International Journal of Medical Microbiology 312 (2022) 151560

Contents lists available at ScienceDirect

International Journal of Medical Microbiology

journal homepage: www.elsevier.com/locate/ijmm

Machine learning on the road to unlocking microbiota's potential for boosting immune checkpoint therapy

Szymon Wojciechowski [a], Monika Majchrzak-Górecka [a], Paweł Biernat [a], Krzysztof Odrzywołek [a,b], Łukasz Pruss [a,d], Konrad Zych [a], Jan Majta [a,c], Kaja Milanowska-Zabel [a,*]

[a] Ardigen, Podole 76, 30-394 Kraków, Poland
[b] Institute of Computer Science, Faculty of Computer Science, Electronics and Telecommunications, AGH University of Science and Technology, Mickiewicza 30, 30-059 Kraków, Poland
[c] Department of Computational Biophysics and Bioinformatics, Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University, Krakow, Poland
[d] Department of Biochemistry, Molecular Biology and Biotechnology, Faculty of Chemistry, Wroclaw University of Science and Technology, 50-373 Wroclaw, Poland

ARTICLE INFO

Keywords:
Microbiome
Gut
Oncology
Immune checkpoint therapy
Machine learning

ABSTRACT

The intestinal microbiota is a complex and diverse ecological community that fulfills multiple functions and substantially impacts human health. Despite its plasticity, unfavorable conditions can cause perturbations leading to so-called dysbiosis, which have been connected to multiple diseases. Unfortunately, understanding the mechanisms underlying the crosstalk between those microorganisms and their host is proving to be difficult. Traditionally used bioinformatic tools have difficulties to fully exploit big data generated for this purpose by modern high throughput screens. Machine Learning (ML) may be a potential means of solving such problems, but it requires diligent application to allow for drawing valid conclusions. This is especially crucial as gaining insight into the mechanistic basis of microbial impact on human health is highly anticipated in numerous fields of study. This includes oncology, where growing amounts of studies implicate the gut ecosystems in both cancerogenesis and antineoplastic treatment outcomes. Based on these reports and first signs of clinical benefits related to microbiota modulation in human trials, hopes are rising for the development of microbiome-derived diagnostics and therapeutics. In this mini-review, we're inspecting analytical approaches used to uncover the role of gut microbiome in immune checkpoint therapy (ICT) with the use of shotgun metagenomic sequencing (SMS) data.

1. The impact of host-microbiota crosstalk on human health

Microorganisms inhabiting the human body fulfill numerous functions crucial for the host (Dominguez-Bello et al., 2019; Hill and Round, 2021; Valdes et al., 2018). Among them are a vast number of ways in which microbiota is involved in maintaining human health, such as controlling pathogen growth or modulating immune system development and functioning (Bäumler and Sperandio, 2016; Gensollen et al., 2016; Levy et al., 2017a; Zheng et al., 2020). However the exact elements that make up the microbiome and determine its specific capabilities are dependent on numerous factors (Berry et al., 2020; Mohajeri et al., 2018; Zeevi et al., 2015). As those are both host-intrinsic and extrinsic (Aleman and Valenzano, 2019; Dong and Gupta, 2019; Keebaugh and Ja, 2017; Kurilshikov et al., 2021, 2017; Reimer, 2019; Scepanovic et al., 2019; Weersma et al., 2020), the microbiota composition and its functional potential differs not only between populations (Deschasaux et al., 2018; Gupta et al., 2017; He et al., 2018), and individuals (Arumugam et al., 2011; Consortium, 2012; Falony et al., 2016; Zhernakova et al., 2016) but for each person it is also subject to change over time (David et al., 2014; Mehta et al., 2018). Consistency can be observed in respect to the presence of the main components of microbial communities at higher taxonomic levels, however, their relative proportions and the exact species present vary markedly between individuals (Consortium, 2012). This results in a great degree of variation in composition, existing between seemingly healthy adults, making it difficult to define a "normal/healthy microbiota". However, a so-called "dysbiosis", can be generally described as a state resulting from an occurrence of conditions beyond an ecosystem's resistance and resilience, causing the disruption of balance within this ecosystem. Dysbiosis may result in the elimination or underrepresentation of some

* Corresponding author.
E-mail address: kaja.milanowska-zabel@ardigen.com (K. Milanowska-Zabel).

https://doi.org/10.1016/j.ijmm.2022.151560
Received 30 August 2021; Received in revised form 15 July 2022; Accepted 31 August 2022
Available online 9 September 2022

commensal species and the outgrowth of potentially pathogenic microorganisms that are otherwise kept contained (Dogra et al., 2020; Levy et al., 2017b; Lozupone et al., 2012). This adverse state has been associated with multiple clinical conditions (Malard et al., 2020; Ruff et al., 2020) and, based on Pubmed entries for the term "dysbiosis", the number of new studies related to this topic has been growing exponentially in the last decade. However, the road from merely observing correlations to really understanding the mode of action and being able to take advantage of this knowledge is not easy.

Just the sheer numbers of microorganisms that inhabit the human gut already make the analysis challenging. The gut microbiota of an average healthy adult is composed of tens of trillions of bacteria, including more than a thousand known bacterial species which translates to few million genes (Claesson et al., 2009; Gilbert et al., 2018; Lozupone et al., 2012). On top of that, not all microbial species have been discovered and characterized due to their challenging culturing requirements (Almeida et al., 2019). However, even the identification of all microbial species composing one's microbiota at a given point in time could still not be enough to understand its functional potential. That is because these microorganisms constantly interact with the host and with each other, responding to changing conditions by swiftly evolving (Rook et al., 2017; Rosenberg and Zilber-Rosenberg, 2018). Thus literature reports pertaining to characteristics of a given bacteria can get outdated quite quickly especially in aspects such as antimicrobial resistance (Chevereau et al., 2015; Woods et al., 2020; Yong et al., 2009). This plasticity also results in multiple microbes that are deemed beneficial or neutral to the host under certain circumstances but still can cause harm under others (Ruff et al., 2020).

2. Complex ecosystems require robust analyses

Microbiota research has its roots in laboratory analysis informed by the use of bioinformatic and statistical tools. However, this traditional approach is time-consuming, labor-intensive and often not enough to grasp the essence of such microbial communities. These kinds of complex and extremely dynamic ecosystems require detailed information to allow for drawing valid conclusions on their functioning (Gerber, 2014).

Currently, collecting such data is becoming possible with the help of technological advancements in the field of high-throughput screens, resulting in an increase of affordability and accessibility of these groundbreaking tools. For example, Shotgun Metagenome Sequencing (SMS) generates data allowing not only for quantification of both high and low-abundant microorganisms (not restricted to bacteria), but also enables functional profiling. Through this the actual potential of these communities can be investigated (Claesson et al., 2017). Although around half of the inferred microbial protein sequences are similar to nothing stored in databases, deep learning tools can be fed with SMS data to gain even more insight into intestinal ecosystem functioning (Odrzywołek et al., 2022). Combining this kind of detailed information on microbiota with multiple, complex laboratory and clinical parameters may hold the keys to solving numerous clinical challenges, such as assessing disease risk, making a diagnosis, prognosing the course of a disease, and personalizing a treatment. However, fully exploiting such data is not trivial as it needs to be diligently analyzed with use of proper data science tools allowing for management and integration of multidimensional big data (Cammarota et al., 2020). As a result, there is a noticeable turn towards using Machine Learning (ML) approaches (Cammarota et al., 2020; Namkung, 2020; Sundh et al., 2021; Wirbel et al., 2021). ML is a class of statistical learning methods, loosely classified as algorithms that improve themselves in concurrent runs based on the evaluation of previous runs. In other words, the algorithm learns by itself, hence the name. For analysis of microbiome data the main advantage is the ability to analyze multidimensional data as a whole (multivariate analysis) and indicate the most important dimensions. Furthermore, the ML has readily available techniques to assess the transferability of the results. Nevertheless, due to the specificity of the SMS data – usually a scarce number of samples and richness of information per sample - carefree use of ML can easily lead to drawing false conclusions.

3. Gut microbiota and cancer

A field with especially high, and so far unmet, needs of harvesting microbiota's potential is oncology (Elkrief et al., 2019; Zipkin, 2021). The impact of some bacteria on cancerogenesis has been investigated for quite some time (Garrett, 2015). However, the true interest in the microbiome got sparked by research published in the mid-2010s showing correlation between its composition and therapy outcomes (Gopalakrishnan et al., 2018; Iida et al., 2013; Matson et al., 2018; Routy et al., 2018; Vetizou et al., 2015; Viaud et al., 2013). Since then, a range of studies has been conducted to explore the role of intestinal communities in multiple aspects, including safety, tolerability and the efficacy of antineoplastic treatments – especially in context of immune checkpoint therapy (ICT) (Helmink et al., 2019; Sepich-Poore et al., 2021; Vivarelli et al., 2019). Recent reports have shown that modulation of the gut microbiome - by fecal microbial transplant - can in fact be a valid mean for increasing response to this therapy (Baruch et al., 2021; Davar et al., 2021). However, fecal microbial transplant itself comes with plenty of technical challenges, safety concerns and biological uncertainties (Cammarota et al., 2019; Ma et al., 2017; Park and Seo, 2021). That is why scientists are looking to determine the microbiome's mode of action - particular actionable features within the microbial haystack. Identification of such features would allow for the precise design of microbiota-based diagnostic tools and therapeutic strategies (Zitvogel et al., 2018). The research on microbiome-derived solutions for improving ICT outcomes is currently gaining momentum, as numerous research projects and commercial ventures are being undertaken (Zipkin, 2021).

Taking into consideration the fact that the success of these endeavors hinges on drawing valid conclusions, it is of the utmost importance that the collected data is properly analyzed. Thus, it is crucial to understand if the tools used for uncovering the role of microbiome in cancer immunotherapy, especially the ones from the ML toolbox, are being applied correctly and therefore to what extent the obtained results are reliable. To our knowledge, until now, no critical analysis of the approaches used for this purpose has been conducted. Thus, we decided to dedicate the following mini-review to inspecting the statistical methods used in analyzing microbiome in cancer immunotherapy.

4. The state of literature

To analyze all published results coming from the application of ML methodology to SMS data, we performed a manual literature search in PubMed with use of terms: ("immune checkpoint" OR immunotherapy) AND (metagenome OR "gut microbiome"), which provided us with 304 results (on 01.04.2021). We included only the studies pertaining to the analysis of patient SMS data, because they reflect the full diversity of the microbiome, enabling full utilization of computational analysis. This resulted in the identification of 10 original papers and 1 meta-analysis of publicly available SMS datasets that were published in the course of the last 4 years, which analyzed the influence of gut bacteria on cancers such as metastatic melanoma (Frankel et al., 2017; Gopalakrishnan et al., 2018; Limeta et al., 2020; Matson et al., 2018; Peters et al., 2019; Wind et al., 2020), renal cell carcinoma (RCC) (Derosa et al., 2020; Routy et al., 2018; Salgia et al., 2020), non-small cell lung cancer (NSCLC) (Cvetkovic et al., 2021; Routy et al., 2018) and gastrointestinal cancers (Peng et al., 2020). Publications discussed in this review are summarized in Table 1 and Table 2.

Authors used either (or both) of the two standard assessment criteria of therapy success: RECIST 1.1 (Response Evaluation Criteria in Solid Tumors 1.1) (Eisenhauer et al., 2009) and survival time. RECIST 1.1 is a set of rules for comparing tumor burden at baseline and in follow-ups in order to classify a patient as having a complete response (CR), partial

Table 1
Summary of the SMS data sets analyzed in the review.

Publication & dataset	Cancer type	Data type & analysis	SMS parameters	Response classification (based on BOR)	Survival data used	Cohort characteristics
Frankel et al., 2017 (63) (PRJNA397906)	Melanoma	Metagenomic (taxonomic & functional analysis) and metabolomic profiles	Illumina HiSeq 2000 (100 bp paired-end reads)	R: CR, PR, SD; NR: PD		39 subjects (23 R / 16NR): anti-PD-1 mono: pembrolizumab (14), nivolumab (1), anti-PD-1 & anti-CTLA-4 combo: nivolumab & ipilimumab (24)
Gopalakrishnan et al., 2018 (51) (PRJEB22893)	Melanoma	Metagenomic (taxonomic & functional) profiles	Illumina HiSeq (2 ×100 bp paired-end reads)	R: CR, PR, SD > 6 mos; NR: SD < 6 mos, PD	PFS	25 subjects (14 R / 11NR): anti-PD-1 mono (25)
Matson et al., 2018 (52) (PRJNA399742)	Melanoma	Metagenomic (taxonomic) profiles	Illumina NextSeq (2 ×150 bp paired-end reads)	R: CR, PR; NR: SD, PD		39 subjects (15 R / 24NR): anti-PD-1 mono: pembrolizumab (30), nivolumab (5), anti-CTLA-4 mono: ipilimumab (4)
Peters et al., 2019 (64) (PRJNA541981)	Melanoma	Metagenomic (taxonomic & functional) and metatranscriptomic profiles	Illumina HiSeq 2500 (2 ×101 bp paired-end reads)	–	PFS	27 subjects: anti-PD-1 mono (14), anti-CTLA-4 (1), anti-PD-1 & anti-CTLA-4 combo (12)
Wind et al., 2020 (65) (SMS data not freely available)	Melanoma	Metagenomic (taxonomic & functional) profiles	Illumina HiSeq (paired-end reads)	R: CR, PR (if confirmed at the next scan), SD > 12 wks; NR: SD < 12 wks SD, PD (except if not confirmed at the second scan)	PFS, OS	25 subjects (12 R / 13 NR): anti-PD-1 mono (23), anti-PD-1 & anti-CTLA-4 combo (2)
Routy et al., 2018 (53) (PRJEB22863)	NSCLC, RCC	Metagenomic (taxonomic) profiles	Ion-proton technology (ThermoFisher; 150 bp single-end reads)	R: CR, PR, SD; NR: PD, dead	PFS	NSCLC (time point 0): 65 subjects (33 R / 32NR): anti-PD-1 mono: nivolumab (65) RCC (time point 0): 62 subjects (42 R / 20 NR): anti-PD-1 mono: nivolumab (62)
Cvetkovic et al., 2021 (67) (SMS data not freely available)	NSCLC	Metagenomic (taxonomic) profiles	Ion-proton technology (ThermoFisher; 150 bp single-end reads)	R: CR, PR, SD > 6 mos; NR: SD < 6 mos, PD	PFS, OS	71 subjects (no info on R / NR no.): anti-PD-1 mono: nivolumab (38), pembrolizumab (24), anti-PD-L1 mono: durwalumab (1), anti-PD-1 & chemotherapy combo: pembrolizumab + chemotherapy (7), anti-PD-L1 & anti-CTLA-4 combo: durwalumab & tremelimumab (1)
Derosa et al., 2020 (68) (SMS data not freely available)	RCC	Metagenomic (taxonomic) profiles	Ion-proton technology (ThermoFisher; 150 bp single-end reads)	R: CR, PR, SD > 6 mos; NR: SD < 6 mos, PD	PFS	69 * subjects (32 R / 37 NR): anti-PD-1 mono: nivolumab * 29 new and 40 subjects already included in Routy et al., 2018 (53)
Salgia et al., 2020 (69) (SMS data not freely available)	RCC	Metagenomic (taxonomic) profiles	Illumina NextSeq High Output (2 ×150 bp paired-end reads)	R: CR, PR, SD > 4 mos; NR: SD < 4 mos, PD		31 subjects (18 R / 13 NR); anti-PD-1 mono: nivolumab (24), anti-PD-1 & anti-CTLA-4 combo: nivolumab & ipilimumab (7)
Peng et al., 2020 (70) (PRJNA615114)	Gastrointestinal cancers (colorectal, esophageal, gastric & other)	Metagenomic (taxonomic & functional) profiles	Illumina NovoSeq 6000 (2 ×150 bp paired-end reads)	R: (CR, PR, SD) > 3 mos; NR: PD < 3 mos	PFS	40 subjects (25 R / 15NR): anti-PD-1 mono, anti-PD-L1 mono, anti-PD-1 & anti-CTLA-4 combo

Abbreviations: BOR, best overall response; combo, combination therapy; CR, complete response; CTLA-4, cytotoxic T-lymphocyte-associated protein 4; mono, monotherapy; NR, non-responder; NSCLC, Non-Small Cell Lung Cancer; OS, overall survival; PD, progressive disease; PD-1, programmed cell death 1; PD-L1, programmed cell death ligand 1; PFS, progression-free survival; PR, partial response; R, responder; RCC, Renal Cell Carcinoma; SD, stable disease; SMS, Shotgun metagenomic sequencing;

Table 2
Summary of analytical methods used in the publications.

Publication	Survival analysis	Statistical Test	Frameworks	Diversity & unsupervised	Supervised machine learning	Details on methods used
Frankel et al., 2017 (63)		●	●	●		– Statistical tests: Welch's t, Wilcoxon signed-rank, Mann-Whitney U – Frameworks: LEfSe – Diversity & unsupervised: Hierarchical clustering
Gopalakrishnan et al., 2018 (51)	●	●	●	●	●	– Survival analysis: KM estimates, Cox regression – Statistical tests: Mann-Whitney U, Fischer's exact test – Frameworks: LEfSe – Diversity & unsupervised: PCoA, Hierarchical clustering – **Supervised ML:** Feature selection, univariate linear regression
Matson et al., 2018 (52)		●		●		– Statistical tests: Permutation, Mann-Whitney U, ANOVA – Diversity & unsupervised: Hierarchical clustering, PCA
Peters et al., 2019 (64)	●	●	●	●	●	– Survival analysis: KM estimates, Cox regression – Statistical tests: Fisher's – Frameworks: MiRKAT-S, OMiSA – Diversity & unsupervised: Hierarchical clustering – **Supervised ML:** Cox regression
Wind et al., 2020 (65)	●	●			●	– Survival analysis: Cox regression – Statistical tests: Kruskal-Wallis, Fisher's exact – **Supervised ML:** Logistic regression, least-squares regression
Routy et al., 2018 (53)	●	●			●	– Survival analysis: KM estimates, Cox regression, Mantel-Cox test – Statistical tests: ANOVA, Wilcoxon, Student t, Cochran-Armitage
Cvetkovic et al., 2021 (67)	●	●	●	●		– Survival analysis: KM estimates – Statistical tests: Student t, Fisher's exact, Mann-Whitney U – Frameworks: DESeq2 – Diversity & unsupervised: nMDS
Derosa et al., 2020 (68)	●	●	●	●	●	– Survival analysis: KM estimates – Statistical tests: Chi-square, Mann-Whitney U, Student t, ANOVA – Frameworks: LEfSe – Diversity & unsupervised: PCoA – **Supervised ML:** PLS, random forest
Salgia et al., 2020 (69)		●	●	●		– Statistical tests: Kruskal-Wallis – Frameworks: LEfSe – Diversity & unsupervised: Hierarchical clustering
Peng et al., 2020 (70)	●	●	●	●	●	– Survival analysis: KM estimates – Statistical tests: Wilcoxon, Mann-Whitney U – Frameworks: Omnibus – Diversity & unsupervised: – **Supervised ML:** SVM, ElasticNet, multiple others
Limeta et al., 2020 (66)	●	●		●	●	– Survival analysis: KM estimates – Statistical tests: Wilcoxon's rank-sum, Fisher's exact – Diversity & unsupervised: Beta-diversity, hierarchical clustering – **Supervised ML:** Random forest

Abbreviations: ANOVA, Analysis of Variance; KM, Kaplan-Meier; LEfSe, Linear Discriminant Analysis, coupled with Effect Size; MiRKAT-S, Microbiome Regression-based Kernel Association Test for censored Survival outcomes; ML, machine learning; nMDS, non-Metric Multidimensional Scaling; OMiSA, Optimal Microbiome-based Survival Analysis; PCA, Principal Component Analysis; PCoA, Principal Coordinate Analysis; PLS, Partial Least Squares; SVM, Support Vector Machine;

response (PR), stable disease (SD) or progressive disease (PD). Based on all patient evaluation(s) during a particular treatment, best overall responses (BOR) - defined as the best response across all time points up to the end of the treatment - can be assigned. All of the studies that recorded BOR used these values for dichotomizing patients into responders (R) and non-responders (NR). This allows for the use of statistical tests to compare the microbial compositions between R/NR groups. It also opens the door to using supervised ML for predicting R or NR status from microbiome composition.

Alternatively, instead of using rigid R/NR groups, some authors based their analyses on one of the two survival time measures, widely used in the context of anti-cancer therapies: progression-free survival (PFS) and overall survival (OS). PFS measures the time between the start of the therapy and the tumor's progression (or death from any cause), whereas OS measures the life length after the initiation of a therapy. While OS analysis is the gold standard for determining therapeutic

efficacy, obtaining this kind of data can be cumbersome. For example, in case of melanoma patients treated with ICIs as a first line therapy, observation spanning over several years must be planned for, if median OS is to be reached (Hamid et al., 2019; Wolchok et al., 2022). Therefore, although they are significantly less accurate than OS, surrogate "early" endpoints - especially PFS and overall response rate (ORR) – are very often used to estimate drug activity (Cooper et al., 2020). Moreover, in phase II ICI clinical trials, ORR - defined as the proportion of patients who experienced CR or PR - was the most frequently used endpoint (Hamada et al., 2018). Thus, it is not surprising that only two of the studies discussed in this review made any use of OS data, but most used PFS to some extent (see Table 1). Of the latter group, only two used PFS to draw conclusions about the microbiome response to treatment: Peters et al., utilized specific values and Derosa et al., dichotomized patients using a cut-off value of 12 months. In other studies, the survival data were used only to verify the results obtained based on BOR. Just like in the case of R/NR, PFS and OS come with their own set of standard analytical methods designed to link changes in microbiome with later death or progression times. The collective name for these methods is survival analysis (Clark et al., 2003) and are also based on either specialized statistical tests or supervised ML methods. Therefore, we decided to create a dedicated section for survival analysis in the detailed method description below.

5. Analytical tools

5.1. Survival analysis

An approachable method to visualize time-to-event, such as death or disease progression, are curves estimated with the Kaplan-Meier estimator (Kaplan and Meier, 1958). They have been used in almost all considered publications (Cvetkovic et al., 2021; Gopalakrishnan et al., 2018; Limeta et al., 2020; Peng et al., 2020; Peters et al., 2019; Routy et al., 2018; Wind et al., 2020), due to its widespread understanding and excellent clarity with regard to visualization of treatment outcomes. Such visualization can be illuminative, but it is still subject to the bias of human perception. To overcome this, Kaplan-Meier estimators can be compared using the log-rank test, also known as Mantel-Cox test (Mantel, 1966), which almost universally accompanied Kaplan-Meier curves in the aforementioned studies (Cvetkovic et al., 2021; Gopalakrishnan et al., 2018; Limeta et al., 2020; Peng et al., 2020; Peters et al., 2019; Routy et al., 2018). Yet another approach to survival modeling is Cox regression (also known as Cox Proportional Hazards model) (Cox, 1972), where confounding variables can be taken into account to form a multivariate model. This approach can not only answer questions like “What is the chance that a patient will survive to time X?”, or “What is patients median survival time?”, but also “What is the impact of other variables, like age, or BMI, on the survival time”. Some authors (Gopalakrishnan et al., 2018; Peters et al., 2019; Wind et al., 2020) resorted to this type of analysis and, for instance, Wind et al., found that carriers of *Bacteroides massiliensis* tend to have longer PFS, whereas carriers of an unclassified *Peptostreptococcaceae* species have, on average, shorter time to progression (Wind et al., 2020). In this case, the findings were confirmed with a statistical test - Wald test (Wald, 1943) that assesses the overall significance of the Cox regression (Wind et al., 2020).

5.2. Statistical tests

In the previous paragraph, we have already introduced some of the most popular tools utilized in the discussed papers - the statistical tests. They are, by far, the most frequently used quantitative methods and every publication contains at least one application. On top of being used in survival analysis, they are also applied whenever there is a binary problem, such as a comparison of two groups of patients. This is natural in the context of ICT, where patients can be split into R and NR. Statistical tests to compare values between such groups include: parametric t-test and its generalized version - Welch’s t-test, the non-parametric Mann-Whitney U test (also known as Wilcoxon rank-sum test) (Mann and Whitney, 1947) and the permutation test (see Table 3). Mann-Whitney U test was applied in more than half of the studies (Cvetkovic et al., 2021; Derosa et al., 2020; Frankel et al., 2017; Gopalakrishnan et al., 2018; Limeta et al., 2020; Matson et al., 2018), whereas the permutation test - just in one of them (Matson et al., 2018). As an example of the Wilcoxon test, one can cite the analysis performed by Limeta et al. (Limeta et al., 2020), where 17 operational taxonomic units (OTUs) were determined as differentially abundant between responders and non-responders. Parametric versions - t-test and Welch’s t-test are also frequently employed (Cvetkovic et al., 2021; Derosa et al., 2020; Frankel et al., 2017; Routy et al., 2018). Frankel et al. (2017), for example, used the latter to determine differentially abundant metabolites. A similar analysis was carried out by Routy et al. (2018), where two groups of patients split by PFS at 6 months, were compared with respect to gene and metagenomic species count, but the name of the test used is not disclosed. This group of tests can be applied, when the feature that is being compared is continuous (for example – the abundance of OTUs). But two groups can also be contrasted with a categorical variable, as was the case in the publication by Derosa et al. (2020), where the impact of antibiotic treatment administered prior to ICT, on the response to the antineoplastic therapy was assessed using chi-square test with the Yates correction. Alternatively, Fisher’s exact test (Fisher, 1922) is preferred when the sample size is low. Its application can be found in half of the publications (Cvetkovic et al., 2021; Gopalakrishnan et al., 2018; Limeta et al., 2020; Peters et al., 2019; Wind et al., 2020).

So far we have been looking at the comparisons between two groups, though, comparison of more than two groups can also be required on some occasions. The authors resort then to tests that can handle such cases: the Cochran–Armitage (Armitage, 1955), Kruskal-Wallis tests (Kruskal and Wallis, 1952) or ANOVA (Analysis Of Variance) (Welch, 1951). The first of these was applied by Routy et al. (2018) to assess patients by the presence of *Akkermansia muciniphila* in their feces with respect to the clinical response (PR, SD, PD). ANOVA is used as a parametric extension of the Mann-Whitney U test, where differences between multiple groups can be assessed jointly. It was employed by Peng et al. (2020), at the univariate feature selection step, by Routy et al. (Routy et al., 2018) to determine whether inclusion of *A. muciniphila* into the PD-1 therapy hinders tumor growth and by both Matson et al. (2018) and Cvetkovic et al. (2021) in a similar context. There are multiple versions of the ANOVA analysis which take into account various factors. One such version, the PERMANOVA (Permutational Multivariate Analysis Of Variance) (Anderson, 2001), has been used to compare clusters of patients (Cvetkovic et al., 2021; Derosa et al., 2020). Finally, the Kruskal-Wallis test is a non-parametric equivalent of ANOVA and an extension of the Mann-Whitney test for multiple groups (Derosa et al., 2020; Frankel et al., 2017; Salgia et al., 2020).

On top of all the tests that were previously noted, we also found the application of the Kolmogorov-Smirnov test (Conover, 1972) for examination of distribution of *Prevotella/Bacteroides* ratio among the patients (Peng et al., 2020).

In multiple scenarios, statistical tests are repeated multiple times, since there are multiple entities that require the same statistical procedure. If this is the case, one runs the risk of accepting a false positive outcome as truth, whereas such an outcome stems only from the repeated nature of the experiment. In the considered publications, a typical scheme where this could happen is comparison of abundances of any taxonomic or functional unit between responders and non-responders. We should then take measures to counteract this risk. In statistics, there exists a group of methods dedicated to alleviate this phenomenon – correction (also called adjustment) for multiple testing (Benjamini and Hochberg, 1995). It has become a good habit and belongs to good practices in the domain. Therefore, many publications apply one of these methods. This is the case in most of the studies

Table 3
Summary of statistical tests used in the publications.

Statistical test	Type	How many groups	Target variable	Measured variable	Comments
(Student's) t-test	Parametric	Two	Categorical	Quantitative	
Welch's (unequal variances) t-test					More reliable then Studen's t-test when the groups have unequal variances
Mann-Whitney U test (Wilcoxon rank-sum test)	Non-parametric				Non-parametric replacements for t-test
(Monte Carlo) Permutation test					
Kolmogorov-Smirnov test		One or two			
(Pearson's) Chi-square test	Non-parametric	Two or more	Categorical	Categorical	
Fisher's exact test					Exact test that can be used in place of (Pearson's) chi-square test
Cochran–Armitage test		Two			Used when outcome variable is ordinal
ANOVA	Parametric	Two or more	Categorical	Quantitative	Extends the Studen's t-test to multiple groups.
PERMANOVA	Non-parametric				Non-parametric version of ANOVA
Kruskal-Wallis test		Three or more			Extends the Mann–Whitney U test to multiple groups.

(Cvetkovic et al., 2021; Derosa et al., 2020; Frankel et al., 2017; Gopalakrishnan et al., 2018; Peters et al., 2019; Wind et al., 2020), however, in some (Limeta et al., 2020; Matson et al., 2018), the values are left unadjusted.

5.3. Analytical frameworks

Although immensely popular, statistical tests are not the only quantitative methods that are at the disposal of microbiome researchers. At times, the authors resort to special packages or tools that were developed to simplify and automate their pipelines. Peters et al. (2019) use two such frameworks: MiRKAT-S (Microbiome Regression-based Kernel Association Test for censored Survival outcomes) (Plantinga et al., 2017) and OMiSA (Optimal Microbiome-based Survival Analysis) (Koh et al., 2018) to link microbial composition with PFS. MiRKAT-S is a tool that implements a statistical test for verification of the association between the human microbiota at community level and survival outcomes. It accomplishes this through kernel functions, based on the distances between the microbial profiles. An advantage here is that any sensible distance used in the domain can be handled, including the most popular ones: Bray-Curtis and UniFrac. Peters et al. (2019) claim that application of this method to their data yields "marginal" significance of the microbiome's association and PFS. Consequently, authors of the OMiSA framework develop on the existing MiRKAT-S, add their own method (MiSALN, Microbiome-based Survival Analysis using Linear and Non-linear bases of OTUs) and by varying between these two, make the whole pipeline more adaptable to the composition of the microbiome. Both of these frameworks are based on the Cox regression. Significant association of 16 S profiles with PFS was, however, not determined using the OMiSA pipeline. Both of these pipelines have their implementation as packages in the R language.

Yet another tool, unrelated to survival analysis, is LEfSE (Linear

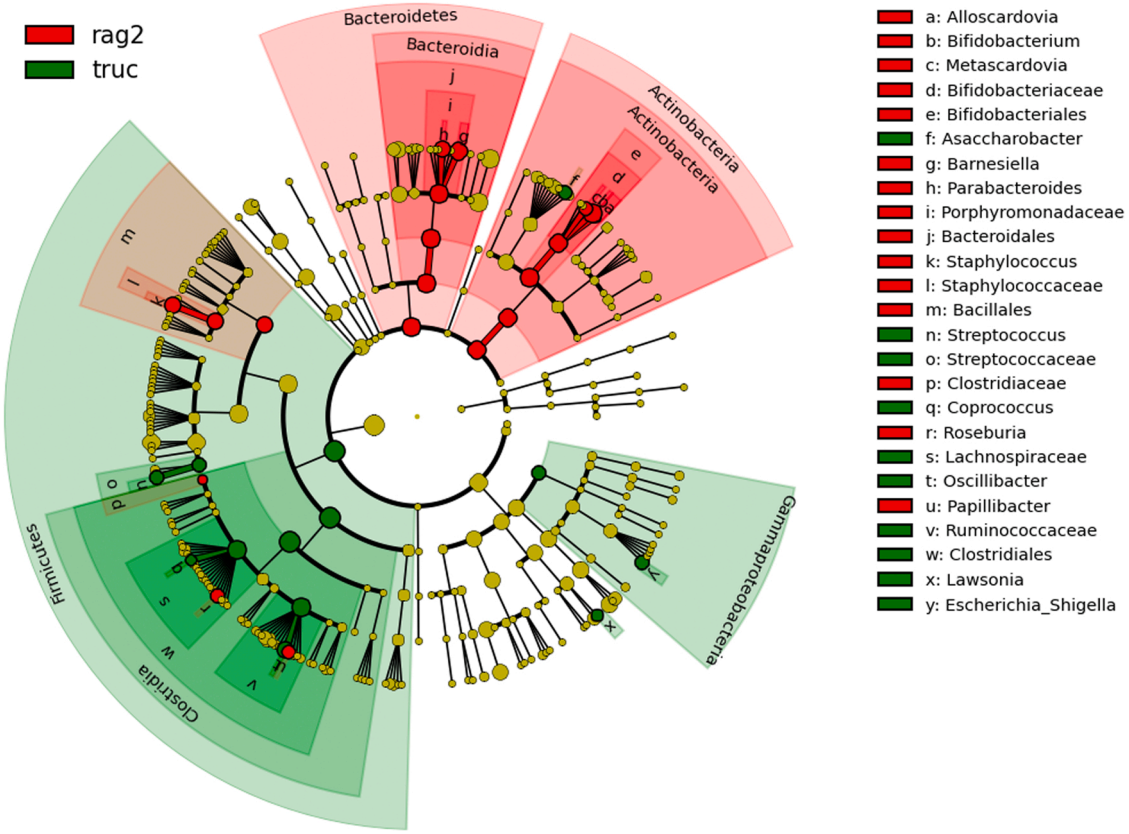

Fig. 1. Visualization of LEfSe results in the cladogram form - interlaying results of the statistical analysis on top of a phylogenetic tree.

Discriminant Analysis, coupled with Effect Size) (Segata et al., 2011). It is a wrapper around a typical pipeline in the biomarker discovery domain that chains three steps (Kruskal-Wallis and Wilcoxon tests and LDA - Linear Discriminant Analysis) and outputs statistically relevant features – either taxonomic or functional. LEfSE enables incorporation of the hierarchical structure of the input data to relate the findings within this structure (Fig. 1). Popularity of this method is underlined with its frequent application (Derosa et al., 2020; Frankel et al., 2017; Gopalakrishnan et al., 2018; Peng et al., 2020). Remarkable feature of the LefSE framework is flexibility to choose multiple groups, classes or subclasses of input to be processed. Although the method uses a multivariate method in the form of LDA, it is primarily based on univariate statistical tests (Kruskal-Wallis and Wilcoxon).

A similar framework – the omnibus method (Chen et al., 2018) - takes its name from the omnibus test, which is employed at one stage of this pipeline. The framework takes jointly into consideration prevalence, abundance and dispersion of the microbiome data and addresses two frequent problems that come up when such data is handled – an excessive number of zeros and outliers. Limeta et al. (2020) took advantage of this method to determine the taxa that significantly differentiate responders and non-responders.

Employing analytical tools can be highly beneficial in microbiome studies as it not only simplifies the whole process but also standardizes the outcomes, which then can be easily compared. However, with a plethora of tools out there, the emphasis shifts to picking the right one. In our opinion, one of the tools that are worth adopting at the crossroad of Microbiome and Immunotherapies is SIAMCAT (Wirbel et al., 2021). The package stands out by providing tools for rigorous multivariate statistical modeling (e.g., LASSO regression, cross validation) with methods to interpret the results. However, one should bear in mind a crucial limitation factor on the utility of SIAMCAT which is the size of the analyzed dataset, as some models chosen by the user could not be well-adjusted for certain sizes.

5.4. Diversity

Clinically actionable results that have the potential to help patients are the most straightforward to obtain at the taxonomic or functional level, since determination of a specific taxon or biological process allows for research on a remedy targeted at that finding. But the authors also investigate bacterial communities jointly via diversity metrics (Whittaker, 1960). Alpha diversity is a function of the number of taxons (at some taxonomic level) present in a given habitat, where in the context of ICT, a habitat is equivalent to the patient's intestines. Comparably low diversity can be related to dysbiosis, which, as mentioned before, has been associated with multiple clinical conditions (Malard et al., 2020; Ruff et al., 2020). The authors, therefore, tend to verify this theory and resort to calculations of some kind of alpha diversity. There is no concordance, however, on what alpha diversity index to choose. Popular choices include richness (Cvetkovic et al., 2021; Peters et al., 2019; Routy et al., 2018), Shannon diversity (Salgia et al., 2020; Wind et al., 2020) or its inverse (Gopalakrishnan et al., 2018; Peng et al., 2020). Two publications took into consideration multiple indices: richness, Shannon, Simpson (Derosa et al., 2020) and Shannon, inverted Shannon, ACE (Abundance-based Coverage Estimator), as well as Chao1 (Limeta et al., 2020). Irrespective of the choice of the index, the values are routinely subjected to statistical analysis using the tests described earlier, especially the tests targeted at comparing two or more cohorts - typically R and NR.

The other prominent set of analyses is oriented at beta-diversity, which compares diversities in two habitats. The notion is largely popular (Cvetkovic et al., 2021; Derosa et al., 2020; Gopalakrishnan et al., 2018; Limeta et al., 2020; Peng et al., 2020; Peters et al., 2019), but its detailed implementation differs in two aspects - the dissimilarity (distance) metric used for visualization and the visualization method itself.

There are a multitude of dissimilarity metrics (popular microbial ecology tool Phyloseq (McMurdie and Holmes, 2013) tool provides 46 different metrics). However, only a handful are routinely used (Cvetkovic et al., 2021; Derosa et al., 2020; Peng et al., 2020, Gopalakrishnan et al., 2018; Limeta et al., 2020; Peters et al., 2019) - Bray-Curtis dissimilarity (BC) (Bray and Curtis, 1957), Jaccard index (Jaccard, 1901) and UniFrac (Lozupone et al., 2007). The main difference between them is that Bray-Curtis dissimilarity (BC) and Jaccard index are purely abundance-based, while UniFrac incorporates relative relatedness of community members (by using the phylogenetic tree). Unifrac can be applied in an unweighted (presence/absence of a branch in the phylogenetic tree) or weighted (taking into account the length of the branch) manner. For equations and details about the metrics please see Table 4. There are no clear guidelines on usage of different metrics and, when construction of the phylogenetic tree is possible, assessing both BC and UniFrac seems to be the best option.

The dissimilarity metrics are used to represent the patients in two dimensions on a scatter plot, which allows for quick inspection of the dataset. In order to do so, the distance matrices need to be reduced to two dimensions, so that graphic libraries can be employed. The authors make heavy use of PCoA (Principal Coordinate Analysis) (Kruskal, 1964) to achieve this (Derosa et al., 2020; Gopalakrishnan et al., 2018; Limeta et al., 2020; Peng et al., 2020; Peters et al., 2019), but in a publication by Cvetkovic et al. (Cvetkovic et al., 2021), non-metric Multidimensional Scaling (CLARKE, 1993) had been applied, whereas in another (Limeta et al., 2020) - t-SNE (t-distributed Stochastic Neighbor Embedding) (Blanche et al., 2013). The choice of methods for dimensionality reduction does not end there. PCA (Principal Component Analysis) was also applied to visualize patients using abundances of 63 OTUs that were previously determined to differentiate responders from non-responders (Matson et al., 2018).

5.5. Unsupervised methods

Dimensionality reduction methods, as exemplified by PCoA, are a part of ML and, more specifically, unsupervised learning. Another prominent class of unsupervised algorithms is used to perform clustering (merging observations into meaningful clusters). This is yet another group of techniques that is favored among researchers. One of the more popular being hierarchical clustering (Derosa et al., 2020; Frankel et al., 2017; Gopalakrishnan et al., 2018; Limeta et al., 2020; Matson et al., 2018; Peng et al., 2020; Peters et al., 2019; Salgia et al., 2020), mostly with complete or Ward linkage. Since this class of models enables drawing a dendrogram plot that visualizes the clusters in an appealing way, such graphs can be found in publications whenever hierarchical clustering has been used. One case when a pattern emerged in hierarchical analysis can be observed in the paper by Peters et al. (2019), where patients were clustered into groups determined by the concordance of Jensen-Shannon dissimilarities, calculated separately on 16S and SMS data.

Table 4
Summary of popular dissimilarity measures.

Dissimilarity measure	Equation	Remarks
Bray-Curtis	$d_{kl} = \frac{\Sigma_{i=1}^{n}\lvert p_{ik} - p_{il}\rvert}{\Sigma_{i=1}^{n}\lvert p_{ik} + p_{il}\rvert}$	p_{ik} is the abundance of taxon i in sample k p_{il} is the abundance of taxon i in sample l
Jaccard	$d_{kl} = \frac{b_{kl}}{(1 + b_{kl})}$	b_{kl} is the Bray-Curtis dissimilarity between sample k and sample l.
unweighted Unifrac	$\frac{\Sigma_{i=1}^{n} b_i\lvert q_{ik} - q_{il}\rvert}{\Sigma_{i=1}^{n} b_i max(q_{ik}, q_{il})}$	q_{ik} is 1 if taxon i is present in sample k and 0 otherwise.
weighted Unifrac	$\sum_{i=1}^{n} b_i\lvert \frac{p_{ik}}{\Sigma_{i=1}^{n} p_{ik}} - \frac{p_{il}}{\Sigma_{i=1}^{n} p_{il}}\rvert$	b_i refers to branch i

5.6. Supervised learning

Supervised learning constitutes the second part of ML and is characterized by the notion of a dependent variable, which is absent in unsupervised learning. It is a variable that the researcher wants to predict for observations outside of their data set. Experiments targeted at the dependent variable open new possibilities that have not yet been absorbed by the microbiome community as evidenced by the relatively low number of experiments where they have been utilized. One technique that has its roots in supervised learning is forward feature selection, which was used by Gopalakrishnan et al. (2018) to select variables to the Cox regression for PFS. The procedure determined two significant variables included into the multivariate model – the abundance of *Faecalibacterium* and prior immunotherapy. In the same publication, the authors also calculated univariate receiver operating characteristic (ROC) area under curve (AUC) values for a number of variables, using the time-dependent ROCs (Van Der Maaten and Hinton, 2008). From this experiment, it turned out that the abundance of *Faecalibacterium* and Bacteroidales reached the highest AUC values.

In Peters et al., authors make use of yet another tool that stems from supervised learning - cross-validation (Peters et al., 2019). Elastic net regularized Cox regression was run within an extensive, 500x-repeated 10-fold cross-validation schema. Elastic net regularization linearly combines the L1 and L2 penalties (Zou and Hastie, 2005). The model included non-penalized covariates, like BMI (body mass index) or antibiotics. With that pipeline, the authors were able to determine a number of significant genera, species and functional pathways that are related to PFS.

Two supervised approaches were implemented in the publication by Wind et al. (2020). The first one assumed implementation of univariate logistic regression models to assess the difference in prevalence of taxa between responders and non-responders. The authors defined prevalence as a binary variable that assigned the value of one, if abundance of a given taxon was greater than zero for a given patient and zero if abundance equaled zero. The results were corrected for multiple tests and no taxon showed significant differentiation. However, this was not the case with the second approach - linear regression trained on relative abundances. The model was run to differentiate responders from non-responders in a univariate fashion, but one of the scenarios was also trained with multiple additional covariates (multivariate model). All of the models were zero-inflated (allowing for frequent observations of zeros) due to an excessive number of zeroes in the data set. Only the second, multivariate model yielded statistically significant outcomes, detecting 68 unique taxa related with response to the ICT.

A well-known ML algorithm – random forest – was implemented by Derosa et al. (2020). With its help, the authors attempted to find out what factors impact the composition of the microbiota the most. This question stems from the observation that in their cohort, a majority of the patients received one previous line of tyrosine kinase inhibitors prior to ICT. Thanks to the random forest analysis, they were able to determine the microbiota-modulating impact of tyrosine kinase inhibitor – axitinib – and antibiotics.

There are, however, publications where a more thorough ML analysis was carried out (Peng et al., 2020). The authors of this publication gathered information and sequenced stool samples from patients suffering from gastrointestinal cancers. 16S data was used twofold. In the first scenario, a number of various ML models was trained with hyperparameter optimization in a stratified, nested cross validation, in order to verify whether the response status can be faithfully predicted within that cohort. Multiple various algorithms were implemented: random forest, extra trees, Support Vector Machine, elastic-net and k-nearest neighbors. All of them achieved accuracy of more than 0.8 with the best surpassing 0.9. The second scenario assumed training of the same models (with hyperparameter optimization) and making predictions for 16S profiles, obtained from patients in two melanoma cohorts (Gopalakrishnan et al., 2018; Matson et al., 2018). The models in this scenario were trained on 90 genera common in all the three data sets. Both of these cohorts were composed of melanoma patients, so the authors were undertaking not only a cross-cohort, but also a cross-indication validation. The results of almost all of these experiments yielded poor results with accuracy of around 0.5 on the combined cohorts. One of the data sets (Gopalakrishnan et al., 2018) performed consistently better, but not to the level of yielding trustworthy results. This data set is unbalanced (29 R/11NR), so the authors decided on measuring the performance of their models using other metrics, as accuracy is not adequate in such cases. Instead, they used the ROC AUC metric to see that one of the models – elastic net – reached AUC of 0.78.

A similar, cross-cohort approach was also assumed by Limeta et al. (2020). Contrarily to Peng et al. (2020), the authors here were using SMS data in their ML experiments and used both taxonomic and functional features jointly. The features that the model was fed with were previously filtered, hence only the differentially abundant ones were retained. Unlike in Peng et al. (2020), this time, only melanoma cohorts were taken into consideration. The training set consisted of patients from three cohorts (Frankel et al., 2017; Gopalakrishnan et al., 2018; Matson et al., 2018) and the fourth cohort served as the validation one (Peters et al., 2019). Prior to the training, the R/NR status definition has been unified across the publications. A random forest of 100'000 trees was trained on the differentially abundant features and used to obtain predictions. The ROC AUC metric on the validation set that was not shown to the model during the training procedure, achieved the value of 0.60, which the authors describe as "modest yet nonrandom".

Every time more than one publication is looked at and the results are in any way transferred from one cohort to another, there exists an indispensable factor that needs to be dealt with - standardization of the dependent variable. Most likely, the dependent variable will assume the R/NR split, but the authors tend to differ in their definition of these groups (see Table 1), even within melanoma research. Some ascribe only CR and PR patients to responders (Matson et al., 2018). Others also included into this category patients which experienced disease stabilization for a given period of time, such as 3 (Wind et al., 2020) or 6 months (Gopalakrishnan et al., 2018). On the other end of the spectrum, in some studies, patients with SD as their BOR have also been considered as responders without any other requirements (Frankel et al., 2017). Finally, some authors resort only to survival analysis and do not provide any fixed labels (Peters et al., 2019) and a rule of assignment has to be devised (Limeta et al., 2020). Training the models on unaligned data may result in lower performance and a lack of generalizability.

6. Concluding remarks

Statistical tests are deeply rooted and universally applied in microbiota studies, including the analyses of its impact on ICT outcomes. Statistical tests are highly-regarded tools in science and their popularity in this domain cannot be treated as a surprise. Unfortunately, their application does not always yield consistent results across studies. For instance, in one publication, *Bacteroides eggerthii* was found to be positively correlated with the response (Wind et al., 2020), whereas in another the correlation was inverted (Matson et al., 2018). The authors of the papers also rarely identify an overlap between their findings and what has already been determined in other studies, which suggests that statistics may not be enough. The cohorts in the publications under review are relatively small, with the maximum of 100 samples (Routy et al., 2018) (not including the pooled cohorts (Limeta et al., 2020)). The statistical significance may therefore be difficult to establish.

ML provides a different approach to microbiota studies and, if applied properly, can lead to more accurate results, which include not only predictive models useful in diagnostics, but also more precise identification of biomarkers or modes of action. Careful use of cross validation offers the potential to validate the findings in-silico (e.g., on a different cohort), leading to early detection of weak models and findings, in turn, saving a lot of costly and tedious laboratory work.

Nonetheless, as data after acquisition might be updated, previously chosen ML models might lead to results inconsistency. Another caveat is connected to the need to manually choose a proper model based on its accuracy, which requires multiple training and testing on data. At the same time, when models are validated to have strong predictive power, one can claim the discovery with much more confidence.

So far, we know that the microbiome is not indifferent in its modulating role of the response to the ICT. Equipped with ML, we may one day be able to determine its exact mode of action.

Funding

This work was part of a project "BioForte Technology for in Silico Identification of Candidates for a New Microbiome-based Therapeutics and Diagnostics" cofunded by European Regional Development Fund (ERDF) as part of Smart Growth Operational Program 2014–2020 (POIR.01.01.01–00–0347/17) awarded to Ardigen S.A.

References

Aleman, F.D.D., Valenzano, D.R., 2019. Microbiome evolution during host aging. PLOS Pathog. 15, e1007727 https://doi.org/10.1371/journal.ppat.1007727.

Almeida, A., Mitchell, A.L., Boland, M., Forster, S.C., Gloor, G.B., Tarkowska, A., Lawley, T.D., Finn, R.D., 2019. A new genomic blueprint of the human gut microbiota. Nature 568, 499–504. https://doi.org/10.1038/s41586-019-0965-1.

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 26, 32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x.

Armitage, P., 1955. Tests for linear trends in proportions and frequencies. Biometrics 11, 375. https://doi.org/10.2307/3001775.

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., Fernandes, G.R., Tap, J., Bruls, T., Batto, J.-M., Bertalan, M., Borruel, N., Casellas, F., Fernandez, L., Gautier, L., Hansen, T., Hattori, M., Hayashi, T., Kleerebezem, M., Kurokawa, K., Leclerc, M., Levenez, F., Manichanh, C., Nielsen, H.B., Nielsen, T., Pons, N., Poulain, J., Qin, J., Sicheritz-Ponten, T., Tims, S., Torrents, D., Ugarte, E., Zoetendal, E.G., Wang, J., Guarner, F., Pedersen, O., de Vos, W.M., Brunak, S., Doré, J., Weissenbach, J., Ehrlich, S.D., Bork, P., 2011. Enterotypes of the human gut microbiome. Nature 473, 174–180. https://doi.org/10.1038/nature09944.

Baruch, E.N., Youngster, I., Ben-Betzalel, G., Ortenberg, R., Lahat, A., Katz, L., Adler, K., Dick-Necula, D., Raskin, S., Bloch, N., Rotin, D., Anafi, L., Avivi, C., Melnichenko, J., Steinberg-Silman, Y., Mamtani, R., Harati, H., Asher, N., Shapira-Frommer, R., Brosh-Nissimov, T., Eshet, Y., Ben-Simon, S., Ziv, O., Khan, M.A.W., Amit, M., Ajami, N.J., Barshack, I., Schachter, J., Wargo, J.A., Koren, O., Markel, G., Boursi, B., 2021. Fecal microbiota transplant promotes response in immunotherapy-refractory melanoma patients. Science 371, 602–609. https://doi.org/10.1126/science.abb5920.

Bäumler, A.J., Sperandio, V., 2016. Interactions between the microbiota and pathogenic bacteria in the gut. Nature 535, 85–93. https://doi.org/10.1038/nature18849.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B 57, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Berry, S.E., Valdes, A.M., Drew, D.A., Asnicar, F., Mazidi, M., Wolf, J., Capdevila, J., Hadjigeorgiou, G., Davies, R., Al Khatib, H., Bonnett, C., Ganesh, S., Bakker, E., Hart, D., Mangino, M., Merino, J., Linenberg, I., Wyatt, P., Ordovas, J.M., Gardner, C.D., Delahanty, L.M., Chan, A.T., Segata, N., Franks, P.W., Spector, T.D., 2020. Human postprandial responses to food and potential for precision nutrition. Nat. Med. 26, 964–973. https://doi.org/10.1038/s41591-020-0934-0.

Blanche, P., Dartigues, J.-F., Jacqmin-Gadda, H., 2013. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Stat. Med. 32, 5381–5397. https://doi.org/10.1002/sim.5958.

Bray, J.R., Curtis, J.T., 1957. An ordination of the upland forest communities of Southern Wisconsin. Ecol. Monogr. 27, 325–349. https://doi.org/10.2307/1942268.

Cammarota, G., Ianiro, G., Kelly, C.R., Mullish, B.H., Allegretti, J.R., Kassam, Z., Putignani, L., Fischer, M., Keller, J.J., Costello, S.P., Sokol, H., Kump, P., Satokari, R., Kahn, S.A., Kao, D., Arkkila, P., Kuijper, E.J., Vehreschild, M.J.G., Pintus, C., Lopetuso, L., Masucci, L., Scaldaferri, F., Terveer, E.M., Nieuwdorp, M., López-Sanromán, A., Kupcinskas, J., Hart, A., Tilg, H., Gasbarrini, A., 2019. International consensus conference on stool banking for faecal microbiota transplantation in clinical practice. Gut 68, 2111–2121. https://doi.org/10.1136/gutjnl-2019-319548.

Cammarota, G., Ianiro, G., Ahern, A., Carbone, C., Temko, A., Claesson, M.J., Gasbarrini, A., Tortora, G., 2020. Gut microbiome, big data and machine learning to promote precision medicine for cancer. Nat. Rev. Gastroenterol. Hepatol. 17, 635–648. https://doi.org/10.1038/s41575-020-0327-3.

Chen, J., King, E., Deek, R., Wei, Z., Yu, Y., Grill, Di, Ballman, K., 2018. An omnibus test for differential distribution analysis of microbiome sequencing data. Bioinformatics 34, 643–651. https://doi.org/10.1093/bioinformatics/btx650.

Chevereau, G., Dravecká, M., Batur, T., Guvenek, A., Ayhan, D.H., Toprak, E., Bollenbach, T., 2015. Quantifying the determinants of evolutionary dynamics leading to drug resistance. PLOS Biol. 13, e1002299 https://doi.org/10.1371/journal.pbio.1002299.

Claesson, M.J., O'Sullivan, O., Wang, Q., Nikkilä, J., Marchesi, J.R., Smidt, H., de Vos, W.M., Ross, R.P., O'Toole, P.W., 2009. Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. PLOS One 4, e6669. https://doi.org/10.1371/journal.pone.0006669.

Claesson, M.J., Clooney, A.G., O'Toole, P.W., 2017. A clinician's guide to microbiome analysis. Nat. Rev. Gastroenterol. Hepatol. 14, 585–595. https://doi.org/10.1038/nrgastro.2017.97.

Clark, T.G., Bradburn, M.J., Love, S.B., Altman, D.G., 2003. Survival analysis part I: basic concepts and first analyses. Br. J. Cancer 89, 232–238. https://doi.org/10.1038/sj.bjc.6601118.

CLARKE, K.R., 1993. Non-parametric multivariate analyses of changes in community structure. Austral. Ecol. 18, 117–143. https://doi.org/10.1111/j.1442-9993.1993.tb00438.x.

Conover, W.J., 1972. A Kolmogorov goodness-of-fit test for discontinuous distributions. J. Am. Stat. Assoc. 67, 591. https://doi.org/10.2307/2284444.

Consortium, T.H.M.P., 2012. Structure, function and diversity of the healthy human microbiome. Nature 486, 207–214. https://doi.org/10.1038/nature11234.

Cooper, K., Tappenden, P., Cantrell, A., Ennis, K., 2020. A systematic review of meta-analyses assessing the validity of tumour response endpoints as surrogates for progression-free or overall survival in cancer. Br. J. Cancer 123, 1686–1696. https://doi.org/10.1038/s41416-020-01050-w.

Cox, D.R., 1972. Regression models and life-tables. J. R. Stat. Soc. Ser. B.

Cvetkovic, L., Régis, C., Richard, C., Derosa, L., Leblond, A., Malo, J., Messaoudene, M., Desilets, A., Belkaid, W., Elkrief, A., Routy, B., Juneau, D., 2021. Physiologic colonic uptake of 18F-FDG on PET/CT is associated with clinical response and gut microbiome composition in patients with advanced non-small cell lung cancer treated with immune checkpoint inhibitors. Eur. J. Nucl. Med. Mol. Imaging 48, 1550–1559. https://doi.org/10.1007/s00259-020-05081-6.

Davar, D., Dzutsev, A.K., McCulloch, J.A., Rodrigues, R.R., Chauvin, J.-M., Morrison, R.M., Deblasio, R.N., Menna, C., Ding, Q., Pagliano, O., Zidi, B., Zhang, S., Badger, J.H., Vetizou, M., Cole, A.M., Fernandes, M.R., Prescott, S., Costa, R.G.F., Balaji, A.K., Morgun, A., Vujkovic-Cvijin, I., Wang, H., Borhani, A.A., Schwartz, M.B., Dubner, H.M., Ernst, S.J., Rose, A., Najjar, Y.G., Belkaid, Y., Kirkwood, J.M., Trinchieri, G., Zarour, H.M., 2021. Fecal microbiota transplant overcomes resistance to anti–PD-1 therapy in melanoma patients. Science 371, 595–602. https://doi.org/10.1126/science.abf3363.

David, L.A., Materna, A.C., Friedman, J., Campos-Baptista, M.I., Blackburn, M.C., Perrotta, A., Erdman, S.E., Alm, E.J., 2014. Host lifestyle affects human microbiota on daily timescales. Genome Biol. 15, 1–15. https://doi.org/10.1186/gb-2014-15-7-r89.

Derosa, L., Routy, B., Fidelle, M., Iebba, V., Alla, L., Pasolli, E., Segata, N., Desnoyer, A., Pietrantonio, F., Ferrere, G., Fahrner, J.-E., Le Chatellier, E., Pons, N., Galleron, N., Roume, H., Duong, C.P.M., Mondragón, L., Iribarren, K., Bonvalet, M., Terrisse, S., Rauber, C., Goubet, A.-G., Daillère, R., Lemaitre, F., Reni, A., Casu, B., Alou, M.T., Alves Costa Silva, C., Raoult, D., Fizazi, K., Escudier, B., Kroemer, G., Albiges, L., Zitvogel, L., 2020. Gut bacteria composition drives primary resistance to cancer immunotherapy in renal cell carcinoma patients. Eur. Urol. 78, 195–206. https://doi.org/10.1016/j.eururo.2020.04.044.

Deschasaux, M., Bouter, K.E., Prodan, A., Levin, E., Groen, A.K., Herrema, H., Tremaroli, V., Bakker, G.J., Attaye, I., Pinto-Sietsma, S.-J., van Raalte, D.H., Snijder, M.B., Nicolaou, M., Peters, R., Zwinderman, A.H., Bäckhed, F., Nieuwdorp, M., 2018. Depicting the composition of gut microbiota in a population with varied ethnic origins but shared geography. Nat. Med. 24, 1526–1531. https://doi.org/10.1038/s41591-018-0160-1.

Dogra, S.K., Doré, J., Damak, S., 2020. Gut microbiota resilience: definition, link to health and strategies for intervention. Front. Microbiol. 11, 2245. https://doi.org/10.3389/fmicb.2020.572921.

Dominguez-Bello, M.G., Godoy-Vitorino, F., Knight, R., Blaser, M.J., 2019. Role of the microbiome in human development. Gut 68, 1108–1114. https://doi.org/10.1136/gutjnl-2018-317503.

Dong, T.S., Gupta, A., 2019. Influence of early life, diet, and the environment on the microbiome. Clin. Gastroenterol. Hepatol. https://doi.org/10.1016/j.cgh.2018.08.067.

Eisenhauer, E.A., Therasse, P., Bogaerts, J., Schwartz, L.H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., Rubinstein, L., Shankar, L., Dodd, L., Kaplan, R., Lacombe, D., Verweij, J., Biologicals, G., 2009. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur. J. Cancer 45, 228–247. https://doi.org/10.1016/j.ejca.2008.10.026.

Elkrief, A., Derosa, L., Zitvogel, L., Kroemer, G., Routy, B., 2019. The intimate relationship between gut microbiota and cancer immunotherapy. Gut Microbes 10, 424–428. https://doi.org/10.1080/19490976.2018.1527167.

Falony, G., Joossens, M., Vieira-Silva, S., Wang, J., Darzi, Y., Faust, K., Kurilshikov, A., Bonder, M.J., Valles-Colomer, M., Vandeputte, D., Tito, R.Y., Chaffron, S., Rymenans, L., Verspecht, C., De Sutter, L., Lima-Mendez, G., D'hoe, K., Jonckheere, K., Homola, D., Garcia, R., Tigchelaar, E.F., Eeckhaudt, L., Fu, J., Henckaerts, L., Zhernakova, A., Wijmenga, C., Raes, J., 2016. Population-level analysis of gut microbiome variation. Science 352, 560–564. https://doi.org/10.1126/science.aad3503.

Fisher, R.A., 1922. On the interpretation of χ 2 from contingency tables, and the calculation of P. J. R. Stat. Soc. 85, 87. https://doi.org/10.2307/2340521.

Frankel, A.E., Coughlin, L.A., Kim, J., Froehlich, T.W., Xie, Y., Frenkel, E.P., Koh, A.Y., 2017. Metagenomic shotgun sequencing and unbiased metabolomic profiling identify specific human gut microbiota and metabolites associated with immune

checkpoint therapy efficacy in melanoma patients. Neoplasia 19, 848–855. https://doi.org/10.1016/j.neo.2017.08.004.
Garrett, W.S., 2015. Cancer and the microbiota. Science 348, 80–86. https://doi.org/10.1126/science.aaa4972.
Gensollen, T., Iyer, S.S., Kasper, D.L., Blumberg, R.S., 2016. How colonization by microbiota in early life shapes the immune system. Science 352, 539–544. https://doi.org/10.1126/science.aad9378.
Gerber, G.K., 2014. The dynamic microbiome. FEBS Lett. 588, 4131–4139. https://doi.org/10.1016/j.febslet.2014.02.037.
Gilbert, J.A., Blaser, M.J., Caporaso, J.G., Jansson, J.K., Lynch, S.V., Knight, R., 2018. Current understanding of the human microbiome. Nat. Med. 24, 392–400. https://doi.org/10.1038/nm.4517.
Gopalakrishnan, V., Spencer, C.N., Nezi, L., Reuben, A., Andrews, M.C., Karpinets, T.V., Prieto, P.A., Vicente, D., Hoffman, K., Wei, S.C., Cogdill, A.P., Zhao, L., Hudgens, C. W., Hutchinson, D.S., Manzo, T., Petaccia de Macedo, M., Cotechini, T., Kumar, T., Chen, W.S., Reddy, S.M., Szczepaniak Sloane, R., Galloway-Pena, J., Jiang, H., Chen, P.L., Shpall, E.J., Rezvani, K., Alousi, A.M., Chemaly, R.F., Shelburne, S., Vence, L.M., Okhuysen, P.C., Jensen, V.B., Swennes, A.G., McAllister, F., Marcelo Riquelme Sanchez, E., Zhang, Y., Le Chatelier, E., Zitvogel, L., Pons, N., Austin-Breneman, J.L., Haydu, L.E., Burton, E.M., Gardner, J.M., Sirmans, E., Hu, J., Lazar, A.J., Tsujikawa, T., Diab, A., Tawbi, H., Glitza, I.C., Hwu, W.J., Patel, S.P., Woodman, S.E., Amaria, R.N., Davies, M.A., Gershenwald, J.E., Hwu, P., Lee, J.E., Zhang, J., Coussens, L.M., Cooper, Z.A., Futreal, P.A., Daniel, C.R., Ajami, N.J., Petrosino, J.F., Tetzlaff, M.T., Sharma, P., Allison, J.P., Jenq, R.R., Wargo, J.A., 2018. Gut microbiome modulates response to anti–PD-1 immunotherapy in melanoma patients. Science 359, 97–103. https://doi.org/10.1126/science.aan4236.
Gupta, V.K., Paul, S., Dutta, C., 2017. Geography, ethnicity or subsistence-specific variations in human microbiome composition and diversity. Front. Microbiol. 8, 1162. https://doi.org/10.3389/fmicb.2017.01162.
Hamada, T., Kosumi, K., Nakai, Koike, K., 2018. Surrogate study endpoints in the era of cancer immunotherapy. Ann. Transl. Med. S27 https://doi.org/10.21037/atm.2018.09.31.
Hamid, O., Robert, C., Daud, A., Hodi, F.S., Hwu, W.J., Kefford, R., Wolchok, J.D., Hersey, P., Joseph, R., Weber, J.S., Dronca, R., Mitchell, T.C., Patnaik, A., Zarour, H. M., Joshua, A.M., Zhao, Q., Jensen, E., Ahsan, S., Ibrahim, N., Ribas, A., 2019. Five-year survival outcomes for patients with advanced melanoma treated with pembrolizumab in KEYNOTE-001. Ann. Oncol. 30, 582–588. https://doi.org/10.1093/annonc/mdz011.
He, Y., Wu, W., Zheng, H.M., Li, P., McDonald, D., Sheng, H.F., Chen, M.X., Chen, Z.H., Ji, G.Y., Zheng, Z.D.X., Mujagond, P., Chen, X.J., Rong, Z.H., Chen, P., Lyu, L.Y., Wang, X., Wu, C.Bin, Yu, N., Xu, Y.J., Yin, J., Raes, J., Knight, R., Ma, W.J., Zhou, H. W., 2018. Regional variation limits applications of healthy gut microbiome reference ranges and disease models. Nat. Med. 24, 1532–1535. https://doi.org/10.1038/s41591-018-0164-x.
Helmink, B.A., Khan, M.A.W., Hermann, A., Gopalakrishnan, V., Wargo, J.A., 2019. The microbiome, cancer, and cancer therapy. Nat. Med. 25, 377–388. https://doi.org/10.1038/s41591-019-0377-7.
Hill, J.H., Round, J.L., 2021. SnapShot: microbiota effects on host physiology. Cell 184. https://doi.org/10.1016/j.cell.2021.04.026.
Iida, N., Dzutsev, A., Stewart, C.A., Smith, L., Bouladoux, N., Weingarten, R.A., Molina, D.A., Salcedo, R., Back, T., Cramer, S., Dai, R.-M., Kiu, H., Cardone, M., Naik, S., Patri, A.K., Wang, E., Marincola, F.M., Frank, K.M., Belkaid, Y., Trinchieri, G., Goldszmid, R.S., 2013. Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. Science 342, 967–970. https://doi.org/10.1126/science.1240527.
Jaccard, P., 1901. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. Bull. Del. la Société Vaud. Des. Sci. Nat. 37, 547–579.
Kaplan, E.L., Meier, P., 1958. Nonparametric estimation from incomplete observations. J. Am. Stat. Assoc. 53, 457. https://doi.org/10.2307/2281868.
Keebaugh, E.S., Ja, W.W., 2017. Breaking down walls: microbiota and the aging gut. Cell Host Microbe 21, 417–418. https://doi.org/10.1016/j.chom.2017.03.013.
Koh, H., Livanos, A.E., Blaser, M.J., Li, H., 2018. A highly adaptive microbiome-based association test for survival traits. BMC Genom. 19, 210. https://doi.org/10.1186/s12864-018-4599-8.
Kruskal, J.B., 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika 29, 1–27. https://doi.org/10.1007/BF02289565.
Kruskal, W.H., Wallis, W.A., 1952. Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. 47, 583–621. https://doi.org/10.1080/01621459.1952.10483441.
Kurilshikov, A., Wijmenga, C., Fu, J., Zhernakova, A., 2017. Host genetics and gut microbiome: challenges and perspectives. Trends Immunol. 38, 633–647. https://doi.org/10.1016/j.it.2017.06.003.
Kurilshikov, A., Medina-Gomez, C., Bacigalupe, R., Radjabzadeh, D., Wang, J., Demirkan, A., Le Roy, C.I., Raygoza Garay, J.A., Finnicum, C.T., Liu, X., Zhernakova, D.V., Bonder, M.J., Hansen, T.H., Frost, F., Rühlemann, M.C., Turpin, W., Moon, J.-Y., Kim, H.-N., Lüll, K., Barkan, E., Shah, S.A., Fornage, M., Szopinska-Tokov, J., Wallen, Z.D., Borisevich, D., Agreus, L., Andreasson, A., Bang, C., Bedrani, L., Bell, J.T., Bisgaard, H., Boehnke, M., Boomsma, D.I., Burk, R. D., Claringbould, A., Croitoru, K., Davies, G.E., van Duijn, C.M., Duijts, L., Falony, G., Fu, J., van der Graaf, A., Hansen, T., Homuth, G., Hughes, D.A., Ijzerman, R.G., Jackson, M.A., Jaddoe, V.W.V., Joossens, M., Jørgensen, T., Keszthelyi, D., Knight, R., Laakso, M., Laudes, M., Launer, L.J., Lieb, W., Lusis, A.J., Masclee, A.A.M., Moll, H.A., Mujagic, Z., Qibin, Q., Rothschild, D., Shin, H., Sørensen, S.J., Steves, C.J., Thorsen, J., Timpson, N.J., Tito, R.Y., Vieira-Silva, S., Völker, U., Völzke, H., Võsa, U., Wade, K.H., Walter, S., Watanabe, K., Weiss, S., Weiss, F.U., Weissbrod, O., Westra, H.-J., Willemsen, G., Payami, H., Jonkers, D.M.A. E., Arias Vasquez, A., de Geus, E.J.C., Meyer, K.A., Stokholm, J., Segal, E., Org, E., Wijmenga, C., Kim, H.-L., Kaplan, R.C., Spector, T.D., Uitterlinden, A.G., Rivadeneira, F., Franke, A., Lerch, M.M., Franke, L., Sanna, S., D'Amato, M., Pedersen, O., Paterson, A.D., Kraaij, R., Raes, J., Zhernakova, A., 2021. Large-scale association analyses identify host factors influencing human gut microbiome composition. Nat. Genet. 53, 156–165. https://doi.org/10.1038/s41588-020-00763-1.
Levy, M., Blacher, E., Elinav, E., 2017a. Microbiome, metabolites and host immunity. Curr. Opin. Microbiol. 35, 8–15. https://doi.org/10.1016/j.mib.2016.10.003.
Levy, M., Kolodziejczyk, A.A., Thaiss, C.A., Elinav, E., 2017b. Dysbiosis and the immune system. Nat. Rev. Immunol. 17, 219–232. https://doi.org/10.1038/nri.2017.7.
Limeta, A., Ji, B., Levin, M., Gatto, F., Nielsen, J., 2020. Meta-analysis of the gut microbiota in predicting response to cancer immunotherapy in metastatic melanoma. JCI Insight 5. https://doi.org/10.1172/jci.insight.140940.
Lozupone, C.A., Hamady, M., Kelley, S.T., Knight, R., 2007. Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial communities. Appl. Environ. Microbiol. 73, 1576–1585. https://doi.org/10.1128/AEM.01996-06.
Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., Knight, R., 2012. Diversity, stability and resilience of the human gut microbiota. Nature 489, 220–230. https://doi.org/10.1038/nature11550.
Ma, Y., Liu, J., Rhodes, C., Nie, Y., Zhang, F., 2017. Ethical issues in fecal microbiota transplantation in practice. Am. J. Bioeth. 17, 34–45. https://doi.org/10.1080/15265161.2017.1299240.
Malard, F., Dore, J., Gaugler, B., Mohty, M., 2020. Introduction to host microbiome symbiosis in health and disease. Mucosal Immunol. 14, 547–554. https://doi.org/10.1038/s41385-020-00365-4.
Mann, H.B., Whitney, D.R., 1947. On a test of whether one of two random variables is stochastically larger than the other. Ann. Math. Stat. 18, 50–60. https://doi.org/10.1214/aoms/1177730491.
Mantel, N., 1966. Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemother. Rep. 50, 163–170.
Matson, V., Fessler, J., Bao, R., Chongsuwat, T., Zha, Y., Alegre, M.-L., Luke, J.J., Gajewski, T.F., 2018. The commensal microbiome is associated with anti–PD-1 efficacy in metastatic melanoma patients. Science 359, 104–108. https://doi.org/10.1126/science.aao3290.
McMurdie, P.J., Holmes, S., 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLOS One 8. https://doi.org/10.1371/JOURNAL.PONE.0061217.
Mehta, R.S., Abu-Ali, G.S., Drew, D.A., Lloyd-Price, J., Subramanian, A., Lochhead, P., Joshi, A.D., Ivey, K.L., Khalili, H., Brown, G.T., DuLong, C., Song, M., Nguyen, L.H., Mallick, H., Rimm, E.B., Izard, J., Huttenhower, C., Chan, A.T., 2018. Stability of the human faecal microbiome in a cohort of adult men. Nat. Microbiol. 3, 347–355. https://doi.org/10.1038/s41564-017-0096-0.
Mohajeri, M.H., Brummer, R.J.M., Rastall, R.A., Weersma, R.K., Harmsen, H.J.M., Faas, M., Eggersdorfer, M., 2018. The role of the microbiome for human health: from basic science to clinical applications. Eur. J. Nutr. 57, 1–14. https://doi.org/10.1007/s00394-018-1703-4.
Namkung, J., 2020. Machine learning methods for microbiome studies. J. Microbiol. 583 (58), 206–216. https://doi.org/10.1007/S12275-020-0066-8.
Odrzywołek, Krzysztof, Karwowska, Zuzanna, Majta, Jan, Byrski, Aleksander, Milanowska-Zabel, Kaja, Kościółek, Tomasz, 2022. Deep embeddings to comprehend and visualize microbiome protein space. Scientific Reports 12, 10332. https://doi.org/10.1038/s41598-022-14055-7.
Park, S.-Y., Seo, G.S., 2021. Fecal microbiota transplantation: is it safe. Clin. Endosc. 54, 157–160. https://doi.org/10.5946/ce.2021.072.
Peng, Z., Cheng, S., Kou, Y., Wang, Z., Jin, R., Hu, H., Zhang, X., Gong, J., Li, J., Lu, M., Wang, X., Zhou, J., Lu, Z., Zhang, Q., Tzeng, D.T.W., Bi, D., Tan, Y., Shen, L., 2020. The gut microbiome is associated with clinical response to Anti–PD-1/PD-L1 immunotherapy in gastrointestinal cancer. Cancer Immunol. Res. 8, 1251–1261. https://doi.org/10.1158/2326-6066.CIR-19-1014.
Peters, B.A., Wilson, M., Moran, U., Pavlick, A., Izsak, A., Wechter, T., Weber, J.S., Osman, I., Ahn, J., 2019. Relating the gut metagenome and metatranscriptome to immunotherapy responses in melanoma patients. Genome Med. 11, 61. https://doi.org/10.1186/s13073-019-0672-4.
Plantinga, A., Zhan, X., Zhao, N., Chen, J., Jenq, R.R., Wu, M.C., 2017. MiRKAT-S: a community-level test of association between the microbiota and survival times. Microbiome 5, 17. https://doi.org/10.1186/s40168-017-0239-9.
Reimer, R.A., 2019. Establishing the role of diet in the microbiota–disease axis. Nat. Rev. Gastroenterol. Hepatol. 16, 86–87. https://doi.org/10.1038/s41575-018-0093-7.
Rook, G., Bäckhed, F., Levin, B.R., McFall-Ngai, M.J., McLean, A.R., 2017. Evolution, human-microbe interactions, and life history plasticity. Lancet 390, 521–530. https://doi.org/10.1016/S0140-6736(17)30566-4.
Rosenberg, E., Zilber-Rosenberg, I., 2018. The hologenome concept of evolution after 10 years. Microbiome 6, 78. https://doi.org/10.1186/s40168-018-0457-9.
Routy, B., Le Chatelier, E., Derosa, L., Duong, C.P.M.M., Alou, M.T., Daillère, R., Fluckiger, A., Messaoudene, M., Rauber, C., Roberti, M.P., Fidelle, M., Flament, C., Poirier-Colame, V., Opolon, P., Klein, C., Iribarren, K., Mondragón, L., Jacquelot, N., Qu, B., Ferrere, G., Clémenson, C., Mezquita, L., Masip, J.R., Naltet, C., Brosseau, S., Kaderbhai, C., Richard, C., Rizvi, H., Levenez, F., Galleron, N., Quinquis, B., Pons, N., Ryffel, B., Minard-Colin, V., Gonin, P., Soria, J.-C., Deutsch, E., Loriot, Y., Ghiringhelli, F., Zalcman, G., Goldwasser, F., Escudier, B., Hellmann, M.D., Eggermont, A., Raoult, D., Albiges, L., Kroemer, G., Zitvogel, L., 2018. Gut microbiome influences efficacy of PD-1–based immunotherapy against epithelial tumors. Science 359, 91–97. https://doi.org/10.1126/science.aan3706.

Ruff, W.E., Greiling, T.M., Kriegel, M.A., 2020. Host–microbiota interactions in immune-mediated diseases. Nat. Rev. Microbiol. 18, 521–538. https://doi.org/10.1038/s41579-020-0367-2.

Salgia, N.J., Bergerot, P.G., Maia, M.C., Dizman, N., Hsu, J., Gillece, J.D., Folkerts, M., Reining, L., Trent, J., Highlander, S.K., Pal, S.K., 2020. Stool microbiome profiling of patients with metastatic renal cell carcinoma receiving anti–PD-1 immune checkpoint inhibitors. Eur. Urol. 78, 498–502. https://doi.org/10.1016/j.eururo.2020.07.011.

Scepanovic, P., Hodel, F., Mondot, S., Partula, V., Byrd, A., Hammer, C., Alanio, C., Bergstedt, J., Patin, E., Touvier, M., Lantz, O., Albert, M.L., Duffy, D., Quintana-Murci, L., Fellay, J., 2019. A comprehensive assessment of demographic, environmental, and host genetic associations with gut microbiome diversity in healthy individuals. Microbiome 7, 130. https://doi.org/10.1186/s40168-019-0747-x.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., Huttenhower, C., 2011. Metagenomic biomarker discovery and explanation. Genome Biol. 12, 1–18. https://doi.org/10.1186/gb-2011-12-6-r60.

Sepich-Poore, G.D., Zitvogel, L., Straussman, R., Hasty, J., Wargo, J.A., Knight, R., 2021. The microbiome and human cancer. Science 371, eabc4552. https://doi.org/10.1126/science.abc4552.

Sundh, J., Tanash, H., Arian, R., Neves-Guimaraes, A., Broberg, K., Lindved, G., Kern, T., Zych, K., Nielsen, H.B., Halling, A., Ohlsson, B., Jönsson, D., 2021. Advanced dental cleaning is associated with reduced risk of COPD exacerbations - a randomized controlled trial. Int. J. Chron. Obstruct. Pulmon. Dis 16, 3203–3215. https://doi.org/10.2147/COPD.S327036.

Valdes, A.M., Walter, J., Segal, E., Spector, T.D., 2018. Role of the gut microbiota in nutrition and health. BMJ 361, 36–44. https://doi.org/10.1136/bmj.k2179.

Van Der Maaten, L., Hinton, G., 2008. Visualizing data using t-SNE. J. Mach. Learn. Res.

Vetizou, M., Pitt, J.M., Daillere, R., Lepage, P., Waldschmitt, N., Flament, C., Rusakiewicz, S., Routy, B., Roberti, M.P., Duong, C.P.M., Poirier-Colame, V., Roux, A., Becharef, S., Formenti, S., Golden, E., Cording, S., Eberl, G., Schlitzer, A., Ginhoux, F., Mani, S., Yamazaki, T., Jacquelot, N., Enot, D.P., Berard, M., Nigou, J., Opolon, P., Eggermont, A., Woerther, P.-L., Chachaty, E., Chaput, N., Robert, C., Mateus, C., Kroemer, G., Raoult, D., Boneca, I.G., Carbonnel, F., Chamaillard, M., Zitvogel, L., 2015. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. Science 350, 1079–1084. https://doi.org/10.1126/science.aad1329.

Viaud, S., Saccheri, F., Mignot, G., Yamazaki, T., Daillere, R., Hannani, D., Enot, D.P., Pfirschke, C., Engblom, C., Pittet, M.J., Schlitzer, A., Ginhoux, F., Apetoh, L., Chachaty, E., Woerther, P.-L., Eberl, G., Berard, M., Ecobichon, C., Clermont, D., Bizet, C., Gaboriau-Routhiau, V., Cerf-Bensussan, N., Opolon, P., Yessaad, N., Vivier, E., Ryffel, B., Elson, C.O., Dore, J., Kroemer, G., Lepage, P., Boneca, I.G., Ghiringhelli, F., Zitvogel, L., 2013. The intestinal microbiota modulates the anticancer immune effects of cyclophosphamide. Science 342, 971–976. https://doi.org/10.1126/science.1240537.

Vivarelli, S., Salemi, R., Candido, S., Falzone, L., Santagati, M., Stefani, S., Torino, F., Banna, G.L., Tonini, G., Libra, M., 2019. Gut microbiota and cancer: from pathogenesis to therapy. Cancers 11, 38. https://doi.org/10.3390/cancers11010038.

Wald, A., 1943. Tests of statistical hypotheses concerning several parameters when the number of observations is large. Trans. Am. Math. Soc. 54, 426. https://doi.org/10.2307/1990256.

Weersma, R.K., Zhernakova, A., Fu, J., 2020. Interaction between drugs and the gut microbiome. Gut 69, 1510–1519. https://doi.org/10.1136/gutjnl-2019-320204.

Welch, B.L., 1951. On the comparison of several mean values: an alternative approach. Biometrika 38, 330. https://doi.org/10.2307/2332579.

Whittaker, R.H., 1960. Vegetation of the Siskiyou mountains, Oregon and California. Ecol. Monogr. 30, 279–338. https://doi.org/10.2307/1943563.

Wind, T.T., Gacesa, R., Vich Vila, A., de Haan, J.J., Jalving, M., Weersma, R.K., Hospers, G.A.P., 2020. Gut microbial species and metabolic pathways associated with response to treatment with immune checkpoint inhibitors in metastatic melanoma. Melanoma Res. 30, 235–246. https://doi.org/10.1097/CMR.0000000000000656.

Wirbel, J., Zych, K., Essex, M., Karcher, N., Kartal, E., Salazar, G., Bork, P., Sunagawa, S., Zeller, G., 2021. Microbiome meta-analysis and cross-disease comparison enabled by the SIAMCAT machine learning toolbox. Genome Biol. 22 https://doi.org/10.1186/S13059-021-02306-1.

Wolchok, J.D., Chiarion-Sileni, V., Gonzalez, R., Grob, J.J., Rutkowski, P., Lao, C.D., Cowey, C.L., Schadendorf, D., Wagstaff, J., Dummer, R., Ferrucci, P.F., Smylie, M., Butler, M.O., Hill, A., Márquez-Rodas, I., Haanen, J.B.A.G., Guidoboni, M., Maio, M., Schöffski, P., Hodi, F.S., 2022. Long-Term Outcomes With Nivolumab Plus Ipilimumab or Nivolumab Alone Versus Ipilimumab in Patients With Advanced Melanoma. J. Clin. Oncol. 40, 127–137. https://doi.org/10.1200/JCO.21.02229.

Woods, L.C., Gorrell, R.J., Taylor, F., Connallon, T., Kwok, T., McDonald, M.J., 2020. Horizontal gene transfer potentiates adaptation by reducing selective constraints on the spread of genetic variation. Proc. Natl. Acad. Sci. USA 117, 26868–26875. https://doi.org/10.1073/pnas.2005331117.

Yong, D., Toleman, M.A., Giske, C.G., Cho, H.S., Sundman, K., Lee, K., Walsh, T.R., 2009. Characterization of a new metallo-β-lactamase gene, bla NDM-1, and a novel erythromycin esterase gene carried on a unique genetic structure in Klebsiella pneumoniae sequence Type 14 from India. Antimicrob. Agents Chemother. 53, 5046–5054. https://doi.org/10.1128/AAC.00774-09.

Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., Ben-Yacov, O., Lador, D., Avnit-Sagi, T., Lotan-Pompan, M., Suez, J., Mahdi, J.A., Matot, E., Malka, G., Kosower, N., Rein, M., Zilberman-Schapira, G., Dohnalová, L., Pevsner-Fischer, M., Bikovsky, R., Halpern, Z., Elinav, E., Segal, E., 2015. Personalized nutrition by prediction of glycemic responses. Cell 163, 1079–1094. https://doi.org/10.1016/j.cell.2015.11.001.

Zheng, D., Liwinski, T., Elinav, E., 2020. Interaction between microbiota and immunity in health and disease. Cell Res. 30, 492–506. https://doi.org/10.1038/s41422-020-0332-7.

Zhernakova, A., Kurilshikov, A., Bonder, M.J., Tigchelaar, E.F., Schirmer, M., Vatanen, T., Mujagic, Z., Vila, A.V., Falony, G., Vieira-Silva, S., Wang, J., Imhann, F., Brandsma, E., Jankipersadsing, S.A., Joossens, M., Cenit, M.C., Deelen, P., Swertz, M. A., Weersma, R.K., Feskens, E.J.M., Netea, M.G., Gevers, D., Jonkers, D., Franke, L., Aulchenko, Y.S., Huttenhower, C., Raes, J., Hofker, M.H., Xavier, R.J., Wijmenga, C., Fu, J., 2016. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. Science 352, 565–569. https://doi.org/10.1126/science.aad3369.

Zipkin, M., 2021. Fecal microbiota potentiate checkpoint inhibitors, unleash microbiome startups. Nat. Biotechnol. 39, 529–532. https://doi.org/10.1038/d41587-021-00002-w.

Zitvogel, L., Ma, Y., Raoult, D., Kroemer, G., Gajewski, T.F., 2018. The microbiome in cancer immunotherapy: diagnostic tools and therapeutic strategies. Science 359, 1366–1370. https://doi.org/10.1126/science.aar6918.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B (Stat. Methodol. 67, 301–320. https://doi.org/10.1111/J.1467-9868.2005.00503.X.