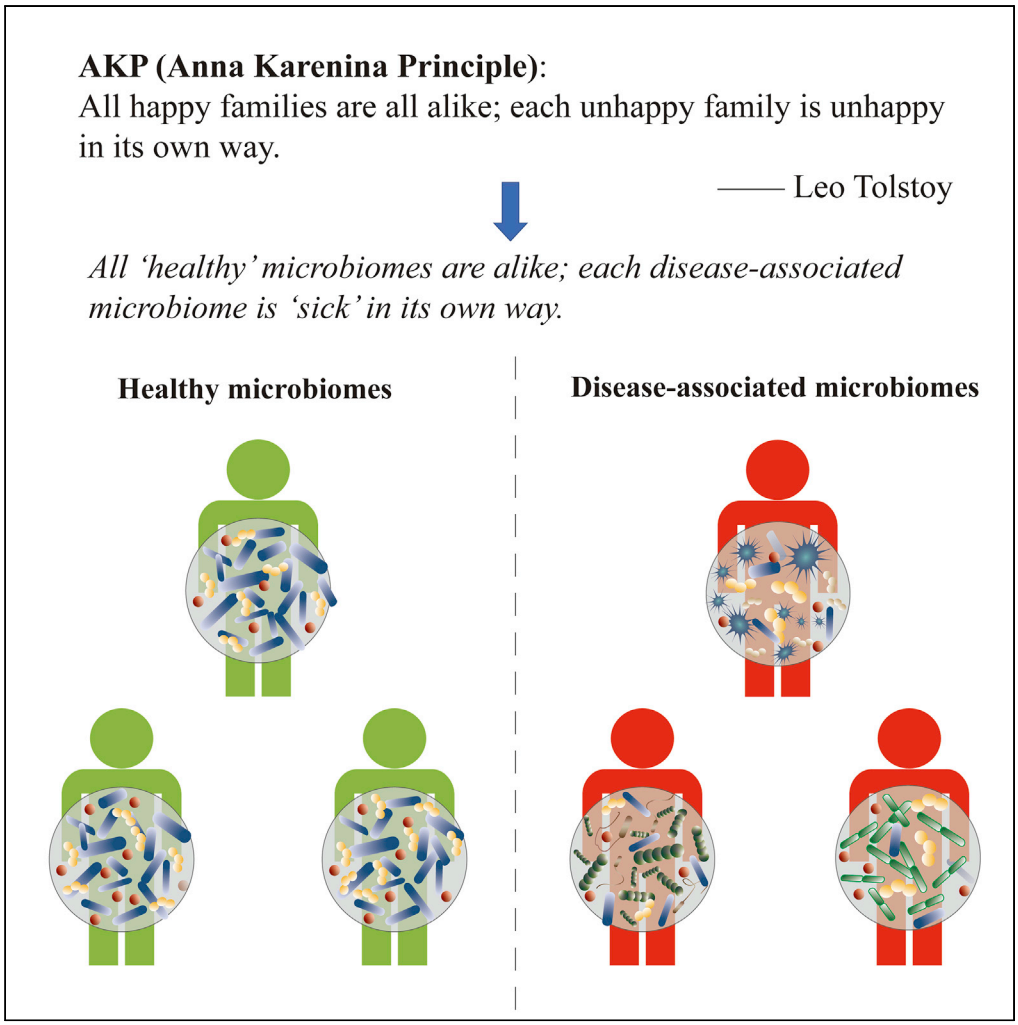
iScience

CellPress

Article

Testing the Anna Karenina Principle in Human Microbiome-Associated Diseases



Zhanshan (Sam) Ma

ma@vandals.uidaho.edu

HIGHLIGHTS

About a half of microbiome-associated diseases follow AKP (Anna Karenina principle)

AKP effects are primarily influenced by highly dominant microbial species

About one-fourth of microbiome-associated diseases follow the anti-AKP

All species appear to play equal roles in influencing the anti-AKP effects

Ma, iScience *23*, 101007
April 24, 2020
https://doi.org/10.1016/j.isci.2020.101007

iScience

CellPress

Article

Testing the Anna Karenina Principle in Human Microbiome-Associated Diseases

Zhanshan (Sam) Ma[1,2,3,*]

SUMMARY

The AKP (Anna Karenina principle), which refers to observations inspired by the opening line of Leo Tolstoy's Anna Karenina, "all happy families are all alike; each unhappy family is unhappy in its own way," predicts that all "healthy" microbiomes are alike and each disease-associated microbiome is "sick" in its own way in human microbiome-associated diseases (MADs). The AKP hypothesis predicts the rise of heterogeneity/stochasticity in human microbiomes associated with dysbiosis due to MADs. We used the beta-diversity in Hill numbers and stochasticity analysis to detect AKP and anti-AKP effects. We tested the AKP with 27 human MAD studies and discovered that the AKP, anti-AKP, and non-AKP effects were exhibited in approximately 50%, 25%, and 25% of the MAD cases, respectively. Mechanistically, AKP effects are primarily influenced by highly dominant microbial species and less influenced by rare species. In contrast, all species appear to play equal roles in influencing anti-AKP effects.

INTRODUCTION

The mostly peaceful coexistence of human microbiomes with hosts and the tolerance by our immune system is still poorly understood in human biology and modern biomedicine. In fact, the interaction between the immune system and microbiome is bidirectional. On the one hand, the immune system certainly plays a critical role in shaping and maintaining the human microbiome; on the other hand, gut microbiome of infants may help to train the full development of their immune systems (Rooks and Garrett, 2016; Levy et al., 2017; Thaiss et al., 2016). There is a hypothesis that animal regulation of immunity evolved to not only defend against pathogens but also carefully regulate symbiotic microbes (Giongo et al., 2010; Zaneveld et al., 2017; Rizzetto et al., 2018; Lotter and Altfeld, 2019). In consideration of the significance of immune system, Zaneveld et al. (2017) argued that the so-termed AKP (Anna Karenina principle) effects in animal microbiomes are ubiquitous and significant and frequently linked to deteriorating host health. The AKP refers to observations derived from the opening line of Leo Tolstoy's Anna Karenina: "*all happy families are all alike; each unhappy family is unhappy in its own way.*" In terms of the microbiome-associated diseases (MADs), it may be translated into a hypothesis: all "healthy" microbiomes (of healthy individuals) are alike; each "diseased" microbiome (of a patient with MAD) is "sick" in its own way.

Studies on animal and human microbiomes have suggested that microbiome stability is a hallmark of healthy host physiology, which is consistent with the evolution of animals (humans) in a sea of microbes (Zaneveld et al., 2017; Ma and Ellison, 2019). For example, it has been found that compromised host immunity can induce AKP effects (e.g., reviews by Williams et al., 2016; Zaneveld et al., 2017) and gut microbiome has been found deeply involved in allergenic and autoimmune disorders (Giongo et al., 2010; Halfvarson et al., 2017). AKP effects imply that microbiome instability (dysbiosis) can only be observed in comparison with normal variation (heterogeneity) (Brüssow, 2016; Zaneveld et al., 2017). Furthermore, a hallmark of AKP effects is the rising heterogeneity and/or stochasticity in community composition and assembly, which can be measured with beta-diversity (Zaneveld et al., 2017). We perform a meta-analysis with a big dataset of 27 MAD case studies that cover all five major microbiome habitats (airway, oral, gut, skin, and vaginal) and include most high-profile MADs such as obesity, inflammatory bowel disease (IBD), diabetes, and neurodegenerative diseases. Methodologically, we take advantages of the Hill numbers, which have been recognized as the most appropriate alpha-diversity metrics, and their multiplicative partition of beta-diversity is found to be superior to other existing beta-diversity measures (Chao et al., 2014, 2019; Ma, 2017; Ma and Li, 2018). In addition, we use a very recent framework for assessing and interpreting ecological stochasticity (Ning et al., 2019) to cross-verify the findings from the beta-diversity measures, given that rising stochasticity is considered as another hallmark of the AKP (Zaneveld et al. 2017). Both the approaches essentially measure similarity/dissimilarity among microbiome samples, each with unique advantages. The Hill numbers present the so-termed diversity *profile*, which offer a series of diversity measures corresponding to different diversity order (q = 0, 1, 2, …), weighted differently by species abundances. Therefore, the diversity

[1]Computational Biology and Medical Ecology Lab, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

[2]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China

[3]Lead Contact

*Correspondence: ma@vandals.uidaho.edu

https://doi.org/10.1016/j.isci.2020.101007

Microbiome	Treatments (Healthy vs. Diseased)	Diversity Order	Average of Healthy (H) Treatments	Average of Diseased (D) Treatments	*p* Value (H≠D)	*p* value (H < D) AKP	*p* value (H > D) Anti-AKP
IBD1	Healthy vs. CD	*q=0*	1.678	1.696	0.036	0.018	0.982
		q=1	1.619	1.607	0.850	0.575	0.425
		q=2	1.667	1.602	0.033	0.983	0.017
		q=3	1.691	1.598	0.012	0.994	0.006
	Healthy vs. UC	*q=0*	1.678	1.812	0.000	0.000	1.000
		q=1	1.619	1.792	0.000	0.000	1.000
		q=2	1.667	1.811	0.000	0.000	1.000
		q=3	1.691	1.823	0.000	0.000	1.000
Obesity	Lean vs. Overweight	*q=0*	1.631	1.608	0.000	1.000	0.000
		q=1	1.605	1.591	0.091	0.955	0.045
		q=2	1.655	1.689	0.037	0.019	0.981
		q=3	1.678	1.736	0.008	0.004	0.996
	Lean vs. Obese	*q=0*	1.631	1.618	0.000	1.000	0.000
		q=1	1.605	1.585	0.000	1.000	0.000
		q=2	1.655	1.675	0.001	0.000	1.000
		q=3	1.678	1.719	0.000	0.000	1.000
CRC	Healthy vs. CRC	*q=0*	1.929	1.946	0.000	0.000	1.000
		q=1	1.954	1.971	0.000	0.000	1.000
		q=2	1.977	1.987	0.000	0.000	1.000
		q=3	1.980	1.989	0.000	0.000	1.000
HIV1	Negative vs. ART	*q=0*	1.635	1.675	0.000	0.000	1.000
		q=1	1.697	1.733	0.000	0.000	1.000
		q=2	1.759	1.787	0.005	0.002	0.998
		q=3	1.780	1.813	0.002	0.001	0.999
	Negative vs. Non-ART	*q=0*	1.635	1.730	0.000	0.000	1.000
		q=1	1.697	1.783	0.000	0.000	1.000
		q=2	1.759	1.825	0.000	0.000	1.000
		q=3	1.780	1.847	0.000	0.000	1.000
Type 1 Diabetes (T1D) and Obesity	Normal Healthy vs. Normal T1D	*q=0*	1.535	1.541	0.007	0.003	0.997
		q=1	1.615	1.605	0.546	0.727	0.273
		q=2	1.724	1.669	0.014	0.993	0.007
		q=3	1.759	1.690	0.006	0.997	0.003

Table 1. Wilcoxon Tests for the AKP Effects in the 27 MAD (Microbiome-Associated Diseases) Case Studies Based on the Beta-Diversity in Hill Numbers, Portion of Results Excerpted from Table S3

(Continued on next page)

Microbiome	Treatments (Healthy vs. Diseased)	Diversity Order	Average of Healthy (H) Treatments	Average of Diseased (D) Treatments	*p* Value (H≠D)	*p* value (H < D) AKP	*p* value (H > D) Anti-AKP
	Obesity Healthy vs. Obesity T1D	*q=0*	1.476	1.577	0.000	0.000	1.000
		q=1	1.539	1.606	0.027	0.014	0.987
		q=2	1.599	1.642	0.010	0.005	0.995
		q=3	1.622	1.650	0.019	0.010	0.990
Gout	Healthy vs. Gout	*q=0*	1.666	1.775	0.000	0.000	1.000
		q=1	1.533	1.682	0.000	0.000	1.000
		q=2	1.596	1.696	0.000	0.000	1.000
		q=3	1.636	1.710	0.000	0.000	1.000
■■■	■■■	■■■	■■■	■■■	■■■	■■■	■■■

Table 1. Continued

CD, Crohn's Disease; UC, Ulcerative Colitis; IBD, Inflammatory Bowel Disease; CRC, Colorectal Cancer; ART, Antiretroviral Therapy.

profile in Hill numbers provides comprehensive diversity metrics, on the whole spectrum of commonness versus rarity in terms of species abundance distribution (SAD) (level). Since the SAD is well known to be a highly skewed long-tail distribution, the Hill numbers hence can comprehensively capture the characteristics of community diversity (similarity) at different sections of SAD and produce not only a more comprehensive but also an accurate assessment than traditional diversity measures such as species richness and Shannon entropy.

Ning et al. (2019) recently developed a new null-model based framework for assessing and interpreting ecological stochasticity. With sophisticated computational procedures and algorithms, the framework was actually presented as a simple index—normalized stochasticity ratio (NSR), which makes its application simple but provides a powerful tool for disentangling the relative importance of stochastic and deterministic forces in shaping community diversity. An underlying principle in devising the NSR framework Ning et al. (2019) adopted is that deterministic forces drive the community more similar to dissimilar than null expectation. In the present study, we take advantage of its similarity metrics to detect the AKP effects since the similarity metrics can directly address our problem—determining whether or not the divergence (difference) between the similarity of intra-healthy individuals and the similarity of intra-diseased individuals in their microbiome compositions is statistically significant. If the similarity of "unhappy families" is significantly lower than that of "happy families," we may declare the presence of an AKP effect. In the Ning et al. (2019) framework, they recommended using the Ružička similarity metrics, which is a true distance function based on species abundance (Ružička, 1958), and we use this similarity metric to cross-verify the AKP test results from the beta-diversity approach in the Hill numbers.

RESULTS

Detecting the AKP Effects with Beta-Diversity in Hill Numbers

Table S1 summarized brief descriptions on the 27 MAD (microbiome-associated disease) studies used in this study. We first computed the pairwise beta-diversity for all samples in the healthy and diseased treatments, respectively, for each study, based on Equations (1–4) (Table S2) and then computed the average beta-diversity in the Hill numbers for each treatment of the 27 MAD cases, as well as the *p* values from Wilcoxon tests for the differences in the beta-diversity between the healthy (H) and diseased (D) treatments for each case study (Table S3).

Specifically, Table S2 listed the pairwise beta-diversity for each treatment (either H or D treatment), i.e., computing a beta-diversity value for each pair of samples within each treatment. The pairwise beta-diversity values of samples within each H treatment represent the heterogeneity within the H treatment (intra-H treatment). The pairwise beta-diversity values of samples within each D treatment represent the heterogeneity within the D treatment (intra-D treatment). By conducting Wilcoxon test with the intra-H treatment beta-diversity and intra-H treatment beta-diversity, we can detect the existence of AKP effects. Table S3 exhibited the summary

Category, Treatment & Statistics			$q = 0$	$q = 1$	$q = 2$	$q = 3$
AKP	Healthy	Mean	1.646	1.584	1.635	1.674
		Std. Err.	0.029	0.037	0.035	0.030
	Diseased	Mean	1.707	1.656	1.714	1.753
		Std. Err.	0.028	0.033	0.029	0.025
Anti-AKP	Healthy	Mean	1.707	1.688	1.776	1.811
		Std. Err.	0.032	0.032	0.030	0.029
	Diseased	Mean	1.641	1.620	1.707	1.731
		Std. Err.	0.033	0.038	0.038	0.041
Non-AKP[a]	Healthy	Mean	1.637	1.664	1.699	1.720
		Std. Err.	0.067	0.044	0.043	0.042
	Diseased	Mean	1.634	1.642	1.695	1.726
		Std. Err.	0.067	0.053	0.049	0.047
Total	Healthy	Mean	1.663	1.632	1.689	1.715
		Std. Err.	0.024	0.024	0.023	0.022
	Diseased	Mean	1.664	1.644	1.707	1.736
		Std. Err.	0.024	0.023	0.021	0.020

The percentage for each of the three categories (AKP, Anti-AKP, and Non-AKP), summarized from Table S3 and computed based on Wilcoxon tests of the differences between the intra-H and intra-D treatments in their beta-diversity with Hill numbers

Category	$q = 0$	$q = 1$	$q = 2$	$q = 3$
Percentage of AKP	38.1	50.0	50.0	57.1
Percentage of Anti-AKP	33.3	33.3	28.6	26.2
Percentage of Non-AKP	28.6	26.2	23.8	23.8

Table 2. The Mean and Standard Error of the *Beta-Diversity* (at Different Diversity Order $q = 0$–3) for the Healthy and Diseased Treatments of the Three Categories (AKP, Anti-AKP, Non-AKP), Respectively, Summarized from Table S3; See the Bottom Section for the Percentages of AKP, anti-AKP, and non-AKP

[a]Table S2 contains the pairwise beta-diversity in Hill numbers for the microbiome samples of the 27 MAD case studies. Table S3was obtained from performing Wilcoxon test based on the beta-diversity listed in Table S2.

statistics (mean, standard error) of beta-diversity and p values from the Wilcoxon test. The threshold of p value = 0.05 is used to determine whether a particular MAD case satisfies the AKP ($D > H$), anti-AKP ($D < H$), or non-AKP ($D \approx H$). Table 1 excerpted portion of the results from Table S3, and Table 2 (Figure 1) summarized the test results from Table S3. From Table 2 and Table S3, we summarize the following three findings:

(1) The proportions (percentages) of the 27 MAD studies exhibiting AKP effects varied from 38.1% to 57.1% depending on the diversity order (q), with $q = 0$ (species richness) having the lowest proportion and $q = 3$ having the highest proportion. This suggests that the AKP effects are more sensitive to highly dominant species ($q = 3$) than to common species ($q = 1$) or total species numbers ($q = 0$). In other words, what matter more are the highly dominant species, rather than rare species, in terms of displaying the AKP effects.

(2) The proportions of the 27 MAD studies exhibiting anti-AKP are relatively stable across diversity orders ($q = 0$–3), ranged from 26.2% to 33.3%, with $q = 3$ (very dominant species) exhibiting the lowest proportion and $q = 1$ (common species) and $q = 0$ (total species or species richness) exhibiting the highest proportion. As expected, the pattern of anti-AKP is opposite to that of the AKP. That is, what matter less are the highly dominant species, rather than rare species, in terms of displaying the anti-AKP effects.

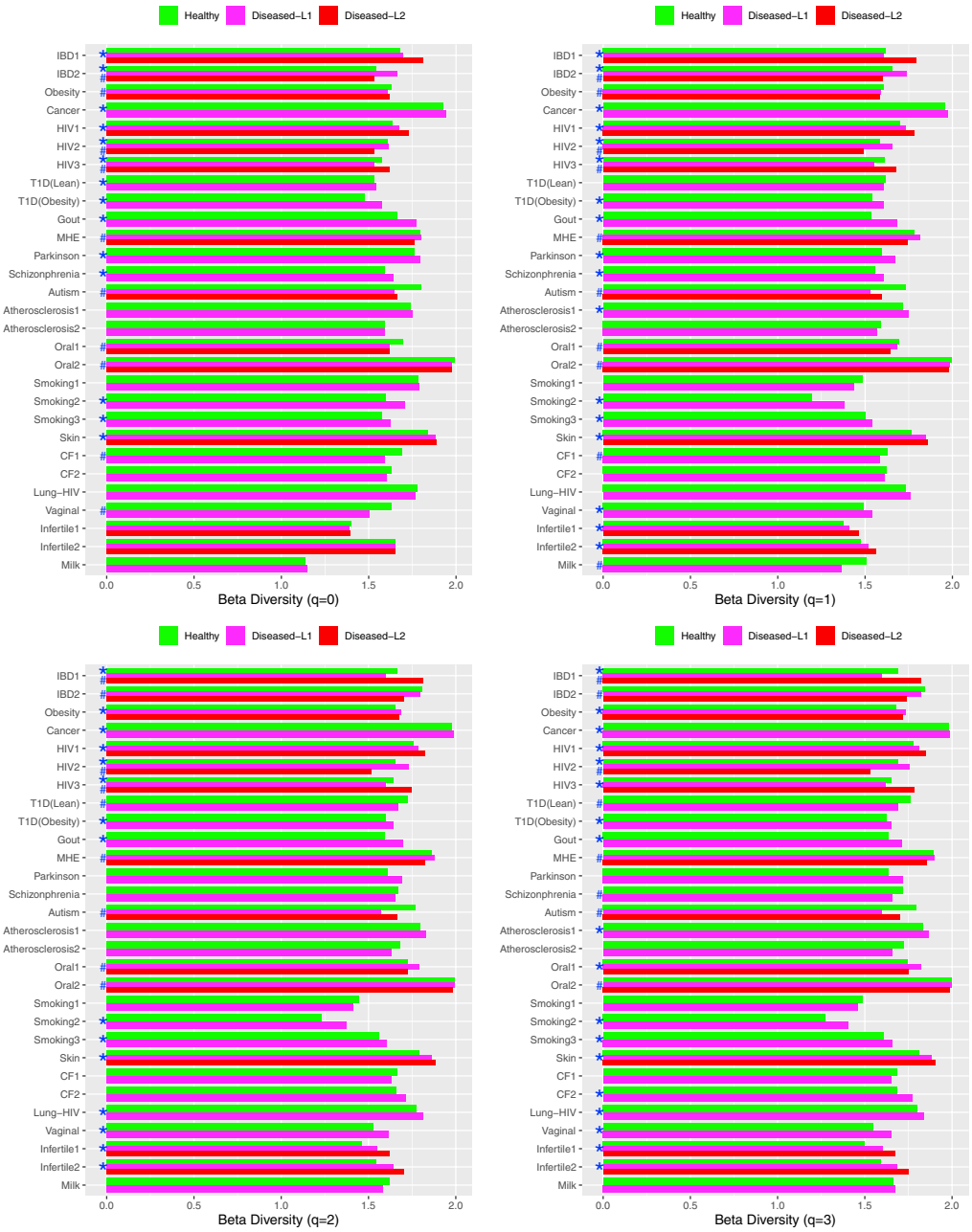


Figure 1. Beta Diversity

The mean beta-diversity (at each diversity order q = 0–3) for each of the 27 MAD (microbiome-associated disease) case studies used for detecting the AKP (Anna Karenina principle): the cases detected with AKP effects are marked with *; the cases detected with Anti-AKP effects are marked with #.

(3) The proportions of the 27 MAD studies exhibiting non-AKP are relatively stable across diversity orders (q = 0–3), ranged from 23.8% to 28.6%, with q = 2–3 (dominant species) exhibiting the equal proportion (23.8%, also the lowest) and q = 0 (species richness) exhibiting the highest proportion. Hence, it appears that the continuum (spectrum) of commonness versus rarity in species abundances does not influence the proportion of non-AKP pattern, given that the continuum determines the weights (of species abundance distribution) used for computing the diversity (Hill numbers) at different diversity orders.

Site	Disease Case Study	Index	Treatments		Similarity (*C*)		*p* Value of Wilcoxon Test		
			Healthy (H)	Diseased (D)	Healthy (H)	Diseased (D)	≠	>	<
Gut	IBD (Inflammatory Bowel Disease)	1	Healthy	CD	0.124	0.128	0.776	0.612	0.388
			Healthy	UC	0.124	0.060	0.000	0.000	1.000
		2	Healthy	CD	0.108	0.065	0.000	0.000	1.000
			Healthy	UC	0.108	0.110	0.001	1.000	0.000
	Obesity	3	Lean	Overweight	0.121	0.125	0.273	0.864	0.136
			Lean	Obesity	0.121	0.132	0.000	1.000	0.000
	Cancer	4	Healthy	Cancer	0.011	0.007	0.000	0.000	1.000
	HIV	5	Negative	Treatment	0.091	0.082	0.004	0.002	0.998
			Negative	Non-treat	0.091	0.064	0.000	0.000	1.000
		6	Negative	Treatment	0.135	0.107	0.000	0.000	1.000
			Negative	Non-treat	0.135	0.173	0.001	0.999	0.001
		7	Negative	Treatment	0.136	0.158	0.012	0.994	0.006
			Negative	Non-treat	0.136	0.095	0.002	0.001	0.999
	T1D (Lean)	8	H	T1D	0.119	0.131	0.274	0.863	0.137
			HO	T1DO	0.156	0.143	0.881	0.560	0.440
	Gout	9	Healthy	Gout	0.164	0.102	0.000	0.000	1.000
	MHE	10	Healthy	MHE	0.062	0.072	0.043	0.979	0.021
			Control	MHE	0.058	0.072	0.000	1.000	0.000
	Parkinson's Disease	11	Healthy	PD	0.131	0.104	0.000	0.000	1.000
	Schizophrenia	12	Healthy	Diseased	0.141	0.117	0.000	0.000	1.000
	Autism	13	Healthy	Autism	0.092	0.160	0.000	1.000	0.000
			Healthy	Neurotypical	0.092	0.132	0.000	1.000	0.000
	Atherosclerosis	14	Healthy	Diseased	0.080	0.071	0.062	0.031	0.969
Intra-H (healthy) treatments				Mean	0.138		NA		
				Std. Err.	0.015				
Intra-D (diseased) treatments				Mean	0.118				
				Std. Err.	0.010				
% With significant differences between intra-H and intra-D treatments					NA		72.5 (29/40)	50.0 (20/40)	30.0 (12/40)
% Without significant differences between intra-H and intra-D treatments							27.5 (11/40)	50.0 (20/40)	70.0 (28/40)
AKP (%)								50% (20/40)	

Table 3. The Means of the Similarity (*C*) for the Intra-healthy Treatment, Intra-diseased Treatment, as well as Wilcoxon Tests for Detecting the AKP Effects: *C*(*H*)>C(*D*) Indicating AKP Effects, *C*(*H*)<C(*D*) Indicating Anti-AKP Effects, *C*(*H*) = C(*D*) Indicating Non-AKP Effects

(Continued on next page)

Site	Disease Case Study	Index	Treatments		Similarity (*C*)		*p* Value of Wilcoxon Test		
			Healthy (H)	Diseased (D)	Healthy (H)	Diseased (D)	≠	>	<
Anti-AKP (%)									30% (12/40)
Non-AKP (%)							27.5% (11/40)		

Table 3. Continued

CD, Crohn's Disease; UC, Ulcerative Colitis; T1D, Type-1 Diabetes; T1DO, Type-1 Diabetes)(Obese); MHE, Minimal Hepatic Encephalopathy; PD, Parkinson's Disease.

Summarized from Table S4.

Detecting the AKP Effects with the Ning et al. Framework for Quantifying Ecological Similarity and Stochasticity

Table S4 listed the results of similarity (*C*) and *p* value from Wilcoxon tests for the differences in the similarity (*C*) between the H and D treatments of 27 MAD case studies. As introduced previously, if the similarity of H treatment is lower (higher) than that of D treatment, then the test indicates the existence of AKP (anti-AKP) effects; otherwise no AKP effects exist. Table 3 is excerpted and summarized from Table S4. It is shown that the percentages of MAD studies with AKP effects and anti-AKP effects were 50% and 30%, respectively. There were approximately 27% of MAD cases without displaying AKP effects. These numbers are rather close to the previous AKP test results from the beta-diversity approach. Both approaches cross-verified each other's findings, but the beta-diversity approach offered more comprehensive results obviously. Hence, in the remainder of this article, we focus on the findings from beta-diversity approach.

Conclusions and Discussion

In summary, our tests based on the beta-diversity profiles (Hill numbers) and a big dataset of 27 MAD studies demonstrate that the AKP effects exist in 50% or more of the MAD cases, except for the species richness ($q = 0$). Furthermore, the effects seem more significant (sensitive) for highly dominant species (OTUs) (57%) and less significant (sensitive) at species richness level (38%). In other words, the far-reaching changes occurred with the very abundant (dominant) species in the microbiome associated with diseases, but the total species number (species richness) is less sensitive to MAD. Therefore, we postulate that the AKP effects can be primarily attributed to highly abundant species and less to rare species. In contrast with the AKP effects, the anti-AKP effects, i.e., lower beta-diversity associated with MAD, are demonstrated in approximately one-fourth of the studied cases. The remaining one-fourth of the MAD cases show no changes in the beta-diversity, namely, the non-AKP cases. Interestingly, the patterns of anti-AKP and non-AKP appear relatively stable across diversity orders. In other words, all species appear to play equal roles in the anti-AKP and non-AKP effects, unlike in the AKP effects, in which highly dominant species can play a more important role.

The approximately 50% or more of positive cases of AKP effects echo the finding in a previous shared-species analysis (SSA) by Ma et al. (2019), which used the same dataset as this study. In the SSA, it was found that, in approximately 50% of the studied cases, shared species between the healthy and diseased microbiome samples were lower than that expected by chance. That is, MADs are associated with increased heterogeneity in species compositions, which is manifested by the reduced number of shared species between the healthy and diseased treatments. In the same study, it was found that, in only approximately one-third of the studied cases, there were significant diversity-disease relationships (DDRs), i.e., significant differences between the healthy and diseased treatments in term of the *alpha*-diversity. In contrast with the 50% of AKP effects, the DDR (measured with *alpha*-diversity) seems less sensitive. This also indicates that AKP effects, measured by beta-diversity, are more sensitive to MADs than the DDR measured with alpha-diversity. Hence, we argue that the AKP effects should be more promising for developing personified diagnosis indicators for diseases, as suggested by Zaneveld et al. (2017), than the routinely computed alpha-diversity indexes in the DDR analysis (Ma et al., 2019).

A key element of the AKP is the bidirectional interactions between the immune system and human microbiome. The notion that stability is associated with the healthy microbiome or host physiology mirrors the notion that dysbiosis (or rising heterogeneity) is often associated with the diseased microbiome or

abnormal host physiology. The disruption or breakdown modes of the bidirectional interactions between the immune system and symbiotic microbiotas can vary from case to case (of diseases), and therefore, AKP effects are unlikely to emerge in all MAD cases. In our opinion, a demonstration of 50% or more positive AKP effects should be a strong piece of evidence supporting the AKP hypothesis in the human MADs.

The AKP hypothesis predicts the rise of stochasticity in animal/human microbiomes due to stress, diseases, and immune system dysfunctions. However, it does not negate the importance of *deterministic* changes in community composition (Zaneveld et al., 2017). In the human microbiome research, the effects of deterministic forces have been well recognized. One such example is the community state types (CSTs) of human vaginal microbiomes (Ravel et al., 2011; Gajer et al., 2012; Doyle et al., 2018). According to the four processes (mechanisms) synthesis of community dynamics (Vellend, 2016; Hanson et al., 2012), which states that it is the four processes (i.e., selection, drift, dispersal, and speciation) that drive the community dynamics, CST should primarily be shaped by deterministic selection forces (such as host genome). Nevertheless, one of the CSTs, the CST-IV that is almost exclusively associated with BV (bacterial vaginosis) disease, showed significant intra-type heterogeneity (variability). For example, the initial classification of the CST-IV (Ravel et al., 2011) was further classified into two sub-types to accommodate more diverse communities associated with BV (Gajer et al., 2012); yet, the classification system could not fully cover more diverse communities (Doyle et al., 2018; Li and Ma, 2019). This example of human vaginal microbiome and associated BV suggests that deterministic selection forces and stochastic disturbances such as BV are often interwoven and it is their joint forces that drive the dynamics of human vaginal microbiome, possibly leading to the rising heterogeneity of community compositions and even displaying AKP effects.

It should be noted that, to increase the robustness of the statistical analyses performed for detecting the AKP, anti-AKP effects, we conducted two additional statistical analyses. One was to apply the FDR (false discovery rate) control to the test results (p values) of the AKP/anti-AKP effects with non-parametric Wilcoxon test exhibited in Table S3, and the FDR-adjusted results were exhibited in Table S5. The FDR control was applied to control the expected proportions of "discoveries" (rejected null hypotheses) that are false (i.e., wrong rejections) during multiple tests. In other words, FDR control was designed to increase the test power to detect true positives, while still controlling the proportion of type I errors (i.e., false positives) at a specified level (Korthauer et al., 2019). Another alternative test we adopted was to use the "Effect Size test" based on Cohen (1988) d-statistic in place of Wilcoxon test, and the results from the alternative test were displayed in Table S6. The advantage of using the effect size test can be to alleviate or even eliminate the potential influence of different sample sizes on the test results. We observed that the differences these two additional tests made were less than 5% on average and did not cause any change of the conclusions we inferred in previous sections.

Limitations of the Study

It is suggested that future tests of AKP theory should be expanded in two frontiers: one is to use temporal data and another is to simultaneously assess and interpret the balance and relative importance of deterministic versus stochastic forces. The former is necessary for investigating the diversity-stability relationship (DSR) (the other side of disease-associated dysbiosis) (Ma and Ellison, 2019), which can be simultaneously utilized to assess the temporal heterogeneity associated with AKP effects. The latter can be much more challenging and requires both methodological innovations and more sophisticated datasets. Obviously, the recent development of the normalized stochasticity ratio framework by Ning et al. (2019) presented a promising methodological advance, and yet, sufficient datasets from longitudinal studies are still rare. Still, as suggested by one anonymous expert reviewer of this article, a third frontier for investigating the AKP theory can be to perform "cross-diseases" tests by asking the question: "do microbiomes associated with different diseases demonstrate greater heterogeneity than control healthy microbiomes across the same studies?" For example, Duvallet et al. (2017) demonstrated that there are both universal and disease-specific signatures in the gut microbiome. With more extensive datasets (e.g., sufficiently large numbers of disease cases for all microbiome types such as gut and skins), the proposed approach used in this study should also be useful for performing cross-diseases AKP tests.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

DATA AND CODE AVAILABILITY

All datasets analyzed in this study are available in public domain and see Table S1 for the access information.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101007.

ACKNOWLEDGMENTS

I appreciate two anonymous expert reviewers and Dr. Stefano Tonzani, the iScience Lead Editor, for their insightful comments and suggestions, which helped to improve my manuscript significantly. I am also indebted to Mr. L.W. Li and Miss Wendy Li of the Chinese Academy of Sciences for their computational support to this study. This study received funding from the following sources: A National Natural Science Foundation of China (NSFC) Grant (No. 31970116) on Medical Ecology of Human Microbiome; The Cloud-Ridge Industry Technology Leader Award; An International Cooperation Grant (YNST) on Genomics & Metagenomics Big Data. The funders played no roles in interpreting the results.

AUTHOR CONTRIBUTIONS

Z.S.M. designed and performed the study and wrote the manuscript.

DECLARATION OF INTERESTS

The author declares no competing interests.

Received: November 21, 2019
Revised: February 26, 2020
Accepted: March 18, 2020
Published: April 24, 2020

REFERENCES

Brüssow, H. (2016). How stable is the human gut microbiota? And why this question matters. Environ. Microbiol. *18*, 2779–2783.

Chao, A., Chiu, C.H., and Jost, L. (2014). Unifying species diversity, phylogenetic diversity, functional diversity and related similarity and differentiation measures through Hill numbers. Annu. Rev. Ecol. Evol. Syst. *45*, 297–324.

Chao, A., Colwell, R.K., Gotelli, N.J., and Thorn, S. (2019). Proportional mixture of two rarefaction/extrapolation curves to forecast biodiversity changes under landscape transformation. Ecol. Lett. *22*, 1913–1922.

Cohen, J. (1988). The Statistical Power Analysis for the Behavioral Sciences (J. of the American Statistical Association).

Doyle, R., Gondwe, A., Fan, Y., Maleta, K., Ashorn, P., Klein, N., and Harris, K. (2018). A lactobacillus-deficient vaginal microbiota dominates postpartum women in rural Malawi. Appl. Environ. Microbiol. *84*, e02150–17.

Duvallet, C., Gibbons, S.M., Gurry, T., Irizarry, R.A., and Alm, E.J. (2017). Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. Nat. Commun. *8*, 1784.

Gajer, P., Brotman, R.M., Bai, G., Sakamoto, J., Schütte, U.M., Zhong, X., Koenig, S.S., Fu, L., Ma, Z.S., Zhou, X., et al. (2012). Temporal dynamics of the human vaginal microbiota. Sci. Transl. Med. *132*, 132ra52.

Giongo, A., Gano, K.A., Crabb, D.B., Mukherjee, N., Novelo, L.L., Casella, G., Drew, J.C., Ilonen, J., Knip, M., Hyöty, H., et al. (2010). Toward defining the autoimmune microbiome for type-1diabetes. ISME J. *5*, 82–91.

Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C., and Martiny, J.B.H. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. Nat. Rev. Microbiol. *10*, 497–506.

Halfvarson, J., Brislawn, C.J., Lamendella, R., Vázquez-Baeza, Y., Walters, W.A., Bramer, L.M., D'Amato, M., Bonfiglio, F., McDonald, D., Gonzalez, A., et al. (2017). Dynamics of the human gut microbiome in inflammatory bowel disease. Nat. Microbiol. *2*, 17004.

Korthauer, K., Kimes, P.K., Duvallet, C., Reyes, A., Subramanian, A., Teng, M., Shukla, C., Alm, E.J., and Hicks, S.C. (2019). A practical guide to methods controlling false discoveries in computational biology. Genome Biol. *20*, 118.

Levy, M., Kolodziejczyk, A.A., Thaiss, C.A., and Elinav, E. (2017). Dysbiosis and the immune system. Nat. Rev. Immunol. *17*, 219–232.

Li, W.D., and Ma, Z.S. (2019). Diversity scaling of human vaginal microbial communities. Zoolog. Res. *40*, 587–594.

Lotter, H., and Altfeld, M. (2019). Sex differences in immunity. Semin. Immunopathol. *41*, 133–135.

Ma, Z.S. (2017). Measuring microbiome diversity and similarity with Hill Numbers. Chapter 8. In Metagenomics, Nagarajan Muniyandi, ed. (Elsevier). https://doi.org/10.1016/B978-0-08-102268-9.00008-2.

Ma, Z.S., and Li, L.W. (2018). Measuring metagenome diversity and similarity with Hill numbers. Mol. Ecol. Resour. *18*, 1339–1355.

Ma, Z.S., Li, L.W., and Gotelli, N.J. (2019). Diversity-disease relationships and shared species analyses for human microbiome-associated diseases. ISME J. *13*, 1911–1919.

Ma, Z.S., and Ellison, A.M. (2019). Dominance network analysis provides a new framework for studying the diversity-stability relationship. Ecol. Monogr. *89*, e01358.

Ning, D., Deng, Y., Tiedje, J.M., and Zhou, J. (2019). A general framework for quantitatively assessing ecological stochasticity. Proc. Natl. Acad. Sci. U S A *116*, 16892–16898.

Ravel, J., Gajer, P., Abdo, Z., Schneider, G.M., Koenig, S.S.K., McCulle, S.L., Karlebach, S., Gorle, R., Russell, J., Tacket, C.O., et al. (2011). Vaginal microbiome of reproductive-age women. Proc. Natl. Acad. Sci. U S A *108* (*Suppl. 1*), 4680–4687.

Rizzetto, L., Fava, F., Tuohy, K.M., and Selmi, C. (2018). Connecting the immune system, systemic

chronic inflammation and the gut microbiome: the role of sex. J. Autoimmun. *92*, 12–34.

Rooks, M.G., and Garrett, W.S. (2016). Gut microbiota, metabolites and host immunity. Nat. Rev. Immunol. *16*, 341–352.

Ružička, M. (1958). Anwendung mathematisch-statistischer methoden in der Geobotanik (Synthetische Bearbeitung von Aufnahmen). Biológia, Bratislava *13*, 647–661.

Thaiss, C.A., Zmora, N., Levy, M., and Elinav, E. (2016). The microbiome and innate immunity. Nature *535*, 65–74.

Vellend, M. (2016). The Theory of Ecological Communities (Princeton University Press).

Williams, B., Landay, A., and Presti, R.M. (2016). Microbiome alterations in HIV infection a review. Cell. Micobiol. *18*, 645–651.

Zaneveld, J.R., Mcminds, R., and Vega Thurber, R. (2017). Stress and stability: applying the Anna Karenina principle to animal microbiomes. Nat. Microbiol. *2*, 17121.

iScience, Volume *23*

Supplemental Information

Testing the Anna Karenina Principle in Human Microbiome-Associated Diseases

Zhanshan (Sam) Ma

Supplemental Information

Supplemental Tables S1, S4 & S6

(Supplemental Tables S2, S3 & S5 are in Excel tables)

Table S1. The 27 datasets of human microbiome associated diseases (MADs) case studies based on the 16s-rRNA amplicon sequencing, utilized for testing the AKP (Anna Karenina principle). Related to the Transparent Methods and Table Figure 1.

Index	Sites	Disease	**Treatments** (the sample size for each treatment is parenthesized)	References*
1	Gut	IBD (Inflammatory Bowel Disease)	Crohn's disease (CD) (18), Ulcerative colitis (UC) (38), Healthy (18)	Papa et al (2012)
2			CD (251) , UC (324), Healthy (62)	Halfvarson et al (2017)
3		Obesity	Lean (61), Obese (196) and Overweight (24)	Turnbaugh et al (2009)
4		CRC	Cancer (46) vs. Healthy (56)	Wang et al (2012)
5		HIV	Non-treatment (7) vs. ART treatment (11)	Neff et al (2018)
6			WithART(14), WithoutART(12), Healthy (22)	Lozupone et al (2013)
7			HIV Negative (20) *vs.* Positive (40)	McHardy et al (2013)
8		Type 1 diabetes and Obesity	Normal T1D (33) vs Normal Healthy (33)	Kim, Jane: https://clinicaltrials.gov/ct2/show/NCT02938806
			Obesity T1D (24) vs. Obesity Healthy (26)	
9		Gout	Disease (41) vs. Healthy (42)	Guo et al (2015)
10		MHE	MHE (25), Cirrhotic (25), Healthy (25)	Zhang et al (2013)
11		Parkinson's Disease	Disease (205) vs. Healthy (133)	Hill-Burns et al 2017)
12		Schizophrenia	Disease (25). vs. Healthy (25)	Dilip Jeste (UC San Diego): https://profiles.ucsd.edu/dilip.jeste
13		Autism, Neurotypical	Autism (88), Neurotypical (41), Healthy (14)	Kang et al (2017)
14		Atherosclerosis	Disease (15) vs. Healthy (15)	Koren et al (2010)
15	Oral	Atherosclerosis	Disease (14) vs. Healthy (15)	Koren et al (2010)
16		Periodontitis	PB (22), PnB (22), Healthy (17)	Abusleme et al (2013)
17			Disease (29), Control (29), Healthy (29)	Griffen et al (2012)
18		Smoking	Smoking (6) vs. Non-smoking (9)	Lazarevic et al (2010)
19			Smoking (74) vs. Non-smoking (72)	Charlson et al (2010)
20	Nostril	Smoking	Smoking (74) vs. Non-smoking (71)	Charlson et al (2010)
21	Skin	Psoriasis	Disease (77), Control (83), Healthy (76)	Alekseyenko et al (2013)
22	Lung	Cystic Fibrosis (CF)	End of Treatment (23) vs. Exacerbation (23)	Fodor et al (2012)
23			Disease (16) vs. Healthy (10)	Blainey et al 2012)
24		HIV	Disease (82) vs. Healthy (77)	Lozupone et al (2013)
25	Vaginal	Bacterial Vaginosis (BV)	BV (Nugent score=7-10) *vs*. healthy (Nugent score=0-6)	Srinivasan et al (2012)
26	Semen	Infertile	Abnormal (33), Subnormal (28), Normal (35). Genus level and species level	Weng et al 2014
27	Milk	Mastitis	Mastitis (4) *vs*. Healthy (16)	Urbaniak (2015)

*References (Cited in Table S1 for Data Sources)

Papa E, Docktor M, Smillie C, et al. 2012. Non-Invasive Mapping of the Gastrointestinal Microbiota Identifies Children with Inflammatory Bowel Disease. *Plos One*, 7(6):e39242.

Halfvarson J, Brislawn CJ, Lamendella R, et al. 2017. Dynamics of the human gut microbiome in Inflammatory Bowel Disease. *Nature Microbiology*, 2:17004.

Turnbaugh PJ, Hamady M, Yatsunenko T, et al. 2009. A core gut microbiome in obese and lean twins. *Nature*, 457(7228):480.

Wang T, Cai G, Qiu Y, et al. 2012. Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *The ISME Journal*, 6(2):320-329.

Neff CP, Krueger O, Xiong K, et al. 2018. Fecal Microbiota Composition Drives Immune Activation in HIV-infected Individuals. *Ebiomedicine*, 30.

Lozupone C, Cota-Gomez A, Palmer B E, et al. 2013. Widespread colonization of the lung by Tropheryma whipplei in HIV infection. *American Journal of Respiratory & Critical Care Medicine*, 187(10):1110-1117.

Mchardy IH, Li X, Tong M, et al. 2013. HIV Infection is associated with compositional and functional shifts in the rectal mucosal microbiota. *Microbiome*, 1(1):26.

Guo Z, Zhang J, Wang Z, et al. 2016. Intestinal Microbiota Distinguish Gout Patients from Healthy Humans. *Sci Rep*, 6:20602.

Zhang Z, Zhai H, Geng J, et al. 2013. Large-scale survey of gut microbiota associated with MHE Via 16S rRNA-based pyrosequencing. *American Journal of Gastroenterology,* 108(10):1601-1611.

Hill-Burns EM, Debelius JW, Morton JT et al. 2017. Parkinson's Disease and PD Medications Have Distinct Signatures of the Gut Microbiome. *Movement Disorders Official Journal of the Movement Disorder Society*, 32(5):739.

Kim, Jane: for Study Case No 8 in Table S1 https://clinicaltrials.gov/ct2/show/NCT02938806

Dilip Jeste (UC San Diego) for Study Case No 12 in Table S1: https://profiles.ucsd.edu/dilip.jeste

Kang DW, Adams JB, Gregory AC, et al. 2017. Microbiota Transfer Therapy alters gut ecosystem and improves gastrointestinal and autism symptoms: an open-label study. *Microbiome*, 5(1):10.

Koren O, Klaenhammer TR. 2011. Human oral, gut, and plaque microbiota in patients with atherosclerosis. *PNAS*, 108(Suppl 1):4592-4598.

Koren O, Klaenhammer TR. 2011. Human oral, gut, and plaque microbiota in patients with atherosclerosis. *PNAS*, 108(Suppl 1):4592-4598.

Abusleme L, Dupuy AK, Dutzan N, et al. 2013. The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. *The ISME Journal*, 7(5):1016-1025.

Griffen AL, Beall CJ, Campbell JH, et al. 2012. Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. *The ISME Journal*, 6(6):1176.

Lazarevic V, Whiteson K, Hernandez D, et al. 2010. Study of inter- and intra-individual variations in the salivary microbiota. *BMC Genomics*, 11(1):1-11.

Charlson ES, Chen J, Custersallen R, et al. 2010. Disordered microbial communities in the upper respiratory tract of cigarette smokers. *Plos One*, 5(12):e15216.

Charlson ES, Chen J, Custersallen R, et al. 2010. Disordered microbial communities in the upper respiratory tract of cigarette smokers. *Plos One*, 5(12):e15216.

Alekseyenko AV, Perezperez GI, Souza AD, et al. 2013. Community differentiation of the cutaneous microbiota in psoriasis. *Microbiome*,1,1(2013-12-23), 1(1):31-31.

Fodor AA, Klem ER, Gilpin DF, et al. 2012. The Adult Cystic Fibrosis Airway Microbiota Is Stable over Time and Infection Type, and Highly Resilient to Antibiotic Treatment of Exacerbations. *Plos One*, 7(9):e45001.

Blainey PC, Milla CE, Cornfield DN, et al. 2012. Quantitative analysis of the human airway microbial ecology reveals a pervasive signature for cystic fibrosis. *Science Translational Medicine*, 4(153):153ra130.

Lozupone C, Cota-Gomez A, Palmer B E, et al. 2013. Widespread colonization of the lung by Tropheryma whipplei in HIV infection. *American Journal of Respiratory & Critical Care Medicine*, 187(10):1110-1117.

Srinivasan S, Hoffman NG, Morgan MT, et al. 2012. Bacterial Communities in Women with Bacterial Vaginosis: High Resolution Phylogenetic Analyses Reveal Relationships of Microbiota to Clinical Criteria. *Plos One*, 7(6):e37818.

Weng SL, Chiu CM, Lin FM, et al. 2014. Bacterial Communities in Semen from Men of Infertile Couples: Metagenomic Sequencing Reveals Relationships of Seminal Microbiota to Semen Quality. *Plos One*, 9(10):e110152.

Urbaniak C, Mcmillan A, Angelini M, et al. 2014. Effect of chemotherapy on the microbiota and metabolome of human milk, a case report. *Microbiome*, 2(1):24.

Table S4. The mean and median of the similarity (*C*) for the intra-healthy treatment, intra-diseased treatment, respectively, as well as Wilcoxon tests for detecting the AKP effects based on *C*: *C*(*H*)>C(*D*) indicating AKP effects, *C*(*H*)<C(*D*) indicating anti-AKP effects, *C*(*H*)=C(*D*) indicating non-AKP effects. Related to Table 3.

Site	Disease Case Study	Index	Treatments		Similarity (*C*) (Mean)		Similarity (*C*) (Median)		*P*-value of Wilcoxon Test		
			Healthy (H)	Diseased (D)	Healthy (H)	Diseased (D)	Healthy (H)	Diseased (D)	≠	>	<
\Gut	IBD (Inflammatory Bowel Diseased)	1	Healthy	CD	0.124	0.128	0.122	0.106	0.776	0.612	0.388
			Healthy	UC	0.124	0.060	0.122	0.034	0.000	0.000	1.000
		2	Healthy	CD	0.108	0.065	0.082	0.044	0.000	0.000	1.000
			Healthy	UC	0.108	0.110	0.082	0.102	0.001	1.000	0.000
	Obesity	3	Lean	Overweight	0.121	0.125	0.112	0.112	0.273	0.864	0.136
			Lean	Obesity	0.121	0.132	0.112	0.126	0.000	1.000	0.000
	Cancer	4	Healthy	Cancer	0.011	0.007	0.007	0.004	0.000	0.000	1.000
	HIV	5	Negative	Treatment	0.091	0.082	0.075	0.073	0.004	0.002	0.998
			Negative	Non-treat	0.091	0.064	0.075	0.032	0.000	0.000	1.000
		6	Negative	Treatment	0.135	0.107	0.118	0.071	0.000	0.000	1.000
			Negative	Non-treat	0.135	0.173	0.118	0.169	0.001	0.999	0.001
		7	Negative	Treatment	0.136	0.158	0.101	0.125	0.012	0.994	0.006
			Negative	Non-treat	0.136	0.095	0.101	0.075	0.002	0.001	0.999
	T1D (Lean)	8	H	T1D	0.119	0.131	0.122	0.118	0.274	0.863	0.137
			HO	T1DO	0.156	0.143	0.141	0.146	0.881	0.560	0.440
	Gout	9	Healthy	Gout	0.164	0.102	0.139	0.073	0.000	0.000	1.000
	MHE	10	Healthy	MHE	0.062	0.072	0.054	0.056	0.043	0.979	0.021
			Control	MHE	0.058	0.072	0.027	0.056	0.000	1.000	0.000
	Parkinson's Diseased	11	Healthy	PD	0.131	0.104	0.119	0.093	0.000	0.000	1.000
	Schizophrenia	12	Healthy	Diseased	0.141	0.117	0.129	0.104	0.000	0.000	1.000
	Autism	13	Healthy	Autism	0.092	0.160	0.073	0.149	0.000	1.000	0.000
			Healthy	Neurotypical	0.092	0.132	0.073	0.101	0.000	1.000	0.000
	Atherosclerosis	14	Healthy	Diseased	0.080	0.071	0.074	0.062	0.062	0.031	0.969
Oral	Atherosclerosis	15	Healthy	Diseased	0.122	0.130	0.096	0.115	0.107	0.947	0.054
	Periodontitis	16	Healthy	PB	0.101	0.094	0.051	0.080	0.002	0.999	0.001
			Healthy	PnB	0.101	0.108	0.051	0.098	0.000	1.000	0.000
		17	Healthy	Diseased	NA	NA	NA	NA	NA	NA	NA
			Control	Diseased	NA	NA	NA	NA	NA	NA	NA
	Smoking	18	Non-Smoker	Smoker	0.184	0.223	0.146	0.163	0.506	0.754	0.253
		19	Non-Smoker	Smoker	0.423	0.287	0.399	0.265	0.000	0.000	1.000
		20	Non-Smoker	Smoker	0.175	0.160	0.167	0.151	0.000	0.000	1.000
Skin	Psoriasis	21	Control	Lesion	0.065	0.038	0.048	0.025	0.000	0.000	1.000
			Normal	Lesion	0.041	0.038	0.027	0.025	0.080	0.040	0.960
Lung	Cystic Fibrosis (CF)	22	Treated	Exacerbation	0.196	0.201	0.052	0.065	0.010	0.995	0.005
		23	Healthy	Diseased	0.124	0.117	0.121	0.108	0.878	0.439	0.563
	HIV	24	Negative	Positive	0.090	0.079	0.033	0.022	0.200	0.100	0.900
Vaginal	BV	25	Healthy	BV	0.251	0.159	0.178	0.129	0.000	0.000	1.000
Semen	Infertile (Genus level)	26	Normal	Subnormal	0.247	0.218	0.253	0.194	0.000	0.000	1.000
			Normal	Abnormal	0.247	0.180	0.253	0.156	0.000	0.000	1.000
	Infertile (Species level)		Normal	Subnormal	0.186	0.163	0.189	0.143	0.000	0.000	1.000
			Normal	Abnormal	0.186	0.137	0.189	0.116	0.000	0.000	1.000
Milk	Mastitis	27	Healthy	Mastitis	0.178	0.234	0.153	0.220	0.098	0.952	0.049
Intra-H (Healthy) Treatments				Mean	0.138						

	Std. Err.	0.015						
Intra-D (Diseased) Treatments	Mean	0.118						
	Std. Err.	0.010						
% With Significant Differences between Intra-H and Intra-D treatments						72.5 (29/40)	50.0 (20/40)	30.0 (12/40)
% Without Significant Differences between Intra-H and Intra-D treatments						27.5 (11/40)	50.0 (20/40)	70.0 (28/40)
AKP (%)							50 (20/40)	
Anti-AKP (%)								30 (12/40)
Non-AKP (%)						27.5 (11/40)		

Table S6. Tests of the AKP effects with the "Effect Size Test" [Cohen (1988) *d*-statistic] based on Ning *et al*. (2019) similarity (*C*) and the beta diversity in Hill numbers*. Related to the Discussion Section and Figure 1.

Site	**Disease Case Study**	**Index**	**Treatments**		**Similarity (*C*)**		**Beta diversity**							
							***q*=0**		***q*=1**		***q*=2**		***q*=3**	
			Healthy (H)	**Diseased (D)**	***d***	***p***	***d***	***p***	***d***	***p***	***d***	***p***	***d***	***p***
Gut	IBD (Inflammatory Bowel Diseased)	1	Healthy	CD	-0.13	0.24	-0.16	0.17	0.07	0.57	0.27	0.02	0.32	0.01
			Healthy	UC	0.91	0.00	-1.14	0.00	-0.93	0.00	-0.58	0.00	-0.46	0.00
		2	Healthy	CD	0.50	0.00	NA	NA	NA	NA	NA	NA	NA	NA
			Healthy	UC	-0.01	0.56	NA	NA	NA	NA	NA	NA	NA	NA
	Obesity	3	Lean	Overweight	-0.08	0.21	0.32	0.00	0.10	0.11	-0.15	0.02	-0.22	0.00
			Lean	Obesity	-0.12	0.00	0.19	0.00	0.16	0.00	-0.08	0.00	-0.14	0.00
	Cancer	4	Healthy	Cancer	0.50	0.00	-0.47	0.00	-0.33	0.00	-0.15	0.00	-0.12	0.00
	HIV	5	Negative	Treatment	0.17	0.00	-0.23	0.00	-0.17	0.00	-0.12	0.04	-0.13	0.03
			Negative	Non-treat	0.44	0.00	-0.58	0.00	-0.39	0.00	-0.28	0.00	-0.29	0.00
		6	Negative	Treatment	0.30	0.02	-0.07	0.60	-0.35	0.01	-0.24	0.05	-0.21	0.10
			Negative	Non-treat	-0.51	0.00	0.92	0.00	0.52	0.00	0.50	0.00	0.50	0.00
		7	Negative	Treatment	-0.25	0.02	0.40	0.00	0.27	0.01	0.16	0.12	0.10	0.35
			Negative	Non-treat	0.37	0.00	-0.39	0.00	-0.31	0.00	-0.36	0.00	-0.43	0.00
	T1D (Lean)	8	H	T1D	-0.08	0.20	-0.04	0.55	0.05	0.41	0.21	0.00	0.24	0.00
			HO	T1DO	0.25	0.00	-0.53	0.00	-0.27	0.00	-0.12	0.14	-0.06	0.44
	Gout	9	Healthy	Gout	0.66	0.00	-1.25	0.00	-0.73	0.00	-0.33	0.00	-0.22	0.00
	MHE	10	Healthy	MHE	-0.28	0.00	0.35	0.00	0.22	0.01	0.18	0.04	0.17	0.04
			Control	MHE	-0.40	0.00	0.56	0.00	0.41	0.00	0.24	0.00	0.18	0.03
	Parkinson's Diseased	11	Healthy	PD	0.39	0.00	NA	NA	NA	NA	NA	NA	NA	NA
	Schizophrenia	12	Healthy	Diseased	0.32	0.00	-0.91	0.00	-0.35	0.00	0.08	0.30	0.21	0.01
	Autism	13	Healthy	Autism	-0.73	0.00	0.75	0.00	0.74	0.00	0.70	0.00	0.68	0.00
			Healthy	Neurotypical	-0.57	0.00	0.68	0.00	0.48	0.00	0.35	0.00	0.31	0.01
	Atherosclerosis	14	Healthy	Diseased	0.29	0.04	-0.23	0.10	-0.33	0.02	-0.24	0.09	-0.23	0.10
Oral	Atherosclerosis	15	Healthy	Diseased	-0.18	0.21	0.03	0.86	0.13	0.37	0.21	0.16	0.24	0.10
	Periodontitis	16	Healthy	PB	-0.18	0.10	0.60	0.00	0.03	0.78	-0.23	0.04	-0.25	0.02
			Healthy	PnB	-0.30	0.01	0.61	0.00	0.19	0.07	0.00	0.97	-0.01	0.89
		17	Healthy	Diseased	-0.25	0.00	0.27	0.00	0.20	0.00	0.14	0.04	0.11	0.13
			Control	Diseased	-0.06	0.37	0.03	0.70	0.09	0.18	0.14	0.04	0.15	0.03
	Smoking	18	Non-Smoker	Smoker	-0.25	0.42	-0.09	0.77	0.22	0.49	0.07	0.82	0.01	0.97
		19	Non-Smoker	Smoker	0.77	0.00	-1.00	0.00	-0.84	0.00	-0.49	0.00	-0.41	0.00
		20	Non-Smoker	Smoker	0.20	0.00	-0.39	0.00	-0.19	0.00	-0.18	0.00	-0.21	0.00
Skin	Psoriasis	21	Control	Lesion	0.62	0.00	-0.59	0.00	-0.64	0.00	-0.51	0.00	-0.45	0.00

			Normal	Lesion	0.05	0.04	-0.04	0.15	-0.07	0.01	-0.13	0.00	-0.14	0.00
Lung	Cystic Fibrosis (CF)	22	Treated	Exacerbation	-0.11	0.20	0.62	0.00	0.11	0.17	0.08	0.31	0.08	0.34
		23	Healthy	Diseased	0.02	0.93	0.20	0.31	0.05	0.79	-0.24	0.21	-0.35	0.08
	HIV	24	Negative	Positive	0.10	0.01	0.06	0.16	-0.09	0.02	-0.13	0.00	-0.13	0.00
Vaginal	BV	25	Healthy	BV	0.18	0.00	0.75	0.00	-0.20	0.00	-0.29	0.00	-0.32	0.00
Semen	Infertile (Genus level)	26	Normal	Subnormal	0.18	0.01	0.12	0.07	-0.20	0.00	-0.32	0.00	-0.34	0.00
			Normal	Abnormal	0.48	0.00	0.10	0.09	-0.48	0.00	-0.60	0.00	-0.59	0.00
	Infertile (Species level)		Normal	Subnormal	0.19	0.00	0.02	0.78	-0.25	0.00	-0.33	0.00	-0.30	0.00
			Normal	Abnormal	0.47	0.00	0.05	0.37	-0.51	0.00	-0.59	0.00	-0.56	0.00
Milk	Mastitis	27	Healthy	Mastitis	-0.86	0.04	-0.23	0.61	1.19	0.01	0.29	0.53	0.01	0.99
% With Significant Differences between Intra-H and Intra-D treatments					76.2 (32/42)		61.5 (24/39)		74.4 (29/39)		74.4 (29/39)		71.8 (28/39)	
% Without Significant Differences between Intra-H and Intra-D treatments					23.8 (10/42)		38.5 (15/39)		25.6 (10/39)		25.6 (10/39)		28.2 (11/39)	
AKP (%)					52.4 (22/42)		28.2 (11/39)		51.3 (20/39)		51.3 (20/39)		48.7 (19/39)	
Anti-AKP (%)					23.8 (10/42)		33.3 (13/39)		23.1 (9/39)		23.1 (9/39)		23.1 (9/39)	
Non-AKP (%)					23.8 (10/42)		38.5 (15/39)		25.6 (10/39)		25.6 (10/39)		28.2 (11/39)	

*Cohen J (1988) The Statistical Power Analysis for the Behavioral Sciences. *J. of the American Statistical Association*.

For the "Effect Size Tests" based on Similarity (C), *P*-value<0.05 indicates there is significant difference between the H (healthy) and D (diseased) treatments*, if d* (similarity)>0 signaling AKP effects, and *d*(similarity)<0 signaling anti-AKP effects. *P*-value>0.05 indicates Non-AKP effects.

For the "Effect Size Tests" based on Beta-diversity, *P*-value<0.05 indicates there is significant difference between the H (healthy) and D (diseased) treatments*, if d* (beta-diversity)<0 signaling AKP effects, and *d*(similarity)>0 signaling anti-AKP effects. *P*-value>0.05 indicates Non-AKP effects.

TRANSPARENT METHODS

Datasets of Microbiome Associated Diseases (MADs)

Table S1 of the online supplementary information (OSI) exhibited a brief description on the metagenomic datasets of the 27 MAD (microbiome associated disease). These datasets of 16s-rRNA sequencing reads cover all five major habitats and two major body fluids of the human microbiomes (gut, oral, lung, skin, vaginal, milk and semen). They include majority of the high-profile MADs such as obesity, IBD, diabetes, gout, HIV, Parkinson's disease, Schizophrenia, autism, periodontitis, BV (bacterial vaginosis), CRC (colorectal cancer), mastitis, and infertility. These publically available datasets represent for state-of-the-art studies in MADs and were previously used in a meta-analysis by Ma *et al*. (2019) for diversity-disease relationship (DDR) analysis.

Alpha-Diversity in Hill Numbers

Hill numbers are derived from Renyi's entropy and defined with the following formula:

$$ {}^{q}D = \left(\sum_{i=1}^{N} p_i^q \right)^{1/(1-q)} \qquad (1) $$

where N is the number of species or OTUs, p_i is the relative abundance of species i, q is the order number of alpha-diversity, ${}^{q}D$ is the alpha-diversity at diversity order q, *i.e.*, the Hill numbers of q-th order. A recent consensus is that Hill numbers offer the most appropriate measures for alpha-diversities and their multiplicative partition is advantageous over other existing definitions for beta-diversities (Chao *et al*. 2014, 2019, Ma 2017).

The Hill number cannot be defined when q=1, but its limit exists as q approaches to *1*:

$$ {}^{1}D = \lim_{q \to 1} {}^{q}D = \exp\left(-\sum_{i=1}^{N} p_i \log(p_i) \right) \qquad (2) $$

The diversity order (q) reflects the sensitivity of Hill numbers to the relative frequencies of species abundances. If q=0, species abundances do not matter at all since p_i^0 =1, and ${}^{0}D$=*N, i.e.,* the number of species in community or *species richness* of community. If q=1, ${}^{1}D$ is the exponential function of Shannon entropy, and measures the number of typical or *common* species. If q=2, ${}^{2}D$ equals the reciprocal of Simpson index, and it measures the number of *dominant* or very abundant species in the community. Using Hill numbers as biodiversity measures has multiple advantages, and here we mentioned two of the most significant benefits. Hill numbers are in the units of species or species equivalents weighted differently at different diversity orders

(q) and the weights are determined by the species abundance distribution (SAD). For this reason, all Hill numbers at the same diversity order are in the same units and fully comparable across communities, and furthermore the Hill numbers at different diversity orders constitute the so-termed diversity profiles, *i.e.*, a series of Hill numbers at different diversity orders. Therefore, Hill numbers provide comprehensive measures for biodiversity in consideration of the whole spectrum of rarity vs. commonness of species abundance distribution. It also makes the issues associated with choosing different existing diversity orders moot because Hill numbers are functions of the major existing diversity measures as explained previously.

Gamma-Diversity in Hill numbers

While the previously defined alpha-diversity is aimed to measure the diversity within a (single) community, the following gamma-diversity is defined to measure the *total* diversity of pooled multiple (N) communities.

Assuming there are M communities, which can be the community samples from M individuals, the gamma diversity of M communities is defined as:

$$ {}^{q}D_{\gamma} = \left(\sum_{i=1}^{N} \left(\overline{p_i} \right)^{q} \right)^{1/(1-q)}, \qquad (3) $$

where $\overline{p}_i$ is the relative abundance of the i-th species ($i=1, 2,..., N$) in the pooled assemblage of M communities. Comparing Eqn. (3) for gamma diversity with Eqn. (1) for alpha diversity indicates that the gamma diversity is the Hill numbers based on the *relative* abundance of the i-th species gene in the pooled assemblage of M communities.

Beta-Diversity in Hill numbers

Traditionally, there are two types of definitions for the beta-diversity: the additive partition or multiplicative partition of gamma diversity into assumingly independent alpha-diversity and beta-diversity. Recent consensus has supported the use of multiplicatively defined beta-diversity in terms of Hill numbers, *i.e.*, by partitioning gamma diversity into the product of alpha and beta diversities, in which both alpha (${}^{q}D_{\alpha}$) and gamma (${}^{q}D_{\gamma}$)diversities are measured with the Hill numbers (Chao et al. 2014, 2019, Ma 2017). That is:

$$ {}^{q}D_{\beta} = {}^{q}D_{\gamma} / {}^{q}D_{\alpha}. \qquad (4) $$

Detecting AKP effects with Beta-Diversity in Hill numbers

We use beta-diversity (in Hill numbers) as metrics for determining the AKP effects because higher beta-diversity reveals rising heterogeneity in community composition, which is

considered as a hallmark of the AKP (Zaneveld *et al*. 2017). We conduct non-parametric significant test (Wilcoxon test) to determine the statistical significance of the differences in beta-diversity between the healthy (H) and diseased (D) treatments. If the pair-wise beta-diversity within H treatment is significantly lower (higher) than pair-wise beta-diversity within D treatment (P<0.05 from Wilcoxon test), then the test reveals AKP (anti-AKP) effects; otherwise, there is no MAD effect.

Ning *et al.* (2019) framework for quantifying ecological similarity and stochasticity

We use a new framework recently developed by Ning *et al*. (2019) for quantifying ecological stochasticity to cross-verify the results from neutral theory modeling. The theoretical basis of their mathematical framework is that deterministic processes should drive ecological communities more similar or dissimilar than null expectation, and they formulated a sophisticated procedure to implement a null model for quantifying stochasticity. In this study, to test the AKP, we only need portion of their framework, their similarity metric. Ning *et al*. (2019) recommended using the Ružička similarity metrics, which is a true distance function based on species abundance (Ružička 1958). Ružička similarity is a generalization of Jaccard binary similarity coefficient. Let C_{ij} represent the observed similarity between the i-th and j-th community,

$$C_{ij} = \frac{\sum_S \min(p_k^i, p_k^j)}{\sum_S \max(p_k^i, p_k^j)} \tag{5}$$

where S is the number of species, p_k^i and p_k^j are the relative abundance of k-th species in the i-th and j-th community.

We use Wilcoxon test to determine the statistical significance of the difference between the healthy and diseased treatments in their intra-treatment sample similarities. If the intra-treatment similarity (C) of the healthy treatment is significantly lower (higher) than that of the intra-treatment similarity (C) of the diseased treatment, it indicates that the MAD follows the AKP (anti-AKP) principle; otherwise, there is no MAD effect.

References (Cited in "Transparent Methods")

Chao, A., Chiu, C.H., Jost, L. (2014). Unifying species diversity, phylogenetic diversity, functional diversity and related similarity and differentiation measures through Hill numbers. *Annual Reviews of Ecology, Evolution, and Systematics*, **45**: 297–324.

Chao, A., Colwell, R.K., Gotelli, N.J., Thorn, S. (2019). Proportional mixture of two rarefaction/extrapolation curves to forecast biodiversity changes under landscape transformation. *Ecology Letters* https://doi.org/10.1111/ele.13322

Ma, Z.S. (2017). Measuring microbiome diversity and similarity with Hill Numbers. Chapter 8, *in* "*Metagenomics*" https://doi.org/10.1016/B978-0-08-102268-9.00008-2, Elsevier

Ma, Z.S., Li, L.W. (2018). Measuring metagenome diversity and similarity with Hill numbers. *Molecular Ecology Resources*, https://doi.org/10.1111/1755-0998.12923

Ma, Z.S., Li, L.W., Gotelli, N.J. (2019). Diversity-disease relationships and shared species analyses for human microbiome-associated diseases, *The ISME Journal*. DOI: 10.1038/s41396-019-0395-y

Ning, D., Deng, Y., Tiedje, J.M., Zhou, J. (2019). A general framework for quantitatively assessing ecological stochasticity. *PNAS*, 116 (34): 16892-16898.

Ružička, M. (1958). Anwendung mathematisch-statistischer Methoden in der Geobotanik (Synthetische Bearbeitung von Aufnahmen). Biológia, Bratislava, 13: 647–661.

Zaneveld, J.R., Mcminds, R., Vega Thurber, R. (2017). Stress and stability: applying the Anna Karenina principle to animal microbiomes. *Nature Microbiology*, 2(9):17121.