Title

Time of sample collection is critical for the replicability of microbiome analyses

Journal

Nature Metabolism, 6(7)

ISSN

2522-5812

Authors

Allaband, Celeste

Lingaraju, Amulya

Flores Ramos, Stephany

et al.

Publication Date

2024-07-01

DOI

10.1038/s42255-024-01064-1

Peer reviewed

HHS Public Access

Author manuscript

Nat Metab. Author manuscript; available in PMC 2024 August 08.

Published in final edited form as:

Nat Metab. 2024 July ; 6(7): 1282–1293. doi:10.1038/s42255-024-01064-1.

Time of sample collection is critical for the replicability of microbiome analyses

Celeste Allaband[1,2,3], **Amulya Lingaraju**[2], **Stephany Flores Ramos**[1,2,3], **Tanya Kumar**[4], **Haniyeh Javaheri**[2], **Maria D. Tiu**[2], **Ana Carolina Dantas Machado**[2], **R. Alexander Richter**[2], **Emmanuel Elijah**[5,6], **Gabriel G. Haddad**[3,7,8], **Vanessa A. Leone**[9], **Pieter C. Dorrestein**[3,5,6,10], **Rob Knight**[3,6,11,12,13], **Amir Zarrinpar**[2,6,13,14,15,✉]

[1]Division of Biomedical Sciences, University of California, San Diego, La Jolla, CA, USA.

[2]Division of Gastroenterology, University of California, San Diego, La Jolla, CA, USA.

[3]Department of Pediatrics, University of California, San Diego, La Jolla, CA, USA.

[4]Medical Scientist Training Program, University of California San Diego, La Jolla, CA, USA.

[5]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA, USA.

[6]Center for Microbiome Innovation, University of California, San Diego, La Jolla, CA, USA.

[7]Department of Neurosciences, University of California, San Diego, La Jolla, CA, USA.

[8]Rady Children's Hospital, San Diego, CA, USA.

[9]Department of Animal and Dairy Sciences, University of Wisconsin-Madison, Madison, WI, USA.

[10]Center for Computational Mass Spectrometry, University of California, San Diego, La Jolla, CA, USA.

[11]Department of Computer Science and Engineering, University of California, San Diego, La Jolla, CA, USA.

[12]Halıcıoğlu Data Science Institute, University of California, San Diego, La Jolla, CA, USA.

✉**Correspondence and requests for materials** should be addressed to Amir Zarrinpar. azarrinpar@health.ucsd.edu.

Author contributions

C.A. and A.Z. conceptualized the work. C.A., E.E., P.C.D., R.K. and A.Z. determined the methodology. C.A., A.L., S.F.R., T.K., H.J., M.D.T., A.C.D.M. and R.A.R. were involved in data investigation. C.A., S.F.R., T.K., H.J., M.D.T., A.C.D.M. and R.A.R. created visualizations. A.Z. acquired funding and was the project administrator. R.K. and A.Z. supervised the work. G.G.H. and V.A.L. provided resources. C.A., A.L., S.F.R., T.K., H.J., M.D.T. and A.Z. wrote the first draft. All authors contributed to the review and editing of the manuscript.

Code availability

All relevant code notebooks are on GitHub at https://github.com/knightlab-analyses/dynamics/notebooks.

Competing interests

A.Z. is a co-founder and a chief medical officer, and holds equity in Endure Biotherapeutics. P.C.D. is an advisor to Cybele and co-founder and advisor to Ometa and Enveda with previous approval from the University of California, San Diego. All other authors declare no competing interests.

Extended data is available for this paper at https://doi.org/10.1038/s42255-024-01064-1.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42255-024-01064-1.

[13]Shu Chien-Gene Lay Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA.

[14]Division of Gastroenterology, Jennifer Moreno Department of Veterans Affairs Medical Center, La Jolla, CA, USA.

[15]Institute of Diabetes and Metabolic Health, University of California, San Diego, La Jolla, CA, USA.

Abstract

As the microbiome field moves from descriptive and associative research to mechanistic and interventional studies, being able to account for all confounding variables in the experimental design, which includes the maternal effect[1], cage effect[2], facility differences[3], as well as laboratory and sample handling protocols[4], is critical for interpretability of results. Despite significant procedural and bioinformatic improvements, unexplained variability and lack of replicability still occur. One underexplored factor is that the microbiome is dynamic and exhibits diurnal oscillations that can change microbiome composition[5–7]. In this retrospective analysis of 16S amplicon sequencing studies in male mice, we show that sample collection time affects the conclusions drawn from microbiome studies and its effect size is larger than those of a daily experimental intervention or dietary changes. The timing of divergence of the microbiome composition between experimental and control groups is unique to each experiment. Sample collection times as short as only 4 hours apart can lead to vastly different conclusions. Lack of consistency in the time of sample collection may explain poor cross-study replicability in microbiome research. The impact of diurnal rhythms on the outcomes and study design of other fields is unknown but likely significant.

The lack of replicability of microbiome studies is a barrier to understanding how host–microbe interactions contribute to physiological homeostasis and pathophysiological processes, including heart disease and cancer[8]. The ability to rapidly and reproducibly characterize the microbiome is critical to the development of new microbiome-mediated therapeutics and diagnostic biomarkers[9]. In early studies, many confounding variables involving model systems, sample collection protocols, and pipeline processing were not routinely accounted for in study design, often resulting in irreproducible, noisy data[8,10]. Great strides were made to improve and standardize experimental protocols and analysis pipelines, but unexplained variability and lack of replicability still plagues microbiome research. An often-overlooked aspect of the gut microbiome is its dynamic nature with changes occurring throughout the day[5–7]. Disruption of microbiome diurnal dynamics[11–15] is associated with metabolic syndrome spectrum diseases (for example, insulin resistance and increased adiposity)[5,16]. The gut microbiome is intimately linked to host peripheral circadian rhythms and influences physiological processes broadly[7,17,18]. Microbiome-depleted mice antibiotic-induced depletion or germ-free mice) have dampened epithelial and hepatic circadian rhythms[11,19,20]. Analysis of the microbiome from human stool samples collected from a multitude of time points[21–23], as well as 24-h salivary collections[24–26], suggest that the human microbiome also has diurnal fluctuations. In addition, loss of diurnal dynamics of the gut microbiome is a risk factor for developing

disease in a longitudinal study of a large patient cohort[27]. We hypothesize that diurnal variation can affect microbiome results and should always be recorded.

To determine whether the time of sample collection is included in experimental methods of microbiome studies, we reviewed over 550 articles published in 2019 from major journals where new 16S or metagenomic datasets were generated. Only 0.32% reported a specific time of sample collection (Extended Data Fig. 1a–c). These findings are consistent with a recent study of biological sciences articles confirming a low percentage of time-of-day information reporting in biology[28]. As microbiome studies do not commonly report the time of sample collection in their methods, we investigated the effects of sample collection on the potential interpretation of a microbiome study using the datasets from our meta-analysis. A targeted literature review, followed by extensive correspondence (Methods), led to the acquisition of six previously published preclinical datasets in a form suitable for reanalysis (Extended Data Fig. 1d)[13,29–33]. We also included analysis from an unpublished study that is unique and rare as it includes two circadian collections over the course of a single experiment.

For circadian studies, standard notation of time of day is Zeit-geber time (ZT), where lights on is ZT0 (such as 6:00 in our vivarium) and lights off is ZT12 (such as 18:00 in our vivarium). To quantify the effect of sample collection time on the microbial population, we used between-condition distance (BCD; based on weighted UniFrac β-diversity) to show how similar or different microbiomes from two cohorts (for example, experimental and control conditions) are to each other at any given time point. We chose weighted UniFrac because it takes into account both phylogeny and abundance of the organisms present and is empirically the most appropriate metric for repeated sampling of the same individuals across a single day. Thus, changes in BCD (either increasing or decreasing) allows us to assess microbiome compositional fluctuations between experimental conditions over time.

To exemplify the BCD metric with a simple example, we used a subset of the data from Caporaso et al. (2011), a well-known commonly used dataset for tutorials and teaching[34]. All known data patterns were reproduced (Extended Data Fig. 2). This confirms that BCD quantifies the dissimilarity between groups, with higher values indicating more differences and lower values indicating more similarities. Next, we demonstrated the practical application of our approach in analysing in silico circadian-like datasets (Fig. 1). In the first example dataset, the control group has diurnal oscillations in microbiome composition, whereas the experimental group does not (Fig. 1a). This example dataset closely resembles what has been reported when comparing mice on normal chow diet (NCD), which have robust diurnal oscillations, to mice on a high-fat diet (HFD), which do not[13]. By employing a simple BCD method, we are able to effectively quantify the differences between groups at various time points (Fig. 1b). In this example, ZT6 (corresponding to noon) had the highest and significantly different values from all other mock time points. Consequently, mock samples collected at ZT6 would show clear differences between groups, whereas those collected early in the morning (ZT0, or 6:00 on the wall clock) may not exhibit any differences at all.

We also investigated the impact of shifting the mock collection time points by 2 h (Fig. 1c). In this case, there is no single time point but there were time points, such as ZT4 and ZT8, where greater changes were observed in comparison to ZT0. As most investigators collect samples at a single time point, our findings from these mock studies suggest that recording the collection time (or time range, such as ZT4 to ZT8) in single time point studies for future comparison would enhance the replicability of the investigator's results.

In the second example dataset, the control group and experimental groups have opposing diurnal oscillations (Fig. 1d). This contrasting pattern changes the relationship between the two groups at the six mock time points, leading to two distinct time points where the groups maximally diverge (Fig. 1e). Considering the widespread omission of both the time of sample collection and confirmation of simultaneous collection of experimental and control samples in most studies (Extended Data Fig. 1), it is crucial to analyse the impact of collecting samples at different time points on microbiomes. To visualize these effects comprehensively, we used a heatmap of the BCD, which illustrates all combinations of the mock sample collection time points (Fig. 1f), and replicates the patterns seen in Fig. 1d,e.

In the third example dataset, the two groups have diurnal oscillations shifted by 6 h (equivalent to 1.5 time points) (Fig. 1g), which can commonly occur in experiments with daily experimental intervention or some other cause that can shift feeding behaviour (for example, alteration in diet or access to diet). This shift leads to altered dynamics and patterns between the groups (Fig. 1h,i). The 4-h (equivalent to one time point) difference leads to significant differences in the dynamics between the groups (Fig. 1h). This example also illustrates another common experimental practice that can affect microbiome composition results. Lengthy collection procedures involving sequential sampling could result in control groups being sampled in the morning (such as 8:00 to 12:00, ZT2–ZT6); and experimental groups being sampled in the afternoon (such as 12:00 to 16:00, ZT6–ZT10), even if the control and experimental condition have the same diurnal oscillation. Consequently, control groups may be sampled when the two groups are more similar and the experimental groups may be sampled when groups are more different (Fig. 1i). Although based on diurnal oscillations, BCD fluctuations can have a wide range of patterns (Fig. 1b,e,h), affecting analyses and conclusions. The lack of sample collection time records (Extended Data Fig. 1a), may explain why many microbiome studies are not easily replicated.

We assessed how sample collection time impacts the conclusions of a microbiome study by applying the knowledge gained from the analysis of mock samples to a real-world dataset. First, we hypothesized that sample collection time influences the conclusion of a study involving a discrete daily intervention known to modulate diurnal oscillations. We started by reanalysing a previous dataset[35] that used individually housed apolipoprotein E knockout mice (*Apoe*$^{-/-}$) mice under intermittent hypoxia and hypercapnia (IHC) conditions, which serve as a model for obstructive sleep apnoea (Fig. 2a). IHC was initiated at ZT6 in the mice under the experimental condition. Though the study reported overall significantly different microbiome compositions (PERMANOVA, all air versus all IHC; $P = 0.005$) (Extended Data Fig. 3a), samples collected at ZT6 were the main driver of these differences (PERMANOVA, $P = 0.035$) (Extended Data Fig. 3b). Within-condition distance (WCD) did

not have significant fluctuations across the day for both control (Extended Data Fig. 3c) and experimental conditions (Extended Data Fig. 3d). Indicating that WCD, or group dispersion, is relatively stable in this experiment across a 24-h period for both groups. There was also no significant difference between WCD for the two groups across the day (Extended Data Fig. 3e).

In contrast, BCD fluctuated greatly over the course of the day, nearly doubling within a 24-h period (Fig. 2c,d). This suggests that compositional assessments from different times would yield radically different results. BCD increased during IHC exposure, with maximal divergence of the two groups occurring at ZT6 (Fig. 2c). Maximal convergence (similarity) occurred at ZT18, a half day after the maximal divergence when both groups were experimentally similar (Fig. 2c).

As demonstrated in our review of microbiome literature (Extended Data Fig. 1), most studies do not report sample collection time, nor confirm that the control and experimental samples have been collected at the same time. Thus, when we consider all possible time point combinations (Fig. 2d), we find that the highest BCD (greatest divergence) between the two groups is air ZT18 and IHC ZT6, which are 12 h apart. The lowest BCD (greatest convergence) between the two groups is air ZT22 and IHC ZT18, both of which occur during the dark phase and are only 4 h apart. Additionally, we found that the only time point where BCD is significantly different from WCD is ZT6 (Fig. 2e).

To further explore whether these different sampling times affect the conclusions of the compositional analysis, we examined log ratios of biologically relevant phyla (Fig. 2f) and families (Fig. 2g) at the time points corresponding to the highest and lowest BCD (Fig. 2d). We used log ratios for our analysis as they account for biases caused by relative abundance comparisons due to the unknown microbial loads for each sample[36]. The log-ratio balance of two experimentally important families (Bacteroidota (formerly Bacteroidetes), in relation to Verrucomicrobiota (formerly Verrucomicrobia)) shifted significantly during maximal BCD, but the balance was similar between experimental groups at the lowest BCD (Fig. 2f). We also found similar results when examining the log ratios of three metabolically important families (Ruminococcaceae and Muribaculaceae (formerly S24-7), in relation to Verrucomicrobiaceae) (Fig. 2g). Thus, the time of sample collection has a significant effect on microbiome composition and affects the experimental conclusions if the time of sample collection is not controlled or at least reported.

Diet and feeding patterns have a pronounced and reproducible effect on the gastrointestinal environment and luminal microniches[37] but reported effects vary among studies. We hypothesized that sample collection time could lead to inconsistent conclusions. We analysed data from a previously published study[13] where the authors examined the impact of diet and feeding patterns on murine host physiology and the diurnal dynamics of the caecal microbiome. In this experiment (Fig. 3a), we assigned wild-type male C57BL/6J mice into three groups (1) NCD with ad libitum food access; (2) HFD with ad libitum access; and (3) HFD under time-restricted feeding (TRF) access. Although the mice were housed in groups, in that experiment, only one mouse from each cage was euthanized at each time point, thereby removing cage-based variability as a confounding factor in the data.

First, we analysed mice on different diets but with similar ad libitum access to food (Fig. 3b). We found that the greatest differences between the two groups would be when they were eating different diets during the dark phase, as mice primarily consume food in the nocturnal period, but HFD ad libitum mice spread their caloric intake throughout the day and night[38,39]. Thus, we would expect the highest BCD to occur during the dark phase (ZT13 to ZT21) when one group is eating NCD and the other HFD; however, the time point of greatest divergence (highest BCD) actually occurs at ZT9 (Fig. 3b), when mice fed NCD are largely fasting and while mice on HFD are likely eating at low to moderate levels. These results suggest that the luminal environment changes caused by feeding pattern differences drives BCD much more strongly than the diet nutrient profile alone.

To test this theory, we assessed whether consolidating HFD feeding to the nocturnal period with TRF changes the BCD pattern observed in mice fed a different diet. As nutrient profile has an outsized effect on the composition of the microbiome, we hypothesized that BCD would be greatest between NCD ad libitum and HFD TRF groups when they are both eating different diets during the dark phase as both groups, but would be similar in the light phase as both groups would be fasting; however, contrary to our hypothesis and similar to what we observed in the comparison to the HFD ad libitum condition, we found that the highest BCD values were during the light phase, especially ZT9 (Fig. 3c). Thus, given the similarities between Fig. 3b and Fig. 3c, diet is a much more powerful driver of time-dependent BCD differences than feeding pattern. Moreover, these results demonstrate that the time point for maximal BCD cannot be presumed by experimental design or luminal similarities alone. Finally, there is a significant drop in BCD between ZT9 and ZT13 in both Fig. 3b and Fig. 3c, demonstrating that as little as a 4-h shift in time can lead to completely different microbiome composition.

To determine whether the feeding pattern had any effect on BCD, we performed our analysis on the two HFD conditions. Based on the design of these experiments, we expect low BCD between ad libitum and TRF mice during ZT17 and ZT21, when both groups have access to food. Conversely, we would expect high BCD (divergence) during ZT1 to ZT13, when one group has access to food and the other group is forced to fast. As expected, the ad libitum to TRF BCD is the highest at ZT13 when the two groups would be the most divergent by experimental design (Fig. 3d). The BCD of ad libitum to TRF is significantly lower at ZT17, just 4 h later. It is notable that the overall BCD between the two HFD conditions is much lower (Fig. 3d mean value = 0.1590) than that between the HFD and the NCD conditions (Fig. 3b mean value = 0.6112 and Fig. 3c mean value = 0.6083), demonstrating the tremendous influence of diet on the composition of the gut microbiome.

Overall, these data demonstrate that even though diet has a large compositional effect on the gut microbiome, it is still subject to cyclical fluctuations that can affect microbiome differences observed at different time points. Feeding patterns can also affect microbiome differences observed in different time points but to a smaller extent.

Though the microbiome of the large and small intestine are quite different[40], the diurnal dynamics of the latter has only recently been characterized[11,33,41]. We hypothesized that the dynamic response to changes in diet are not the same between these two gastrointestinal

regions. We pursued this hypothesis by analysing the results from two of our previously published studies that investigated the diurnal dynamics in different gastrointestinal (GI) regions[13,33] (Fig. 3a). The Zarrinpar et al. 2014 study[13] (used for the caecum analysis presented in Fig. 3b–d) later had 16S rRNA amplicon sequencing performed on ileal samples from the same mice as a study published in 2022 (ref. 33) (Fig. 3e–g). These results revealed different daily patterns for each of the three different sets of comparisons in the caecum (Fig. 3b–d) compared to the ileum (Fig. 3e–g). Though TRF did not affect which time points had the highest and lowest BCDs when comparing HFD to NCD in the caecum (Fig. 3c,f), the nadir shifted to an earlier time point in the ileum (from ZT21 in Fig. 3c to ZT13 in Fig. 3f). The ileal results between the HFD and the NCD conditions (Fig. 3e mean value = 0.3718 and Fig. 3f mean value = 0.3714) show a decreased impact of diet on the composition of the gut microbiome for the region. Moreover, comparing feeding patterns between the two HFD groups showed higher BCD (Fig. 3g mean value = 0.2099 > Fig. 3d mean value = 0.1590), particularly at ZT13, the time that we had hypothesized would have the biggest difference. Thus, the ileal data suggest that feeding pattern had a much larger effect on the BCD differences than what we observed in the caecum (Fig. 3e).

Direct comparisons between ileum and caecum microbiome dynamics require samples processed and analysed with the same pipeline. Leone et al.[11], groups of mice had ad libitum access to a NCD or high milk-fat diets (MFDs) (Fig. 3h). The microbial dynamics in the caecum and ileum was generally similar, but they only significantly differed in BCD at one time point. Caecal samples collected at ZT6, in the middle of the light phase, demonstrated a low BCD, whereas the ileal samples were much higher (Fig. 3i). Circadian stool collections can thus inform experimentalists when the greatest and least microbiome differences can be observed. One additional study with a different experimental design also demonstrated large shifts in BCD between the patterns seen across the day in different gastrointestinal regions (Extended Data Fig. 4). Moreover, a recent study showed differences in micro-niche sites (luminal contents versus mucosally adherent bacteria) within a single gastrointestinal region can have their own dynamic changes as well[42–44] (Extended Data Fig. 5).

The studies presented thus far show experimental exposure of a week (Fig. 2) to months (Fig. 3); however, it is not clear whether the dynamic fluctuations observed early in an experiment are the same as those seen much later when phenotypic studies are performed. To better understand this, we reanalysed previously published data on the IHC intervention from our laboratory (Fig. 3)[27]. In this longitudinal IHC experiment, we collected stool biweekly at the same time (ZT3 to ZT5) for 10 weeks until the phenotype was detectable[37] (Fig. 4a). This experiment was performed in the same facility, by the same people, using the same protocols as the experiment presented in Fig. 2. The BCD between IHC and air conditions increased over the course of the experiment (Fig. 4b), indicating that the groups diverged significantly as the phenotype developed (week 10.5 compared to week 19.5, $P = 2.56 \times 10^{-8}$, paired Wilcoxon rank-sum, test statistic of 1,126). There was a significant positive coefficient in the linear regression analysis ($P = 6.72 \times 10^{-56}$; equation $y = 0.016x + 0.119$). By holding the time of collection constant, we were able to observe the compositional shift over time.

Many microbiome experiments do not report time of day of sample collection; however, many also do not report the duration of interventions before sample collection. Stool specimens often are collected when a phenotype is present, although this is often not explicitly stated in the manuscripts. We hypothesized that the length of longitudinal exposure affects BCD and optimal sampling time points. In a forthcoming manuscript examining the effect of TRF on atherosclerotic phenotype development in *Ldlr* knockout ($Ldlr^{-/-}$) mice, we reanalysed the data to determine whether BCD changes over the course of a long study (Fig. 4c). In this rare dataset with two 24-h faecal sample collections (every 4 h for 24 h) on the same mice, we collected samples after 1 week ('early'; pre-atherosclerotic phenotype development) and after 20 weeks ('late'; post-atherosclerotic phenotype development) of daily TRF interventions.

As with the previous examples, the time of sample collection during the day affects the BCD dynamics between the control (ad libitum) and experimental (TRF) groups (Fig. 4d); however, we did not observe any differences in the ad libitum to TRF BCD at any time point between the early and late phase of the experiment (Fig. 4d–f). The peak-to-trough ratios were not significantly different between the early (Fig. 4e) and late collection (Fig. 4f). Generally, these results demonstrate that longitudinal measures of BCD within a single experiment are relatively consistent over time.

In conclusion, since 2014 there has been unequivocal and reproducible research from multiple laboratories demonstrating diurnal fluctuations in the composition of the gut microbiome[5,6,11–15,45,46]. Yet neither sample collection time nor the rationale for the selection of this time is reported outside studies that are focused on diurnal fluctuations of the microbiome. Here, we show that the conclusions of a microbiome research study are greatly dependent on the time of sample collection, and that experimental and control groups undergo a cycle of diverging and converging microbiome composition depending on the nature and timing of experimental interventions.

Our findings suggest compositional fluidity sensitive to host factors including environmental exposures, diet, gut region and luminal micro-niche. One notable limitation of our analyses is a lack of available studies in female mice for comparison and additional analyses to address that gap are needed in the future. By subsetting the distance matrix for only BCD values, we can quantify the difference between control and experimental groups across time, optimizing sample collection for even single time point studies. BCD analysis confirms that peak and trough distances can be as little as 4 h apart. This timescale may still be an overestimate; available datasets did not collect stool samples at less than 4-h intervals. Because >90% of microbiome studies do not report when samples are collected, laboratories may unknowingly be collecting at suboptimal time points. Our recommendations, including recording lighting conditions as well as the sample collection time window in ZT notation, are summarized in Box 1. We encourage funders and journal editors to promote these guidelines as our recommendations offer clear directives for new researchers in the field and ensure that studies can be more easily compared and built upon. We argue that not controlling for host circadian rhythm time is like trying to measure sea level rise without understanding that tides exist.

Methods

This is a retrospective multi-study analysis. All data were collected and animal experiments conducted in accordance with the guidelines of each study's respective Institutional Animal Care and Use Committee committees. Please refer to source papers for additional details. Sample sizes for each study were determined by various methods and can be found in their source papers, shown in the subsections below. According to the methods sections from each of the source papers, mice were randomly assigned to groups. Data collection and analysis were not performed blind to the conditions of the experiments. Statistical methods that did not assume normality and accounted for the dependent nature of the respective analyses (Wilcoxon and Mann–Whitney–Wilcoxon) were used. All data included in the source dataset were used, unless excluded by rarefaction depth, as indicated in the source paper. While studies with female mice were not excluded, all datasets that met the inclusion criteria contained only male mice.

Literature review

As shown in Extended Data Fig. 1a–c, we used the advanced search option from the four main journal groups, including the American Society for Microbiology (https://msystems.asm.org), Science (https://search.sciencemag.org), Nature (https://www.nature.com) and Cell Press (https://www.cell.com). Searching for the term 'microbiome' in all search fields (abstract, title and main text) during the year 2019 (1 Jan 2019, to 31 Dec 2019) resulted in 586 articles from nine journals; mSystems (American Society for Microbiology), Science Translational Medicine (Science), Science Signalling (Science), Science Advances (Science), Science Immunology (Science), Nature (Nature), Nature Microbiology (Nature), Nature Communications (Nature), Cell Host Microbe (Cell), Cell (Cell), Cell Reports (Cell) and Cell Metabolism (Cell). Our collection sheet includes a total of 16 columns: journal group, journal, year, article title, DOI, PMID, first author, last author, microbiome (yes/no), vivarium (yes/no), vivarium setting, sample host, sample type, collection time, time note and collection time reason. Notation of collection time was recorded as follows: explicitly stated ('yes'; 8:00, ZT4, etc.), implicitly stated ('relative'; 'before surgery', 'in the morning', etc.) or unstated ('not provided'; 'daily', 'once a week', etc.).

Systematic review

(Extended Data Fig. 1d) When searching for the keywords 'circadian microbiome' AND 'mice' in PubMed (https://pubmed.ncbi.nlm.nih.gov/) for articles published over an 8-year period (from 2014 to 2021), we found 79 articles that met our initial criteria. Only 66 of those were research articles and, of the remainder, we found only 14 articles that contained 16S amplicon sequencing samples collected for more than three time points within a 24 or 48-h period. Of these 14 studies, 4 had complete publicly available data at the European Nucleotide Archive(ENA)/European Bioinformatics Institute (EBI). Of the remainder, four had incomplete datasets on ENA/EBI[12,15,32,47] and the rest were not publicly available. We then contacted the authors of all studies with missing or incomplete data and got the following responses: four were unable to locate the missing data[12,15,48,49], three could not provide data in a format suitable for reanalysis[47,50,51] and three did not respond to repeated

inquiries[52–54]. This resulted in the acquisition of five previously published datasets in a form suitable for reanalysis[13,29–32].

Microbiome

All of the data in this paper are a reanalysis of previously published 16S studies, except for the data shown in Fig. 4a–d (C.F.A & A.Z., manuscript in preparation). Please refer to the respective source papers for detailed methods, including sample handling and preliminary processing. Raw data were procured from the respective data repositories as stated in the source paper, typically the ENA. This data were then run through a standard QIIME2 pipeline (v.2021.8)[55] as follows: samples demultiplexed, denoised via deblur[56] into the amplicon sequence variant table and feature table underwent rarefaction (as stated in the source paper, see individual methods sections), representative sequences underwent fragment insertion on Greengenes_13_8 via SATé-enabled phylogenetic placement[57] to create the phylogenetic tree and weighted UniFrac distances[58] were calculated. The resulting weighted UniFrac distance matrix was filtered for only BCDs as relevant to each study. Thus, using BCD values will show how similar the microbiomes from the two conditions are to each other at any given time point. As BCD values are a subset of the weighted UniFrac distance matrix values, both conditions (control and experimental) are taken into account with each distance value shown. Changes in BCD will demonstrate convergence (decreasing distance and increased similarity) or divergence (increasing distance and decreased similarity) of the microbiome composition between two groups. Circadian time notation is used throughout the paper to denote when samples were collected: ZT were lights on at ZT0. Data were visualized using custom Python scripts, which can be found at https://github.com/knightlab-analyses/dynamics.

Figure 1 shows mock abundance data, created by generating a unit cosine wave over 24 h, applying appropriate time offset, then adding 10% random jitter to each mock time point measurement over four replicates. Amplitude of waves was fixed to 0.5 for cycling taxa or 0 for non-cycling taxa.

Figure 2 shows, in brief, two groups of 10-week-old male *Apoe*$^{-/-}$ mice on C57BL/6J background (002052; The Jackson Laboratory) were individually housed in a 12-h light–dark (12:12 L:D) vivarium. All mice were given an atherosclerotic-promoting diet (1.25% cholesterol, 21% milk fat; 4.5 kcal g^{-1}; TD.96121; Envigo-Teklad Madison) starting at 10 weeks of age until the end of the study. Mice in the experimental group were exposed to IHC conditions that consisted of 4 min of synchronized O_2 reduction from 21% to 8% and synchronized elevation of CO_2 from 0.5% to 8%, followed by alternating periods of 4 min of normoxia and normocapnia with 1–2-min ramp intervals. IHC conditions were administered in a computer-controlled atmosphere chamber (OxyCycler, Reming Bioinstruments) for 10 h per day during the lights on phase (ZT2 to ZT12) when mice were sleeping, for 10 weeks. Mice in the control group were exposed to normal room air (21% O_2 and 0.5% CO_2) during that same time period. After 6 days, faecal samples were collected every 4 h for 24 h (n = 4 per group). Samples were frozen at −80 °C as soon as they were collected, which was within 4 h of sample production; however, environmental factors and biases in molecular techniques likely confer greater amounts of variation to microbial communities

than differences in short-term storage conditions[36,37,59–61]. The 16S amplicon sequencing was performed on the V4 region using standard protocols (http://www.earthmicrobiome.org/emp-standard-protocols/). Rarefaction was set at 12,000 reads to control for sequencing effort. The source paper provides additional details[28].

Figure 3a–g shows that, in brief, wild-type specific-pathogen-free C57BL/6 group-housed male mice (three mice per cage, six cages total) were provided either NCD (LabDiet 5001, 13.5% calories from fat and crude fibre 5.1%) or a HFD (61% fat) and were fed in either an ad libitum manner, with access to food at all times or fed in a TRF manner. TRF mice were allowed unrestricted access to HFD from ZT13 to ZT21. Mice on an NCD ad libitum diet (controls) typically fast during the light phase and consume >80% of their diet during the dark phase[60,61]; however, mice on the HFD ad libitum diet (diet-induced obesity) lose this diurnal feeding pattern and spread their caloric intake throughout both the dark and light phase[62,34]. TRF of HFD consolidates feeding to the nocturnal period by providing access to food in a narrow time window, from ZT13 to ZT21 in this experiment, and is known to prevent the dysmetabolic effects of HFD consumption[13,62]. After 8 weeks under these dietary conditions, three mice from different cages were killed every 4 h for 24 h and intestinal contents collected (n = 3 mice per condition per time point from separate cages; six time points). At ZT13, fasted mice were killed before feeding. The 16S amplicon sequencing was performed on the V1–V3 region using the 454 platform for caecal data. The 16S amplicon sequencing was performed on the V4 region using Illumina primers for ileal data. For both regions, rarefaction was set to 1,000 reads to control for sequencing effort. The source paper provides additional details[13,33].

Figure 3h,i shows a study performed on 8–10-week-old male C57BL/6J specific-pathogen-free mice that were maintained in a 12:12 L:D cycle vivarium. The mice were fed ad libitum with either NCD (Harlan Teklad 2018S, 18% calories from fat and 3.5% crude fibre) or a 37.5% saturated MFD (Harlan Teklad TD.97222 customized diet). As shown in Fig 3, after 5 weeks of being on the NCD or MFD diet, the mice were killed and the caecal and ileal contents were collected every 4 h for 24 h (n = 3 mice per treatment); organ contents were flash frozen and stored at −80 °C. As shown in Extended Data Fig. 3, after 5 weeks of being on the NCD or MFD diet, faecal pellets were collected every 4 h for 24 h (n = 3 mice per treatment); the faecal samples were stored at −80 °C. The 16S amplicon sequencing was performed on the V4–V5 region using standard protocols (https://earthmicrobiome.org/protocols-and-standards/) in a High-Throughput Genome Analysis Core (Institute for Genomics and Systems Biology) at Argonne National Laboratory. Rarefaction was set at 10,000 reads to control for sequencing effort. The source paper provides additional details[30].

As shown in Fig. 4a,b, in brief, two groups of 10-week-old male *Apoe*$^{-/-}$ mice on a C57BL/6J background (002052; The Jackson Laboratory) were kept in a 12:12 L:D vivarium fed NCD (Teklad Rodent Diet 8604, 14% calories from fat and 4% crude fibre) for 2 weeks before they were switched to an atherosclerotic-promoting diet containing 1.25% cholesterol and 21% milk fat (4.5 kcal g^{-1}; TD.96121; Envigo-Teklad Madison) starting at 10 weeks of age until the end of the study. Mice in the experimental group were exposed to IHC conditions as described in Fig. 1 and were administered in a computer-controlled atmosphere chamber (OxyCycler, Reming Bioinstruments) for 10 h per day during the lights

on phase (ZT2-ZT12) for 10 weeks. Mice in the control group were exposed to normal room air (21% O_2 and 0.5% CO_2) during that same time period. Faecal samples were collected twice a week for the duration of the study[27].

As shown in Fig. 4c–f, this study was performed on 10-week-old male *Ldlr*$^{-/-}$ mice (The Jackson Laboratory). After a 2-week acclimatization period, they were fed a high-fat, high cholesterol diet (Research Diets D12109i; Clinton/Cybulsky high-fat rodent diet, regular casein, 1.25% added cholesterol and 0.5% sodium cholate). During the experiment, mice were maintained in 12:12 L:D reverse light-cycled cabinets (Phenome Technologies) at 70 °F and 51% humidity. Experimental and control groups were both on an AD, but one group was fed ad libitum and the other, TRF. In TRF, mice were only allowed to eat for 8 h per day during the dark phase of the day between ZT13 and ZT21. Faecal samples were collected every 4 h for 24 h (n = 6 mice per condition) after 1 week (early; pre-atherosclerotic phenotype) and after 20 weeks (late; post-atherosclerotic phenotype). The 16S rRNA sequencing was performed on the V4 region using the Earth Microbiome standard protocol (https://earthmicrobiome.org/protocols-and-standards/). Rarefaction was set at 11,498 reads to control for sequencing effort. These animal experiments were conducted in accordance with the guidelines of the Institutional Animal Care and Use Committee of the University of California, San Diego.

As shown in Extended Data Fig. 2, we used the ‘Moving Pictures’ Qiime2 tutorial data (de-identified human samples) (https://docs.qiime2.org/2022.11/tutorials/moving-pictures/) for this example. In brief, 16S amplicon sequencing was performed on the V4 region using standard protocols (http://www.earthmicrobiome.org/emp-standard-protocols/). Rarefaction was set at 1,103 reads to control for sequencing effort. Metadata underwent minor modification. The original source paper provides further details[34].

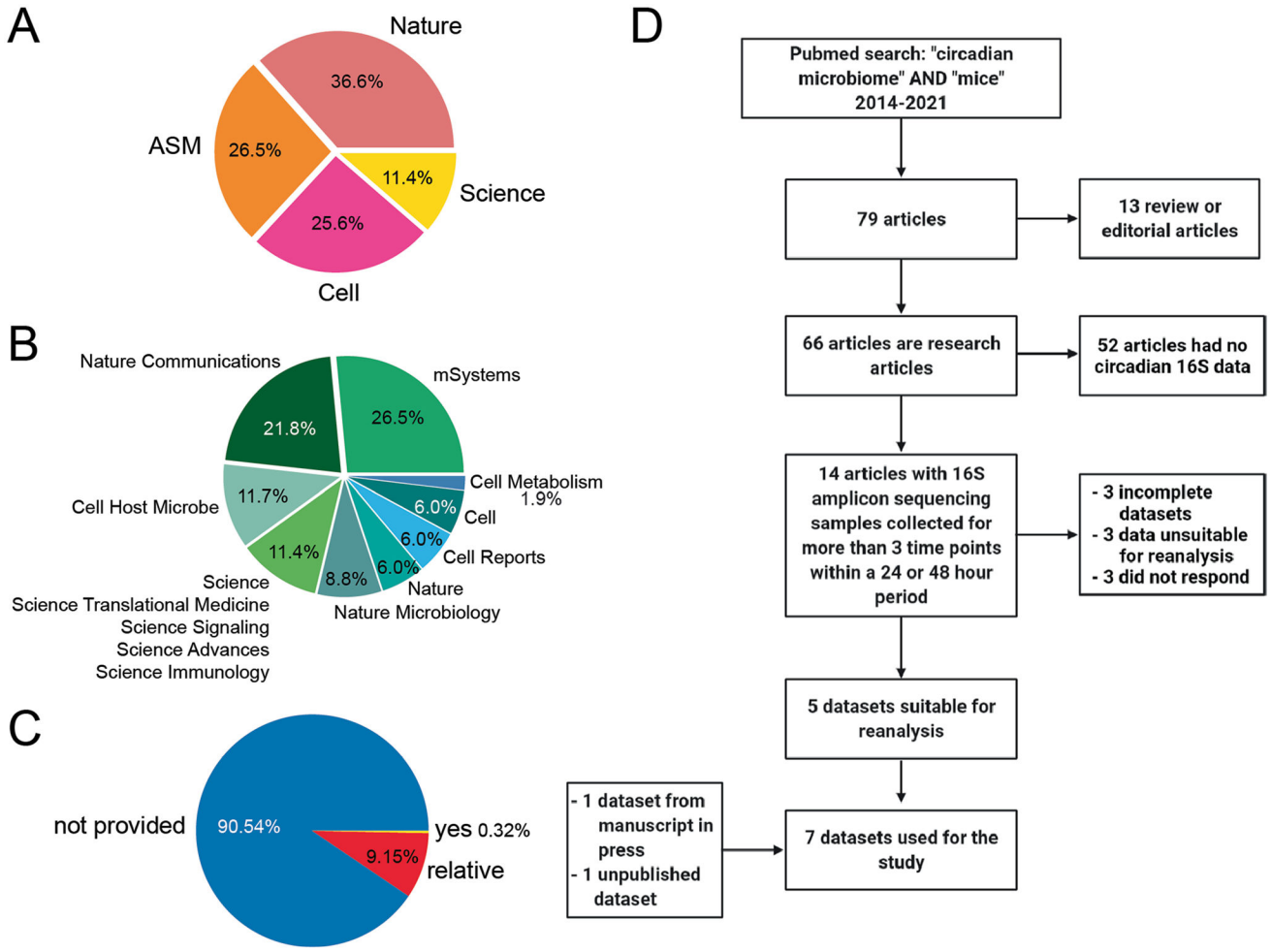
Extended Data Fig. 4 shows that, in brief, two groups of 5-week-old male *Balb/c* mice were kept in either a 12:12 L:D or 0:24 L:D vivarium fed a NCD (unspecified in methods) ad libitum. After 2 weeks on the condition, mice were anaesthetized and killed every 4 h for 24 h (n = 4–5 mice per group per time point). Samples from the intestinal lumen, mucous layer, epithelial layer and caecal contents were collected. The phenol–chloroform method was used for DNA extraction. The 16S rRNA amplicon sequencing was performed on the V4 region. Rarefaction was set to 1,085 reads to control the sequencing effort, as performed in the source paper. The source paper provides a detailed study design and associated protocols[32].

Extended Data Fig. 5 shows a study that was performed on 8–12-week-old wild-type male C57BL/6 mice that were maintained in a 12:12 L:D cycle vivarium. The mice were fed NCD (Harlan Teklad 2018S; 18% calories from fat and 3.5% crude fibre) ad libitum for 4 weeks before sample collection. The mice were killed and the luminal and mucosal small intestinal samples were collected every 4 h for 24 h (except for ZT8, n = 4–5 mice per time point). The samples were frozen and stored at −80 °C. The 16S amplicon sequencing was performed on the V4 region of the genome. Rarefaction was set to 4,200 reads to control the sequencing effort. The source paper provides additional details[31].

Reporting summary

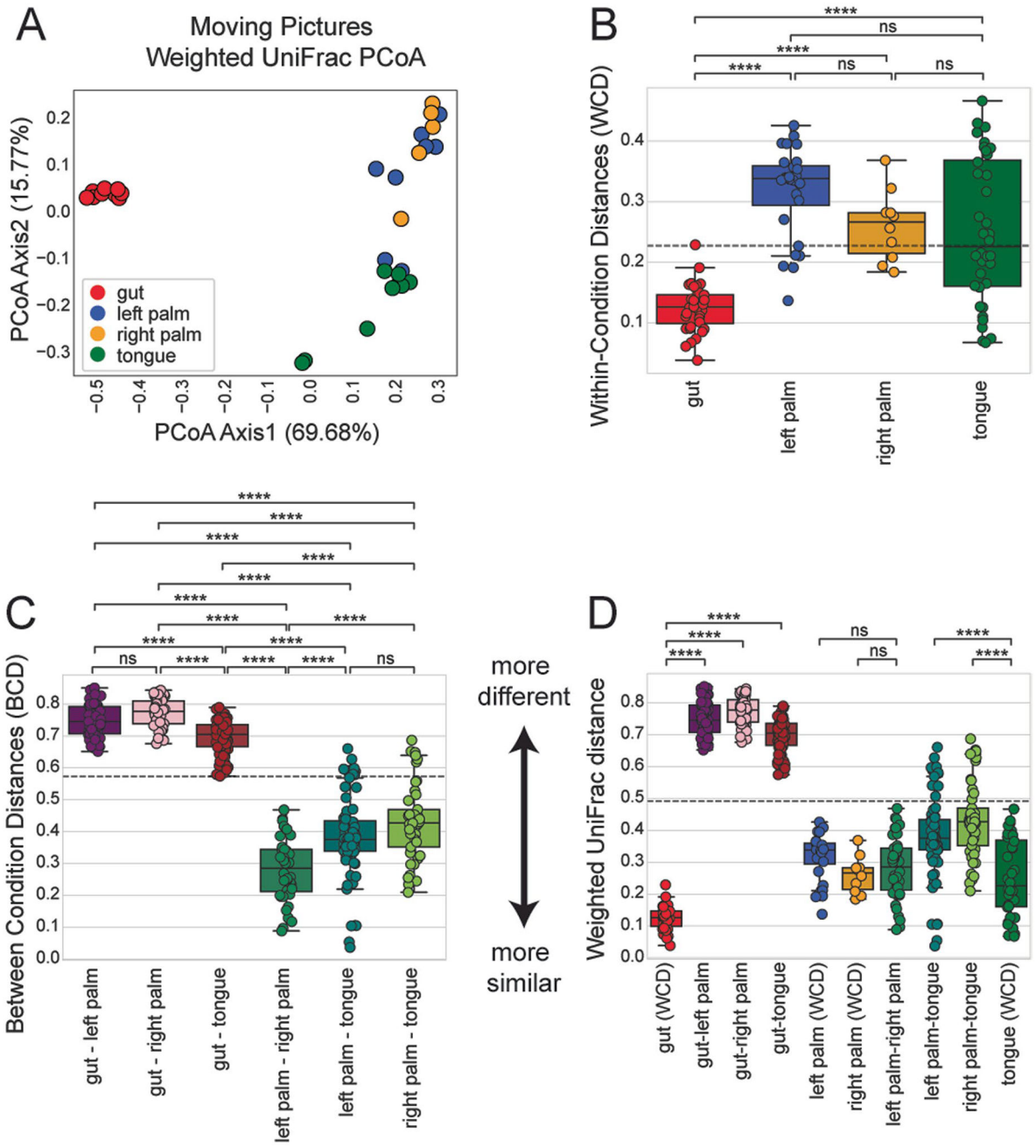
Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Extended Data



Extended Data Fig. 1 |. Microbiome Literature Review.

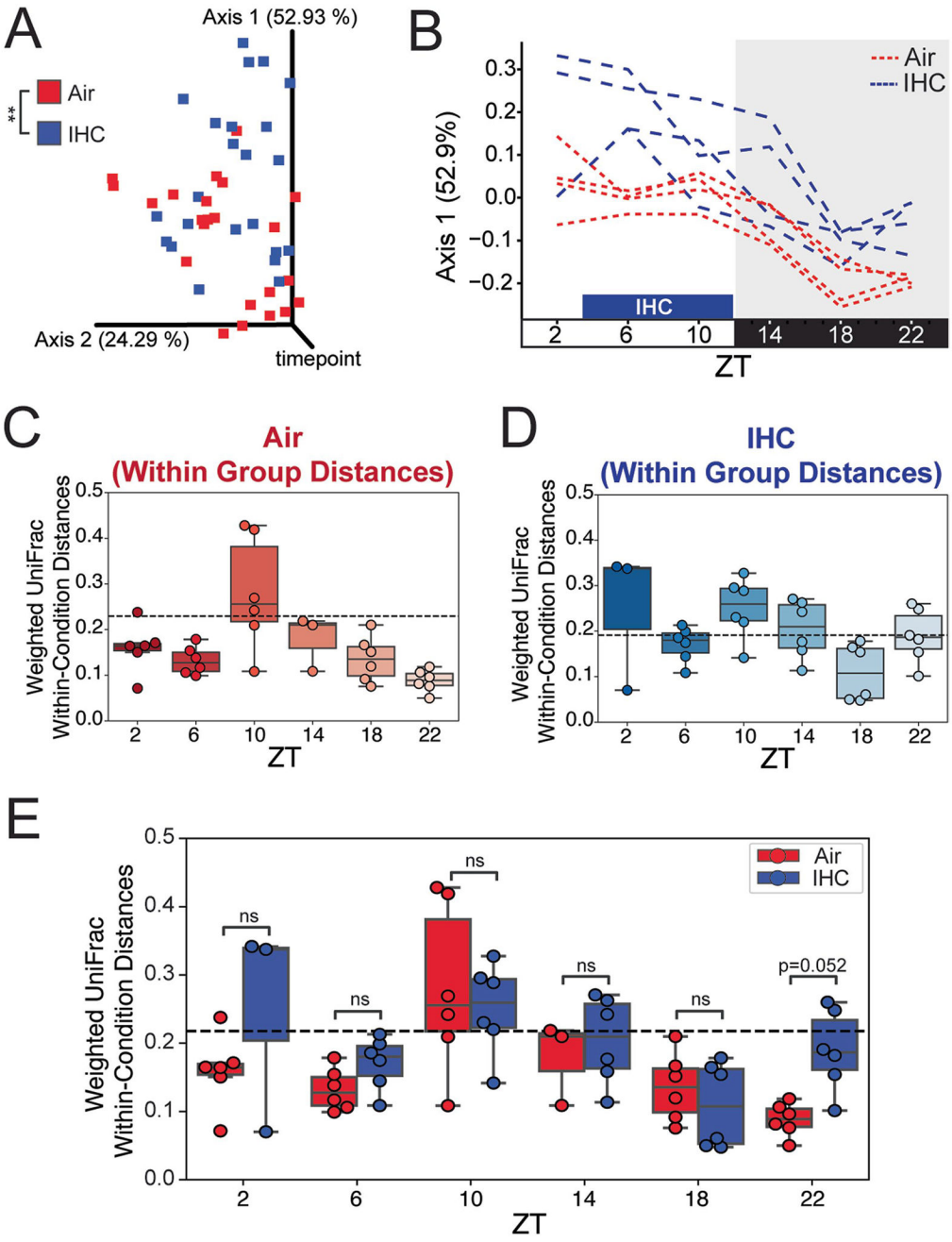
A) 2019 Literature Review Summary. Of the 586 articles containing microbiome (16 S or metagenomic) data, found as described in the methods section, the percentage of microbiome articles from each of the publication groups. **B**) The percentage of microbiome articles belonging to each individual journal in 2019. Because the numerous individual journals from Science represented low percentages individually, they were grouped together. **C**) The percentage articles where collection time was explicitly stated (yes: 8 AM, ZT4, etc.), implicitly stated (relative: 'before surgery', 'in the morning', etc.), or unstated (not provided: 'daily', 'once a week', etc.). **D**) Meta-Analysis Inclusion Criteria Flow Chart. Literature review resulting in the five previously published datasets for meta-an alysis[11,13,28–30].



Extended Data Fig. 2 |. Single Time Point (Non-Circadian) Example.

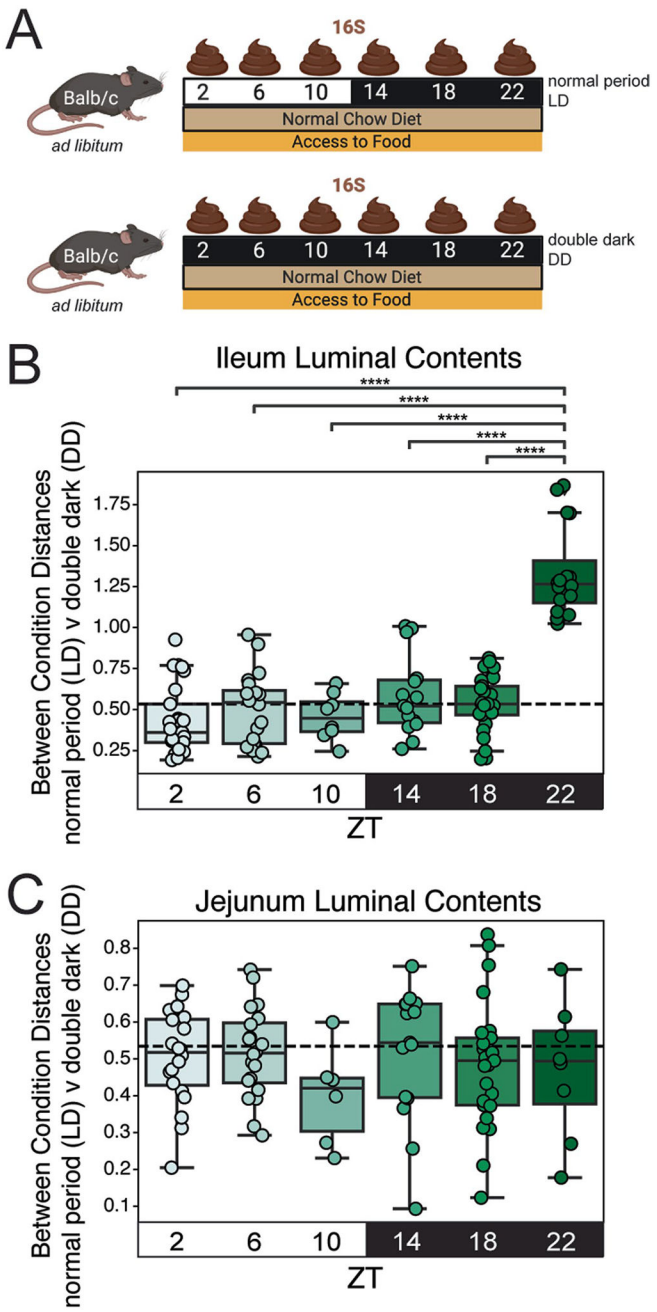
A) Weighted UniFrac PCoA Plot - modified example from Moving Pictures Qiime2 tutorial data [https://docs.qiime2.org/2022.11/tutorials/moving-pictures/]. Each point is a sample. Points were coloured by body site of origin. There are 8 gut, 8 left palm, 9 right palm, and 9 tongue samples. **B**) Within-Condition Distances (WCD) boxplot/stripplot for each body site (n = 8–9 mouse per group per time point). **C**) Between Condition Distances (BCD) boxplot/stripplot for each unique body site comparison (n = 8–9 mouse per group per time point). **D**) All pairwise grouping comparisons, both WCD and BCD, are shown in the boxplots/stripplots (n = 8–9 mouse per group per time point). Only WCD to BCD

statistical differences are shown. Boxplot centre line indicates median, edges of boxes are quartiles, error bars are min and max values. Significance was determined using a paired Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction. Notation: ns (not significant) = $p > 0.05$, * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$, **** = $p < 0.00001$.



Extended Data Fig. 3 |. Additional Analysis of Apoe−/− Mice Exposed to IHC Conditions.
A) Weighted UniFrac PCoA stacked view (same as Fig. 2b but different orientation). Good for assessing overall similarity not broken down by time point. Significance determined by PERMANOVA ($p = 0.005$). **B**) Weighted UniFrac PCoA of only axis 1 over time.

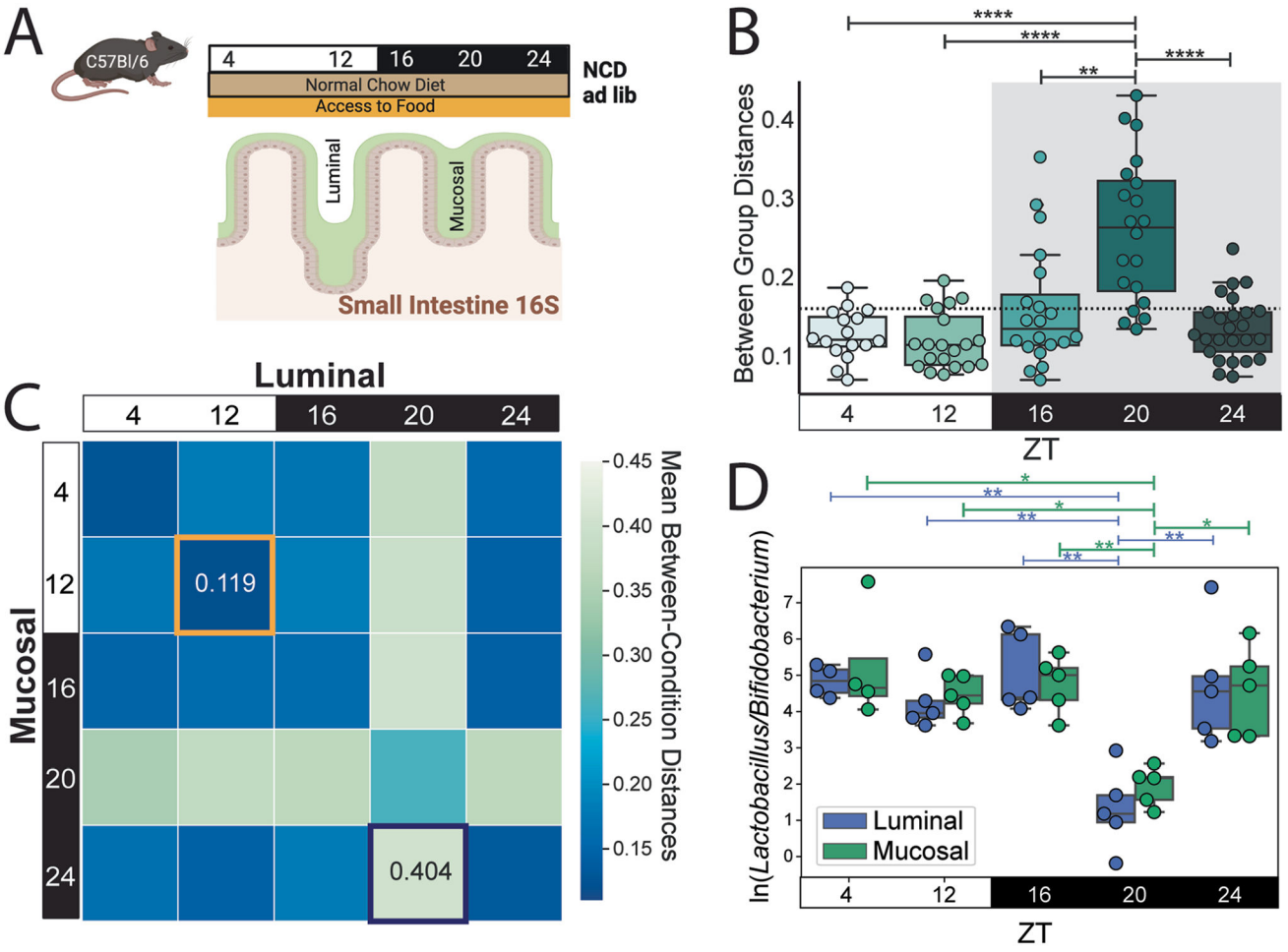
C) Boxplot/scatterplot of within-group weighted UniFrac distance values for the control group (Air, n = 3–4 samples per time point). Unique non-zero values in the matrix were kept. Dotted line indicates the mean of all values presented. No significant differences ($p > 0.05$) found. **D**) Boxplot/scatterplot of within-group weighted UniFrac distance values for the experimental group (IHC, n = 3–4 samples per time point)). Unique non-zero values in the matrix were kept. Dotted line indicates the mean of all values presented. No significant differences ($p > 0.05$) found. **E**) Boxplot/scatterplot of within-group weighted UniFrac distance values for both control (Air) and experimental (IHC) groups [n = 3–4 samples per group per time point]. Mann-Whitney-Wilcoxon test with Bonferroni correction used to determine significant differences between groups. Boxplot centre line indicates median, edges of boxes are quartiles, error bars are min and max values. Notation: ns = not significant, $p > 0.05$; * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$.



Extended Data Fig. 4 |. Irregular differences in diurnal rhythm patterns leads to generally minor shifts in BCD when comparing LD vs DD mice.

A) Experimental design. Balb/c mice were fed NCD ad libitum under 0:24 L:D (24 hr darkness, DD) experimental conditions and compared to 12:12 L:D (LD) control conditions. After 2 weeks, mice from each group were euthanized every 4 hours for 24 hours (N = 4–5 mice/condition) and samples were collected from the proximal small intestine ('jejunum') and distal small intestine ('ileum') contents. **B**) BCD for luminal contents of proximal small intestine samples comparing LD to DD mice (N = 4–5 mice/condition). Dotted line is the average of all shown weighted UniFrac distances. Significance was determined using a paired Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction; notation: ****

= p < 0.00001. **C**) BCD for luminal contents of distal small intestine samples comparing LD to DD mice (N = 4–5 mice/condition). Dotted line is the average of all shown weighted UniFrac distances. Boxplot centre line indicates median, edges of boxes are quartiles, error bars are min and max values.



Extended Data Fig. 5 |. Localized changes in BCD between luminal and mucosal contents. **A**) Experimental design and sample collection for a local site study. Small intestinal samples were collected every 4 hours for 24 hours (N = 4–5 mice/condition, skipping ZT8). Mice were fed ad libitum on the same diet (NCD) for 4 weeks before samples were taken. **B**) BCD for luminal vs mucosal conditions (N = 4–5 mice/condition). The dotted line is the average of all shown weighted UniFrac distances. Significance is determined using the Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction. **C**) Heatmap of mean BCD distances comparing luminal and mucosal by time point (N = 4–5 mice/condition). Highest value highlighted in navy, lowest value highlighted in gold. Boxplot centre line indicates median, edges of boxes are quartiles, error bars are min and max values. Significance was determined using a paired Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction. Notation: * = p < 0.05; ** = p < 0.01; *** = p < 0.001, **** = p < 0.00001. **D**) Experimentally relevant log ratio, highlighting the changes seen at ZT20 (N = 4–5 mice/condition). Boxplot center line indicates median, edges of boxes are

quartiles, error bars are min and max values. Significance was determined using a paired Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction. Notation: * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$, **** = $p < 0.00001$.

Acknowledgements

C.A. was supported by NIH T32 OD017863. S.F.R. is supported by the Soros Foundation. A.L. is supported by the AHA Postdoctoral Fellowship grant. T.K. is supported by NIH T32 GM719876. A.C.D.M. is supported by R01 HL148801-02S1. G.G.H. and A.Z. are supported by NIH R01 HL157445. A.Z. is further supported by the VA Merit BLR&D Award I01 BX005707 and NIH grants R01 AI163483, R01 HL148801, R01 EB030134 and U01 CA265719. All authors receive institutional support from NIH P30 DK120515, P30 DK063491, P30 CA014195, P50 AA011999 and UL1 TR001442.

Data availability

Literature review data are at https://github.com/knightlab-analyses/dynamics/data/. Figure 1, mock data are at https://github.com/knightlab-analyses/dynamics/data/MockData. Figure 2 (Allaband/Zarrinpar 2021) data are under EBI accession ERP110592. Figure 3 data (longitudinal IHC) are under EBI accession ERP110592 and (longitudinal circadian TRF) EBI accession ERP123226. Figure 4 data (Zarrinpar/Panda 2014) are in the Supplementary Excel file attached to the source paper[13]; (Leone/Chang 2015) figshare for the 16S amplicon sequence data are at https://doi.org/10.6084/m9.figshare.882928 (ref. 63). Extended Data Fig. 2 data (Caporaso/Knight 2011) are at MG-RAST project mgp93 (IDs mgm4457768.3 and mgm4459735.3). Extended Data Fig. 3 data (Wu/Chen 2018) are under ENA accession PRJEB22049. Extended Data Fig. 4 data (Tuganbaev/Elinav 2021) are under ENA accession PRJEB38869.

References

1. Schloss PD Identifying and overcoming threats to reproducibility, replicability, robustness, and generalizability in microbiome research. mBio 9, e00525–18 (2018). [PubMed: 29871915]
2. Gilbert JA et al. Current understanding of the human microbiome. Nat. Med 24, 392–400 (2018). [PubMed: 29634682]
3. Knight R et al. Best practices for analysing microbiomes. Nat. Rev. Microbiol 16, 410–422 (2018). [PubMed: 29795328]
4. Ley RE et al. Obesity alters gut microbial ecology. Proc. Natl Acad. Sci. USA 102, 11070–11075 (2005). [PubMed: 16033867]
5. Deloris Alexander A et al. Quantitative PCR assays for mouse enteric flora reveal strain-dependent differences in composition that are influenced by the microenvironment. Mamm. Genome 17, 1093–1104 (2006). [PubMed: 17091319]
6. Friswell MK et al. Site and strain-specific variation in gut microbiota profiles and metabolism in experimental mice. PLoS ONE 5, e8584 (2010). [PubMed: 20052418]
7. Sinha R et al. Assessment of variation in microbial community amplicon sequencing by the Microbiome Quality Control (MBQC) project consortium. Nat. Biotechnol 35, 1077–1086 (2017). [PubMed: 28967885]
8. Alvarez Y, Glotfelty LG, Blank N, Dohnalová L & Thaiss CA The microbiome as a circadian coordinator of metabolism. Endocrinology 161, bqaa059 (2020).
9. Frazier K & Chang EB Intersection of the gut microbiome and circadian rhythms in metabolism. Trends Endocrinol. Metab 31, 25–36 (2020). [PubMed: 31677970]
10. Heddes M et al. The intestinal clock drives the microbiome to maintain gastrointestinal homeostasis. Nat. Commun 13, 6068 (2022). [PubMed: 36241650]

11. Leone V et al. Effects of diurnal variation of gut microbes and high-fat feeding on host circadian clock function and metabolism. Cell Host Microbe 17, 681–689 (2015). [PubMed: 25891358]
12. Thaiss CA et al. Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. Cell 159, 514–529 (2014). [PubMed: 25417104]
13. Zarrinpar A, Chaix A, Yooseph S & Panda S Diet and feeding pattern affect the diurnal dynamics of the gut microbiome. Cell Metab 20, 1006–1017 (2014). [PubMed: 25470548]
14. Liang X, Bushman FD & FitzGerald GA Rhythmicity of the intestinal microbiota is regulated by gender and the host circadian clock. Proc. Natl Acad. Sci. USA 112, 10479–10484 (2015). [PubMed: 26240359]
15. Thaiss CA et al. Microbiota diurnal rhythmicity programs host transcriptome oscillations. Cell 167, 1495–1510 (2016). [PubMed: 27912059]
16. Yu F et al. Deficiency of intestinal Bmal1 prevents obesity induced by high-fat feeding. Nat. Commun 12, 5323 (2021). [PubMed: 34493722]
17. Leone VA et al. Atypical behavioral and thermoregulatory circadian rhythms in mice lacking a microbiome. Sci. Rep 12, 14491 (2022). [PubMed: 36008471]
18. Thaiss CA, Zeevi D, Levy M, Segal E & Elinav E A day in the life of the meta-organism: diurnal rhythms of the intestinal microbiome and its host. Gut Microbes 6, 137–142 (2015). [PubMed: 25901892]
19. Mukherji A, Kobiita A, Ye T & Chambon P Homeostasis in intestinal epithelium is orchestrated by the circadian clock and microbiota cues transduced by TLRs. Cell 153, 812–827 (2013). [PubMed: 23663780]
20. Weger BD et al. The mouse microbiome is required for sex-specific diurnal rhythms of gene expression and metabolism. Cell Metab 29, 362–382 (2019). [PubMed: 30344015]
21. Kaczmarek JL, Musaad SM & Holscher HD Time of day and eating behaviors are associated with the composition and function of the human gastrointestinal microbiota. Am. J. Clin. Nutr 106, 1220–1231 (2017). [PubMed: 28971851]
22. Skarke C et al. A pilot characterization of the human chronobiome. Sci. Rep 7, 17141 (2017). [PubMed: 29215023]
23. Jones J, Reinke SN, Ali A, Palmer DJ & Christophersen CT Fecal sample collection methods and time of day impact microbiome composition and short chain fatty acid concentrations. Sci. Rep 11, 13964 (2021). [PubMed: 34234185]
24. Collado MC et al. Timing of food intake impacts daily rhythms of human salivary microbiota: a randomized, crossover study. FASEB J 32, 2060–2072 (2018). [PubMed: 29233857]
25. Kohn JN et al. Differing salivary microbiome diversity, community and diurnal rhythmicity in association with affective state and peripheral inflammation in adults. Brain. Behav. Immun 87, 591–602 (2020). [PubMed: 32061904]
26. Takayasu L et al. Circadian oscillations of microbial and functional composition in the human salivary microbiome. DNA Res 24, 261–270 (2017). [PubMed: 28338745]
27. Reitmeier S et al. Arrhythmic gut microbiome signatures predict risk of type 2 diabetes. Cell Host Microbe 28, 258–272 (2020). [PubMed: 32619440]
28. Allaband C et al. Intermittent hypoxia and hypercapnia alter diurnal rhythms of luminal gut microbiome and metabolome. mSystems 10.1128/mSystems.00116-21 (2021).
29. Tuganbaev T et al. Diet diurnally regulates small intestinal microbiome-epithelial-immune homeostasis and enteritis. Cell 182, 1441–1459 (2020). [PubMed: 32888430]
30. Wu G et al. Light exposure influences the diurnal oscillation of gut microbiota in mice. Biochem. Biophys. Res. Commun 501, 16–23 (2018). [PubMed: 29730287]
31. Nelson RJ et al. Time of day as a critical variable in biology. BMC Biol 20, 142 (2022). [PubMed: 35705939]
32. Dantas Machado AC et al. Diet and feeding pattern modulate diurnal dynamics of the ileal microbiome and transcriptome. Cell Rep 40, 111008 (2022). [PubMed: 35793637]
33. Morton JT et al. Establishing microbial composition measurement standards with reference frames. Nat. Commun 10, 2719 (2019). [PubMed: 31222023]

34. Caporaso JG et al. Moving pictures of the human microbiome. Genome Biol 12, R50 (2011). [PubMed: 21624126]
35. Bisanz JE, Upadhyay V, Turnbaugh JA, Ly K & Turnbaugh PJ Meta-analysis reveals reproducible gut microbiome alterations in response to a high-fat diet. Cell Host Microbe 26, 265–272.e4 (2019). [PubMed: 31324413]
36. Kohsaka A et al. High-fat diet disrupts behavioral and molecular circadian rhythms in mice. Cell Metab 6, 414–421 (2007). [PubMed: 17983587]
37. Hatori M et al. Time-restricted feeding without reducing caloric intake prevents metabolic diseases in mice fed a high-fat diet. Cell Metab 15, 848–860 (2012). [PubMed: 22608008]
38. Baker F Normal rumen microflora and microfauna of cattle. Nature 149, 220 (1942).
39. Zhang L, Wu W, Lee Y-K, Xie J & Zhang H Spatial heterogeneity and co-occurrence of mucosal and luminal microbiome across swine intestinal tract. Front. Microbiol 9, 48 (2018). [PubMed: 29472900]
40. Klymiuk I et al. Characterization of the luminal and mucosa-associated microbiome along the gastrointestinal tract: results from surgically treated preterm infants and a murine model. Nutrients 13, 1030 (2021). [PubMed: 33806771]
41. Kim D et al. Comparison of sampling methods in assessing the microbiome from patients with ulcerative colitis. BMC Gastroenterol 21, 396 (2021). [PubMed: 34686128]
42. Tripathi A et al. Intermittent hypoxia and hypercapnia reproducibly change the gut microbiome and metabolome across rodent model systems. mSystems 4, e00058–19 (2019). [PubMed: 31058230]
43. Uhr GT, Dohnalová L & Thaiss CA The Dimension of Time in Host-Microbiome Interactions. mSystems 4, e00216–e00218 (2019).
44. Voigt RM et al. Circadian disorganization alters intestinal microbiota. PLoS ONE 9, e97500 (2014). [PubMed: 24848969]
45. McDonald D et al. American gut: an open platform for citizen science microbiome research. mSystems 3, e00031–18 (2018). [PubMed: 29795809]
46. Borodulin K et al. Cohort profile: the National FINRISK Study. Int. J. Epidemiol 47, 696–696i (2018). [PubMed: 29165699]
47. Ren B et al. Methionine restriction improves gut barrier function by reshaping diurnal rhythms of inflammation-related microbes in aged mice. Front. Nutr 8, 746592 (2021). [PubMed: 35004799]
48. Beli E, Prabakaran S, Krishnan P, Evans-Molina C & Grant MB Loss of diurnal oscillatory rhythms in gut microbiota correlates with changes in circulating metabolites in type 2 diabetic db/db mice. Nutrients 11, E2310 (2019).
49. Wang L et al. Methionine restriction regulates cognitive function in high-fat diet-fed mice: roles of diurnal rhythms of SCFAs producing- and inflammation-related microbes. Mol. Nutr. Food Res 64, e2000190 (2020). [PubMed: 32729963]
50. Guo T et al. Oolong tea polyphenols ameliorate circadian rhythm of intestinal microbiome and liver clock genes in mouse model. J. Agric. Food Chem 67, 11969–11976 (2019). [PubMed: 31583884]
51. Mistry P et al. Circadian influence on the microbiome improves heart failure outcomes. J. Mol. Cell. Cardiol 149, 54–72 (2020). [PubMed: 32961201]
52. Shao Y et al. Effects of sleeve gastrectomy on the composition and diurnal oscillation of gut microbiota related to the metabolic improvements. Surg. Obes. Relat. Dis 14, 731–739 (2018). [PubMed: 29680673]
53. Bolyen E et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol 37, 852–857 (2019). [PubMed: 31341288]
54. Amir A et al. Deblur rapidly resolves single-nucleotide community sequence patterns. mSystems 2, e00191–16 (2017). [PubMed: 28289731]
55. Mirarab S, Nguyen N & Warnow T in Biocomputing 2012, 247–258 (World Scientific, 2011).
56. Lozupone C, Lladser ME, Knights D, Stombaugh J & Knight R UniFrac: an effective distance metric for microbial community comparison. ISME J 5, 169–172 (2011). [PubMed: 20827291]
57. Lauber CL, Zhou N, Gordon JI, Knight R & Fierer N Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples: Influence of

short-term storage conditions on microbiota. FEMS Microbiol. Lett 307, 80–86 (2010). [PubMed: 20412303]

58. Marotz C et al. Evaluation of the effect of storage methods on fecal, saliva, and skin microbiome composition. mSystems 6, e01329–20 (2021). [PubMed: 33906915]
59. Song SJ et al. Preservation methods differ in fecal microbiome stability, affecting suitability for field studies. mSystems 1, e00021–16 (2016).
60. Wu GD et al. Sampling and pyrosequencing methods for characterizing bacterial communities in the human gut using 16S sequence tags. BMC Microbiol 10, 206 (2010). [PubMed: 20673359]
61. Piedrahita JA, Zhang SH, Hagaman JR, Oliver PM & Maeda N Generation of mice carrying a mutant apolipoprotein E gene inactivated by gene targeting in embryonic stem cells. Proc. Natl Acad. Sci. USA 89, 4471–4475 (1992). [PubMed: 1584779]
62. Chaix A, Zarrinpar A, Miu P & Panda S Time-restricted feeding is a preventative and therapeutic intervention against diverse nutritional challenges. Cell Metab 20, 991–1005 (2014). [PubMed: 25470547]
63. Gibbons S Diel Mouse Gut Study (HF/LF diet). figshare 10.6084/m9.figshare.882928 (2015).

Box 1 Recommendations for microbiome sample collection. These recommendations should assist researchers in accounting for host temporal dynamics in their experimental design

- Because both in silico and in vivo data indicate that circadian rhythms impact microbiome composition, we propose that investigators include in the methods section of the manuscript the vivarium lighting conditions, confirmation that fresh samples were collected and the sample collection time in ZT notation.
- Investigators should explicitly confirm that the control and experimental condition samples were collected at the same approximate ZT.
- Our findings highlight the substantial impact of diet, feeding patterns and gastrointestinal region on microbiome compositional fluctuations. Therefore, we recommend developing a best-practice guide for microbiome sample collection that accounts for diurnal rhythms across different subfields or preclinical models. This will provide clear guidelines for researchers and facilitate the comparison and replication of studies.
- In longitudinal studies, ensure samples are collected consistently at the same time of day to account for time's influence on data. Our findings suggest that differences between control and experimental groups remain stable over 20 weeks; however, further research within each subfield/preclinical model with varied designs and extended durations is necessary to fully assess this observation.
- To enhance research integrity and reproducibility, we urge the National Institutes of Health (NIH) and journal editors to promote these guidelines for microbiome sample collection. Their leadership in encouraging these practices will be crucial in standardizing methodologies and ensuring the reliability of research findings across the field.

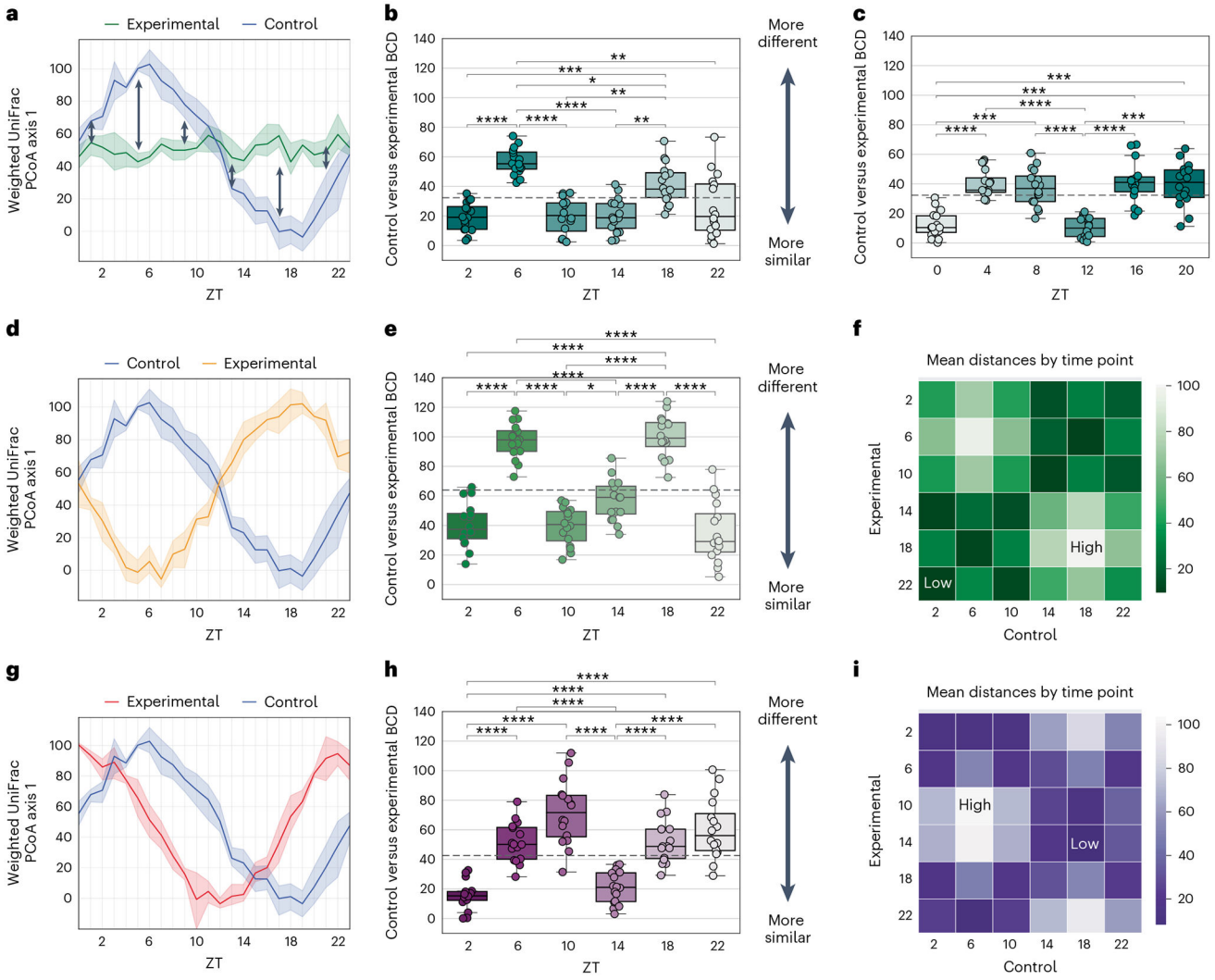


Fig. 1 |. Mock circadian data to explain and exemplify new metric BCD.

Example 1: diurnal oscillations compared to non-oscillatory. **a**, PCoA axis 1 over the course of a day (ZT0 at lights on). Arrows indicate what BCD is measuring. Blue indicates the control group, green indicates the experimental group. Shaded regions indicate 95% confidence interval. **b**, Box-plot/strip-plot of control-to-experimental BCD ($n = 4$ per group) at six different time points (denoted by blue-green colour). **c**, Box-plot/strip-plot of control-to-experimental BCD ($n = 4$ per group) at six alternate time points (shifted by 2 h). Example 2: opposing diurnal oscillations. **d**, PCoA axis 1 over the course of a day. Blue indicates control group, yellow indicates the experimental group and shaded regions indicate 95% confidence intervals. **e**, Box-plot/strip-plot of control-to-experimental BCD ($n = 4$ per group) at six different time points (denoted by green colour). **f**, Heatmap of mean ($n = 4$ per group) mock β diversity distance values by the six selected time points, calculated using only control group samples. Highest and lowest values are labelled. Example 3: offset (by 6 h/1.5 time points) diurnal oscillations. **g**, PCoA axis 1 over the course of a day. Blue indicates control group, red indicates experimental group and shaded regions indicate 95% confidence interval. **h**, Box-plot/strip-plot of control-to-experimental BCD ($n = 4$ per group)

at six different time points (denoted by purple colour). **i**, Heatmap of mean (n = 4 per group) mock β diversity distance values by the six selected time points, calculated using only control group samples. Highest and lowest values are labelled. Box-plot centre line in all panels indicates median, edges of boxes represent quartiles and error bars show min and max values. Significance for all was determined using a paired, two-sided Mann–Whitney–Wilcoxon test with Bonferroni correction. *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.00001.

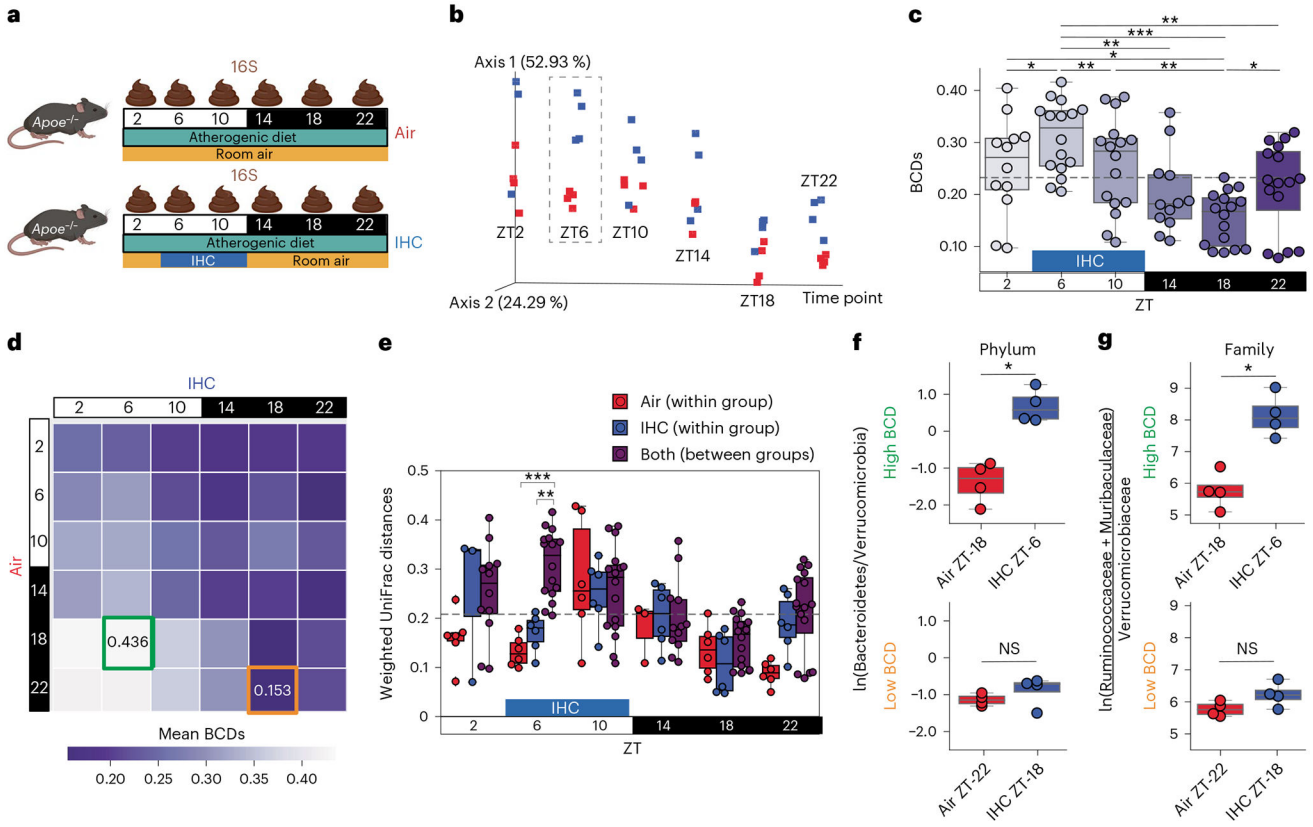


Fig. 2 |. Microbiome analysis of *Apoe*$^{-/-}$ mice exposed to IHC conditions show different outcomes depending on time point of sample collection.

a, Experimental design. $n = 4$ mice per group per time point; ZT2 one IHC sample rarefied out, ZT14 one air sample rarefied out. Image was created with BioRender.com. **b**, Weighted UniFrac PCoA lateral view, with time points as one axis. **c**, BCD, a subset of weighted UniFrac β-diversity distances ($n = 4$ mice per group per time point). Significance was determined using paired Wilcoxon rank-sum test. The BCD values in this experiment were oscillating in a circadian fashion (MetaCycle, JTK method, $P < 0.001$). **d**, BCD heatmap by time point. Replicates were collapsed by taking the mean. Highest are highlighted in green and lowest are highlighted in orange. **e**, Box-plot/scatter-plot of weighted UniFrac distance values for WCD for control (air, red), WCDs for experimental (IHC, blue) and BCDs (both, purple) groups ($n = 4$ mice per group per time point). Mann–Whitney–Wilcoxon test with Bonferroni correction used to determine significant differences between the three groups. **f**,**g**, At the peak and trough time points identified in **d**, the logarithmic ratios of differentially abundant key phyla of interest ($n = 4$ mice per group per time point) (**f**) and the logarithmic ratios of differentially abundant key families of interest ($n = 4$ mice per group per time point) (**g**). A Mann–Whitney–Wilcoxon test with Bonferroni correction was used to determine significant differences. The box-plot centre line in all panels indicates the median, edges of boxes show quartiles and error bars represent min and max values. NS, not significant; *$P <$ 0.05; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.00001$.

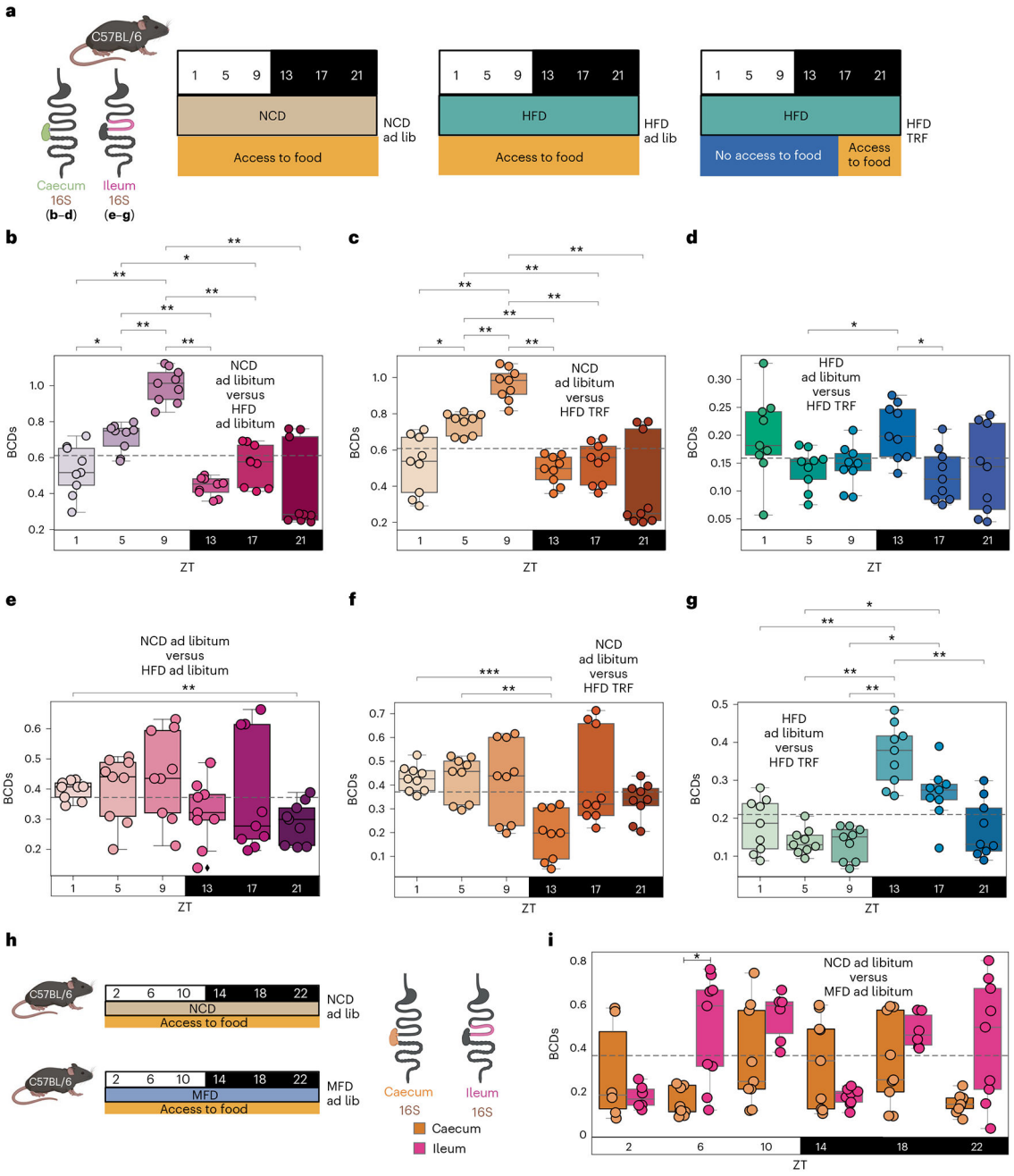


Fig. 3 |. Gastrointestinal regions have individual time dynamics that are influenced by diet and feeding patterns.

a, Experimental design. TRF mice were restricted to eating only between ZT13–ZT21. Time point ZT13 was collected before access to the diet was given and, thus, mice were fasted at this time point. Time points were taken every 4 h for 24 h (n = 3 mice per condition per time point from separate cages; six total time points). Every point on the box-plot + swarm-plot represents the calculated β diversity distance between a control and experimental mouse. Image was created with BioRender.com. **b**–**d**, Caecum. BCD box-plot/swarm-plot for NCD ad libitum (control) versus HFD ad libitum (**b**), NCD ad libitum (control) versus HFD TRF

(**c**) and HFD ad libitum versus HFD TRF (**d**). **e**–**g**, Ileum. BCD box-plot/swarm-plot for NCD ad libitum versus HFD ad libitum (**e**), NCD ad libitum versus HFD TRF (**f**) and HFD ad libitum versus HFD TRF (**g**). The dotted line is the average/mean of all shown weighted UniFrac distances. Significance was determined using Mann–Whitney–Wilcoxon test two-sided with Bonferroni correction. **h**, Experimental design. Mice were fed ad libitum with either NCD or high MFD. After 5 weeks, caecal and ileal samples were collected every 4 h for 24 h ($n = 3$ mice per condition per time point). Image was created with BioRender.com. **i**, BCD for both ileal and caecal samples comparing NCD versus high MFD The dotted line is the average of all shown weighted UniFrac distances. Ileal versus caecal pairwise significance was determined using Mann–Whitney–Wilcoxon test two-sided with Bonferroni correction. Box-plot centre line indicates median, edges of boxes show quartiles and error bars indicate min and max values. Significance was determined using a two-sided Mann–Whitney–Wilcoxon test with Bonferroni correction. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.00001$.

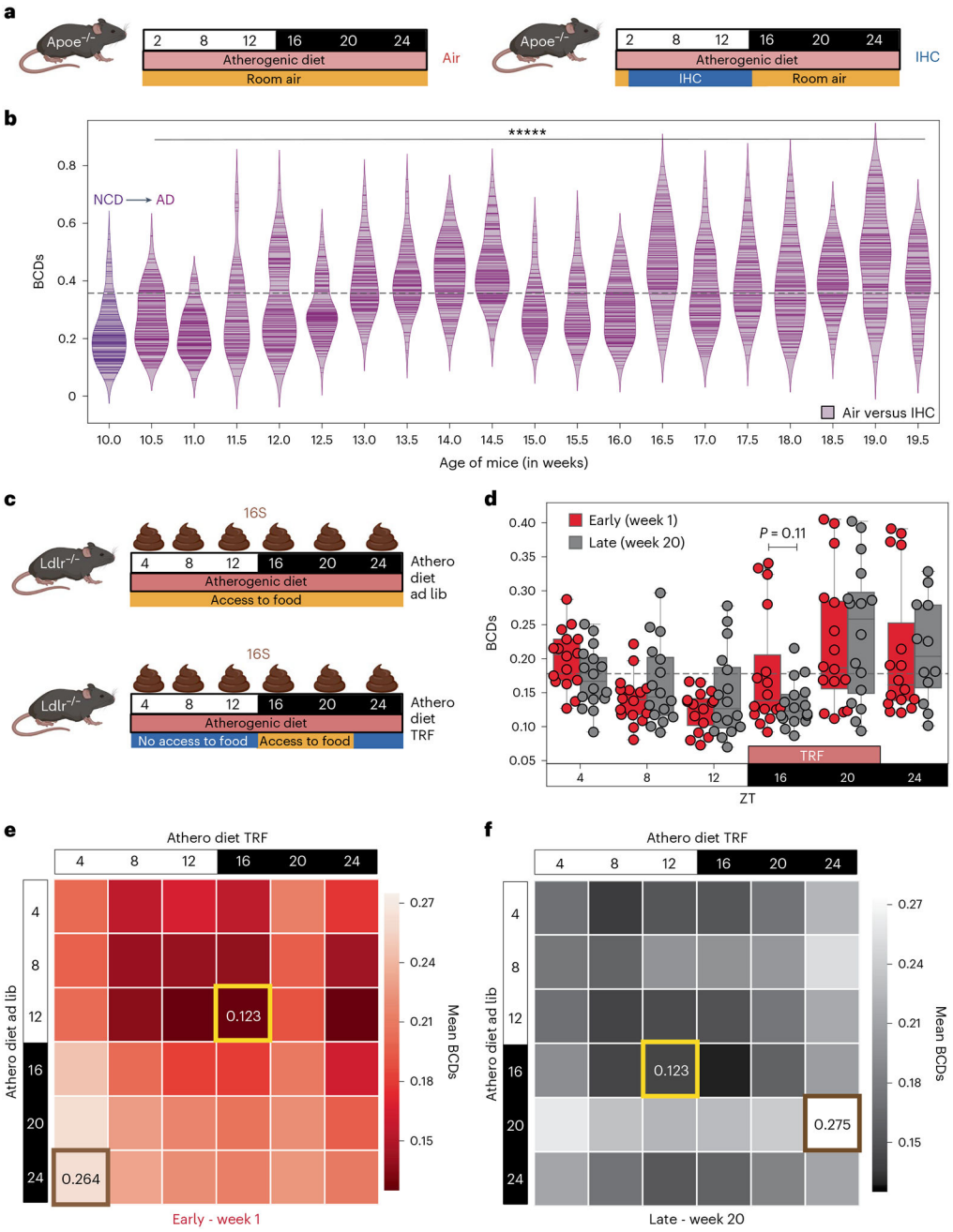


Fig. 4 |. Longitudinal data are also susceptible to the influence of time.

a, Experimental design and sample collection for longitudinal IHC study. During the 10 weeks of exposure to either normal room air or IHC conditions, samples were collected between ZT3 and ZT5 every 3–4 days for the duration of the study (n = 12 mice per condition). Image was created with BioRender.com. **b**, Faecal BCD violin plot over the course of the IHC longitudinal study. Diet switch was from NCD to an atherogenic diet (AD) and occurred on day 1. Dotted line shows the mean of all data shown. While most of the other time points are significantly different from each other, the only comparison shown is between age 10.5 weeks and 19.5 weeks, which is most relevant to our

discussion. Significance was determined using a paired, two-sided Wilcoxon rank-sum test. **c**, Experimental design and sample collection for TRF study. Mice were fed AD either ad libitum or TRF. Samples were collected every 4 h for 24 h (n = 6 mice per condition per time point) after 1 week (early; pre-phenotype) and after 20 weeks (late; post-phenotype). Image was created with BioRender.com. **d**, Faecal BCD box-plot/scatter-plot for ad libitum versus TRF conditions at the early (week 1) and late (week 20) time points. Dotted line is the average of all of the weighted UniFrac distances. Significance was determined using a paired, two-sided Wilcoxon rank-sum test. **e**,**f**, BCD heatmap for early samples (**e**) and late samples (**f**). Replicates were collapsed by taking the mean. The highest value is highlighted in tan and the lowest value is highlighted in yellow. Box-plot centre line indicates median, edges of boxes show quartiles and error bars indicate the min and max values. *P < 0.05; **P < 0.01; ***P < 0.001; *****P < 0.00001.