientist at HORAIZON Technology BV, for assisting with the power calculations.

Author contributions
Conceptualization: A.M., E.J.R., E.L., A.S., F.S., R.M. Methodology: A.M., E.J.R., B.V., E.L., R.M. Data curation: B.V., E.L. Formal analysis: B.V., E.L., M.C.d.G, N.P. Writing-original draft preparation A.M., B.V. Writing-review and editing, E.L., M.C.d.G., N.P., F.S., A.S., J.A.M.L., J.B., R.M. Supervision, E.L., J.A.M.L., J.B., R.M. All authors have read and approved the final manuscript.

Funding
No external funding was received.

Availability of data and materials
Sequence data that support the findings of this study have been deposited in the European nucleotide archive: https://www.ebi.ac.uk/ena/browser/view/PRJEB70928.

Declarations

Ethics approval and consent to participate
All animals were housed in accordance with Dutch law and international ethical and scientific standards and guidelines (EU Directive 63/2010). All procedures and husbandry were compliant with the above standards and legislations. No interventions other than required for veterinary care were performed on these animals, therefore, no approval from the competent authorities was required. Nevertheless, additional approval was obtained by the institutional animal welfare body (IvD 018A). The Biomedical Primate Research Centre (BPRC) is accredited by the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC) International.

Consent for publication
Not applicable.

Competing interests
Kasper Faunafood (Woerden, The Netherlands) and Altromin Spezialfutter GmbH (Lage, Germany) kindly sponsored their diets used in this study. The sponsors had no role in the design of the study; in the sample collection, analyses or data interpretation; in writing of the manuscript, or in the decision to publish the results.

Author details
[1]Biomedical Primate Research Centre, Lange Kleiweg 161, 2288 GJ Rijswijk, The Netherlands. [2]Department Population Health Sciences, Animals in Science and Society, Faculty of Veterinary Medicine, Utrecht University, Heidelberglaan 8, 3584 CM Utrecht, The Netherlands. [3]HORAIZON Technology BV, Marshallaan 2, 2625 GZ Delft, The Netherlands. [4]Department of Vascular Medicine, Amsterdam UMC, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands. [5]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton CB10 1SA, UK. [6]Department of Microbiology and Systems Biology, Organization for Applied Scientific Research (TNO), Sylviusweg 71, 2333 BE Leiden, The Netherlands.

Received: 2 April 2024 Accepted: 6 September 2024
Published online: 23 September 2024

References

1. Holmberg CA, Leininger R, Wheeldon E, Slater D, Henrickson R, Anderson J. Clinicopathological studies of gastrointestinal disease in macaques. Vet Pathol I. 1982;9:163–70.
2. Muñoz-Zanzi CA, Thurmond MC, Hird DW, Lerche NW. Effect of weaning time and associated management practices on postweaning chronic diarrhea in captive rhesus monkeys (*Macaca mulatta*). Comp Med. 1999;49(6):617–21.
3. Rout WR, Formal SB, Dammin GJ, Giannella RA. Pathophysiology of salmonella diarrhea in the rhesus monkey: intestinal transport, morphological and bacteriological studies. Gastroenterology. 1974;67(1):59–70.
4. Kanthaswamy S, Elfenbein HA, Ardeshir A, Ng J, Hyde D, Smith DG, Lerche N. Familial aggregation of chronic diarrhea disease (CDD) in rhesus macaques (*Macaca mulatta*). Am J Primatol. 2014;76(3):262–70.
5. Prongay K, Park B, Murphy SJ. Risk factor analysis may provide clues to diarrhea prevention in outdoor-housed rhesus macaques (*Macaca mulatta*). Am J Primatol. 2013;75(8):872–82.
6. Johnson AL, Keesler RI, Lewis AD, Reader JR, Laing ST. Common and Not-So-Common pathologic findings of the gastrointestinal tract of rhesus and cynomolgus macaques. Toxicol Pathol. 2022;50(5):638–59.
7. Russell RG, Sarmiento JI, Fox J, Panigrahi P. Evidence of reinfection with multiple strains of campylobacter jejuni and campylobacter coli in macaca nemestrina housed under hyperendemic conditions. Infect Immun. 1990;58(7):2149–55.
8. Schneider NJ, Prather EC, Lewis AL, Scatterday JE, Hardy AV. Enteric bacteriological studies in a large colony of primates. Ann N Y Acad Sci. 1960;85(3):935–41.
9. Sestak K, Merritt CK, Borda J, Saylor E, Schwamberger SR, Cogswell F, Didier ES, Didier PJ, Plauche G, Bohm RP, et al. Infectious agent and immune response characteristics of chronic enterocolitis in captive rhesus macaques. Infect Immun. 2003;71(7):4079–86.
10. Delwart E, Tisza MJ, Altan E, Li Y, Deng X, Hartigan-O'connor DJ, Ardeshir A. Idiopathic chronic diarrhea in rhesus macaques is not associated with enteric viral infections. Viruses. 2021;13(12):2503.
11. Wang KY, Christe KL, Yee J, Roberts JA, Ardeshir A. Rotavirus is associated with decompensated diarrhea among young rhesus macaques (*Macaca mulatta*). Am J Primatol. 2019;81(1):e22948.
12. Wang Y, Tu X, Humphrey C, McClure H, Jiang X, Qin C, Glass RI, Jiang B. Detection of viral agents in fecal specimens of monkeys with diarrhea. J Med Primatol. 2007;36(2):101–7.
13. Bethune MT, Borda JT, Ribka E, Liu MX, Phillippi-Falkenstein K, Jandacek RJ, Doxiadis GG, Gray GM, Khosla C, Sestak K. A non-human primate model for gluten sensitivity. PLoS One. 2008;3(2):e1614.
14. Hart NA, Reeves J, Dalgard DW, Adamson RH. Lactose malabsorption in two sibling rhesus monkeys (*Macaca mulatta*). J Med Primatol. 1980;9(5):309–13.
15. Mustafa T, Li Q, Kelly LE, Gibbon A, Ryan I, Roffey K, Simonds S, Cowley MA, Sleeman MW. Food hypersensitivity-induced chronic gastrointestinal inflammation in a non-human primate model of diet-induced obesity. Plos one. 2019;14(4):e0214621.
16. Sestak K, Conroy L, Aye PP, Mehra S, Doxiadis GG, Kaushal D. Improved xenobiotic metabolism and reduced susceptibility to cancer in gluten-sensitive macaques upon introduction of a gluten-free diet. PLoS One. 2011;6(4):e18648.
17. Wen CP, Antonowicz I, Tovar E, McGandy RB, Gershoff SN. Lactose feeding in lactose-intolerant monkeys. Am J Clin Nutr. 1973;26(11):1224–8.
18. Sebastio G, Villa M, Sartorio R, Guzzetta V, Poggi V, Auricchio S, Boll W, Mantei N, Semenza G. Control of lactase in human adult-type hypolactasia and in weaning rabbits and rats. Am J Hum Genet. 1989;45(4):489–97.
19. Wang Y, Harvey CB, Hollox EJ, Phillips AD, Poulter M, Clay P, Walker-Smith JA, Swallow DM. The genetically programmed down-regulation of lactase in children. Gastroenterology. 1998;114(6):1230–6.
20. Francavilla R, Calasso M, Calace L, Siragusa S, Ndagijimana M, Vernocchi P, Brunetti L, Mancino G, Tedeschi G, Guerzoni E, et al. Effect of lactose on gut microbiota and metabolome of infants with cow's milk allergy. Pediatr Allergy Immunol. 2012;23(5):420–7.
21. Hammer HF, Ana CAS, Schiller LR, Fordtran JS. Studies of osmotic diarrhea induced in normal subjects by ingestion of polyethylene glycol and lactulose. J Clin Investig. 1989;84(4):1056–62.
22. Misselwitz B, Butter M, Verbeke K, Fox MR. Update on lactose malabsorption and intolerance: pathogenesis, diagnosis and clinical management. Gut. 2019;68(11):2080–91.
23. Suarez F, Levitt M. Lactose malabsorption and diarrhea. Nutrition. 1997;13(1):53–4.
24. Christopher NL, Bayless TM. Role of the small bowel and colon in lactose-induced diarrhea. Gastroenterology. 1971;60(5):845–52.
25. Cheng J, Zhang LT. Simulation of vessel tissue remodeling with residual stress: an application to in-stent restenosis. Int J Smart Nano Mater. 2019;10(1):11–27.
26. Nagpal R, Wang S, Solberg Woods LC, Seshie O, Chung ST, Shively CA, Register TC, Craft S, McClain DA, Yadav H. Comparative microbiome signatures and short-chain fatty acids in mouse, rat, non-human primate, and human feces. Front Microbiol. 2018;9:2897.
27. Newman TM, Shively CA, Register TC, Appt SE, Yadav H, Colwell RR, Fanelli B, Dadlani M, Graubics K, Nguyen UT, et al. Diet, obesity, and the gut microbiome as determinants modulating metabolic outcomes in a non-human primate model. Microbiome. 2021;9(1):100.
28. Hou K, Wu ZX, Chen XY, Wang JQ, Zhang D, Xiao C, Zhu D, Koya JB, Wei L, Li J, Chen ZS. Microbiota in health and diseases. Signal Transduct Target Therapy. 2022;7(1):1–28.
29. Kap YS, Bus-Spoor C, van Driel N, Dubbelaar ML, Grit C, Kooistra SM, Fagrouch ZC, Verschoor EJ, Bauer J, Eggen BJL, et al. Targeted diet modification reduces multiple sclerosis–like disease in adult marmoset monkeys from an outbred colony. J Immunol. 2018;201(11):3229–43.
30. Larroya A, Pantoja J, Codoñer-Franch P, Cenit MC. Towards tailored gut microbiome-based and dietary interventions for promoting the development and maintenance of a healthy brain. Front Pediatr. 2021;9:705859.
31. Vatanen T, Franzosa EA, Schwager R, Tripathi S, Arthur TD, Vehik K, Lernmark Å, Hagopian WA, Rewers MJ, She JX, et al. The human gut microbiome in early-onset type 1 diabetes from the TEDDY study. Nature. 2018;562(7728):589–94.

32. Vos WMD, Tilg H, Hul MV, Cani PD. Gut microbiome and health: mechanistic insights. Gut. 2022;71(5):1020–32.
33. Zheng P, Li Z, Zhou Z. Gut microbiome in type 1 diabetes: a comprehensive review. Diabet/Metab Res Rev. 2018;34(7):e3043.
34. Compo NR, Mieles-Rodriguez L, Gomez DE. Fecal bacterial microbiota of healthy free-ranging, healthy corralled, and chronic diarrheic corralled rhesus macaques (*Macaca mulatta*). Comp Med. 2021;71(2):152–65.
35. McKenna P, Hoffmann C, Minkah N, Aye PP, Lackner A, Liu Z, Lozupone CA, Hamady M, Knight R, Bushman FD. The macaque gut microbiome in health, lentiviral infection, and chronic enterocolitis. PLoS Pathog. 2008;4(2): e20.
36. Westreich ST, Ardeshir A, Alkan Z, Kable ME, Korf I, Lemay DG. Fecal metatranscriptomics of macaques with idiopathic chronic diarrhea reveals altered mucin degradation and fucose utilization. Microbiome. 2019;7:1–7.
37. Yang S, Liu Y, Yang N, Lan Y, Lan W, Feng J, Yue B, He M, Zhang L, Zhang A, Price M. The gut microbiome and antibiotic resistome of chronic diarrhea rhesus macaques (*Macaca mulatta*) and its similarity to the human gut microbiome. Microbiome. 2022;10(1):29.
38. Stols-Gonçalves D, Mak AL, Madsen MS, van der Vossen EWJ, Bruinstroop E, Henneman P, Mol F, Scheithauer TPM, Smits L, Witjes J, et al. Faecal microbiota transplantation affects liver DNA methylation in Non-alcoholic fatty liver disease: a multi-omics approach. Gut Microbes. 2023. https://doi.org/10.1080/19490976.2023.2223330.
39. van der Vossen EWJ, Bastos D, Stols-Gonçalves D, de Goffau MC, Davids M, Pereira JPB, Li Yim AYF, Henneman P, Netea MG, de Vos WM, et al. Effects of fecal microbiota transplant on DNA methylation in subjects with metabolic syndrome. Gut Microbes. 2021. https://doi.org/10.1080/19490976.2021.1993513.
40. Voermans B, De Goffau M, Nieuwdorp M, Levin E: Microbe2Pixel: Taxonomy informed deep-learning models and explanations.
41. Qin N, Yang F, Li A, Prifti E, Chen Y, Shao L, Guo J, Le Chatelier E, Yao J, Wu L, et al. Alterations of the human gut microbiome in liver cirrhosis. Nature. 2014;513(7516):59–64.
42. Vangay P, Hillmann BM, Knights D. Microbiome learning repo (ML Repo): a public repository of microbiome regression and classification tasks. Gigascience. 2019. https://doi.org/10.1093/gigascience/giz042.
43. Wastyk HC, Fragiadakis GK, Perelman D, Dahan D, Merrill BD, Feiqiao BY, Topf M, Gonzalez CG, Van Treuren W, Han S, Robinson JL. Gut-microbiota-targeted diets modulate human immune status. Cell. 2021;184(16):4137–53.
44. Maaskant A, Janssen I, Wouters IM, van Eerdenburg FJCM, Remarque EJ, Langermans JAM, Bakker J. Assessment of indoor air quality for group-housed macaques (Macaca spp.). Animals. 2022;12(14):1750.
45. Cavett CL, Tonero M, Marks SL, Winston JA, Gilor C, Rudinsky AJ. Consistency of faecal scoring using two canine faecal scoring systems. J Small Anim Pract. 2021;62(3):167–73.
46. Moxham G. Waltham feces scoring system-A tool for veterinarians and pet owners: how does your pet rate. Waltham focus. 2001;11(2):24–5.
47. Chen S, Zhou Y, Chen Y, Gu J: Fastp: An ultra-fast all-in-one FASTQ preprocessor. In: Bioinformatics: 2018; 2018.
48. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biology. 2019;20(1):1.
49. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Connor R, Funk K, Kelly C, Kim S, et al. Database resources of the national center for biotechnology information. Nucl Acids Res. 2022;50(D1):D20–6.
50. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in metagenomics data. PeerJ Comput Sci. 2017;2017(1):14082.
51. Dale R, Grüning B, Sjödin A, Rowe J, Chapman BA, Tomkins-Tinch CH, Valieris R, Batut B, Caprez A, Cokelaer T, et al. Bioconda: sustainable and comprehensive software distribution for the life sciences. Nat Methods. 2018;15(7):475–6.
52. Chen T, Guestrin C: XGBoost: A scalable tree boosting system. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: 2016; 2016.
53. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Mach Lear. 2006;63(1):3–42.
54. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A: Catboost: Unbiased boosting with categorical features. In: Advances in Neural Information Processing Systems: 2018; 2018.
55. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. Scikit-learn: machine learning in python. J Mach Learn Res. 2011;12:2825–30.
56. Abu-Mostafa YS, Magdon-Ismail M, Lin HT. Learning from Data: A Short Course. New York: AMLBook.com; 2012.
57. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, et al. SciPy 1.0: fundamental algorithms for scientific computing in python. Nat Methods. 2020;17(3):352.
58. Graw S, Henn R, Thompson JA, Koestler DC. PwrEWAS: a user-friendly tool for comprehensive power estimation for epigenome wide association studies (EWAS). BMC Bioinf. 2019. https://doi.org/10.1186/s12859-019-2804-7.
59. Wei ZY, Rao JH, Tang MT, Zhao GA, Li QC, Wu LM, Liu SQ, Li BH, Xiao BQ, Liu XY, et al. Characterization of changes and driver microbes in gut microbiota during healthy aging using a captive monkey model. Genom Proteom Bioinform. 2022;20(2):350–65.
60. Chen T, Li Y, Liang J, Li Y, Huang Z. Gut microbiota of provisioned and wild rhesus macaques (*Macaca mulatta*) living in a limestone forest in southwest Guangxi China. MicrobiologyOpen. 2020. https://doi.org/10.1002/mbo3.981.
61. Jia T, Chang WS, Marcelino VR, Zhao S, Liu X, You Y, Holmes EC, Shi M, Zhang C. Characterization of the gut microbiome and resistomes of wild and zoo-captive macaques. Front Vet Sci. 2022;8:778556.
62. Boolchandani M, Blake KS, Tilley DH, Cabada MM, Schwartz DJ, Patel S, Morales ML, Meza R, Soto G, Isidean SD, et al. Impact of international travel and diarrhea on gut microbiome and resistome dynamics. Nat Commun. 2022. https://doi.org/10.1038/s41467-022-34862-w.
63. Mättö J, Maunuksela L, Kajander K, Palva A, Korpela R, Kassinen A, Saarela M. Composition and temporal stability of gastrointestinal microbiota in irritable bowel syndrome—a longitudinal study in IBS and control subjects. FEMS Immunol Med Microbiol. 2005;43(2):213–22.
64. Huang P, Dong Q, Wang Y, Tian Y, Wang S, Zhang C, Yu L, Tian F, Gao X, Guo H, et al. Gut microbial genomes with paired isolates from China illustrate probiotic and cardiometabolic effects. Cell Genom. 2024;4(6): 100559.
65. Firrman J, Liu L, Mahalak K, Hu W, Bittinger K, Moustafa A, Jones SM, Narrowe A, Tomasula P. An in vitro analysis of how lactose modifies the gut microbiota structure and function of adults in a donor-independent manner. Front Nutr. 2022;9:1040744.
66. Chen P, Qiu Y, Liu G, Li X, Cheng J, Liu K, Qu W, Zhu C, Kastelic JP, Han B, et al. Characterization of Streptococcus lutetiensis isolated from clinical mastitis of dairy cows. J Dairy Sci. 2021;104(1):702–14.
67. Hodge HM, Sherman JM. Streptococcus equinus. J Bacteriol. 1937;33(3):283–9.
68. Kaneuchi C, Seki AM, Komagata K. Production of succinic acid from citric acid and related acids by lactobacillus strains. Appl Environ Microbiol. 1988;54(12):3053–3053.
69. Tsukahara T, Ushida K. Succinate accumulation in pig large intestine during antibiotic-associated diarrhea and the constitution of succinate-producing flora. J Gen Appl Microbiol. 2002;48(3):143–54.
70. Wei YH, Ma X, Zhao JC, Wang XQ, Gao CQ. Succinate metabolism and its regulation of host-microbe interactions. Gut Microbes. 2023;15(1):2190300.
71. Scheiman J, Luber JM, Chavkin TA, MacDonald T, Tung A, Pham LD, Wibowo MC, Wurth RC, Punthambaker S, Tierney BT, et al. Meta-omics analysis of elite athletes identifies a performance-enhancing microbe that functions via lactate metabolism. Nat Med. 2019;25(7):1104–9.
72. Xue H, Zhang M, Ma J, Chen T, Wang F, Tang X. Lactose-induced chronic diarrhea results from abnormal luminal microbial fermentation and disorder of ion transport in the colon. Front Physiol. 2020;11:877.
73. Cardelle-Cobas A, Fernández M, Salazar N, Martínez-Villaluenga C, Villamiel M, Ruas-Madiedo P, Reyes-Gavilán CGDL. Bifidogenic effect and stimulation of short chain fatty acid production in human faecal slurry cultures by oligosaccharides derived from lactose and lactulose. J Dairy Res. 2009;76(3):317–25.
74. Binder HJ. Role of colonic short-chain fatty acid transport in diarrhea. Ann Rev Physiol. 2010;72(1):297–313.
75. Zhou X, Liu Y, Xiong X, Chen J, Tang W, He L, Zhang Z, Yin Y, Li F. Intestinal accumulation of microbiota-produced succinate caused by loss of

microRNAs leads to diarrhea in weanling piglets. Gut Microbes. 2022. https://doi.org/10.1080/19490976.2022.2091369.
76. JanssenDuijghuijsen L, Looijesteijn E, van den Belt M, Gerhard B, Ziegler M, Ariens R, Tjoelker R, Geurts J. Changes in gut microbiota and lactose intolerance symptoms before and after daily lactose supplementation in individuals with the lactase nonpersistent genotype. Am J Clin Nutr. 2024;119(3):702–10.
77. Shaukat A, Levitt MD, Taylor BC, MacDonald R, Shamliyan TA, Kane RL, Wilt TJ. Systematic review: effective management strategies for lactose intolerance. Ann Intern Med. 2010;152(12):797–803.
78. Jia J, Frantz N, Khoo C, Gibson GR, Rastall RA, McCartney AL. Investigation of the faecal microbiota associated with canine chronic diarrhea. FEMS Microbiol Ecol. 2010;71(2):304–12.
79. Perez-Muñoz ME, Sugden S, Harmsen HJ, A't Hart B, Laman JD, Walter J. Nutritional and ecological perspectives of the interrelationships between diet and the gut microbiome in multiple sclerosis: insights from marmosets. iScience. 2021;24(7):102709.
80. Amato KR, Metcalf JL, Song SJ, Hale VL, Clayton J, Ackermann G, Humphrey G, Niu K, Cui D, Zhao H, et al. Using the gut microbiota as a novel tool for examining colobine primate GI health. Glob Ecol Conserv. 2016;7:225–37.
81. Harmsen HJ, Wildeboer-Veloo AC, Grijpstra J, Knol J, Degener JE, Welling GW. Development of 16S rRNA-based probes for the coriobacterium group and the atopobium cluster and their application for enumeration of coriobacteriaceae in human feces from volunteers of different age groups. Appl Environ Microbiol. 2000;66(10):4523–7.
82. Teofani A, Marafini I, Laudisi F, Pietrucci D, Salvatori S, Unida V, Biocca S, Monteleone G, Desideri A. Intestinal taxa abundance and diversity in inflammatory bowel disease patients: an analysis including covariates and confounders. Nutrients. 2022;14(2):260.
83. Schmidt TSB, Raes J, Bork P. The human gut microbiome: from association to modulation. Cell. 2018;172(6):1198–215.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.