Special Proceedings: ISBN#: 978-81-950383-1-2
24th Annual Conference, 23–27 February 2022; pp 205-213

Are Some People More Likely Than Others to Develop Disease When Exposed to an Infection?

Shyamal Peddada
Biostatistics and Bioinformatics Branch
Eunice Shriver National Institute of Child Health and Human Development (NICHD), NIH
Bethesda, MD

Received: 14 August 2022; Revised: 20 August 2022; Accepted: 24 August 2022

Abstract

When a pandemic strikes a population, not everyone is similarly affected. Some people are more susceptible to a disease than others, and the question is: why? A person with a "good" innate immune system potentially has better defense mechanisms to confront or respond to an infection than those who do not have good immune system. Understanding the underlying heterogeneity in immunity is a question of scientific interest. There are potentially numerous reasons for heterogeneity within a population, such as, genetics, environment, socioeconomics, and so on. It is well-established that microbiome plays an important role in inflammation and immune response. In this article we summarize the findings of Chen *et al*. (2021), on the role of microbiome in the HIV infection and AIDS, who demonstrated that changes in gut microbiome takes place months before the onset of HIV infection and years before the HIV patients progressed to develop AIDS. Given the findings of that study, one may speculate a similar phenomenon for other infectious diseases, such as the novel coronavirus (COVID-19).

Key Words: Alpha Diversity; Beta Diversity: Differential Abundance; HIV infection; Infectious diseases; Microbiome.

1. Background

Human health is dependent on complex interactions between our genome, the external factors, and the internal environment over time. The external environment broadly includes chemicals to which we are exposed, the air we breathe, our water supply, diet, physical activity, and other factors. The internal environment includes hormones, the microbiome and microbial biproducts such as the short chain fatty acids, various metabolites, and so on. The interaction of gene by external factors over time as been extensively studied in the literature. During the past decade, researchers began to recognize the important role played by the microbiome on human health. Humans are estimated to have 45.6 million bacterial genes in oral and gut microbiome alone (*cf*., Tierney *et al*., 2019), which is about 2000-fold more than human genes. Therefore, the microbiome is sometimes referred to as the "second genome", or another "organ" of human body

Corresponding Author: Shyamal Peddada
Email: shyamal.peddada@nih.gov

(*cf*., O'Hara and Shanahan, 2006, Relman and Falkow, 2017, Hurst, 2017). It is now well-established that the microbiome is involved in metabolism, immune response, and inflammation. Gut microbiome also regulate mood (Sampson and Mazmanian, 2015, Steenbergen *et al*., 2015), anxiety, cognition, pain, aging, and a host of other factors of human health and behavior. It is not surprising that numerous diseases such as allergies, asthma, obesity (Turnbaugh *et al*., 2009), Crohn's disease (Gevers *et al*. 2014), inflammatory bowel diseases, and HIV are associated with alterations in the microbiome (Lozupone *et al*., 2013).

The role of the microbiome on human health and disease has been demonstrated for many other diseases. For example, in the general population, there is greater microbial diversity among people with no asthma than those with asthma, with a greater abundance of taxa such as *Faecalibacterium prausnitzii, Sutterella wadsworthensis* and *Bacteroides stercoris*, and microbial byproducts such as short chain fatty acids, acetate and butyrate (Wang *et al*., 2018). Conversely, there is a greater abundance of pro-inflammatory bacteria such as *Clostridium bolteae*, *Clostridium ramosum, Clostridium spiroforme* and *Eggerthella lenta* among people with asthma (Wang *et al*., 2018). Gut microbiome is also involved in the production of immunoglobulin E (IgE) (Cahenzli *et al*., 2013) and are associated with lung functions such as forced expiratory volume (FEV) (Begley et *al*., 2018). Furthermore, as seen in recent literature (Vila *et al*., 2020, Weersma and Zhernakova, 2020), medications potentially affect the gut microbiome, which in turn impacts the efficacy of treatments (Weersma and Zhernakova, 2020).

We use the terms "taxa", "bacteria", and "microbe" interchangeably. Often the terms "microbiome" and "microbiota" are used in the literature interchangeably although these two are distinct terms. Microbiota refers to the taxa describing various organisms whereas microbiome is a broader term that includes microbiota and their genes.

Given the recent, and ongoing COVID-19 pandemic, some natural questions to ask are: What is the effect of the microbiome on infectious diseases such as COVID-19, HIV, etc.? What caused some people to be more susceptible to acquiring a disease than others? Are there generally significant differences between the gut microbiome of people who acquire a disease and those who do not?

The role of the microbiome in infectious diseases is well-documented in the literature, with several review articles and Perspectives written on this subject in recent years, e.g., Harris *et al*., (2017), Libertucci and Young (2019), Cai *et a*l., (2021), Giovanni *et al*. (2021), Harper *et al*. (2021), and Hussain *et al*. (2021). Interactions between the human microbiome and pathogens, and the role of microbiome in stimulating the host immune system to defend against pathogens and hence protect against infections is well-characterized in these review articles. Interest in this area has increased with the recent pandemic to understand the mechanistic role of the microbiome for developing prebiotics, probiotics, fecal microbiome transplantation (FMT), and other treatments for infectious diseases. For example, according to Yeoh *et al*. (2021) the gut microbiota plays an important role in modulating markers of host immune system such as the cytokines and inflammatory markers, and thus plays a crucial part in lessening the severity of COVID-19.

Although the existing literature suggests a change in the gut microbial composition after the onset of a disease among infected people, the question remains whether people with gut dysbiosis

are prone to develop a disease if exposed to an infection. This question is difficult to answer because it is not a common practice to collect stool samples in the general population to investigate who in the future gets a disease and who does not. The prospective Multicenter AIDS Cohort Study (MACS) provided an opportunity for Chen *et al.* (2021) to answer the above question in the context of HIV infection and AIDS. During the HIV pandemic of the 1980's, the MACS was established at four centers in the United States, namely, Pittsburgh, Baltimore, Chicago, and Los Angeles. The study recruited men who had sex with men (MSM) before any of the recruited individuals seroconverted, *i.e*., before becoming HIV positive (HIV+). This cohort allowed Chen *et al.* (2021) to investigate differences in gut microbial compositions between a group of men who became HIV+, or developed AIDS after becoming HIV+, with those who remained HIV negative (HIV-). In Section 2 we briefly describe the microbiome data and methodologies for analyzing those data, and we summarize in Section 3 the findings of Chen *et al*. (2021). Concluding remarks are provided in Section 4.

2. A Brief Overview of Statistical Methodology

Observed microbiome data are count data derived from sequencing a specimen obtained from an ecosystem, in the present case the human gut. Two popular technologies used to generate microbiome data are the technology based on 16s ribosomal RNA (16s rRNA) and the shotgun metagenomics. Since 16s rRNA is highly conserved in almost all bacteria and its function has not evolutionarily changed (*cf.* Janda and Abbott, 2007), it is commonly used by researchers conducting microbiome surveys. Microbial count data obtained by using 16s rRNA are commonly referred to as "16s data." Although 16s rRNA technology is specific for bacterial profiling, the shotgun metagenomics surveys not only bacteria but also sequences all genomic DNA. Thus, in addition to high taxonomic resolution, at the level of species and strain, the shotgun metagenomics allows host DNA inference, functional profiling, and metabolic pathway analysis. Since the shotgun metagenomics method currently is far more expensive than 16s rRNA, some researchers apply informatic tools, such as PICRUSt (Langille *et al*., 2013), to 16s data for functional profiling.

In the following description, our focus is on the analysis of 16s data. The observed microbiome data are a matrix of counts, with rows representing various taxa and columns representing the samples. Two important characteristics of these microbiome data are that
(1) typically, a very large proportion of the entries of this matrix are zero. The zero entries may arise for several reasons as detailed in Kaul *et al*. (2017); and (2) for reasons explained below, the observed counts are compositional, *i.e*., reside in a simplex.

A variety of statistical parameters are considered when comparing two or more experimental groups. Common parameters of interest are alpha diversity, beta diversity, taxon abundance, and taxon relative abundance. The alpha diversity parameter measures the diversity within samples. Numerous measures of alpha diversity appear in the microbiome literature; some examples include Shannon's entropy, the Gini-Simpson index, and the Chao1 index. The beta diversity parameter measures diversity in taxa between samples or between groups and, similar to alpha diversity, the microbiome literature provides a variety of measures of beta diversity due to differing concepts of distances between samples. For a review of these measures of diversities, we refer to Weiss *et al*. (2017) and references therein.

In addition to measures of diversity, researchers are interested in identifying taxa that are differentially abundant between experimental groups. This area of research, known as *differential abundance analysis*, is centered on testing the null hypothesis of equality of (relative) abundance of a taxon between two or more groups against various alternative hypothesis.

Suppose there are G experimental groups and n_i subjects in the ith experimental group $i = 1,2,\ldots,G$, In a unit volume of an ecosystem of the kth subject, $k = 1,2,..,n_i$ in the ith experimental group $i = 1,2,\ldots,G$, let A_{ijk} denote the unobservable true abundance of the jth taxon, $j = 1,2,..,m$. Using the 16s or shotgun metagenomics technologies we obtain O_{ijk}, the observed counts of the jth taxon, on the kth subject in the ith experimental group.

Define $T_{ik} = \sum_{j=1}^{m} O_{ijk}$, sometimes called the *library size* of the k^{th}subject in the ith experimental group. Due to sample collection methods and technology, T_{ik}, in practice, is highly variable among subjects. Also, within each subject k, as the sample collection and preparations change, the observed counts O_{ijk} are assumed to change proportionally. Thus, the observed counts O_{ijk} within each subject are compositional and hence are in a simplex. Statisticians often convert these observed counts to relative abundances, $R_{ijk} = O_{ijk}/T_{ik}$, so that $\sum_{j=1}^{m} R_{ijk} = 1$. One may view the observed counts O_{ijk} as an unknown fraction (or multiple) c_{ik} of the true abundance A_{ijk}. The population abundance parameter of interest is $\mu_{ij} = E\left(A_{ijk}\right), i = 1, 2, \ldots, m, j = 1,2,\ldots,G$. Unfortunately, this parameter μ_{ij}cannot be estimated unbiasedly unless the bias due to the nuisance parameter c_{ik} is eliminated. Suppose that $\lambda_{ijk} = A_{ijk}/\sum_{j=1}^{m} A_{ijk}$ denotes the relative abundance of the jth taxon, $j = 1,2,\ldots,m$, on the kth subject, $k = 1,2,\ldots,n_i$, in the i^{th}experimental group, $i = 1,2,\ldots,G$, then λ_{ijk} can be estimated by the observed relative abundance of R_{ijk}. For this reason, as an alternative to μ_{ij}, researchers sometimes are interested in making inferences about the mean relative abundance $\theta_{ij} = E(\lambda_{ijk})$. Although it is natural to study the relative abundance because it does not involve the nuisance parameter c_{ik}, but from a clinical or scientific point of view, the relative abundance parameter may be difficult to interpret. Consider the two ecosystems in the toy example provided in Tables 1a and 1b. Table 1a consists of the abundances of five taxa in the two ecosystems. The counts of the first four taxa are identical across the two ecosystems, with only Taxon 5 differentially abundant between the two ecosystems. Often researchers are interested in identifying Taxon 5, the differentially abundant taxon. However, if one were to consider relative abundances (Table 1b), all five taxa have differential relative abundances between the ecosystems. Although it is mathematically correct that the relative abundances differ between the ecosystems, clinically or scientifically it may not be a useful piece of information. Thus, there are reasons to prefer to test for equality of abundances rather than the equality of relative abundances of taxa between ecosystems. However, it is a challenging problem to test for equality of abundances between two or more ecosystems because of the nuisance parameter mentioned above. Several methods have been proposed in the literature to eliminate the bias due to the unknown nuisance parameter. Some methods commonly used in the literature include

Table 1a: Abundances of taxa

Taxon	Ecosystem 1	Ecosystem 2
Taxon1	1	1
Taxon2	4	4
Taxon3	10	10
Taxon4	20	20
Taxon5	65	100
Sum	100	135

Table 1b: Relative abundance of taxa

Taxon	Ecosystem 1	Ecosystem 2
Taxon1	0.01	0.01
Taxon2	0.04	0.03
Taxon3	0.1	0.07
Taxon4	0.2	0.15
Taxon5	0.65	0.74
Sum	1	1

ALDEx2 (Fernandes *et al*., 2014), ANCOM (Mandal *et al*., 2015), ANCOM-BC (Lin and Peddada, 2020), and RNA-seq-based methods such as edgeR (Robinson *et al*., 2010), DESeq2 (Love *et al*., 2014).

Although the above methods are among the popular methods for conducting differential abundance analysis, they are rapidly getting outdated with several new methods introduced in the literature on a regular basis (Zhou *et al*., 2021, Hu *et al*., 2022). Recently, Nearing *et al*. (2022) conducted an exhaustive numerical study involving 38 different 16s data sets to evaluate different available methods for differential abundance analysis and they concluded that ALDEx2 and ANCOM-II produce the most consistent results. However, these authors did not include LiNDA (Zhou *et al*., 2021) or LOCOM (Hu *et al*., 2022) which were perhaps not available at the time Nearing *et al*. (2022) was published.

3. The Findings of Chen *et al*. (2021)

Using the stool and blood samples collected from men during their first clinical visit in the 1980's by MACS, Chen *et al*. (2021) investigated the differences in microbial compositions of men who developed HIV infection at a future time and those who did not. Furthermore, they also investigated differences in the microbial compositions among those who developed AIDS at different time points in the future.

The study consisted of 265 participants who were HIV negative (*i.e*., did not seroconvert, denoted as negative controls (NC)) at the beginning of the study, and among these 156 remained HIV- but 109 seroconverted, *i.e*., became HIV+, within about six months after the first samples were collected (denoted as seroconverters (SC)). Of the 109 who seroconverted, 32 of them developed AIDS within 5 years, 31 developed AIDS between 5 to 10 years and 46 took more than 10 years to develop AIDS. The data are summarized in the schematic provided in Figure 1. Chen *et al*. (2021) compared the SC and the NC groups using their microbiome data collected at the first visit when all of them were HIV-, the SC group did not yet seroconvert. Similarly, they compared the different AIDS groups (G1, G2 and G3) using their microbiome data collected at the first visit when all of them were HIV-.

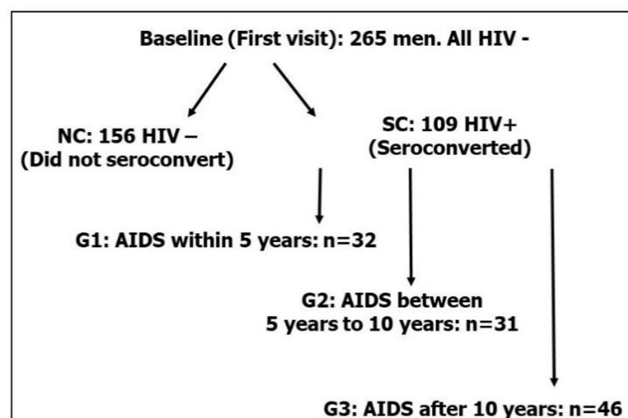

Figure 1: Study participants in the Multicenter AIDS Cohort Study (MACS)

Using ANCOM-BC for differential abundance analysis of microbiome and some standard regression-based methods for other data collected in the study, such as alpha diversity, cytokines data, ratio of CD4/CD8 counts, and short chain fatty acids data, Chen *et al*. (2021) made several interesting and important discoveries. At the baseline, or the first visit, when all 265 men in the sample are HIV negative, not surprisingly Chen *et al*. (2021) did not find differences in the ratio of CD4/CD8 between those who became HIV+ in the future versus those who stayed HIV-.

However, as expected, by the next visit when some became HIV+, the CD4/CD8 ratio was significantly different between the two groups of men at the second visit. Although there was no significant difference in the alpha diversity at the baseline between men who later became HIV+ and those who remained HIV-., very interestingly, Chen *et al*. (2021) did find differential abundance of various pro and anti-inflammatory taxa at the baseline between the two groups of men. Thus, they discovered intestinal dysbiosis months before men developed HIV infection, characterized by increase in pro-inflammatory taxa such as *Prevotella Stercorea* and a reduction in commensal bacteria such as *Bacteroides spp*, *Akkermansia Muciniphila, Alistipes spp, and Ruminococcus Spp*. Not only did they see such differences at the baseline, but Chen *et al*. (2021) also discovered that ratio of *Prevotellaceae to Bacteroidaceae* was highly correlated with HIV infection at a later time point. In view of these findings, it is not surprising that they also discovered elevated levels of circulating cytokines IL-6, LBP, sCD14, sCD163, before developing HIV infection. Increased levels of these cytokines suggest immune response to changes in the bacterial composition. There is growing evidence in the literature demonstrating the important roles played by short chain fatty acids produced by the gut microbiota. Propionate is one such short chain fatty acid which Chen *et al*. (2021) found to be positively correlated with the levels of CD4/CD8 counts. Taken all these findings together, it appears that differences in gut microbial composition could be in the pathway of a person developing HIV infection. Chen *et al.* (2021) also found significant increase in pro-inflammatory bacteria and a significant decrease in the anti-inflammatory commensal bacteria at baseline among those who developed AIDS soon after becoming HIV+ compared to those who were either slow to develop or never developed AIDS.

4. Conclusions

Although the focus of Chen *et al.* (2021) was on understanding the role of microbiome in HIV infection and the development of AIDS, their work together with emerging literature cited in this paper, suggests that gut microbiome may potentially play an important role in other infectious diseases, including COVID-19. There appear to be differences in the composition of gut microbiome in people who later became HIV+ than those remained HIV-, and those who rapidly developed AIDS versus those who did not. Before one can assert about other infectious diseases, more carefully planned studies similar to MACS' HIV/AIDS study are needed. The MACS study provided a unique opportunity because, at the beginning of the study none of the men were HIV+, but over time some became HIV+, and the comparisons of microbiome data were performed before anyone became HIV+. Such studies are not easy to conduct, unless it becomes a common practice to collect microbiome samples routinely, for example during the annual physical exams. Another possibility is to collect stool samples from subjects when they have an infectious disease and again obtain stool samples after they are fully recovered from their disease. Of course, such a design assumes that the subjects gut microbiome did not change permanently once a person is infected.

From a statistical perspective, these data and this line of research provide opportunities to develop statistical methods for analyzing these complex data.

Acknowledgement

The author's research was funded by the intramural research program of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD, NIH).

References

Begley L., Madapoosi S., Opron K., Ndum O., Baptist A., Rysso K., Erb-Downward J. R. and Huang Y. J. (2018). Gut microbiota relationships to lung function and adult asthma phenotype: a pilot study. *BMJ Open Respiratory Research*, **5**(**1**), e000324.

Cahenzli J., Köller Y., Wyss M., Geuking M. B. and McCoy K. D. (2013). Intestinal microbial diversity during early-life colonization shapes long-term IgE levels. *Cell Host Microbe*, **14**(**5**), 559e570.

Cai Y., Chen L., Zhang S., Zeng L. and Zeng G. (2021). The role of gut microbiota in infectious diseases. *WIREs Mechanisms of Disease*. DOI: 10.1002/wsbm.1551.

Chen Y., Lin H., Cole M., Morris A., Martinson J., Mckay H., Mimiaga M., Margolick J., Fitch A., Methe B., Srinivas V. R., Peddada S. D. and Rinaldo C. R. (2021). Signature changes in gut microbiome are associated with increased susceptibility to HIV-1 infection in MSM. *Microbiome*. doi: 10.1186/s40168-021-01168-w.

Fernandes, A. D. *et al*. (2014). Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. *Microbiome*, **2**, 15.

Gevers D., Kugathasan S., Denson L. A., Vazquez-Baeza Y., Treuren W. V., Ren B., Schwager E., Knights D., Song S. J., Yassour M., *et al*. (2014). The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host and Microbe*, **15**(**3**), 382-392.

Giovanni M. Y., Schnieder J. S., Calder T. and Fauci A. S. (2021). Refocusing human microbiota research in infectious and immune-mediated diseases: Advancing to the next stage. *Journal of Infectious Diseases*. DOI: 10.1093/infdis/jiaa706.

Harper A., Vijaykumar V., Ouwehand A. C., ter Haar J., Obis D., Espadaler J., Binda S., Desiraju S. and Day R. (2021). Viral infections, the microbiome, and probiotics. *Frontiers in Cellular Infections and Microbiology*. https://doi.org/10.3389/fcimb.2020.596166.

Harris V. C., Haak B. W., van Hensbroek M. B. and Wiersinga W. J. (2017). The intestinal microbiome in infectious diseases: The clinical relevance of a rapidly emerging field. *Open Forum Infectious Diseases*, doi: 10.1093/ofid/ofx144.

Hu Y., Satten G., Hu Y. -J. (2022). LOCOM: A logistic regression model for testing differential abundance in compositional microbiome data with false discovery rate control. https://doi.org/10.1073/pnas.2122788119.

Hurst G. D. D. (2017). Extended genomes: symbiosis and evolution. *Interface Focus*, **7**(**5**), 20170001.

Hussain I., Cher G. L. Y., Abid M. A. and Abid M. B. (2021). Role of gut microbiome in COVID-19: An insight into pathogenesis and therapeutic potential. *Frontiers in Immunology*. doi: 10.3389/fimmu.2021.765965.

Janda J. M. and Abbott S. L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, Perils, and Pitfalls. *Journal of Clinical Microbiology*, **45**, 2761-2764.

Kaul A., Mandal S., Davidov O. and Peddada S. D. (2017). Analysis of microbiome data in the presence of excess zeros. *Frontiers in Microbiology*. doi: 10.3389/fmicb.2017.02114

Langille M. G. I., Zaneveld J., Caporaso J. G., McDonald D., Knights D., Reyes J. A., Clemente J. C., Burkepile D. E., Vega Thurber R. L., Knight R., Beiko R. G. and Huttenhower C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology*, **31**, 814–821.

Libertucci J. and Young V. B. (2019). The role of the microbiota in infectious diseases. *Nature Microbiology*, https://doi.org/10.1038/s41564-018-0278-4

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**, 550.

Lozupone C. A., Li M., Campbell T. B., Flores S. C., Linderman D., Gebert M. J., Knight R., Fontenot A. P. and Palmer B. E. (2013). Alterations in the gut microbiota associated with HIV-1 infection. *Cell Host and Microbe*, **14(3)**, 29 - 339.

Mandal S., Van Treuren W., White R. A., Eggesbø M., Knight R. and Peddada S. D. (2015). Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microbial Ecology in Health and Disease,* **26***,* 1–7.

Nearing, J. T., Douglas, G. M., Hayes, M. G. *et al.* (2022). Microbiome differential abundance methods produce different results across 38 datasets. *Nature Communications*, **13,** 342. https://doi.org/10.1038/s41467-022-28034-z

O'Hara A. M. and Shanahan F. (2006). The gut organ as a forgotten organ. *EMBO Reports*, **7**(**7**), 688-693.

Relman D. A. and Falkow S. (2017). The meaning and impact of the human genome sequence for microbiology. *Trends in Microbiology*, **9**(**5**), 206-208.

Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140.

Sampson T. R. and Mazmanian S. K. (2015) Control of brain development, function, and behavior by the microbiome. *Cell Host Microbe*, **17**(**5**), 565–576.
https://doi.org/10.1016/j.chom.2015.04.011

Steenbergen L., Sellaro R., van Hemert S., Bosch J. A. and Colzato L. S. (2015). A randomized controlled trial to test the effect of multispecies probiotics on cognitive reactivity to sad mood. *Brain*, *Behaviour and Immunity*, **48**, 258–264.

Tierney B. T., Yang Z., Luber J. M., Beaudin M., Wibowo M. C., Baek C., Mehlenbacher E., Patel C. J. and Kostic A. D. (2019). The landscape of genetic content in the gut and oral human microbiome. *Cell Host and Microbe*, **26**(**2**), 283-295.

Turnbaugh P. J., Hamady M., Yatsunenko T., Cantarel B. L., Duncan A., Ley R. E., Sogin M. L., Jones W. J., Roe B. A., Affourtit J. P., Henrissat B., Heath A. C., Knight R. and Gordon J. I. (2009). A core gut microbiome in obese and lean twins. *Nature*, **457**(**7228**), 480.

Vila A. V., Collij V., Sanna S., Sinha T., Imhann F., Bourgonje A. R., Mujagic Z., Jonkers D. M. A. E., Masclee A. A. M., Fu J., Kurilshikov A., Wijmenga C., Zhernakova A.and Weersma R. K. (2020). Impact of commonly used drugs on the composition and metabolic function of the gut microbiota. *Nature Communications*, **11**, 362. https://doi.org/10.1038/s41467-019-14177-z.

Weersma R. K., Zhernakova A. and Fu J. (2020). Interaction between drugs and the gut microbiome. *Gut*, **0**, 1–10.

Wang Q. I., Li F., Liang B., Liang Y., Chen S., Mo X., Ju Y., Zhao H., Jia H., Spector T. D., Xie H. and Guo R. (2018). A metagenome-wide association study of gut microbiota in asthma in UK adults. *BMC Microbiology*, **18**(**1**), 114.

Weiss S., Xu Z. Z., Peddada S., Amir A., Bittinger K., Gonzalez A. *et al*. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome*, **5**, 27. doi: 10.1186/s40168-017-0237-y.

Wu H. J. and Wu E. (2012) The role of gut microbiota in immune homeostasis and autoimmunity. *Gut Microbes*, **3**(**1**), 4–14.

Yeoh Y. K., Zuo T., Lui G. C.-Y. *et al*. (2021). Gut microbiota composition reflects disease severity and dysfunctional immune responses in patients with COVID-19. *Gut*, doi:10.1136/gutjnl-2020-323020.

Zhou, H., He, K., Chen, J. *et al.* (2022). LinDA: Linear models for differential abundance analysis of microbiome compositional data. *Genome Biology*, **23,** 95. https://doi.org/10.1186/s13059-022-02655-5.