

UC Berkeley

UC Berkeley Previously Published Works

Title

Relative abundance data can misrepresent heritability of the microbiome

Permalink

<https://escholarship.org/uc/item/0fz5p3qn>

Journal

Microbiome, 11(1)

ISSN

2049-2618

Authors

Bruijning, Marjolein

Ayroles, Julien F

Henry, Lucas P

et al.

Publication Date

2023

DOI

10.1186/s40168-023-01669-w

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

RESEARCH

Open Access



Relative abundance data can misrepresent heritability of the microbiome

Marjolein Bruijning^{1,2*}, Julien F. Ayroles^{2,3}, Lucas P. Henry^{2,3,4}, Britt Koskella⁵, Kyle M. Meyer⁵ and C. Jessica E. Metcalf²

Abstract

Background Host genetics can shape microbiome composition, but to what extent it does, remains unclear. Like any other complex trait, this important question can be addressed by estimating the heritability (h^2) of the microbiome—the proportion of variance in the abundance in each taxon that is attributable to host genetic variation. However, unlike most complex traits, microbiome heritability is typically based on relative abundance data, where taxon-specific abundances are expressed as the proportion of the total microbial abundance in a sample.

Results We derived an analytical approximation for the heritability that one obtains when using such relative, and not absolute, abundances, based on an underlying quantitative genetic model for absolute abundances. Based on this, we uncovered three problems that can arise when using relative abundances to estimate microbiome heritability: (1) the interdependency between taxa can lead to imprecise heritability estimates. This problem is most apparent for dominant taxa. (2) Large sample size leads to high false discovery rates. With enough statistical power, the result is a strong overestimation of the number of heritable taxa in a community. (3) Microbial co-abundances lead to biased heritability estimates.

Conclusions We discuss several potential solutions for advancing the field, focusing on technical and statistical developments, and conclude that caution must be taken when interpreting heritability estimates and comparing values across studies.

Keywords Absolute abundance, Compositional data, Genetic variance, Host-microbe associations, Microbiota, Phenotypic variance

Background

The number of host phenotypes known to be impacted by the microbiome is ever-growing, from metabolism to behavior, including its influence on a range of disease risk factors [1–3]. However, we are only beginning to understand the contribution of host genetics in shaping microbiome composition [4–8], with a growing body of research exploring such associations and their mechanisms (direct, or indirect, e.g., mediated by genetic associations with diet). This interest stems not only from our desire to understand how evolution and coevolution has shaped host-microbiome interactions over both shorter and longer (i.e., macroevolution) timescales [1, 9], but identifying a genetic basis of host-microbe associations

*Correspondence:

Marjolein Bruijning
m.bruijning@uva.nl

¹ Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, 1090 GE Amsterdam, The Netherlands

² Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA

³ Lewis-Sigler Institute for Integrative Genomics, Princeton, NJ 08544, USA

⁴ Center for Genomics and Systems Biology, Department of Biology, New York University, New York City 10003, USA

⁵ Department of Integrative Biology, University of California, Berkeley, CA 94720, USA



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

also has important applied health implications [10]. Critical to address these questions, is our ability to correctly measure the relative importance of hereditary and environmental influences on microbiome composition.

Heritability is a central parameter in quantitative genetics that quantifies a key aspect of the genetic basis for resemblance between parents and offspring. The heritability of a phenotypic trait is a statistical property, defined as the proportion of the phenotypic variance in a population that is attributable to genetic variation [11]. When estimating microbiome heritability, the focal phenotypic trait is typically either a measure of community composition or the abundance of a given taxon [12]. While a consensus is emerging that the heritability of most microbiome members is relatively low, specific estimates of the importance of host genetic variation in shaping microbiome composition vary widely across studies (Table 1). For example, a recent study found that 97% of the gut microbes in baboons has a significant non-zero heritability [13], while a different study concluded that host genetic background only plays a minor role in shaping microbiome composition in humans [14]. How should we interpret such substantial differences among studies: do these really reflect biological differences?

The use of heritability as a metric is ubiquitous in genetics, yet what it really measures and how it is interpreted remains the source of much confusion. Heritability is by definition a population-specific estimate, and subject to the influence from the environment and the genetic structure of the population. Moreover, detection of a non-zero microbiome heritability does not tell us anything about the mechanisms that cause related individuals to have, on average, more similar microbiomes. Several mechanisms are possible. Microbes might be vertically transmitted from the parents (and typically the mother) to offspring, for example via transfer during vaginal delivery, or via breast milk [35, 36]; or horizontally transmitted from other family members, perhaps simply due to their proximity [37]. Both effects could result in tight connections between host and microbial genotypes, and thus inflate heritability. Alternatively, host genotype might directly influence the types of microbes that can establish, as shown in species of woodrats [38], and this mechanism will also yield high estimates of heritability. Conversely, if heritability is estimated to be zero this need not mean that there is no vertical transmission (or horizontal transmission from relatives), it might simply mean that the effects of the environment are much larger and overwhelm these transmission effects, as has been found in marine sponges [39].

A methodological complexity when estimating microbiome heritability, is that the absolute microbial abundances are typically unknown. It is therefore common practice to calculate relative abundances by setting the sum in each sample to 1, generating so-called ‘compositional data.’ The inherent problems with compositional data have been acknowledged for some time, and they are known to lead to spurious correlations between variables, even when there exists no correlation at all [40]. This has more recently been discussed in the context of microbial data, for example when testing for differentially abundant microbes across treatment groups (e.g., host disease status) [41–46]. The estimation of microbiome heritability is rooted in comparison of differential abundances among host genotypes, and could therefore be subject to similar issues. However to date, studies reporting microbe heritability estimates, have not explicitly considered the potential problems associated with the use of compositional data.

We present an approximation of the taxon-specific heritability that one obtains when using relative abundances (we call this estimate φ^2). We show that this metric differs from traditional h^2 estimates: φ^2 is not simply a function of host genetic and phenotypic variance, but also depends on various other properties of the focal microbe and the rest of the community. Based on this, we identify three main problems that can arise when using relative abundance data to estimate taxon heritabilities. First, as relative abundances inherently covary, a heritable signal for some microbes can lead to spurious heritability estimates of non-heritable microbes or, vice versa, non-heritable microbes can mask a genetic signal in heritable microbes. This problem is most apparent for dominant taxa, and the impact of the issue diminishes for low abundance taxa, where the two heritability estimates (h^2 and φ^2) converge. However, a related second problem remains: while the estimated heritability of a non-heritable microbe can become close to zero, it may never completely reach zero. When a large number of host are sampled, even such a very weak (spurious) heritable signal can be highly significant, reflecting greater statistical power. When considering many microbial taxa in a community, the result is a considerable overestimation of the overall proportion of heritable microbes. Third, microbial taxa that covary in abundance (for instance caused by shared niches or microbial interactions), can result in large discrepancies between h^2 and φ^2 , systematically biasing heritability estimates. Depending on the nature and sign of the covariance, this can either mask or inflate true heritability signals. After deriving our approximation for φ^2 , we detail each of these problems. We show that our analytical results match results when we estimate heritability by fitting statistical models to simulated datasets. We then

Table 1 Summary of the studies estimating heritabilities of the abundance of microbial taxa, sorted by sample size. More details in Additional file 1: Appendix S4 on methodology per study

Number	Host system	# Samples	# Taxa	# Heritable taxa	Average non-zero heritability	Host genetic relatedness based on ^a	Normalization/transformation	Reference
1	Mice	32	43	NA ^b	0.47 ^b	Lineage	Total sum scaling	[15]
2	Chickens	56	23	0	NA	Pedigree	Log-transformation and scaling	[16]
3 ^c	Humans	93	116	14	0.35	SNPs	Quantile normalization	[17]
		91	104	10	0.37			
		127	102	13	0.26			
4	Humans	108	221	0	NA	Twins	Box-Cox transformation	[5] data from [18]
5	Humans	126	2,933	0	NA	Twins	Box-Cox transformation	[5] data from [19]
6 ^d	Humans	244	3	1	0.35	Twins	Arcsine square root transformation	[20]
		88	3	0	NA			
7	Humans	250	109	11	NA ^e	Twins	Box-Cox transformation	[21]
8	Humans	270	249	26	0.58	Pedigree	Inverse normal transformation	[22]
9	Switchgrass	383	110	21	0.24	SNPs	Total sum scaling	[23]
10 ^f	Cows	650	512	39	NA ^e	SNPs	Quantile normalization	[24]
		200	512	3				
11	Humans	485	91	42	0.34	Twins	Log transformation and scaling	[25]
12	Humans	542	369	85	0.27	Twins	Inverse normal transformation	[26]
13	Mice	592	43	NA ^b	0.51 ^b	SNPs	Total sum scaling	[27]
14	Sorghum	600	1189	443	0.22	Lineage	Cumulative sum scaling	[28]
15	Humans	655	85	52	0.24	Twins	Inverse normal transformation	[29]
16	Humans	1068	21	6	0.40	SNPs	Box-Cox transformation	[30]
17	Humans	1081	909	10	0.29	Twins	Box-Cox transformation	[5]
18	Humans	1176	209	11	0.31	Twins	Inverse rank-sum transformation	[31]
19 ^g	Pigs	1205	1678	170	0.056	Lineage	Total sum scaling	[32]
		1295	1678	261	0.078			
		1283	1678	366	0.099			
20 ^h	Maize	4866	792	143	0.17	Lineage	Log transformation	[33]
		45	2557	5	0.45			
21	Humans	3261	945	52	0.30	Twins	Box-Cox transformation	[6]
22	Humans	4745	242	31	0.20	Pedigree	Centered log-ratio transformation	[34]
23	Baboons	16,234	283	273	0.068	Pedigree	Total sum scaling	[13]

^a Type of host genetic data to estimate heritability. Pedigree: take into account pedigree to estimate narrow-sense h^2 . SNPs: incorporate genetic relatedness matrix based on SNPs, to calculate SNP heritability. Lineage: genotype/lineage as random effect, estimates broad-sense H^2 . Twins: compare MZ with DZ twins, to estimate broad-sense H^2

^b No significance measures are provided. Average heritability is therefore calculated using all estimates

^c Analyses are done for winter, summer and both seasons combined

^d European and African ancestry

^e Heritability estimates per taxon are not provided

^f Two different breeds

^g Three time points during host development

^h 2010 and 2015 field study

discuss empirical heritability estimates obtained from published studies in the light of our results. In the discussion, we outline some solutions that may partly solve the here described issues. We conclude that caution must be taken when interpreting heritability estimates based on relative abundances and comparing values across studies, and that approximations of microbial absolute abundances may help remedy this issue.

The heritability of a taxon’s abundance

When estimating the heritability of a taxon, one relies on a quantitative genetic model, considering a taxon’s abundance as a quantitative phenotypic trait of the host. The absolute abundance of taxon *i* in host *j* (P_{ij}) can be written as

$$P_{ij} = \alpha_i + G_{ij} + E_{ij} \tag{1}$$

where α_i is the average absolute abundance of microbe *i*, G_{ij} is the breeding value or host genetic contribution (for simplicity, we assume no genetic dominance or epistasis), E_{ij} is the environmental contribution (residual), and we assume no $G \times E$ interactions. Equation 1 can be extended by including additional factors that affect taxon abundance, such as host age, sex, or season.

Across host individuals, the absolute abundance of microbe *i* is assumed to follow a normal distribution with mean α_i and variance V_{Pi} (non-normal data requires transformations or link functions, see e.g., [47], but this is beyond the scope of this manuscript). This variance can be decomposed into a genetic and environmental contribution (assuming no genotype-environment covariance):

$$V_{Pi} = V_{Gi} + V_{Ei} \tag{2}$$

Following the definition of the heritability, the heritability of taxon *i* is

$$h_i^2 = \frac{V_{Gi}}{V_{Pi}} \tag{3}$$

When the absolute abundances are known, one can simply estimate the taxon heritability by quantifying the proportion of the total variance that is attributable to host genetic variation (e.g., by fitting a mixed effects model [48]). In case that all assumptions are met (e.g., normally distributed phenotypes, no $G \times E$ interactions), the heritability can be estimated (note that in this case, as we assume no dominance or epistasis, the broad-sense and narrow-sense heritability are identical). Both direct genetic effects and indirect genetic effects, e.g., through other microbes or behavior that is under genetic control, can result in a non-zero heritability.

However, we typically do not know the absolute microbial abundances. Instead, most of the time we quantify how the relative abundance of taxon *i* varies across host individuals and estimate the heritability as the proportion of the variance in relative abundance that is attributable to genetic variation. Below we derive an equation for the obtained heritability when one uses relative, and not absolute abundances, based on the underlying model shown in Eqs. 1–2.

An approximation of the heritability based on relative abundances

As outlined above, the absolute abundance of microbe taxon *i* is distributed across host individuals as

$$P_i \sim \text{Normal}(\alpha_i, V_{Pi}) \tag{4}$$

The distribution of relative abundances not only depends on the focal microbe, but also on the absolute abundance of the entire community, consisting of *M* taxa. The community absolute abundance *C* (where $C = \sum_j^M P_j$) is also a normally distributed variable, where its mean equals the sum of the average abundances over all *M* taxa. The variance depends on the variance in each taxon, plus the sum of each phenotypic covariance between microbial pair, so that

$$C \sim \text{Normal}\left(\sum_{j=1}^M \alpha_j, \sum_{j=1}^M V_{Pj} + 2 \sum_{1 \leq j < k \leq M} \text{cov}_P(j, k)\right) \tag{5}$$

The relative abundance of focal microbe *i* (which we call fraction f_{Pi}) is calculated as the absolute abundance of focal taxon *i*, divided by the entire community abundance, and therefore is distributed as the ratio between Eq. 4 and Eq. 5:

$$f_{Pi} \sim \frac{\text{Normal}(\alpha_i, V_{Pi})}{\text{Normal}(\sum \alpha, \sum V_P + 2 \sum \text{cov}_P(j, k))} \tag{6}$$

We are interested in quantifying $\text{var}(f_{Pi})$, as this gives us the total variance in the relative abundance, analogous to V_{Pi} . Similarly, we can obtain how relative abundances vary between host genotypes, by replacing V_{Pi} and V_P , by V_{Gi} and V_G , respectively, and considering genetic covariances cov_G between each pair of microbes:

$$f_{Gi} \sim \frac{\text{Normal}(\alpha_i, V_{Gi})}{\text{Normal}(\sum \alpha, \sum V_G + 2 \sum \text{cov}_G(j, k))} \tag{7}$$

The proportion of the variance in relative abundance explained by host genetic variation (i.e., the heritability based on relative abundances or fractions, from now on called ϕ^2) is then

$$\varphi^2 = \frac{\text{var}(f_{Gi})}{\text{var}(f_{Pi})} \tag{8}$$

In other words, Eq. 8 gives the heritability that one obtains when using relative, and not absolute, abundances. Ideally, if relative abundances are used as a proxy for absolute abundance, the heritability measure is the same when using absolute and relative abundances, i.e., one hopes that $h^2 = \varphi^2$.

An approximation of the heritability of taxon i that one obtains when using relative abundance data, and not absolute abundance data, is given by

$$\varphi^2 \approx \frac{A^2 V_G + \alpha^2 \omega - 2\alpha(A\gamma - \alpha\kappa)}{A^2 V_P + \alpha^2 z - 2\alpha(A(\gamma + \epsilon) - \alpha(v + \kappa))} \tag{9}$$

(see Additional file 1: Appendix S1). Heritability φ^2 is a function of properties of the focal taxon, with parameters V_G and V_P describing the genetic and phenotypic variance in absolute abundances, and α describing the average absolute abundance (to improve readability, we omit subscripts i). It follows from Eq. 9 that φ^2 is also a function of the summed genetic and environmental covariances between focal taxon i and each of the other taxa in the community (γ and ϵ , respectively). Finally, φ^2 is a function of various properties of the background community (excluding the focal taxon): A is the average absolute abundance of the background community, ω and z are the total host genetic and phenotypic variance in absolute abundances of the background community (i.e., the variances summed over all taxa), and κ and ν are the sums of the genetic and environmental covariances between each pair of background community members.

Notice the difference between Eq. 3 and Eq. 9: whereas h^2 is (by definition) only a function of V_G and V_P , the heritability estimate that one obtains when using relative abundances, depends on various additional properties of the focal microbe (α), the entire community (A , ω , z , κ , ν) and interactions between the focal microbe and the community (γ , ϵ). Depending on the biology of the host-microbiome system as well as on properties of the data, we identified three problems that can arise as a consequence.

Problem 1: interdependency between taxa leads to imprecise heritability estimates

As relative abundances are not independent, heritable variation in some microbes can lead to spurious non-zero heritabilities, in other microbes. Or vice versa, non-heritable microbes can mask a genetic signal in heritable microbes. Consider the extreme scenario with only two equally abundant microbes, where microbe A

has a heritability of 1, and microbe B has a heritability of 0 (Fig. 1a). Because abundances are scaled to relative abundances, it would still seem that variation in microbe B abundance is shaped by host genetics (Fig. 1a). Moreover, expressing both abundances as relative abundances partly obscures the host genetic effect on microbe A. This results in a heritability estimate of 0.5 for both species, which is wrong in both cases, and leads to the incorrect conclusion that both microbes are heritable.

This can be formalized using Eq. 9, which, in the absence of genetic and environmental covariances, simplifies to

$$\varphi^2 \approx \frac{\alpha^2 \omega + A^2 V_G}{\alpha^2 z + A^2 V_P} \tag{10}$$

It follows that for a focal taxon with a very low average abundance (i.e., $\alpha \ll A$), the estimated heritability approaches the same value as when based on absolute abundances (Eq. 3):

$$\lim_{\alpha/A \rightarrow 0} \varphi^2 \approx \frac{V_G}{V_P} = h^2 \tag{11}$$

However, for a very dominant taxon ($\alpha \gg A$) it becomes more difficult to retrieve the true heritability h^2 , approaching:

$$\lim_{\alpha/A \rightarrow \infty} \varphi^2 \approx \frac{\omega}{z} \tag{12}$$

Remember that ω and z are the total genetic and phenotypic variance of the entire background community (summed over all microbes, excluding the focal microbe). Thus, for a highly dominant microbe, the estimated heritability approaches the heritability of the background community, and is not shaped at all by the genetic and phenotypic variance of the focal microbe.

This implies that depending on properties of both the focal microbe and the rest of the community, heritability estimates can be biased in different directions (Fig. 1b): we will underestimate the heritability of an abundant microbe when it is harbored by a non-heritable community (black line in Fig. 1b). On the other hand, an abundant microbe with no host genetic signal, will still appear heritable when it occurs in the background of a heritable community (grey line in Fig. 1b). As a result, the error in the heritability (i.e., the absolute difference between φ^2 and h^2), depends on both the heritability of the focal microbe, as well as on the heritability of the background community, and in general increases with an increasing abundance relative to the background community (Fig. 1c). When $\frac{\alpha}{\alpha+A} < 0.05$ (for instance, in the case of 20 equally abundant taxa in a community), the expected absolute error will be less than 10% for

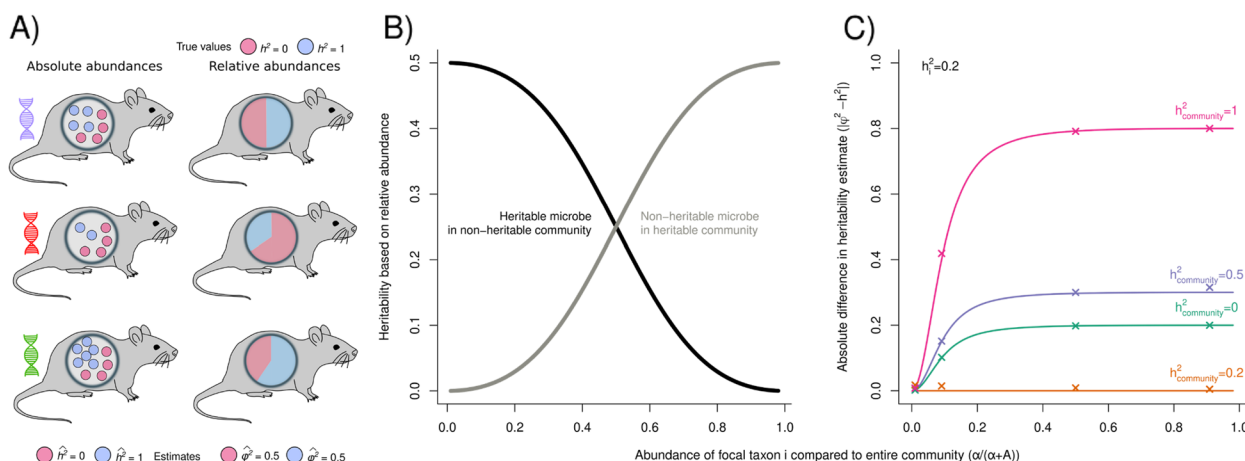


Fig. 1 As relative microbial abundances are interdependent, a heritable signal in one microbe can lead to a spurious heritable signal in a second microbe that is not heritable, or mask a genetic signal in a heritable microbe. **A** As an example we show three host (mouse) genotypes with two microbes, where one microbe is fully heritable (blue, $h^2 = 1$), and one microbe is not heritable (red, $h^2 = 0$). As a consequence, the average absolute abundance of microbe Blue differs among genotypes, while the average abundance of microbe Red is constant. Using the absolute abundances (and with enough host replicates), heritabilities can correctly be estimated. However, as relative abundances are not independent, a host genetic signal in the abundance of the heritable microbe, will also create a host genetic signal in the second microbe, creating variation in relative abundance among genotypes. This leads to an incorrect heritability estimate $\hat{\varphi}^2 = 0.5$ for both microbes. **B** When based on relative abundances, properties of both the focal microbe and of the entire community shape the heritability estimates. Here, we vary the average absolute abundance of the focal microbe (α) compared to the absolute abundance of the rest of the community (A) (x-axis shows $\frac{\alpha}{\alpha+A}$). Black line: focal microbe has a heritability of 0.5; the background community is not heritable ($A = 1$; $z = \left(\frac{1}{6}\right)^2$; $\omega = 0$; $V_P = \left(\frac{1}{6}\right)^2$; $V_G = 0.5\left(\frac{1}{6}\right)^2$). Grey line: focal microbe is not heritable, but the rest of the community has an average heritability of 0.5 ($A = 1$; $z = \left(\frac{1}{6}\right)^2$; $\omega = 0.5\left(\frac{1}{6}\right)^2$; $V_P = \left(\frac{1}{6}\right)^2$; $V_G = 0$). **C** Difference in heritability estimates when based on absolute or relative abundances (y axis) when varying α compared to A (x axis). When the focal microbe has a low average absolute abundance compared to the total average abundance of the rest of the community (for instance, in the case of many microbial taxa), the difference between φ^2 and h^2 becomes smaller. h^2 of the focal taxon i is 0.2, and colored lines show varying heritabilities of the background community ($h^2_{\text{community}} = \frac{\omega}{z}$). $A = 100$; $z = 100\left(\frac{1}{6}\right)^2$; $V_P = \alpha\left(\frac{1}{6}\right)^2$. Crosses show results when we estimate heritabilities by fitting a mixed effects model on simulated relative abundance data. To this end, we simulated a population of hosts (500 genotypes \times 1000 replicates within each genotype), with microbial communities consisting of 100 taxa (more details in Additional file 1: Appendix S2.1–2.3)

all conditions shown in Fig. 1c. Here, we note that the error not only depends on the total abundance of the background community (A) compared to the abundance of the focal microbe (α), but also on how variances z and V_P scale with A and α , respectively (in Fig. 1c, V_P is kept proportional to α).

Problem 2: large sample size leads to high false discovery rates

Microbes that are not heritable can still show a genetic signal when abundance measurements are relative, due to the interdependency of the relative abundances. Using Eq. 9 and in the absence of environmental covariances, it follows that the estimated heritability of a non-heritable microbe (by setting $V_G = 0$) is

$$\varphi^2 \approx \frac{\alpha^2 \omega}{\alpha^2 z + A^2 V_P} \tag{13}$$

Unless the entire background community is not heritable (i.e., $\omega = 0$), Eq. 13 will be larger than 0.

Although φ^2 approaches zero when α becomes small compared to A , it might never reach zero.

Even low φ^2 values can appear significant with enough statistical power. We performed a power analysis using the R-package *simr* [49], based on a log likelihood ratio test comparing a model with and without host genetics, to calculate the probability that the null hypothesis ($H_0: \varphi^2 = 0$) is (wrongly) rejected (Additional file 1: Appendix S3 for details). Results again depend on both properties of the focal microbe and of the rest of the community (Fig. 2), but in general, larger sample sizes increase the chance that non-heritable microbes are considered heritable. With a large enough dataset, statistical power reaches 100% (Fig. 2).

As a consequence, the number of heritable microbes in a community can be strongly overestimated, especially with a high sample size (more details in results in Additional file 1: Appendix S2.4). It is important to note that the high false discovery rates are not a problem of, for instance, sampling error or confounding factors, and increasing data collection

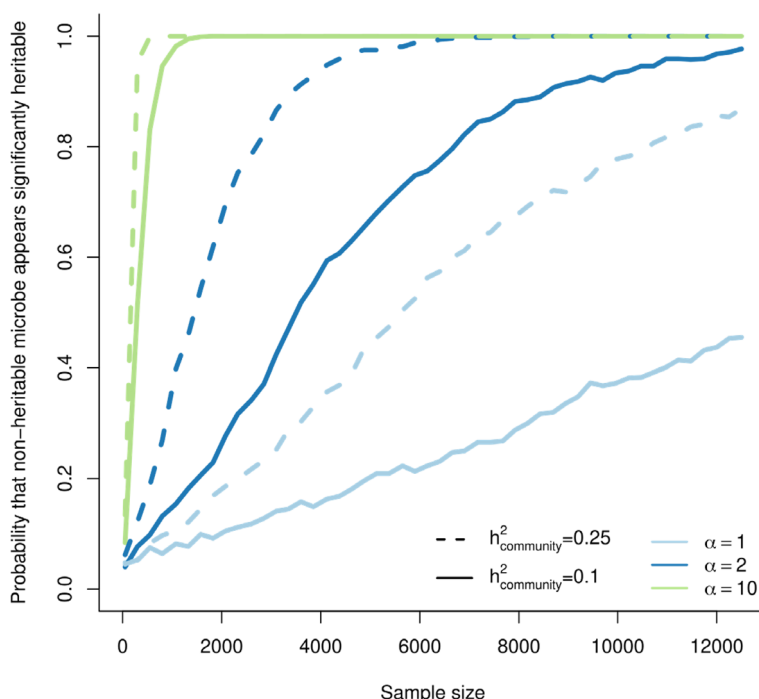


Fig. 2 The probability that the heritability of a non-heritable microbe ($V_G = 0$) wrongly appears significant ($\alpha < 0.05$) increases with sample size, based on a power analysis using the R-package *simr* (28). Results depend both on properties of the focal microbe, and on the rest of the community: colors show different abundances of the focal microbe (α) while keeping the background community abundance constant. Line type shows the heritability of the background community (solid lines: $\frac{\alpha}{z} = 0.1$; dotted lines: $\frac{\alpha}{z} = 0.25$). $V_p = \left(\frac{1}{6}\right)^2$; $A = 100$; $z = 100 \cdot \left(\frac{1}{6}\right)^2$

efforts or quality alone will not resolve these issues. Similarly, more advanced modeling approaches such as cross-validation, permutation analysis and correcting for multiple testing are unlikely to fully solve this. This is because the problem is inherent to the use of relative abundances: there really *is* a host genetic signal in the relative abundances of non-heritable microbes (i.e., it is not a type 1 error; as Eq. 13 shows, φ^2 really *is* larger than 0).

Problem 3: microbial co-abundances lead to biased heritability estimates

Up to this point, we assumed that the covariance terms in Eq. 9 (i.e., γ , ϵ , ν and κ) were zero. We will now show that relaxing this assumption leads to biased heritability estimates.

Non-zero covariance terms reflect the co-abundance of microbial taxa. In our framing, there are two processes that can cause microbial abundances to covary: host genetic correlations and environmental correlations. The first creates microbial co-abundances at the level of the host genotypes: e.g., a host genotype with an—on average—higher abundance of microbe A, also has a higher abundance of microbe B. The second creates co-abundances at the individual host level, by creating

correlated environmental (residual) terms. Note that, as is general practice in quantitative genetics, we use the term ‘environment’ to capture everything outside of genetics: it is essentially a residual term. In the case of the microbiome, it captures the effect of ecological environmental factors on microbial abundances, such as temperature or soil, where shared niches among microbes can result in environmental correlations. The residual term also captures effects of the environment inside and shaped by the host, the abundance of other microbes within a host, or simply unexplained noise. One biological process acting within hosts that would lead to the environmental terms being correlated, is microbial interactions. Strong mutualistic interactions, e.g., as a result of cross-feeding or public good production, result in positive environmental correlations. Antagonistic interactions, on the other hand, result in negative environmental correlations.

Non-zero covariances can change heritability estimates in different directions, depending on the nature of the covariance (i.e., genetic or environmental), and whether the covariance involves the focal taxon (γ , ϵ) and/or the background community (ν , κ). For the results presented here, we assume that each microbial pair (including focal

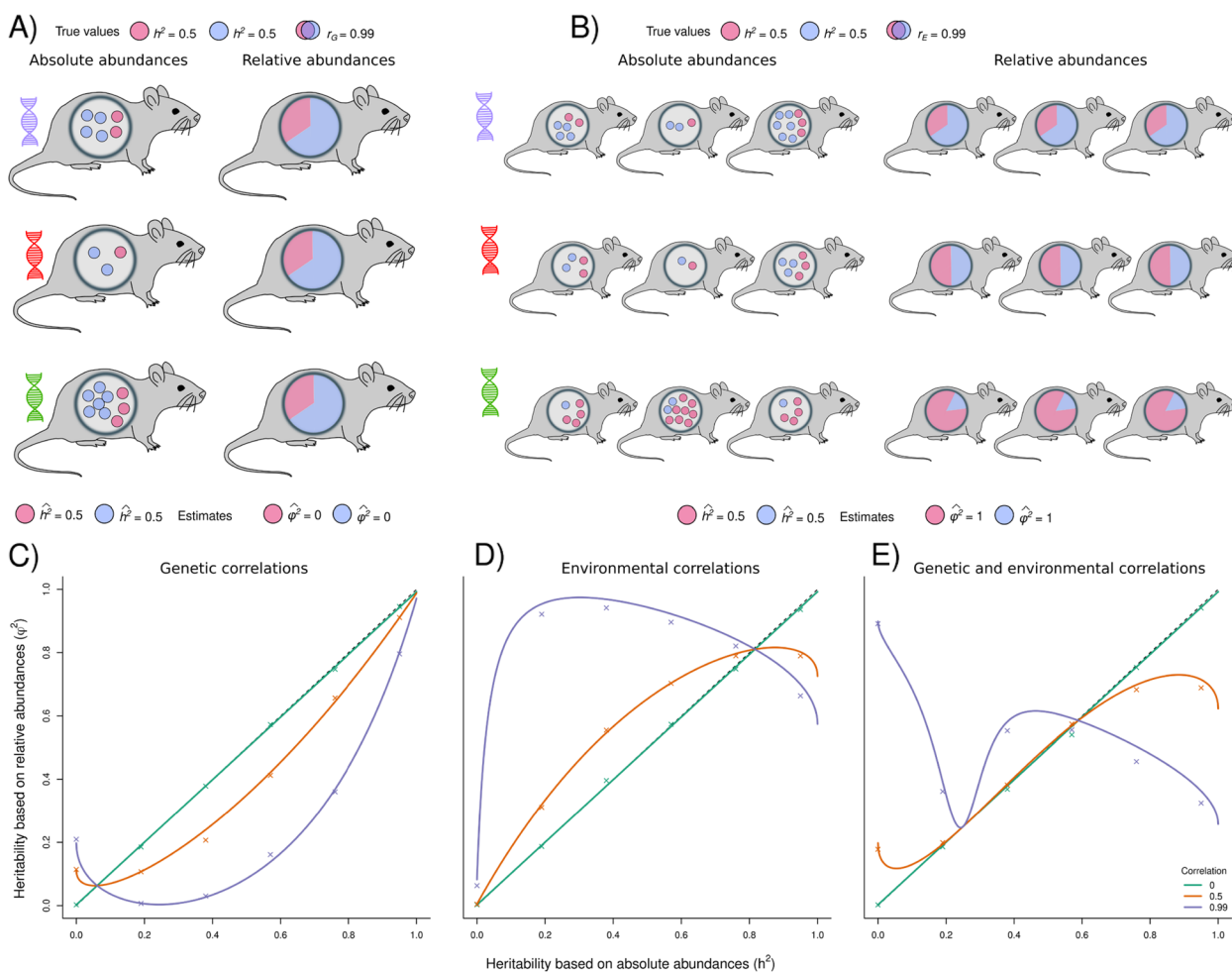


Fig. 3 The use of relative abundances leads to biased heritability estimates when there exists host genetic and/or environmental correlations between microbes. **A** Illustrates the effects of genetic correlations. As an example, we show three host genotypes and two microbes that are both partly heritable ($h^2 = 0.5$), and with a strong genetic correlation ($r_G = 0.99$). This implies that host breeding values for the two microbes are strongly correlated. As a consequence, the average absolute abundance of both microbes varies in the same way across host genotypes. Heritabilities can accurately be estimated when using these absolute abundances (estimates for both microbes: $\hat{h}^2 = 0.5$). When calculating the relative abundances, however, any variation across host genotypes disappears. This leads to an incorrect heritability estimate $\hat{\varphi}^2 = 0$ for both microbes, completely masking the host genetic signal. **B** Illustrates the effects of environmental correlations. We here show three host genotypes and two microbes that show a strong environmental correlation ($r_E = 0.99$). As a result, this decreases the amount of variation within genotypes. Heritabilities can be accurately estimated when using the absolute abundances. However, because variation in relative abundance within each genotype is greatly reduced, one obtains a wrong heritability estimate $\hat{\varphi}^2 = 1$ for both species. **C–E** Comparison of heritability estimates when based on absolute and relative abundances, varying the environmental correlation (**C**), the genetic correlation (**D**) or both (**E**). $\alpha = 1$; $A = 100$; $V_P = \left(\frac{1}{6}\right)^2$; $z = 100 \cdot V_P$; $\frac{\omega}{z} = 0.25$. Crosses show results when we estimate heritabilities by fitting a mixed effects model on simulated relative abundance data. To this end, we simulated a population of hosts (500 genotypes \times 500 replicates within each genotype) (more details in Additional file 1: Appendix S2.5)

and background community members) has the same genetic and environmental correlation.

In a community with positive genetic covariances, the heritabilities are generally biased downwards (Fig. 3c). This is because positive genetic covariances have a relatively larger (negative) effect on the numerator than on the denominator (Eq. 9). To make this intuitive, consider the scenario where two equally-abundant microbes both

have a heritability of 0.5, and also have a strong genetic correlation ($r_G = 0.99$). Such a strong genetic correlation implies that the host genetic effects for the two microbes covary, so that two microbes show co-abundance at the host genotype level. As a consequence, the absolute abundances vary across host genotypes for both microbes, but they vary in exactly the same way (Fig. 3a). When calculating relative abundances, variation in abundance across

genotypes completely disappears, which leads to the incorrect conclusion that none of the microbes show a heritable signal.

The exception is when the true heritability is close to zero: now, positive genetic covariances lead to an overestimation of the true heritability (Fig. 3c). This occurs when $A\gamma < a\kappa$, causing the covariance term in Eq. 9 to become negative (thereby increasing the numerator). Since γ is the total genetic covariance between taxon i and each of the other microbes, it becomes small when V_G is close to zero. As a result, $A\gamma < a\kappa$, leading to an overestimated heritability.

Positive environmental covariances in a community (for instance, a highly mutualistic community) has largely opposite effects, by (negatively) affecting the denominator but not the numerator (Eq. 9). Whereas positive host genetic correlations between microbes tend to decrease variation in relative abundance *between genotypes*, positive environmental correlations tend to decrease the amount of variation *within genotypes* (Fig. 3b). When variation within each genotype is reduced, this creates more unique microbiomes to each genotype, suggestive of microbe heritability. As a result, positive environmental covariances lead to a general upward bias in the heritabilities (Fig. 3c). Only if $A\epsilon < av$, the true heritabilities are underestimated. This happens, for instance, if there

is little environmental variance in the focal taxon (i.e., a high heritability), causing ϵ to be low.

Finally, when both positive genetic and environmental correlations exist in a community, the relationship between the two heritability measures can become highly non-linear, making it essentially impossible to predict h^2 based on φ^2 (Fig. 3e).

Framing the current empirical range of estimates

Our results provide additional context in considering the range of estimates of heritabilities published to date. First, our results indicate that estimates of the taxon heritabilities can be precise if each focal taxon has low abundance compared to the total community abundance (and assuming no microbial co-abundances) (Fig. 1c). Our review of the literature indicates that the median number of taxa included in a study is 221 (Table 1). Since most taxa therefore are likely to have low relative abundances, heritability estimates of most individual (low-abundance) taxa may be quite accurate. There is, however, also a wide range in the number of included taxa across studies (varying between 3 and 2933 taxa), and furthermore, human microbiomes are often characterized by a few dominant taxa [50], and this may be the case for many host species. Our results indicate that for studies that only include a few taxa (for instance in low-diverse communities or in communities where most taxa are rare and thus excluded in the

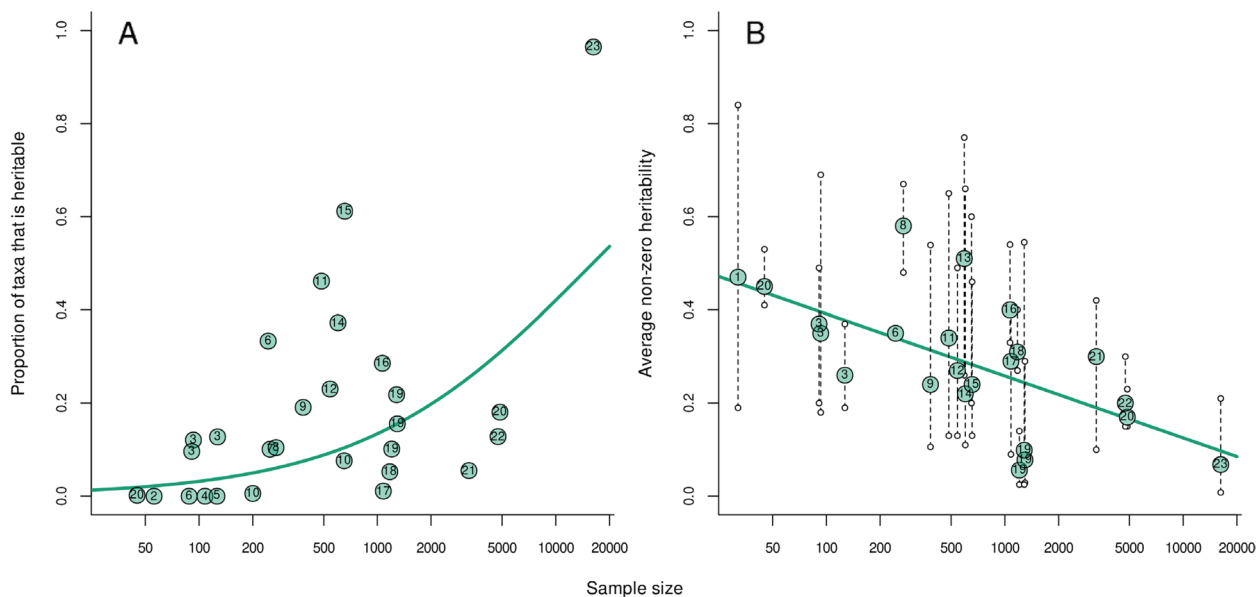


Fig. 4 Empirical estimates of the proportion of heritable taxa (**A**) and the average taxon heritability, including all significantly heritable microbes (**B**), plotted against sample size, i.e., number of hosts sampled (note the log scale on the x axis). Dots depict values given in Table 1, where the numbers in each dot correspond to the column ‘Number’ in Table 1). Teal lines show the mean prediction based on **A** a binomial regression (here the number of trials is the number of taxa), and **B** a linear regression. In **B**, dotted lines connect average heritabilities to the lowest and highest significant heritability found in each study, shown as open dots

filtering steps), or where microbiome communities are characterized by a few highly dominant taxa, precise heritability estimates will be challenging to obtain.

We identified a second problem that is related to the number of sampled hosts: the proportion of heritable microbes can be considerably overestimated due to high false discovery rates. Empirical estimates of the proportion of heritable microbes, show a positive association with the number of hosts sampled (Fig. 4a; binomial regression: p value < 0.0001). Of course, larger sample sizes always lead to more significant results, as higher sample sizes lead to more power to detect small effects. The challenge here is that without knowing more about the underlying community, we cannot establish how much of this inflation is 'real' and how much is due to false discovery. Every microbe may eventually appear significantly heritable with enough statistical power (Fig. 2), even if its absolute abundance is not shaped at all by host genetics, and false discovery rate approaches (such as Benjamini-Hochberg) do not solve this issue (Additional file 1: Appendix S2.4). This is due to the interdependency microbiome members will have with other, truly heritable, microbes; and a positive relationship between sample size and the proportion of heritable microbes will emerge even if the true proportion heritable is constant across populations (Additional file 1: Appendix S2.4).

Shifting the focus from the proportion of the taxa that is heritable to considering heritability of taxa, this quantity (including only taxa with a significant heritable signal) varies widely within as well as between studies (Fig. 4b), nearly covering the entire 0–1 range. Across studies, the lowest and highest reported significant heritabilities are 0.008 and 0.84, respectively. The average significant heritability in a community is 0.30, and ranges between 0.056 and 0.58 across studies. It is notable that empirical estimates suggest a negative correlation between sample size and the average heritability, where studies that include a higher number of host individuals report lower average heritabilities (Fig. 4b; linear regression: p value = 0.002). This could be due to publication bias in smaller studies, in favor of higher heritability estimates, which could suggest that the true microbiome heritabilities may be lower than sometimes reported. However, it could also be that studies with a larger sample size include an increased number of taxa with a low estimated heritability, that can either truly be heritable or reflect false positives, thereby decreasing the average heritability.

The included studies clearly differ in many aspects other than sample size, both biological (e.g., host system, population and tissue, taxonomic levels, any other covariates) and methodological (e.g., data collection, significance measure, statistical model). There is no

reason to expect that the true proportion of heritable microbes or the average heritability is the same across studies—to the contrary. Further, there clearly is variation that is not explained by sample size, indicating that other factors (likely both biological as well as methodological) also play a role. Yet, variation in sample size alone explains considerable variation across studies in both the proportion of heritable microbes (pseudo- $R^2 = 37\%$) and in the average heritability ($R^2 = 39\%$).

Finally, our results indicate that bias in φ^2 relative to h^2 depends on both the magnitude of h^2 and the underlying pattern of genetic and environmental correlations (Fig. 3). Since little is known about the nature and strength of correlations (in absolute abundance) among microbes, it is hard to interpret the impact of this bias on published results to date. Yet, these results do underscore the importance of further efforts to estimate the co-abundance patterns.

Discussion

Despite the common usage of microbial relative abundance data as a proxy for absolute abundance to estimate microbial heritabilities, few studies have considered the inherent problems that can result from statistical analysis of relative abundances. By their nature, relative abundance data are not independent, creating correlations between variables (microbial abundances) that do not exist in absolute terms. Here we argue that (1) this can lead to imprecise estimates of heritabilities, especially for microbiomes with fewer taxa and/or highly abundant taxa. (2) Large sample sizes can drive overestimates of the proportion of heritable microbes by increasing the false discovery rate. (3) Patterns of microbial co-abundance, likely to be common in most biological systems, will further bias heritability estimates. Similar challenges have been demonstrated for microbial differential abundance analysis, where it is challenging to control high false discovery rates [42, 46, 51–54]. Characterizing the mechanisms underlying these issues helps identify when they might occur, and the direction of bias expected given the number of microbial taxa and their abundance, the number of hosts, and patterns of co-abundance. It is important to note that heritabilities based on relative abundances are potentially misleading only if one wishes to make inferences on host genetic control over *absolute* abundances, i.e., if relative abundances are used as a proxy for absolute abundances. If the metric of interest is, in fact, the heritability of relative abundance, the true value of φ^2 is directly accessible using available relative abundance data. However, recent studies have proposed that absolute abundances may provide more biological information [55, 56], for instance showing an association between total bacterial load and Crohn's disease [57].

Also, as φ^2 is a function of both properties of the focal microbe and of the entire community (Eq. 9), its biological interpretation is potentially challenging. Unfortunately, there does not seem to be a simple solution to fully address the problems described here, but below we discuss several potential approaches for advancing the field.

One solution that would clearly solve the issue of interdependent relative abundance data, is quantifying taxon (or group) absolute abundances. In cases where specific microbial taxa are of interest, such taxa can be directly quantified using such targeted approaches to abundance estimates as quantitative PCR (qPCR), droplet digital PCR (ddPCR), or flow cytometry [43, 57–59]. Additionally, for microbes that are readily cultivable, counts of colony forming units (CFUs) from culturing serve as a method to estimate absolute abundance. However, these approaches remain challenging for microbiome-wide studies that are concerned with the hundreds to thousands of taxa that comprise a given microbiome. One possible solution is to integrate microbial relative abundance data with estimates of the total microbial load of the sample. For instance, if a given taxon represents 1% of the 16S rRNA gene reads in a sample, multiplying that 1% by the total number of 16S rRNA gene amplicons (derived, e.g., from qPCR estimates using the same primers, ng of DNA, and PCR cycle numbers), can provide an estimate of that taxon's absolute abundance. To further improve such an approach, researchers could target known single-copy genes, rather than the 16S rRNA gene, e.g., *rpoB* [60]. Studies that compare inferences when using absolute vs. relative abundances are beginning to emerge [43, 61], although we are not aware of any study that addresses this in the context of microbiome heritability.

In addition to laboratory techniques, new data analysis approaches could prove beneficial. There exists an extensive body of literature on how to analyze compositional data (pioneered by Aitchison [1982]), with relevance to microbiome studies (but also genomics (gene expression), geology (mineral composition) and chemistry (chemical composition)). It is beyond the scope of this paper to provide a comprehensive overview of all available methods, but we refer the interested reader to [41, 51, 62, 63] for studies applying such methods to microbial data. Here, we briefly explain the main intuition behind these approaches, and how these may help to improve the accuracy of heritability estimates.

Data normalization is a first solution for obtaining better proxies of the absolute abundances. Instead of dividing the number of reads per taxon by the total number of reads in a sample, one divides the total number of reads by some normalization factor. This

involves choosing an appropriate 'reference' value, i.e., deciding what the appropriate comparison is within each sample. The advantage of comparing the number of reads for each taxon to a set reference, is that it makes abundances less sensitive to the other taxa that are in the sample. In addition, it is recommended to express compositional data as log ratios. This transforms data from a simplex to real space, making it more suitable for standard statistical tests [64, 65].

The merit of different normalization and transformation methods critically depends on the chosen reference. If there are 'reference' taxa, known to have constant abundance across samples, one could divide each sample by the number of reads for these reference taxa, thus transforming the relative abundance in each sample into comparable abundances across samples (this is similar to using reference genes to normalize gene expression data), and calculate the log-ratios (additive log-ratio transformation) [52, 65]. One could also calculate the log-ratio between each taxon and the geometric mean of all taxa (centered log-ratio transformation) [66]. Alternatively, if only a small number of microbial taxa is thought to be differentially abundant across samples, one could calculate a normalization factor based on some quantile (e.g., median) of each sample's count distribution (cumulative-sum scaling) [67].

If there truly is a known reference taxon with a constant abundance, or if the average abundance truly is identical in all samples, one could successfully correct for sample coverage differences by applying the appropriate normalization/transformation, and retrieve the true heritabilities (Additional file 1: Appendix S5). However, while some studies on microbiome heritabilities apply data transformations (e.g., centered log-ratio transformation [13, 34], Box-Cox [5] or inverse normal transformation [29]), we lack a validation that such transformations are justified and remedy any existing issues. There is currently little empirical data to guide us in choosing appropriate normalization factors.

It could be more fruitful to focus on the actual heritability estimates, than to focus on the number of significantly heritable taxa. Focusing exclusively on p values, with some arbitrary threshold for results to be 'significant', has been criticized [68, 69], and dichotomizing results into 'significant' and 'not significant' may be particularly problematic for microbiome heritabilities. That is because relative abundances are interdependent: an increase in the abundance of one taxon will inevitably decrease the relative abundance of other taxa. This implies that host genetic variation for the absolute abundance in few microbes, might also lead to genetic variation for other,

non-heritable, microbes. Therefore, the null hypothesis (i.e., that there is no host genetic signal in the relative abundances of microbiome members) might rarely be true. With a large enough sample size, this will lead to a statistically significant effect [69] (Fig. 2), even if the effects may be biologically meaningless.

By focusing on effect sizes, we can delineate the heritable taxa that are biologically most relevant. Our results indicate that, unless the focal microbe has a very high abundance compared to the rest of the community (Fig. 1) or microbial abundances covary (Fig. 3), taxon-specific heritability estimates based on relative abundances are unbiased. One could (a priori) set a threshold heritability, and only consider heritabilities exceeding this threshold to be biologically relevant. For instance, Goodrich et al. (2016) only present results of taxa that have an estimated heritability > 0.2 .

In addition to focusing on effect sizes, assessing the cumulative evidence for specific microbial taxa will help to identify microbes that are truly heritable and biologically relevant. Grieneisen et al. (2021) found a correlation between their heritability estimates and estimates from previously reported studies (although their effect sizes are much lower). Also, Goodrich et al. (2016) pinpointed various taxa with consistent non-zero heritabilities across studies and across hosts systems. Looking for such consistent results will indicate which taxa merit more detailed study, especially for microbes associated with host performance. Multiple studies have reported high heritabilities for members of the *Christensenellaceae* family, with estimates ranging between 30 and 60% [5, 6, 22, 29, 70]. Members of the *Christensenellaceae* have been linked to several host metabolic traits [70]; for example, a higher relative abundance has been associated with a lower body mass index [5].

In this study, we specifically focused on the consequences of using relative abundances, where the sum in each sample is set to 1, or 100%. The analysis of real-world microbiome datasets comes with additional challenges. First of all, variation across samples not only results in unknown absolute abundances, it also implies different levels of uncertainty. For example, 100 counts of a given taxon in a sample with 10,000 reads, clearly allows for more robust statistical inference than 1 count in a sample with 100 reads, even though the relative abundance in both cases is the same (1%). This information gets lost when converting data into relative abundances. One solution here is to use regression analysis that allow to model these counts directly [71].

Second, variation in sampling extent has other important implications. We do not know the extent to which we have sampled a host's microbiome,

i.e., what fraction of an individual microbiome was collected for sampling? Knowing the fraction of a microbiome that a sample comprises is crucial to extrapolate absolute abundances to the level of the microbiome [72]. In addition, we do not know how thoroughly a sample was assessed, i.e., was the number of sequences sufficient to reveal all of a sample's taxa, or would additional sequencing reveal more taxa? Variation in sampling extent influences the expected number of sampled taxa, where more sequencing reads increases the expected observed microbial richness up to the point of complete assessment [73]. Solutions to address this include rarefying [74], but this is not without criticism [75]. An excess of zero counts results in zero-inflated data, violating the assumption of normally distributed residuals that underlies many parametric statistical tests. Some studies therefore perform log-based transformations to normalize data. However, as we know from community ecology, log transforming count data leads to biased and imprecise estimates, and it involves choosing an arbitrary offset [76]. Further, log-based transformations can lead to incorrect microbiome community-level comparisons, for example resulting in poor estimates of Bray–Curtis dissimilarities [77].

Third, due to the cumulative effect of systematic errors in each experimental step (e.g., extraction, amplification, sequencing), the measured microbiome composition can be highly biased [78]. Finally, as with any model, a quantitative genetic model has a set of assumptions, and the accuracy of microbiome heritability estimates will depend on the extent these assumptions are met. For instance, an important assumption that is often made, is that the variance due to GxE interactions contributes little to the total variance and can therefore be ignored, even though GxE interactions vary in magnitude across populations [79].

How these additional complications further influence the robustness of our microbiome heritability estimates, on top of the issues we describe here, remains to be investigated. With this study, we hope to make researchers aware of the challenges associated with the estimation of microbiome heritabilities. We urge researchers to be careful in interpreting estimates of the heritability of individual taxa, as well as in interpreting the overall proportion of heritable microbes. A focus on consistent results across studies, as well as continued investment in both technical and statistical developments to obtain better approximations of absolute abundances, will likely improve our ability to study the microbiome members that are the most intimately associated with their hosts.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40168-023-01669-w>.

Additional file 1: Appendix S1. Approximating the heritability when based on relative abundances. **Appendix S2.** Numerical simulation results. **Appendix S3.** Power analysis. **Appendix S4.** Summary of published results. **Appendix S5.** Data transformations.

Acknowledgements

Not applicable.

Authors' contributions

MB conceived and performed the analysis, with strong input from CJEM. MB wrote the initial draft of the manuscript, and JFA, LPH, KMM, BK and CJEM contributed substantially to reviewing and editing the manuscript.

Funding

MB is funded by the Dutch Research Council (NWO) (Rubicon grant number 019.192EN.017). BK, KMM and CJEM are funded by the US National Science Foundation (Award number 1754494). LH is funded by the US National Science Foundation (Graduate Research Fellowship Program, #DGE1656466).

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 31 May 2022 Accepted: 13 September 2023

Published online: 09 October 2023

References

- Lynch JB, Hsiao EY. Microbiomes as sources of emergent host phenotypes. *Science*. 1979;2019(365):1405–9.
- Zheng D, Liwinski T, Elinav E. Interaction between microbiota and immunity in health and disease. *Cell Res*. 2020;30:492–506.
- Cho I, Blaser MJ. The human microbiome: at the interface of health and disease. *Nat Rev Genet*. 2012;13:260–70.
- Davenport ER. Elucidating the role of the host genome in shaping microbiome composition. *Gut Microbes*. 2016;7:178–84.
- Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blehman R, et al. Human genetics shape the gut microbiome. *Cell*. 2014;159:789–99.
- Goodrich JK, Davenport ER, Beaumont M, Jackson MA, Knight R, Ober C, et al. Genetic determinants of the gut microbiome in UK twins. *Cell Host Microbe*. 2016;19:731–43.
- Lopera-Maya EA, Kurilshikov A, van der Graaf A, Hu S, Andreu-Sánchez S, Chen L, et al. Effect of host genetics on the gut microbiome in 7,738 participants of the Dutch Microbiome Project. *Nat Genet*. 2022;54:143–51.
- Sanna S, Kurilshikov A, van der Graaf A, Fu J, Zhernakova A. Challenges and future directions for studying effects of host genetics on the gut microbiome. *Nat Genet*. 2022;54:100–6.
- Dethlefsen L, McFall-Ngai M, Relman DA. An ecological and evolutionary perspective on human–microbe mutualism and disease. *Nature*. 2007;449:811–8.
- Hall AB, Tolonen AC, Xavier RJ. Human genetic variation and the gut microbiome in disease. *Nat Rev Gen*. 2017;18:690–9.
- Fisher RA. The correlation between relatives on the supposition of Mendelian inheritance. *Transact Royal Soc Edinburgh*. 1918;52:399–433.
- van Opstal EJ, Bordenstein SR. Rethinking heritability of the microbiome. *Science*. 1979;2015(349):1172–3.
- Grieneisen L, Dasari M, Gould TJ, Björk JR, Grenier J-C, Yotova V, et al. Gut microbiome heritability is nearly universal but environmentally contingent. *Science*. 1979;2021(373):181–6.
- Rothschild D, Weissbrod O, Barkan E, Kurilshikov A, Korem T, Zeevi D, et al. Environment dominates over host genetics in shaping human gut microbiota. *Nature*. 2018;555:210.
- O'Connor A, Quizon PM, Albright JE, Lin FT, Bennett BJ. Responsiveness of cardiometabolic-related microbiota to diet is influenced by host genetics. *Mamm Genome*. 2014;25:583–99.
- Zhao L, Wang G, Siegel P, He C, Wang H, Zhao W, et al. Quantitative genetic background of the host influences gut microbiomes in chickens. *Sci Rep*. 2013;3:1163.
- Davenport ER, Cusanovich DA, Michelini K, Barreiro LB, Ober C, Gilad Y. Genome-wide association studies of the human gut microbiota. *PLoS ONE*. 2015;10:e0140301.
- Turnbaugh PJ, Hamady M, Yatsunenkov T, Cantarel BL, Duncan A, Ley RE, et al. A core gut microbiome in obese and lean twins. *Nature*. 2009;457:480–4.
- Yatsunenkov T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreiras M, et al. Human gut microbiome viewed across age and geography. *Nature*. 2012;486:222–7.
- Wright ML, Fettweis JM, Eaves LJ, Silberg JL, Neale MC, Serrano MG, et al. Vaginal microbiome *Lactobacillus crispatus* is heritable among European American women. *Commun Biol*. 2021;4:1–6.
- Xie H, Guo R, Zhong H, Feng Q, Lan Z, Qin B, et al. Shotgun metagenomics of 250 adult twins reveals genetic and environmental impacts on the gut microbiome. *Cell Syst*. 2016;3:572–84.
- Turpin W, Espin-Garcia O, Xu W, Silverberg MS, Kevans D, Smith MI, et al. Association of host genome with intestinal microbial composition in a large healthy cohort. *Nat Genet*. 2016;48:1413–7.
- Sutherland J, Bell T, Trexler RV, Carlson JE, Lasky JR. Host genomic influence on bacterial composition in the switchgrass rhizosphere. *Mol Ecol*. 2022;31(14):3934–50.
- Wallace RJ, Sasson G, Garnsworthy PC, Tapio I, Gregson E, Bani P, et al. A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions. *Sci Adv*. 2019;5:8391.
- Gomez A, Espinoza JL, Harkins DM, Leong P, Saffery R, Bockmann M, et al. Host genetic control of the oral microbiome in health and disease. *Cell Host Microbe*. 2017;22:269–78.
- Si J, You HJ, Yu J, Sung J, Ko G. *Prevotella* as a hub for vaginal microbiota under the influence of host genetics and their association with obesity. *Cell Host Microbe*. 2017;21:97–105.
- Org E, Parks BW, Joo JWJ, Emert B, Schwartzman W, Kang EY, et al. Genetic and environmental control of host-gut microbiota interactions. *Genome Res*. 2015;25:1558–69.
- Deng S, Caddell DF, Xu G, Dahlen L, Washington L, Yang J, et al. Genome wide association study reveals plant loci controlling heritability of the rhizosphere microbiome. *ISME J*. 2021;15:1–14.
- Lim MY, You HJ, Yoon HS, Kwon B, Lee JY, Lee S, et al. The effect of heritability and host genetics on the gut microbiota and metabolic syndrome. *Gut*. 2017;66:1031–8.
- Ishida S, Kato K, Tanaka M, Odamaki T, Kubo R, Mitsuyama E, et al. Genome-wide association studies and heritability analysis reveal the involvement of host genetics in the Japanese gut microbiota. *Commun Biol*. 2020;3:1–10.
- Kurilshikov A, Medina-Gomez C, Bacigalupe R, Radjabzadeh D, Wang J, Demirkan A, et al. Large-scale association analyses identify host factors influencing human gut microbiome composition. *Nat Genet*. 2021;53:156–65.
- Bergamaschi M, Maltecca C, Schillebeeckx C, McNulty NP, Schwab C, Shull C, et al. Heritability and genome-wide association of swine gut microbiome features with growth and fatness parameters. *Sci Rep*. 2020;10:1–12.
- Walters WA, Jin Z, Youngblut N, Wallace JG, Sutter J, Zhang W, et al. Large-scale replicated field study of maize rhizosphere identifies heritable microbes. *Proc Natl Acad Sci*. 2018;115:7368–73.
- Gacesa R, Kurilshikov A, Vila AV, Sinha T, Klaassen MAY, Bolte LA, et al. The Dutch Microbiome Project defines factors that shape the healthy gut microbiome. *bioRxiv*. 2020;2020:11.

35. Ferretti P, Pasolli E, Tett A, Asnicar F, Gorfer V, Fedi S, et al. Mother-to-infant microbial transmission from different body sites shapes the developing infant gut microbiome. *Cell Host Microbe*. 2018;24:133–45.
36. Bäckhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, et al. Dynamics and stabilization of the human gut microbiome during the first year of life. *Cell Host Microbe*. 2015;17:690–703.
37. Tung J, Barreiro LB, Burns MB, Grenier J-C, Lynch J, Grieneisen LE, et al. Social networks predict gut microbiome composition in wild baboons. *Elife*. 2015;4: e05224.
38. Weinstein SB, Martinez-Mota R, Stapleton TE, Klure DM, Greenhalgh R, Orr TJ, et al. Microbiome stability and structure is governed by host phylogeny over diet and geography in woodrats (*Neotoma* spp.). *Proc Natl Acad Sci*. 2021;118:2108787118.
39. Björk JR, Diez-Vives C, Astudillo-García C, Archie EA, Montoya JM. Vertical transmission of sponge microbiota is inconsistent and unfaithful. *Nat Ecol Evol*. 2019;3:1172–83.
40. Pearson K. Mathematical contributions to the theory of evolution.—on a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proc Royal Soc London*. 1897;60:489–98.
41. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome datasets are compositional: and this is not optional. *Front Microbiol*. 2017;8:2224.
42. Zhou H, Zhang X, He K, Chen J. LinDA: linear models for differential abundance analysis of microbiome compositional data. *Genome Biol*. 2022;23:1 arXiv preprint arXiv:210400242. 2021.
43. Rao C, Coyte KZ, Bainter W, Geha RS, Martin CR, Rakoff-Nahoum S. Multi-kingdom ecological drivers of microbiota assembly in preterm infants. *Nature*. 2021;591:633–8.
44. Carr A, Diener C, Baliga NS, Gibbons SM. Use and abuse of correlation analyses in microbial ecology. *ISME J*. 2019;13:2647–55.
45. Gloor GB, Reid G. Compositional analysis: a valid approach to analyze microbiome high-throughput sequencing data. *Can J Microbiol*. 2016;62:692–703.
46. Nearing JT, Douglas GM, Hayes MG, MacDonald J, Desai DK, Allward N, et al. Microbiome differential abundance methods produce different results across 38 datasets. *Nature Commun*. 2022;13:1–16.
47. de Villemereuil P, Gimenez O, Doligez B. Comparing parent–offspring regression with frequentist and Bayesian animal models to estimate heritability in wild populations: a simulation study for Gaussian and binary traits. *Methods Ecol Evol*. 2013;4:260–75.
48. Wilson AJ, Réale D, Clements MN, Morrissey MM, Postma E, Walling CA, et al. An ecologist's guide to the animal model. *J Anim Ecol*. 2010;79:13–26.
49. Green P, MacLeod CJ. simr: an R package for power analysis of generalised linear mixed models by simulation. *Methods Ecol Evol*. 2016;7:493–8.
50. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. *Nature*. 2011;473:174–80.
51. Hawinkel S, Mattiello F, Bijmans L, Thas O. A broken promise: microbiome differential abundance methods do not control the false discovery rate. *Brief Bioinform*. 2019;20:210–21.
52. Mandal S, Van Treuren W, White RA, Eggesbø M, Knight R, Peddada SD. Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb Ecol Health Dis*. 2015;26:27663.
53. Morton JT, Sanders J, Quinn RA, McDonald D, Gonzalez A, Vázquez-Baeza Y, et al. Balance trees reveal microbial niche differentiation. *mSystems*. 2017;2:e00162-16.
54. Weiss S, Van Treuren W, Lozupone C, Faust K, Friedman J, Deng Y, et al. Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. *ISME J*. 2016;10:1669–81.
55. Maghini DG, Dvorak M, Dahlen A, Roos M, Kuersten S, Bhatt AS. Quantifying bias introduced by sample collection in relative and absolute microbiome measurements. *Nat Biotechnol*. 2023;2023:1–11.
56. Tkacz A, Hortala M, Poole PS. Absolute quantitation of microbiota abundance in environmental samples. *Microbiome*. 2018;6:1–13.
57. Vandeputte D, Kathagen G, D'hoë K, Vieira-Silva S, Valles-Colomer M, Sabino J, et al. Quantitative microbiome profiling links gut community variation to microbial load. *Nature*. 2017;551:507–11.
58. Reese AT, Phillips SR, Owens LA, Venable EM, Langergraber KE, Machanda ZP, et al. Age patterning in wild chimpanzee gut microbiota diversity reveals differences from humans in early life. *Curr Biol*. 2021;31:613–20.
59. Barlow JT, Bogatyrev SR, Ismagilov RF. A quantitative sequencing framework for absolute abundance measurements of mucosal and luminal microbial communities. *Nat Commun*. 2020;11:1–13.
60. Case RJ, Boucher Y, Dahllof I, Holmstrom C, Doolittle WF, Kjelleberg S. Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Appl Environ Microbiol*. 2007;73:278–88.
61. Vandeputte D, De Commer L, Tito RY, Kathagen G, Sabino J, Vermeire S, et al. Temporal variability in quantitative human gut microbiome profiles and implications for clinical research. *Nat Commun*. 2021;12:1–13.
62. Lin H, Peddada Das S. Analysis of microbial compositions: a review of normalization and differential abundance analysis. *NPJ Biofilms Microb*. 2020;6:1–13.
63. Quinn TP, Erb I, Gloor G, Notredame C, Richardson MF, Crowley TM. A field guide for the compositional analysis of any-omics data. *Gigascience*. 2019;8:giz107.
64. Aitchison J. The statistical analysis of compositional data. *J Roy Stat Soc: Ser B (Methodol)*. 1982;44:139–60.
65. Greenacre M, Martínez-Álvarez M, Blasco A. Compositional data analysis of microbiome and any-omics datasets: a validation of the additive Logratio transformation. *Front Microbiol*. 2021;12:2625.
66. Fernandes AD, Reid JN, Macklaim JM, McMurrugh TA, Edgell DR, Gloor GB. Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. *Microbiome*. 2014;2:1–13.
67. Paulson JN, Stine OC, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. *Nat Methods*. 2013;10:1200–2.
68. Halsey LG, Curran-Everett D, Vowler SL, Drummond GB. The fickle P value generates irreproducible results. *Nat Methods*. 2015;12:179–85.
69. Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol Rev*. 2007;82:591–605.
70. Waters JL, Ley RE. The human gut bacteria Christensenellaceae are widespread, heritable, and associated with health. *BMC Biol*. 2019;17:1–11.
71. Martin BD, Witten D, Willis AD. Modeling microbial abundances and dysbiosis with beta-binomial regression. *Ann Appl Stat*. 2020;14:94–115.
72. Lin H, Peddada Das S. Analysis of compositions of microbiomes with bias correction. *Nat Commun*. 2020;11:1–11.
73. Willis AD. Rarefaction, alpha diversity, and statistics. *Front Microbiol*. 2019;10:2407.
74. Sanders HL. Marine benthic diversity: a comparative study. *Am Nat*. 1968;102:243–82.
75. McMurdie PJ, Holmes S. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol*. 2014;10: e1003531.
76. O'Hara RB, Kotze DJ. Do not log-transform count data. *Methods Ecol Evol*. 2010;1:118–22.
77. McKnight DT, Huerlimann R, Bower DS, Schwarzkopf L, Alford RA, Zenger KR. Methods for normalizing microbiome data: an ecological perspective. *Methods Ecol Evol*. 2019;10:389–400.
78. McLaren MR, Willis AD, Callahan BJ. Consistent and correctable bias in metagenomic sequencing experiments. *Elife*. 2019;8:e46923.
79. Saltz JB, Bell AM, Flint J, Gomulkiewicz R, Hughes KA, Keagy J. Why does the magnitude of genotype-by-environment interaction vary? *Ecol Evol*. 2018;8:6342.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.