Article

Extension of PERMANOVA to Testing the Mediation Effect of the Microbiome

Ye Yue and Yi-Juan Hu *

Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA 30322, USA; ye.yue@emory.edu
* Correspondence: yijuan.hu@emory.edu

Abstract: Recently, we have seen a growing volume of evidence linking the microbiome and human diseases or clinical outcomes, as well as evidence linking the microbiome and environmental exposures. Now comes the time to assess whether the microbiome mediates the effects of exposures on the outcomes, which will enable researchers to develop interventions to modulate outcomes by modifying microbiome compositions. Use of distance matrices is a popular approach to analyzing complex microbiome data that are high-dimensional, sparse, and compositional. However, the existing distance-based methods for mediation analysis of microbiome data, MedTest and MODIMA, only work well in limited scenarios. PERMANOVA is currently the most commonly used distance-based method for testing microbiome associations. Using the idea of inverse regression, here we extend PERMANOVA to test microbiome-mediation effects by including both the exposure and the outcome as covariates and basing the test on the product of their F statistics. This extension of PERMANOVA, which we call PERMANOVA-med, naturally inherits all the flexible features of PERMANOVA, e.g., allowing adjustment of confounders, accommodating continuous, binary, and multivariate exposure and outcome variables including survival outcomes, and providing an omnibus test that combines the results from analyzing multiple distance matrices. Our extensive simulations indicated that PERMANOVA-med always controlled the type I error and had compelling power over MedTest and MODIMA. Frequently, MedTest had diminished power and MODIMA had inflated type I error. Using real data on melanoma immunotherapy response, we demonstrated the wide applicability of PERMANOVA-med through 16 different mediation analyses, only 6 of which could be performed by MedTest and 4 by MODIMA.

Keywords: inverse regression; global test; community-level test; microbial community; distance matrix

1. Introduction

Microbiome research has proliferated in the last decade due to booming interests in the scientific community, increasing power of high-throughput sequencing, and rapid advancement of data analytics. To date, most microbiome studies have been focused on bivariate associations between the microbiome and the covariates of interest. We have seen a rapidly growing volume of evidence linking the microbiome and human diseases such as obesity [1] or clinical outcomes such as responses to immunotherapy [2]. Similarly, we have seen the relationship between the microbiome and environmental exposures such as diet [3]. Meanwhile, many of these environmental exposures have well-established effects on clinical outcomes. We believe the time has come to assess whether the microbiome plays a mediating role between exposures and outcomes, as depicted in Figure 1a. Identifying such a mediating role of the microbiome enables researchers to develop interventions to modulate the outcomes by modifying the microbiome composition. Accordingly, there is an urgent need for statistical methods that are designed specifically for mediation analysis of microbiome data.
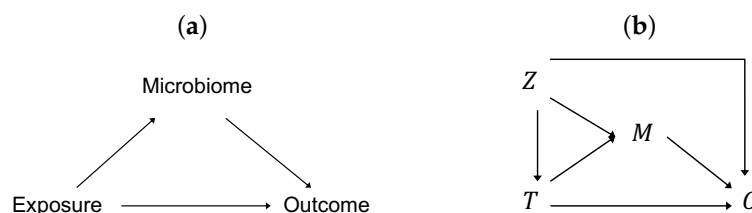
Figure 1. (**a**) Some effect of an exposure on an outcome is mediated through the microbiome. (**b**) *T* denotes the exposure, *M* the microbiome (mediator), *O* the outcome, and *Z* the confounding covariates.

There are a number of special challenges in mediation analysis of microbiome data. Microbiome read count data from 16S amplicon or metagenomic sequencing are typically summarized in a taxa count table, and have unique and complex features. They are high-dimensional (with typically many more taxa than samples), sparse (having 50–90% zero counts), compositional (measuring relative abundances that sum to one), and highly overdispersed. In addition, microbiome studies can be conducted in various observational or clinical settings, and tend to have diverse attributes. The exposure and outcome variables may be continuous, binary, or even multivariate (comprising multiple components such as multiple indicators for a categorical variable). In particular, many clinical outcomes are in the form of times to event (survival times) with possibly censored values. There generally exist confounders (e.g., sex and antibiotic use), as a microbial community is easily modifiable. Finally, the sample sizes are usually small (e.g., 50–100) and the study designs can be complex (e.g., clustered samples [4] or matched sets [5]).

To circumvent the complexities of microbiome count data, a popular approach is to first summarize the taxon-level data into a distance (dissimilarity) matrix that measures the pairwise dissimilarity in the microbiome profiles, and then base the analysis of microbiome data on the distance matrix [6–10]. This approach provides results at the community level, which is usually the first step in an analytical pipeline. Numerous distance measures, with different properties, have been proposed to detect diverse patterns in microbiome data; the most commonly used ones include Jaccard [11], Bray–Curtis [12], and weighted and unweighted UniFrac [13,14]. It is well acknowledged that the optimal choice of a distance measure depends on the underlying variation pattern in a particular dataset, which is unknown a priori. Therefore, it is a common practice to construct an omnibus test that combines the results from analyzing different distance matrices.

Two existing methods, MedTest [15] and MODIMA [16], adopted such a distance-based approach to mediation analysis of microbiome data. Specifically, MedTest uses the principal components (PCs) of a given distance matrix as multiple mediators and tests their joint mediation effects. However, the assumption that the exposure–microbiome association and the microbiome–outcome association coincide at the same set of PCs may be overly optimistic. Furthermore, the PCs may not capture mediation effects at rare taxa. Moreover, MedTest does not accommodate multivariate exposures and outcomes in its current form. MODIMA calculates distance matrices from the exposure, the microbiome, and the outcome, separately, and employs the distance correlation [17,18] for characterizing the exposure–microbiome association and the partial distance correlation [19] for the microbiome–outcome association conditional on the exposure. The distance matrices for the exposures and outcomes naturally accommodate multivariate variables. However, MODIMA does not allow adjustment of confounders and does not provide an omnibus test. Finally, neither MedTest nor MODIMA can handle censored survival times.

PERMANOVA [7] is currently the most commonly used distance-based method in analysis of microbiome data. Although it was originally developed for testing microbiome associations, we find that we can extend PERMANOVA to testing microbiome mediation effects by using the idea of inverse regression and including both the exposure and the outcome as covariates whose *F* statistics capture the exposure–microbiome association and the microbiome–outcome association conditional on the exposure, respectively. This extension of PERMANOVA would naturally inherit all the features of PERMANOVA, some

of which have been a focus of recent development, including adjustment of confounders [4], test of multivariate covariates, test of censored survival times [20], and an omnibus test of multiple distance matrices [21]. Thus, the extension of PERMANOVA would be very appealing to researchers who routinely use PERMANOVA.

In this article, we present PERMANOVA-med, the extension of PERMANOVA to testing the community-level mediation effect of the microbiome. We base PERMANOVA-med on our implementation of PERMANOVA through the function "permanovaFL" in our R package LDM [4], which differs from the "adonis2" implementation in the R package vegan in the permutation scheme and outperformed adonis2 in many situations [4,5,22]. In the methods section, we first motivate the use of inverse regression and then show how to extend PERMANOVA to PERMANOVA-med. In this process, we provide an overview of PERMANOVA, as well as overviews of MedTest and MODIMA to facilitate comparison with PERMANOVA-med. In the results section, we present extensive simulation studies in which we numerically compared PERMANOVA-med to MedTest and MODIMA. We demonstrate the wide applicability of PERMANOVA-med through 16 different mediation analyses of the real data on melanoma immunotherapy response. We also applied PERMANOVA-med to the real data on dietary fiber intake and BMI that were used in the MedTest paper [15]. We conclude with a discussion section.

2. Materials and Methods

2.1. Motivation toward Inverse Regression

The relationships among the exposure (T), mediator (M), outcome (O), and confounders (Z) are depicted in Figure 1b. Assuming a continuous outcome and a continuous mediator and further assuming no exposure–mediator interaction and no unmeasured confounding, the classical mediation model [23] specifies a linear model for the mediator and a linear model for the outcome:

$$\mathrm{E}(M|Z,T) = \alpha_0 + \alpha_Z^{\mathrm{T}} Z + \alpha_T T, \tag{1}$$

$$\mathrm{E}(O|Z,T,M) = \theta_0 + \theta_Z^{\mathrm{T}} Z + \theta_T T + \theta_M M. \tag{2}$$

Note that α_T characterizes the effect of T on M given Z, and θ_M characterizes the effect of M on O given Z and T. Then, it can be shown that the mediation effect is given by $\alpha_T \theta_M$ [24]. However, it is unclear how to use the microbiome composition data, which are represented by a distance matrix here, as a mediator. Furthermore, the forward outcome model (2) is not easily generalizable to an outcome variable that is discrete, multivariate, or censored survival time.

These limitations motivated us to adopt the inverse regression model that exchanges the positions of the outcome and the mediator in model (2). Inverse regression is a commonly used approach to testing associations [25–27]. It has a key advantage of accommodating different types of outcome variables including multivariate variables. In what follows, we show that, by proper orthogonalization of the non-microbiome variables, the inverse regression model we consider "merges" both models (1) and (2) into one regression model, which fits nicely into the framework of PERMANOVA that takes the distance matrix as the response variable.

To be specific, we first sequentially orthogonalize variables Z, T, and O, and denote the residual of T after orthogonalizing against Z by T_r and denote the residual of O after orthogonalizing against (Z, T) by O_r. Then, we consider the following inverse regression model:

$$\mathrm{E}(M|Z,T,O) = \beta_0 + \beta_Z^{\mathrm{T}} Z + \beta_T T_r + \beta_O O_r. \tag{3}$$

For now, we view M as a univariate continuous variable, just as in (1) and (2). Model (3) implies that $\mathrm{E}(M|Z,T) = \beta_0 + \beta_Z^{\mathrm{T}} Z + \beta_T T_r$, which is exactly model (1) after replacing T by T_r. Thus, we easily obtain that $\beta_T = \alpha_T$. Although it is well known that $\beta_O \neq \theta_M$, we see that $\beta_O = 0$ and $\theta_M = 0$ coincide as they both capture the microbiome–outcome association given (Z, T). As a result, testing $\beta_T \beta_O = 0$ is equivalent to testing $\alpha_T \theta_M = 0$,

i.e., whether there exists a mediation effect through M. We find that model (3) fits nicely into the PERMANOVA framework, in which we view M as a distance matrix and the linear regression as a partition of M into additive components corresponding to the orthogonal factors (Z, T_r, O_r).

2.2. Overview of PERMANOVA

PERMANOVA is based on a linear model of covariates that partition a given distance matrix along each covariate. In particular, when the Euclidean distance measure is used on the relative abundance data, it is the total variance of relative abundance data across all taxa that is partitioned into variance explained by each covariate. Following our implementation in permanovaFL [4], we denote the design matrix of all covariates by X and group the columns of X into K submodels, i.e., $X = (X_1, X_2, \ldots, X_K)$. Each submodel includes components that will be tested jointly, such as a single covariate, multiple covariates, or multiple indicators for a categorical covariate. The submodels are first processed into sequentially orthogonal, unit vectors by the Gram–Schmidt process, so that the partition of the distance matrix is unambiguous. This requires that the covariates in X follow a scientifically meaningful order; for example, the confounders should enter first. Let D denote the $n \times n$ distance matrix calculated among n samples, which is often Gower-centered [28] to become $\Delta = -0.5(I - n^{-1}11^T)D^2(I - n^{-1}11^T)$, where D^2 is the element-wise squared D, I is the identity matrix, and 1 is a vector of n ones. The "residual" distance matrix after projecting off all submodels, except the kth one, takes the form $\widetilde{\Delta}_k = \left(I - \sum\limits_{k'=1,\ldots,K,k' \neq k} X_{k'} X_{k'}^T\right) \Delta \left(I - \sum\limits_{k'=1,\ldots,K,k' \neq k} X_{k'} X_{k'}^T\right)$ by noting that $X_k X_k^T$ is the hat matrix for the kth submodel. Then, PERMANOVA tests the effect of the kth submodel by using the F statistic

$$F_k \propto \frac{\text{Tr}\left[X_k X_k^T \widetilde{\Delta}_k X_k X_k^T\right]}{\text{Tr}\left[\left(I - \sum\limits_{k'=1}^{K} X_{k'} X_{k'}^T\right) \widetilde{\Delta}_k \left(I - \sum\limits_{k'=1}^{K} X_{k'} X_{k'}^T\right)\right]},$$

where $\text{Tr}(\cdot)$ is the trace operation. PERMANOVA assesses the significance of the F statistic via permutation, particularly the Freedman–Lane permutation scheme [29] as implemented in permanovaFL. The Supplementary Materials of [4] showed that the Freedman–Lane scheme is equivalent to forming the following statistic for the bth permutation replicate:

$$F_k^{(b)} \propto \frac{\text{Tr}\left[X_k^{(b)} X_k^{(b)T} \widetilde{\Delta}_k X_k^{(b)} X_k^{(b)T}\right]}{\text{Tr}\left[\left(I - \sum\limits_{k'=1}^{K} X_{k'}^{(b)} X_{k'}^{(b)T}\right) \widetilde{\Delta}_k \left(I - \sum\limits_{k'=1}^{K} X_{k'}^{(b)} X_{k'}^{(b)T}\right)\right]}, \tag{4}$$

where $X_k^{(b)}$ is a row-permuted version of X_k and thus the columns of $X_k^{(b)}$ remain orthogonal. Note that the residual distance matrices $\widetilde{\Delta}_k$s do not need to be recalculated for each replicate. In contrast, the permutation scheme implemented in adonis2 replaces all $\widetilde{\Delta}_k$s in F_k and $F_k^{(b)}$ by the raw distance matrix Δ.

PERMANOVA is very versatile. It can handle censored survival times. As proposed in [20], the survival times and censoring statuses are first fit by a Cox model (including non-microbiome risk predictors as covariates) to be converted into the Martingale or deviance residuals, which are then used as a generic continuous covariate in PERMANOVA. Because PERMANOVA bases its inference on permutation, it is robust to small sample sizes. The permutation replicates can also be readily used to construct an omnibus test of multiple distance matrices, which uses the minimum of the p-values obtained from analyzing each distance matrix as the final test statistic and uses the corresponding minima from the permutation replicates to simulate the null distribution. In addition, the permutation can be conducted in ways that preserve the correlation found in the original data, so PERMANOVA

can accommodate certain structures of samples such as clustered samples [4] and matched sets [5]. All the features that PERMANOVA supports were summarized in Figure 2.
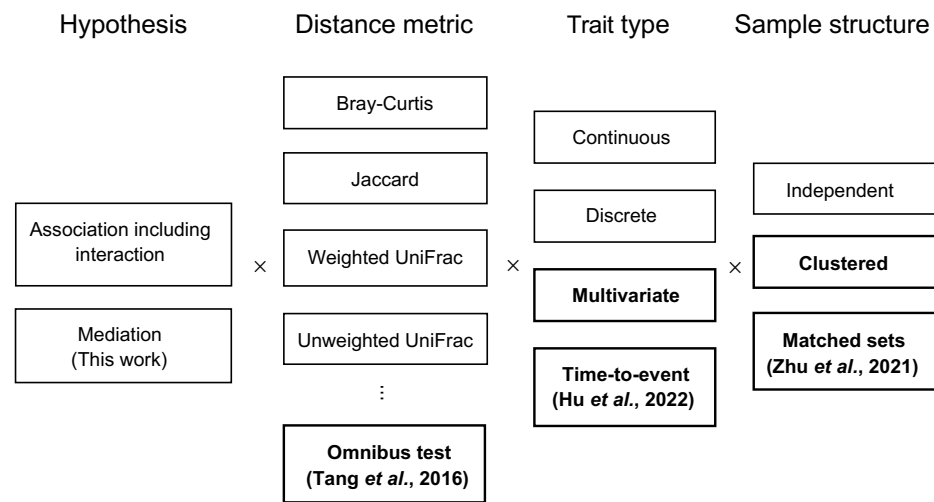

Figure 2. Analyses supported by permanovaFL. Analysis types without a citation were introduced in the original LDM paper [4]. "Clustered" refers to analyses of clustered data where traits of interest vary by cluster or vary both by and within clusters (some analyses may require special structure or additional assumptions). "Matched sets" is a special type of clustered data in which all traits of interest vary within sets. The boxes in bold are key advantages of PERMANOVA-med over MedTest and MODIMA.

2.3. PERMANOVA-med: Extension of PERMANOVA to Mediation Analysis

Under model (3), we set submodels $X_1 = Z$, $X_2 = T_r$, and $X_3 = O_r$, and denote the PERMANOVA F statistics for testing microbiome associations with T_r and O_r by F_T and F_O, respectively. Then, we propose to test the existence of a mediation effect by the microbiome, i.e., $H_0 : \beta_T \beta_O = 0$, using the test statistic

$$\mathbb{U}_{\text{PERMANOVA-med}} = F_T F_O.$$

To claim a mediation effect by the microbiome, both the exposure–microbiome and microbiome–outcome associations (given the exposure) are required to be significant. Thus, the null hypothesis of no mediation is a composite null that consists of no exposure–microbiome association, no microbiome–outcome association, or neither. Accordingly, we construct the following statistic for the bth permutation replicate:

$$\mathbb{U}^{(b)}_{\text{PERMANOVA-med}} = \max \left\{ F_T^{(b)} F_O, \ F_T F_O^{(b)}, \ F_T^{(b)} F_O^{(b)} \right\},$$

where the three product terms correspond to the statistics under the three types of null hypotheses. Then, the p-value is obtained as the proportion of $\mathbb{U}^{(b)}_{\text{PERMANOVA-med}}$ that are equal to or larger than the observed statistic $\mathbb{U}_{\text{PERMANOVA-med}}$. Note that all the F statistics needed for calculating the p-value are directly available from PERMANOVA. As a result, our mediation analysis implemented in the PERMANOVA framework naturally inherits all the features in PERMANOVA.

2.4. Overview of MedTest and MODIMA

MedTest considers microbiome "features" to be the eigenvectors of the Gower-centered distance matrix Δ, denoted by $u_1, u_2, \ldots, u_L$, that are associated with the L positive eigenvalues, denoted by $\lambda_1, \lambda_2, \ldots, \lambda_L$. MedTest assumes that these microbiome features are the units through which the microbiome exert the mediation effect. Thus, it adopts a

test statistic that is a sum of feature-specific mediation effects, each weighted by λ_l (the percentage of variance explained by that feature):

$$\mathbb{U}_{\text{MedTest}} = \sum_{l=1}^{L} \lambda_l |u_l^{\mathsf{T}} T_r| |u_l^{\mathsf{T}} O_r|,$$

where $|.|$ is the absolute value function. Note that $u_l^{\mathsf{T}} T_r$ and $u_l^{\mathsf{T}} O_r$ are the sample Pearson correlation coefficients that measure the associations between the lth feature and the exposure and the outcome, respectively; the sample Pearson correlation coefficient does not easily accommodate multivariate exposure or outcome variables. Similar to PERMANOVA-med, MedTest calculates the maximum of the statistics corresponding to the three types of null hypotheses for the bth permutation replicate:

$$\mathbb{U}_{\text{MedTest}}^{(b)} = \max \left\{ \sum_{l=1}^{L} \lambda_l |u_l^{\mathsf{T}} T_r^{(b)}| |u_l^{\mathsf{T}} O_r|, \ \sum_{l=1}^{L} \lambda_l |u_l^{\mathsf{T}} T_r| |u_l^{\mathsf{T}} O_r^{(b)}|, \ \sum_{l=1}^{L} \lambda_l |u_l^{\mathsf{T}} T_r^{(b)}| |u_l^{\mathsf{T}} O_r^{(b)}| \right\},$$

where $T_r^{(b)}$ and $O_r^{(b)}$ are permuted vectors of T_r and O_r, respectively. Finally, the p-value is obtained as the proportion of $\mathbb{U}_{\text{MedTest}}^{(b)}$ that are equal to or larger than the observed statistic $\mathbb{U}_{\text{MedTest}}$. The power of MedTest may critically depend on whether the exposure–microbiome association and the microbiome–outcome association coincide at the same set of PCs. Furthermore, when the true mediators in the community are rare taxa, the PCs may not effectively capture the variation at these mediators.

In addition to the distance matrix D from the microbiome profiles, MODIMA also requires the $n \times n$ distance matrices (usually the Euclidean distance) being calculated from the exposure data and the outcome data, separately, which we denote by D_T and D_O, respectively. These distance matrices naturally accommodate multivariate variables. Then, MODIMA uses the distance correlation [18], $\text{dCor}(D_T, D)$, for measuring the exposure–microbiome association, which parallels the Pearson correlation with the major difference being that the centered product moment transformation is applied to the distance matrices rather than data vectors. MODIMA uses the partial distance correlation [19], $\text{pdCor}(D_O, D | D_T)$, for measuring the microbiome–outcome association conditional on the exposure, which parallels the Pearson partial correlation. MODIMA adopts the following test statistic:

$$\mathbb{U}_{\text{MODIMA}} = \text{dCor}(D_T, D) \times \text{pdCor}(D_O, D | D_T),$$

and the following statistic for the bth permutation replicate:

$$\mathbb{U}_{\text{MODIMA}}^{(b)} = \begin{cases} \text{dCor}(D_T^{(b)}, D) \times \text{pdCor}(D_O, D | D_T), & \text{if } \text{dCor}(D_T, D) \leq \text{pdCor}(D_O, D | D_T) \\ \text{dCor}(D_T, D) \times \text{pdCor}(D_O^{(b)}, D | D_T), & \text{if } \text{dCor}(D_T, D) > \text{pdCor}(D_O, D | D_T), \end{cases}$$

where $D_T^{(b)}$ and $D_O^{(b)}$ are obtained by permuting both rows and columns of the D_T and D_O matrices, respectively. Although this way of constructing the null statistic appears different from those in PERMANOVA-med and MedTest, they seem asymptotically equivalent. Finally, the p-value is calculated as the proportion of $\mathbb{U}_{\text{MODIMA}}^{(b)}$ that are equal to or larger than the observed statistic $\mathbb{U}_{\text{MODIMA}}$. Note that, in this process, the confounding covariate Z cannot be adjusted. Furthermore, the MODIMA paper pointed out a lack of correspondence between conditional independence and zero partial distance correlation, e.g., a non-zero partial correlation in scenarios with conditionally independent variables. It implies that MODIMA may generate false positive findings under the null hypothesis of no mediation, especially when there is a strong direct effect of the exposure on the outcome (θ_T in model (2)).

2.5. Availability and Implementation

PERMANOVA-med has been added to the existing function "permanovaFL" in our R package LDM, which is available on GitHub at https://github.com/yijuanhu/LDM (accessed on 1 May 2022).

3. Results

3.1. Simulation Studies

Our simulations were based on data on 856 taxa of the upper respiratory tract (URT) microbiome [30], and the mediator model (1) and the forward outcome model (2) as generative models. We considered both binary and continuous exposure variables, continuous outcome variables, and 100 or 200 sample size (n); note that both MedTest and MODIMA papers considered continuous exposures only. In what follows, we number the taxa by decreasing relative abundance so that taxon 1 is the most abundant. We considered 3 mediation mechanisms, in which we assumed the mediating taxa were the top 5 most abundant taxa (taxa 1–5), 100 relatively rare taxa (taxa 51–150), and a mixture of abundant and relatively rare taxa (taxa 4, 5, 51, and 52), which are referred to as M-common, M-rare, and M-mixed, respectively. We further assumed that the mediating taxa played the role through their relative abundances in M-common and M-mixed and through their presence–absence (0/1) statuses in M-rare.

Specifically, for a binary exposure T_i, we assigned half of the samples $T_i = 1$ and the other half $T_i = 0$. For a continuous exposure T_i, we sampled T_i from the Beta(2, 2) distribution. We initially set the baseline relative abundances of all taxa for all samples to the population means that were estimated from the real data, which we denote by $\overline{\pi}_i = (\overline{\pi}_{i1}, \overline{\pi}_{i2}, \ldots, \overline{\pi}_{iJ})$. To induce the effects of the exposure on the mediating taxa, we decreased $\overline{\pi}_{ij}$ by the percentage $\beta_{TM} T_i$ ($\in [0, 1]$) for taxa 3–5 in M-common and taxa 5 and 51 in M-mixed, and then redistributed the decreased amount evenly over the remaining mediating taxa, i.e., taxa 1–2 in M-common and taxa 4 and 52 in M-mixed. In M-rare, we set $\overline{\pi}_{ij}$ for the mediating taxa to 0 with the probability $\beta_{TM} T_i$ independently, and increased $\overline{\pi}_{ij}$ of the most abundant taxon by the total mass that had been set to 0 (which did not affect the presence–absence statuses of the most abundant taxon as it was always present). This way of modifying $\overline{\pi}_i$ did not change the relative abundances of non-associated taxa (except for the most abundant taxon in M-rare) and the modified $\overline{\pi}_i$ still satisfied the compositional constraint (unit sum). Note that β_{TM} characterizes the exposure–microbiome (T-M) association and $\beta_{TM} = 0$ corresponds to no T-M association. Next, we drew the sample-specific composition $\pi_i = (\pi_{i1}, \pi_{i2}, \ldots, \pi_{iJ})$ from the Dirichlet distribution $Dir(\overline{\pi}_i, \theta)$, where the overdispersion parameter θ was set to 0.02 (as estimated from the real data). Then, we generated the read count data using the Multinomial distribution with mean π_i and library size (sequencing depth) sampled from $N(10{,}000; (10{,}000/3)^2)$ and truncated at 2000. Finally, we scaled each read count by the library size to obtain the observed relative abundance, denoted by M_{ij} for taxon j in sample i.

In M-common and M-mixed, we generated the continuous outcome O_i from the following model that allows different directions for the effects of different taxa on the outcome:

$$O_i = \beta_{TO} T_i + \beta_{MO} \text{scale}\left(\sum_{j \in \mathcal{A}_1} M_{ij} - \sum_{j \in \mathcal{A}_2} M_{ij} \right) + \epsilon_i, \tag{5}$$

where $\mathcal{A}_1$ and $\mathcal{A}_2$ are the "increasing" and "decreasing" subsets of mediating taxa as determined above and $\epsilon_i \sim N(0, 0.5^2)$. In M-rare, we define $\mathcal{A}_1$ and $\mathcal{A}_2$ to include taxa 51–100 and taxa 101–150, respectively, and replaced M_{ij} in (5) by $I(M_{ij} \neq 0)$. We also considered a modification of the microbiome–outcome (M-O) association by restricting $\mathcal{A}_1$ and $\mathcal{A}_2$ to a subset of originally selected taxa, i.e., taxa 4 and 5 in M-common, taxa 51 and 52 in M-mixed, and taxa 101–150 in M-rare.

We simulated a binary confounder Z_i in settings with a binary exposure. Note that a confounder is associated with the exposure, the microbiome, and the outcome simul-

taneously (Figure 1b). First, we generated $Z_i = 1$ with probability 0.7 among samples with $T_i = 1$ and with probability 0.3 among those with $T_i = 0$. Then, we used the same operation as used for simulating the T-M association, except that we replaced $\beta_{\mathrm{TM}} T_i$ by $\gamma_{\mathrm{ZM}} Z_i$ with $\gamma_{\mathrm{ZM}} = 0.6$, to further modify $\overline{\pi}_{ij}$ based on Z_i for the mediating taxa that had been modified based on T_i. Finally, we added the term $\gamma_{\mathrm{ZO}} Z_i$ with $\gamma_{\mathrm{ZO}} = 0.7$ to model (5).

We applied PERMANOVA-med and compared it to MedTest and MODIMA, for testing the mediation effect of the microbiome in the simulated data. In M-common and M-mixed, all tests were based on the Bray–Curtis distance. In M-rare, all tests were based on the Jaccard distance. The type I error and power of all tests were assessed at the nominal level 0.05 based on 10,000 and 1000 replicates of data, respectively.

3.2. Simulation Results

We first present results for the simulated data without a confounder. The power of the PERMANOVA-med, MedTest, and MODIMA with varying values of β_{MO}, β_{TM}, β_{TO}, and sample size n are displayed in Figures 3–5 for M-common, M-mixed, and M-rare, respectively. The numerical values of the type I error rates (when $\beta_{\mathrm{MO}} = 0$) shown in these figures are also listed in Table 1.

In M-common with a binary exposure, when the same abundant taxa (taxa 1–5) were used to generate both the T-M and M-O associations (Figure 3a), MedTest was slightly more powerful than PERMANOVA-med, possibly because the top PCs used by MedTest effectively captured both the T-M and M-O associations. When a subset of taxa (taxa 4 and 5) were used for generating the M-O association (Figure 3b), the power of MedTest declined much more quickly than the power of PERMANOVA-med, as the PCs that captured the T-M association (e.g., PC1) may not coincide with the PCs that captured the M-O association (e.g., PC2). MODIMA seemed to be very powerful in some cases (e.g., Figure 3a), but its performance was sensitive to the value of β_{TO}. In particular, MODIMA generated inflated type I error when β_{TO} was enlarged to 0.8 and especially when n was also increased from 100 to 200.

In M-common with a continuous exposure, which tended to result in more complex variation patterns in the data than a binary exposure, MedTest (and MODIMA) lost the advantage in power to PERMANOVA-med, even when taxa 1–5 were used for both the T-M and M-O associations (Figure 3c). Again, MedTest lost further, considerable power to PERMANOVA-med when taxa 4 and 5 were used for the M-O association (Figure 3d) and MODIMA yielded inflated type I error when β_{TO} and n were both large.

As expected, PERMANOVA-med always had significantly higher power than MedTest in M-mixed (Figure 4), and the power difference was more pronounced in M-rare (Figure 5), since PCs became less efficient in capturing variations in less abundant taxa. In M-rare, MODIMA was uniformly less powerful than PERMANOVA-med, even its type I error was clearly inflated.

Finally, when a confounder was added to the simulated data, MODIMA, without the capability to adjust for the confounding effect, produced very inflated type I error (Table 2). Note that, PERMANOVA-med and MedTest always controlled the type I error below the nominal level, with (Table 2) or without (Table 1) the confounder.
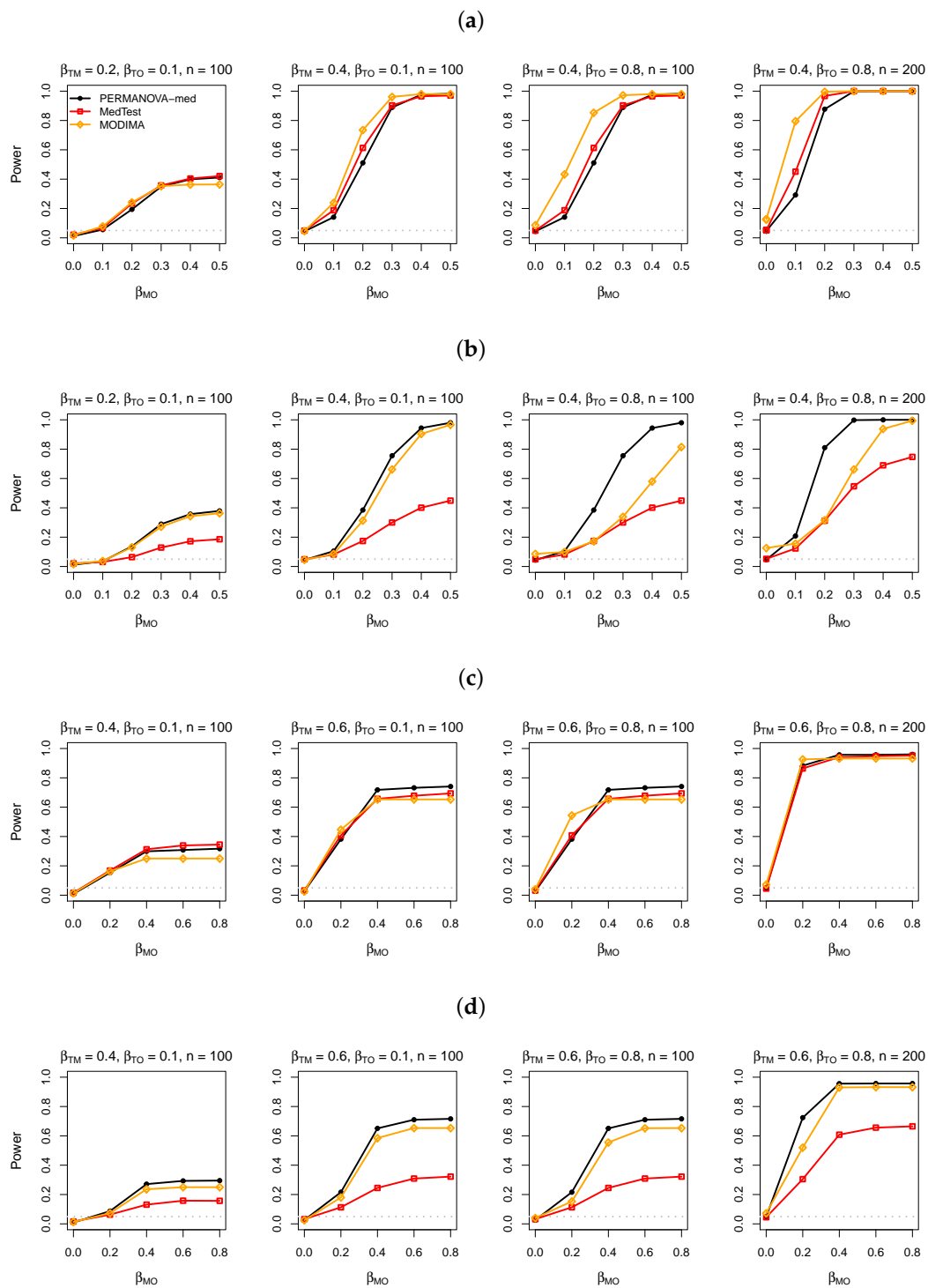
(a)

(b)

(c)

(d)

Figure 3. Simulation results in analysis of simulated data under M-common. (**a**) Binary exposure, taxa 1–5 for the M-O association. (**b**) Binary exposure, taxa 4 and 5 for the M-O association. (**c**) Continuous exposure, taxa 1–5 for the M-O association. (**d**) Continuous exposure, taxa 4 and 5 for the M-O association.
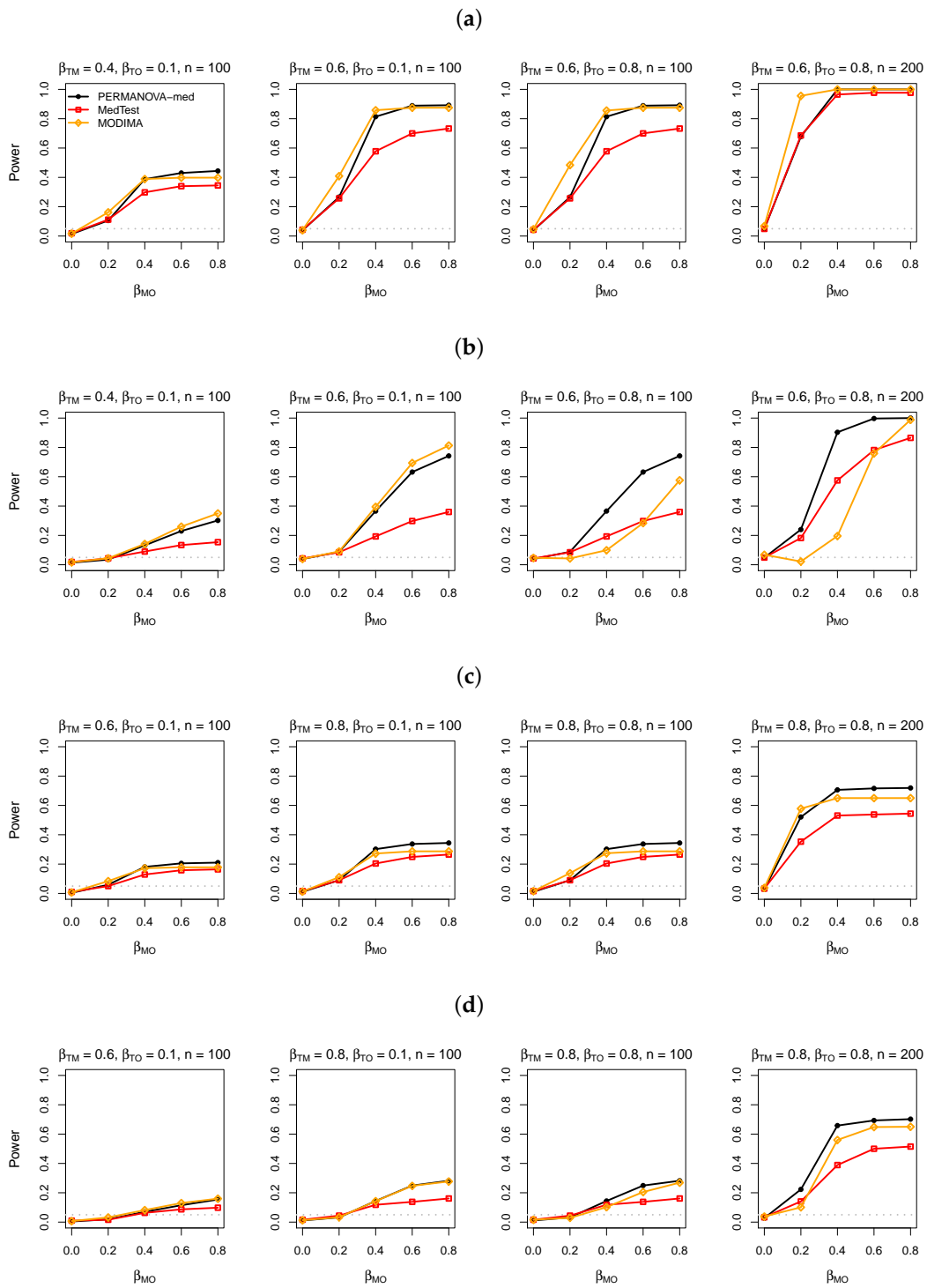
Figure 4. Simulation results in analysis of simulated data under M-mixed. (**a**) Binary exposure, taxa 4, 5, 51, and 52 for the M-O association. (**b**) Binary exposure, taxa 51 and 52 for the M-O association. (**c**) Continuous exposure, taxa 4, 5, 51, and 52 for the M-O association. (**d**) Continuous exposure, taxa 51 and 52 for the M-O association.
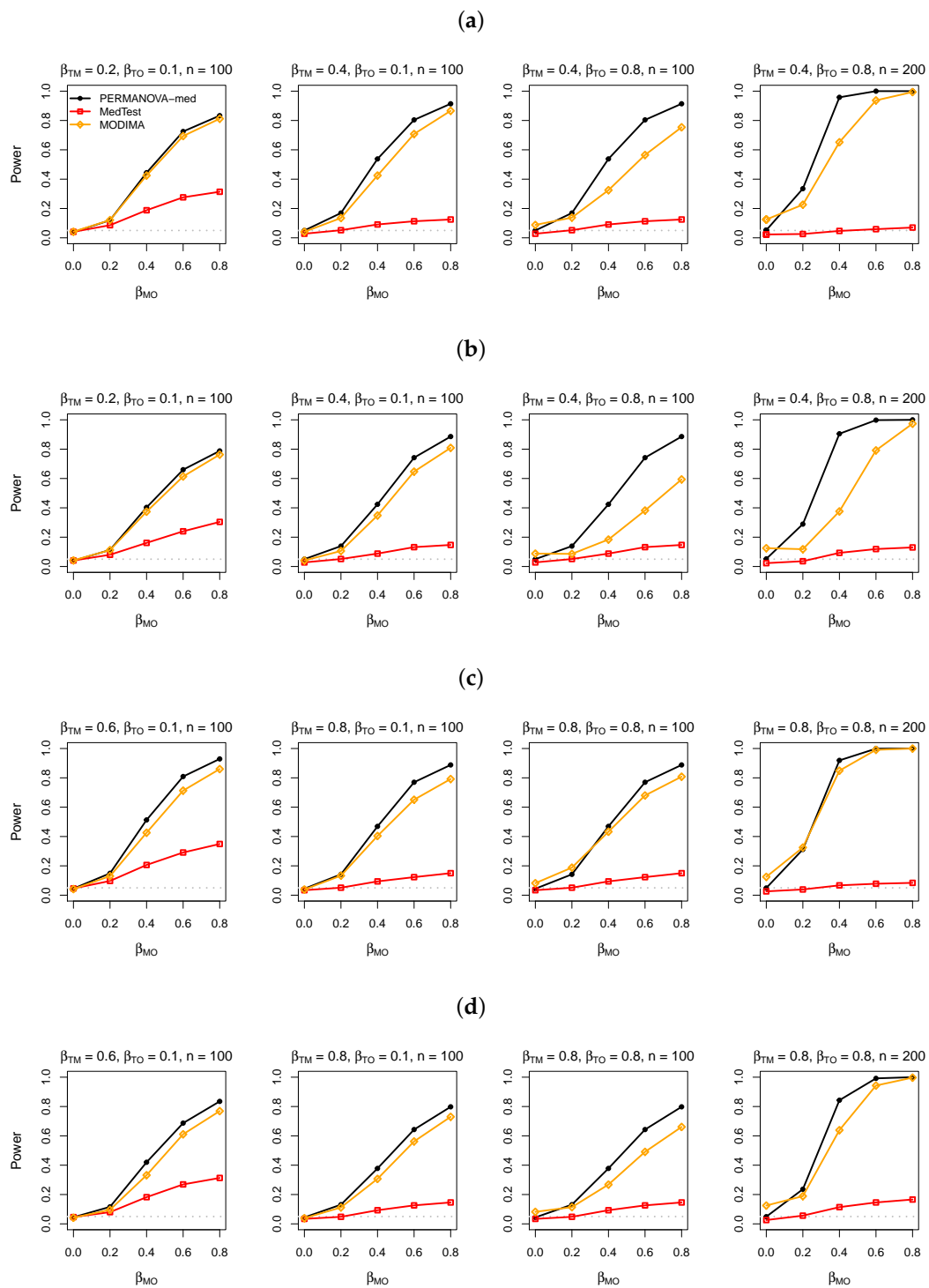
Figure 5. Simulation results in analysis of simulated data under M-rare. (**a**) Binary exposure, taxa 51–150 for the M-O association. (**b**) Binary exposure, taxa 101–150 for the M-O association. (**c**) Continuous exposure, taxa 51–150 for the M-O association. (**d**) Continuous exposure, taxa 101–150 for the M-O association.

Table 1. Type I error (at the level 0.05) in analysis of simulated data without a confounder.

Scenario	Exposure	β_{TM}	β_{TO}	n	PERMANOVA-med	MedTest	MODIMA
M-common	Binary	0.2	0.1	100	0.012	0.021	0.017
		0.4	0.1	100	0.044	0.049	0.046
		0.4	0.8	100	0.044	0.049	0.086
		0.4	0.8	200	0.046	0.052	0.126
	Continuous	0.4	0.1	100	0.009	0.016	0.013
		0.6	0.1	100	0.026	0.032	0.025
		0.6	0.8	100	0.026	0.032	0.040
		0.6	0.8	200	0.048	0.045	0.072
M-mixed	Binary	0.4	0.1	100	0.014	0.019	0.017
		0.6	0.1	100	0.039	0.043	0.040
		0.6	0.8	100	0.039	0.043	0.047
		0.6	0.8	200	0.048	0.049	0.068
	Continuous	0.6	0.1	100	0.004	0.010	0.007
		0.8	0.1	100	0.011	0.016	0.013
		0.8	0.8	100	0.011	0.016	0.016
		0.8	0.8	200	0.027	0.033	0.038
M-rare	Binary	0.2	0.1	100	0.039	0.041	0.042
		0.4	0.1	100	0.050	0.028	0.041
		0.4	0.8	100	0.050	0.028	0.088
		0.4	0.8	200	0.052	0.023	0.125
	Continuous	0.6	0.1	100	0.045	0.046	0.042
		0.8	0.1	100	0.044	0.034	0.039
		0.8	0.8	100	0.044	0.034	0.082
		0.8	0.8	200	0.049	0.026	0.125

Note: The type I error results were generated at $\beta_{MO} = 0$ (i.e., no M-O association), and thus the same for datasets using different sets of taxa for generating the M-O association. β_{TM} characterizes the exposure-microbiome association. β_{TO} characterizes the direct effect of the exposure on the outcome.

Table 2. Type I error (at the level 0.05) in analysis of simulated data with a binary exposure and a binary confounder.

Scenario	β_{TM}	PERMANOVA-med	MedTest	MODIMA
M-common	0.2	0.008	0.014	0.242
	0.4	0.035	0.040	0.385
M-mixed	0.4	0.006	0.015	0.056
	0.6	0.020	0.029	0.103
M-rare	0.2	0.026	0.036	0.279
	0.4	0.046	0.035	0.238

Note: $\beta_{TO} = 0.1$, $\beta_{MO} = 0$, and $n = 100$. MODIMA does not allow adjustment of confounders.

3.3. Real Data on Melanoma Immunotherapy Response

The real data [31] we used were generated from a cohort of 167 melanoma patients, who received immune checkpoint blockade (ICB) treatment and were classified as 106 responders and 61 non-responders. Their progression-free survival times (in days) were observed for 61 patients, censored for 49 patients, and missing for 57 patients. Their gut microbiome were profiled via shotgun metagenomic sequencing to generate a taxa count table including 225 taxa (lowest taxon known for a feature, up to species). These patients were further asked to complete a lifestyle survey, which included assessment of dietary fiber intake and use of probiotic supplements within the past month; 110 provided data for probiotic use, 94 provided data for dietary fiber intake, and 89 provided data for both.

Spencer et al. [31] found in this dataset that higher dietary fiber intake was associated with significantly improved progression-free survival, with the most pronounced benefit

observed in patients with sufficient dietary fiber intake and no probiotic use. They also found marginal significance for the association of dietary fiber intake and response to ICB. In addition, the influence of the gut microbiome on immunotherapy response has been demonstrated in numerous human cohorts as well as in preclinical models [2,32], and the human gut microbiome is itself shaped by diet [3] and medication use [33]. Given this interplay between diet and medication use, gut microbiome, and immunotherapy response, a natural question that arose was then whether some effect of dietary fiber intake and probiotic use on immunotherapy response in this dataset was mediated through the gut microbiome.

We performed a variety of mediation analyses using this dataset. For the outcome, we considered both the progression-free survival and the response to ICB, the former of which is a possibly censored survival time variable and the latter is a binary variable. For the exposure, we considered the dietary fiber intake (sufficient or insufficient), the probiotic use (no/yes), and the four-level categorical variable defined by both dietary fiber intake and probiotics use. Following [31], we additionally compared patients with sufficient dietary fiber intake and no probiotic use to all other three groups. We selected body mass index (BMI), prior treatment, lactate dehydrogenase level (LDH), and stage as potential confounders based on our analysis of bivariate associations, and we performed each mediation analysis with and without adjustment of these confounders. In all 16 mediation analyses, we applied PERMANOVA-med, MedTest, and MODIMA whenever they were applicable. For each method, we constructed tests based on the Bray–Curtis and Jaccard distance measures separately, as well as the omnibus test of both distance measures (except for MODIMA).

All *p*-values were summarized in Table 3. None of the *p*-values were significant at the 0.05 level, possibly due to the small sample sizes. Nevertheless, Table 3 demonstrated the wide applicability of PERMANOVA-med and the limited capabilities of MedTest and MODIMA. Specifically, neither MedTest nor MODIMA can handle censored survival times (the progression-free survival); MODIMA cannot adjust confounders (BMI et al.) nor provide an omnibus test (that combines Bray–Curtis and Jaccard); MedTest cannot handle multivariate exposures (the four-level categorical variable).

Table 3. *p*-values from 16 mediation analyses of the data on melanoma immunotherapy response.

Outcome	Exposure	*n*	PERMANOVA-med			MedTest			MODIMA		
			BC	J	Omni	BC	J	Omni	BC	J	Omni
No adjustment of covariates											
Progression-free survival	Fiber intake	89	0.808	0.965	0.958	-	-	-	-	-	-
	Probiotics	110	0.913	0.716	0.899	-	-	-	-	-	-
	Fiber + probiotics (4 levels)	89	0.777	0.975	0.953	-	-	-	-	-	-
	Sufficient fiber + no probiotics	89	0.717	0.965	0.910	-	-	-	-	-	-
Response to ICB	Fiber intake	94	0.727	0.955	0.903	0.624	0.636	0.837	0.384	0.935	-
	Probiotics	110	0.888	0.589	0.794	0.978	0.698	0.898	0.915	0.381	-
	Fiber + probiotics (4 levels)	89	0.620	0.980	0.827	-	-	-	0.430	0.947	-
	Sufficient fiber + no probiotics	89	0.490	0.955	0.697	0.276	0.626	0.455	0.441	0.947	-
Adjusting for BMI, prior treatment, LDH, stage											
Progression-free survival	Fiber intake	89	0.786	0.990	0.936	-	-	-	-	-	-
	Probiotics	110	0.983	0.788	0.947	-	-	-	-	-	-
	Fiber + probiotics (4 levels)	89	0.770	0.995	0.935	-	-	-	-	-	-
	Sufficient fiber + no probiotics	89	0.725	0.980	0.903	-	-	-	-	-	-
Response to ICB	Fiber intake	94	0.870	0.920	0.975	0.832	0.935	0.966	-	-	-
	Probiotics	110	0.973	0.433	0.630	0.911	0.539	0.773	-	-	-
	Fiber + probiotics (4 levels)	89	0.760	0.975	0.928	-	-	-	-	-	-
	Sufficient fiber + no probiotics	89	0.644	0.925	0.850	0.453	0.973	0.682	-	-	-

Note: BC: Bray-Curtis; J: Jaccard; Omni: the omnibus test that combines the results from analyzing the Bray-Curtis and Jaccard distances; *n*: sample size; -: not applicable.

3.4. Real Data on Dietary Fiber Intake and BMI

We also used the real data that were generated from a cross-sectional study of 98 healthy human subjects, to test whether the effect of dietary fiber intake on BMI was mediated by the gut microbiome. This dataset was previously used by the MedTest paper [15] and described in more detail there. In this dataset, the operational taxonomic unit (OTU) table generated from 16S amplicon sequencing has been rarefied to 2387 read counts per sample and consisted of 4290 OTUs (with at least one count in at least one sample); dietary fiber intake (as assessed by percent calories from dietary fiber) and BMI were both continuous variables. Following the MedTest paper, we constructed the Bray–Curtis, Jaccard, and unweighted, weighted, and generalized UniFrac distance matrices from this OTU table, and we tested the mediation effect by the gut microbiome in the effect of dietary fiber intake on BMI without adjustment of any confounder. This was a simple scenario in which PERMANOVA-med, MedTest, and MODIMA were all applicable. We thus applied the three methods to each distance matrix separately, and we obtained results for the omnibus test of all five distance matrices from PERMANOVA-med and MedTest.

In addition, we noted that the uniform library size (2387) was much lower than the number of OTUs (4290), which implies that there must exist a large number of extremely rare OTUs. It is well known that extremely rare OTUs tend to be errors due to sequencing, misclassification or contamination, and a common practice is to filter out OTUs that have non-zero counts in fewer than five samples [34]. This filter is particularly important for analysis that is based on presence–absence data (e.g., the Jaccard and unweighted UniFrac distance matrices), which is more sensitive to any non-zero counts than analysis based on relative abundance data. Therefore, we also repeated the aforementioned procedure using the filtered OTU table, which consisted of 885 OTUs.

All p-values were summarized in Table 4. In analysis of unfiltered data, we obtained MedTest p-values that resembled those reported in the MedTest paper (their Table 2 [15]); the minor variations can be attributed to stochastic randomness in the permutation procedure. Specifically, MedTest yielded small p-values, 0.004 and 0.0739 for the Jaccard and unweighted UniFrac distances, respectively, suggesting the existence of mediation effect at the presence–absence scale. For the two distances, PERMANOVA-med yielded p-values 0.032 and 0.049, and MODIMA yielded 0.008 and 0.056, all of which were consistent with the MedTest results. After filtering out extremely rare OTUs, the results of PERMANOVA-med and MODIMA stayed largely unchanged, which was expected as extremely rare OTUs should not have a large influence. However, the MedTest p-values for Jaccard and unweighted UniFrac both became very nonsignificant (0.114 and 0.766), possibly due to the change of PCs. For the omnibus test, PERMANOVA-med produced stable and marginally significant p-values, 0.0859 before filtering and 0.0929 after filtering; MedTest produced highly variable p-values, 0.011 before and 0.335 after; MODIMA did not provide results for such a test.

Table 4. p-values from testing the mediation effect by the gut microbiome in the effect of dietary fiber intake on BMI.

	Method	BC	J	UniFrac	WUniFrac	GUniFrac	Omni
No filter	PERMANOVA-med	0.304	0.032	0.0490	0.597	0.235	0.0859
	MedTest	0.530	0.00400	0.0739	0.792	0.521	0.0110
	MODIMA	0.087	0.00800	0.0560	0.451	0.250	-
With filter	PERMANOVA-med	0.289	0.036	0.0769	0.586	0.252	0.0929
	MedTest	0.526	0.114	0.766	0.706	0.438	0.335
	MODIMA	0.084	0.005	0.077	0.351	0.193	-

Note: BC: Bray-Curtis; J: Jaccard; UniFrac: unweighted UniFrac; WUniFrac: weighted UniFrac; GUniFrac: generalized UniFrac [14]; Omni: the omnibus test that combines the results from analyzing all five distances. The filter excluded taxa that were found in fewer than five samples.

4. Discussion

We presented PERMANOVA-med, an extension of PERMANOVA to mediation analysis of microbiome data. Through extensive simulation studies, we observed that PERMANOVA-med did not uniformly outperform MedTest. However, the scenarios in which PERMANOVA-med did outperform seemed more realistic and more general, e.g., scenarios with a mixture of abundant and less abundant mediating taxa, relatively rare mediating taxa, or different sets of taxa associated with the exposure and the outcome. Even in the single scenario that PERMANOVA-med lost power to MedTest (Figure 3a), the power loss was relatively small. The power comparison between PERMANOVA-med and MODIMA was more difficult, as MODIMA often lost control of the type I error. Nevertheless, there were many more scenarios in which PERMANOVA-med had higher power than MODIMA than scenarios when it was the opposite.

The main advantage of PERMANOVA-med over MedTest and MODIMA is its wide applicability to a variety of mediation analyses of microbiome data, which was achieved by using our existing function permanovaFL. Through analysis of the simulated data and the real data, we have illustrated most features in Figure 2 that are supported by permanovaFL, such as multivariate exposures, survival outcomes, and omnibus tests of multiple distance measures. Although we did not cover clustered or matched-set data in this article, these types of data are emerging rapidly in recent years and may also call for mediation analysis. PERMANOVA-med is well positioned to accommodate such data in its current form. Further, PERMANOVA-med is not constrained to analysis of microbiome data but applicable to any high-dimensional data (e.g., genomic, epigenomic, metabolomic, proteomic, and cytokine data) that can be summarized into distance matrices.

Caution is required in interpreting results from PERMANOVA-med (as well as MedTest and MODIMA). Strictly speaking, a significant *p*-value from PERMANOVA-med only means that the microbiome are associated with both the exposure and the exposure-adjusted outcome. External information on causal direction is needed to declare that the microbiome truly mediate the effect of the exposure on the outcome. Although the causal directions in the exposure–outcome and exposure–microbiome relationships may be evident in many cases, the causal direction between the microbiome and the outcome is often less clear because the change of microbiome may well be a consequence of the change of outcome rather than a cause.

PERMANOVA-med is limited to testing the mediation effect by the microbiome at the community level. We have previously developed the linear decomposition model (LDM) that unifies the community-level and taxon-level tests into one framework in the context of testing microbiome associations with traits of interest [4]. Using the idea of inverse regression, we have also extended the LDM, called LDM-med, to testing microbiome mediation at both the community and individual taxon levels [35]; some of those results mirror the results we obtained here. Aside from the capability of LDM-med to detect individual mediating taxa, a major difference between the two works is how we define the mediation effect by the microbiome. In the current work, we declare a mediation effect whenever the exposure perturbs some part of the microbial community and some part of the community influence the outcome; the two parts do not necessarily overlap (e.g., involving different taxa). This definition is reasonable here because, in distance-based analysis, a microbial community is viewed as a whole interconnected entity. The definition of microbiome mediation in [35] was more stringent. Because the main focus there was to detect individual taxa that act as mediators, only taxa that are first affected by the exposure and then influence the outcome were declared be mediating taxa, and only a community that has mediating taxa in it was declared to have a global mediation effect. In practice, how to choose between PERMANOVA-med and LDM-med depends on what type of mediation is of most interest.

Author Contributions: Y.-J.H. conceived the problem, developed the method, and co-wrote the manuscript; Y.Y. implemented the method in R code, performed simulation studies and data analysis,

References

1. Bai, J.; Hu, Y.; Bruner, D. Composition of gut microbiota and its association with body mass index and lifestyle factors in a cohort of 7–18 years old children from the American Gut Project. *Pediatr. Obes.* **2019**, *14*, e12480. [CrossRef] [PubMed]
2. Routy, B.; Le Chatelier, E.; Derosa, L.; Duong, C.P.; Alou, M.T.; Daillère, R.; Fluckiger, A.; Messaoudene, M.; Rauber, C.; Roberti, M.P.; et al. Gut microbiome influences efficacy of PD-1–based immunotherapy against epithelial tumors. *Science* **2018**, *359*, 91–97. [CrossRef] [PubMed]
3. McDonald, D.; Hyde, E.; Debelius, J.W.; Morton, J.T.; Gonzalez, A.; Ackermann, G.; Aksenov, A.A.; Behsaz, B.; Brennan, C.; Chen, Y.; et al. American Gut: An open platform for citizen science microbiome research. *Msystems* **2018**, *3*, e00031-18. [CrossRef] [PubMed]
4. Hu, Y.J.; Satten, G.A. Testing hypotheses about the microbiome using the linear decomposition model (LDM). *Bioinformatics* **2020**, *36*, 4106–4115. [CrossRef]
5. Zhu, Z.; Satten, G.A.; Mitchell, C.; Hu, Y.J. Constraining PERMANOVA and LDM to within-set comparisons by projection improves the efficiency of analyses of matched sets of microbiome data. *Microbiome* **2021**, *9*, 1–19. [CrossRef]
6. Legendre, P.; Anderson, M.J. Distance-based redundancy analysis: Testing multispecies responses in multifactorial ecological experiments. *Ecol. Monogr.* **1999**, *69*, 1–24. [CrossRef]
7. McArdle, B.H.; Anderson, M.J. Fitting multivariate models to community data: A comment on distance-based redundancy analysis. *Ecology* **2001**, *82*, 290–297. [CrossRef]
8. Zhao, N.; Chen, J.; Carroll, I.M.; Ringel-Kulka, T.; Epstein, M.P.; Zhou, H.; Zhou, J.J.; Ringel, Y.; Li, H.; Wu, M.C. Testing in microbiome-profiling studies with MiRKAT, the microbiome regression-based kernel association test. *Am. J. Hum. Genet.* **2015**, *96*, 797–807. [CrossRef]
9. Alekseyenko, A.V. Multivariate Welch t-test on distances. *Bioinformatics* **2016**, *32*, 3552–3558.
10. Zhang, Y.; Han, S.W.; Cox, L.M.; Li, H. A multivariate distance-based analytic framework for microbial interdependence association test in longitudinal study. *Genet. Epidemiol.* **2017**, *41*, 769–778. [CrossRef]
11. Jaccard, P. The distribution of the flora in the alpine zone. *New Phytol.* **1912**, *11*, 37–50. [CrossRef]
12. Bray, J.R.; Curtis, J.T. An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* **1957**, *27*, 326–349. [CrossRef]
13. Lozupone, C.; Knight, R. UniFrac: A new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **2005**, *71*, 8228–8235. [CrossRef]
14. Chen, J.; Bittinger, K.; Charlson, E.S.; Hoffmann, C.; Lewis, J.; Wu, G.D.; Collman, R.G.; Bushman, F.D.; Li, H. Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics* **2012**, *28*, 2106–2113. [CrossRef]
15. Zhang, J.; Wei, Z.; Chen, J. A distance-based approach for testing the mediation effect of the human microbiome. *Bioinformatics* **2018**, *34*, 1875–1883. [CrossRef] [PubMed]
16. Hamidi, B.; Wallace, K.; Alekseyenko, A.V. MODIMA, a method for multivariate omnibus distance mediation analysis, allows for integration of multivariate exposure-mediator-response relationships. *Genes* **2019**, *10*, 524. [CrossRef]
17. Székely, G.J.; Rizzo, M.L.; Bakirov, N.K. Measuring and testing dependence by correlation of distances. *Ann. Stat.* **2007**, *35*, 2769–2794. [CrossRef]
18. Székely, G.J.; Rizzo, M.L. Brownian distance covariance. *Ann. Appl. Stat.* **2009**, *3*, 1236–1265. [CrossRef]
19. Székely, G.J.; Rizzo, M.L. Partial distance correlation with methods for dissimilarities. *Ann. Stat.* **2014**, *42*, 2382–2412. [CrossRef]
20. Hu, Y.; Satten, G.A.; Hu, Y.J. Testing microbiome associations with censored survival outcomes at both the community and individual taxon levels. *bioRxiv* **2022**. [CrossRef]
21. Tang, Z.Z.; Chen, G.; Alekseyenko, A.V. PERMANOVA-S: Association test for microbial community composition that accommodates confounders and multiple distances. *Bioinformatics* **2016**, *32*, 2618–2625. [CrossRef] [PubMed]
22. Hu, Y.J.; Satten, G.A. A rarefaction-without-resampling extension of PERMANOVA for testing presence-absence associations in the microbiome. *bioRxiv* **2021**. [CrossRef]
23. Baron, R.M.; Kenny, D.A. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J. Personal. Soc. Psychol.* **1986**, *51*, 1173. [CrossRef]
24. VanderWeele, T.J.; Vansteelandt, S. Conceptual issues concerning mediation, interventions and composition. *Stat. Its Interface* **2009**, *2*, 457–468. [CrossRef]
25. O'Reilly, P.F.; Hoggart, C.J.; Pomyen, Y.; Calboli, F.C.; Elliott, P.; Jarvelin, M.R.; Coin, L.J. MultiPhen: Joint model of multiple phenotypes can increase discovery in GWAS. *PLoS ONE* **2012**, *7*, e34861. [CrossRef]

26. Wu, B.; Pankow, J.S. Statistical methods for association tests of multiple continuous traits in genome-wide association studies. *Ann. Hum. Genet.* **2015**, *79*, 282–293. [CrossRef]

27. Majumdar, A.; Witte, J.S.; Ghosh, S. Semiparametric allelic tests for mapping multiple phenotypes: Binomial regression and Mahalanobis distance. *Genet. Epidemiol.* **2015**, *39*, 635–650. [CrossRef]

28. Gower, J.C. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* **1966**, *53*, 325–338. [CrossRef]

29. Freedman, D.; Lane, D. A nonstochastic interpretation of reported significance levels. *J. Bus. Econ. Stat.* **1983**, *1*, 292–298.

30. Charlson, E.S.; Chen, J.; Custers-Allen, R.; Bittinger, K.; Li, H.; Sinha, R.; Hwang, J.; Bushman, F.D.; Collman, R.G. Disordered microbial communities in the upper respiratory tract of cigarette smokers. *PLoS ONE* **2010**, *5*, e15216. [CrossRef]

31. Spencer, C.N.; McQuade, J.L.; Gopalakrishnan, V.; McCulloch, J.A.; Vetizou, M.; Cogdill, A.P.; Khan, M.A.W.; Zhang, X.; White, M.G.; Peterson, C.B.; et al. Dietary fiber and probiotics influence the gut microbiome and melanoma immunotherapy response. *Science* **2021**, *374*, 1632–1640. [CrossRef] [PubMed]

32. Matson, V.; Fessler, J.; Bao, R.; Chongsuwat, T.; Zha, Y.; Alegre, M.L.; Luke, J.J.; Gajewski, T.F. The commensal microbiome is associated with anti–PD-1 efficacy in metastatic melanoma patients. *Science* **2018**, *359*, 104–108. [CrossRef] [PubMed]

33. Maier, L.; Pruteanu, M.; Kuhn, M.; Zeller, G.; Telzerow, A.; Anderson, E.E.; Brochado, A.R.; Fernandez, K.C.; Dose, H.; Mori, H.; et al. Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature* **2018**, *555*, 623–628. [CrossRef] [PubMed]

34. Cao, Q.; Sun, X.; Rajesh, K.; Chalasani, N.; Gelow, K.; Katz, B.; Shah, V.H.; Sanyal, A.J.; Smirnova, E. Effects of rare microbiome taxa filtering on statistical analysis. *Front. Microbiol.* **2021**. [CrossRef] [PubMed]

35. Yue, Y.; Hu, Y.J. A new approach to testing mediation of the microbiome at both the community and individual taxon levels. *Bioinformatics* **2022**. [CrossRef]