



REVIEW

# The influence of host genetics on the microbiome [version 1; peer review: 2 approved]

Alexandra Tabrett, Matthew W. Horton

Plant and Microbial Biology, University of Zurich, Zurich, CH-8008, Switzerland

**v1** **First published:** 05 Feb 2020, 9(F1000 Faculty Rev):84 (<https://doi.org/10.12688/f1000research.20835.1>)  
**Latest published:** 05 Feb 2020, 9(F1000 Faculty Rev):84 (<https://doi.org/10.12688/f1000research.20835.1>)

**Abstract**

It is well understood that genetic differences among hosts contribute to variation in pathogen susceptibility and the ability to associate with symbionts. However, it remains unclear just how influential host genes are in shaping the overall microbiome. Studies of both animal and plant microbial communities indicate that host genes impact species richness and the abundances of individual taxa. Analyses of beta diversity (that is, overall similarity), on the other hand, often conclude that hosts play a minor role in shaping microbial communities. In this review, we discuss recent attempts to identify the factors that shape host microbial communities and whether our understanding of these communities is affected by the traits chosen to represent them.

**Keywords**

Host microbiome, plant microbiome, beta diversity, microbes, microbial communities, genetic influence, Microbiome traits, metabolism, plant, plant genetics

**Open Peer Review**

**Reviewer Status**

	Invited Reviewers	
	1	2
<b>version 1</b> 05 Feb 2020		

**F1000 Faculty Reviews** are written by members of the prestigious **F1000 Faculty**. They are commissioned and are peer reviewed before publication to ensure that the final, published version is comprehensive and accessible. The reviewers who approved the final version are listed with their names and affiliations.

- 1 **Maggie R. Wagner**, University of Kansas, Lawrence, USA
- 2 **Rachael E. Antwis**, University of Salford, Salford, UK

Any comments on the article can be found at the end of the article.

**Corresponding author:** Matthew W. Horton ([matthew.horton@botinst.uzh.ch](mailto:matthew.horton@botinst.uzh.ch))

**Author roles:** **Tabrett A:** Writing – Original Draft Preparation, Writing – Review & Editing; **Horton MW:** Conceptualization, Funding Acquisition, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** AT was supported by a ThinkSwiss Scholarship (from the swissnex Network). MWH was funded by the University of Zurich's URPP Evolution in Action

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2020 Tabrett A and Horton MW. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Tabrett A and Horton MW. **The influence of host genetics on the microbiome [version 1; peer review: 2 approved]** F1000Research 2020, 9(F1000 Faculty Rev):84 (<https://doi.org/10.12688/f1000research.20835.1>)

**First published:** 05 Feb 2020, 9(F1000 Faculty Rev):84 (<https://doi.org/10.12688/f1000research.20835.1>)

## Introduction

Microbiome studies often focus on bacteria. However, microbial communities encompass all of the microorganisms in a particular environment. These can include yeasts, filamentous fungi, oomycetes, bacteria, archaea, algae, protists, viruses, nematodes, and even small arthropods. A microbiome, and the interactions within these communities, can directly or indirectly affect a host's health, development, and physiology. In plants, for example, microbes influence important fitness and developmental traits ranging from disease resistance<sup>1</sup> to flowering time<sup>2</sup>. In animals, the microbiome has been shown to influence nutrient uptake<sup>3</sup>, abiotic stress tolerance<sup>4</sup>, and even the development of the central nervous system<sup>5</sup>.

Given the role of microbiota in host health and phenotypic variation, several studies have sought to understand how environmental factors, interactions among microbes, and host genetic differences shape these communities. Whereas some studies have concluded that genetic differences among hosts influence microbiota<sup>6–13</sup>, others have concluded that hosts play at most a minor role<sup>14–16</sup>. The discrepancy in results may arise from differences in perspective: is the proverbial glass (here, the glass contains the heritability of the microbiome) half empty or half full? Differences in study design and methodology also appear to sway results. Perhaps tellingly, studies of model organisms reared in environmentally controlled conditions regularly conclude that microbial communities are under some level of host control<sup>7,11,17,18</sup>; the relationship becomes less clear, however, in environmentally complex field settings<sup>13,16</sup>.

Here, we discuss recent research focused on understanding whether, and to what extent, host microbial communities are under the influence of host genes. Overall, the results from these studies often appear to depend on the approaches used to characterize microbial communities. Therefore, we provide an overview of the most commonly used microbial traits and the possible pitfalls of using each phenotype.

## Microbiome traits

Microbial communities are typically summarized by using one or more of four possible methods. Analyses of beta diversity, alpha diversity, or the abundances of individual taxa are usually conducted after the polymerase chain reaction (PCR) and sequencing of phylogenetically conserved regions, known as marker genes. The most commonly sequenced regions include sections of ribosomal RNA (rRNA) genes or the internal transcribed spacers (ITSs) found in eukaryotic (for example, fungi) DNA. The fourth strategy, characterizing microbial metabolism, is less common; this is likely due to the costs of characterizing microbial communities by using shotgun sequencing or metatranscriptomic approaches.

## Beta diversity

Beta-diversity metrics quantify the overall similarity among samples due to spatial differentiation or other mechanisms, and, given the multivariate nature of microbiome data, these measures have become widely used to understand the factors that influence microbial communities. Perhaps the best known

index is the “semi-metric” developed by the ecologists Bray and Curtis<sup>19</sup>. Alternatively, one can use the phylogenetic distance metric UniFrac, which measures the evolutionary divergence among microbial communities by using a phylogenetic tree constructed from a multiple sequence alignment<sup>20</sup>. Jaccard, Kulczynski, Euclidean, Hellinger, and chi-squared–based distances are other commonly used measures<sup>21</sup>.

Each of these indices has strengths and weaknesses. As an example, phylogenetic distance methods, such as UniFrac, require high-quality sequence alignments that are difficult to generate and curate during large-scale sequencing projects. Notably, high-quality alignments are especially difficult to generate using marker genes that contain sequence-length polymorphisms, such as the ITS regions found in eukaryotic (for example, fungal) DNA. Many (but not all<sup>22</sup>) metrics are also sensitive to differences in the number of sequencing reads among samples<sup>23</sup>, making it necessary to either normalize the data (for example, by expressing the abundances of data as a proportion or a log-ratio) or resample the data to a given read count by using the observed probability distribution within the sample. Identifying robust beta-diversity metrics, and the optimal preprocessing steps, is an area of ongoing research.

Primer mismatches and variation in the number of marker genes among species<sup>24</sup> pose additional problems (for all microbial analyses) and have been implicated in poor reproducibility among sequencing runs<sup>25,26</sup>. Yet another challenge is that zeros are common in species data; that is, many taxa are observed in only one or a few samples. The problem of sparse species data has been discussed in the community ecology literature for decades<sup>27</sup>, but the zeros in microbiome data are particularly difficult to address because they may reflect the real absence of an organism in a sample (that is, a true zero) or they may be the consequence of undersampling and sequencing artifacts.

Some of the most widely used measures of beta diversity were developed by botanists<sup>19,28</sup> investigating whether differences among plant communities across field sites could be attributed to different aspects of the environment. Unfortunately, it is unclear whether the species concept can be applied to prokaryotes, which regularly acquire DNA from the environment through horizontal gene transfer (HGT) mechanisms that blur species boundaries<sup>29</sup>. The frequency of HGT differs across spatial scales<sup>30</sup>, but it is common enough among prokaryotes that pairwise genome-wide DNA similarity and 16S rRNA similarity are poorly correlated<sup>31</sup>. This suggests that when HGT does occur, the phylogenetic markers used to study prokaryotes will poorly represent microbial metabolism and thus the real (dis)similarity among microbiomes.

Despite all of these challenges, beta-diversity measures are regularly used to identify the factors that shape microbiota. The earliest attempts to understand microbiome assembly used gel electrophoresis–based community fingerprinting methods, such as terminal-restriction fragment length polymorphism (T-RFLP) analysis or (automated) ribosomal intergenic spacer analysis (RISA and ARISA). Using RISA and T-RFLP, for

example, Micallef and colleagues<sup>6</sup> found that inbred accessions of *Arabidopsis thaliana* host distinct bacterial communities in the rhizosphere. In similar analyses, Bodenhausen and colleagues<sup>8</sup> and other authors<sup>32</sup> have shown that the leaf microbiome of *A. thaliana* also differs among inbred lines and that mutants defective in the leaf cuticle host altered microbial communities. Over the last decade, as high-throughput sequencing approaches have become more common, these gel-based approaches have largely been replaced by the sequencing of multiplexed amplicon libraries.

Laboratory studies have demonstrated that host genes influence microbial assembly<sup>6–8,17</sup>, which raises the question of whether host effects can also be detected in natural settings, where organisms simultaneously confront abiotic and biotic stresses. Indeed, Redford and colleagues<sup>33</sup> found that trees of the same species have more similar leaf bacteriomes than trees of different species. Consistent with the results of Micallef and colleagues<sup>6</sup> and others<sup>8,9</sup>, inbred accessions of *A. thaliana* have likewise been shown to host distinct microbial communities in field settings. Specifically, both bacteria and fungi in the leaf<sup>34</sup> and root<sup>18</sup> microbiome of *Arabidopsis* are influenced by host genes. Moreover, it is now clear that genome-wide association studies (GWASs) can be used to shed light on the processes shaping microbiome assembly. For example, several promising candidate genes have been identified by using the sample coordinates from principal components analysis as phenotypes in GWASs. The top single-nucleotide polymorphisms (SNPs) associated with leaf fungi (PC1), for example, fall within *GLUCAN SYNTHASE LIKE 11 (GSL11)*, a homolog of a gene (*GSL5*) that deposits callose in the cell wall in response to fungal infection<sup>35</sup>. Similarly, fungi in the roots (PC1) of *A. thaliana* are associated with promising candidate loci, including *PECTIN METHYLESTERASE 26 (PME26)* and its neighbor *PME3*<sup>18</sup>. Like cellulose, pectin and callose are polysaccharides that play critical roles in cell-wall integrity.

Much of what we know about animal microbial communities comes from diverse human microbiome projects. For example, Blekman and colleagues<sup>36</sup> used data from the Human Microbiome Project to explore microbiome habitats distributed across the body. This revealed that host genome similarity, which was cleverly estimated by using host contaminant reads, is moderately correlated ( $R^2 \sim 0.19$ ) with beta diversity in the stool microbiome. Unfortunately, the small sample size of this study ( $n = 93$  individuals) and the failure to correct for population structure have made it difficult to interpret the results from their GWAS. However, the gut microbiomes of much larger human diversity panels are now being characterized. As an example, Wang and colleagues<sup>11</sup> used a classic epidemiological approach to investigate the gut microbiomes of a large northern European cohort ( $n = 1812$ ). In that study, the authors were able to attribute approximately 10% of the variation in the microbiota to genetic differences among individuals. Differences in age, gender, body mass index, and smoking status explained about 9% of the overall variation, and controlling for these covariates in GWASs helped fine-map vitamin D receptor (*VDR*) among the top candidate genes. Follow-up research

with a mouse model (*Vdr*<sup>-/-</sup>) has confirmed the role of *VDR* and helped in understanding its role in host-microbiome cross-talk through bile-acid sensing. Wang and colleagues have established the Microbiome Genome consortium<sup>37</sup> to further investigate the role of human genetics in shaping gut microbiota and the role of the microbiome in human disease. This project now has data for over 19,000 subjects for which rich metadata (for example, smoking status, age, and weight) are also being generated. The size of this cohort, and the accompanying resources, should further aid in understanding the genetic architecture of the microbiome.

Analyses of beta diversity for both plant and animal microbiomes have greatly improved our understanding of the factors that shape microbiota. Despite clear progress, however, the emerging view from this research is that environmental factors are more influential than host genes in shaping microbiomes (for example, as discussed in 16). In particular, it is difficult to detect host effects when environmental factors are not explicitly controlled for or taken into account during analysis<sup>11,38,39</sup>. It is arguable that dispersal limitation and HGT should also be considered when designing experiments. Indeed, one should question the biological relevance of many microbiome traits (to begin with) before sequencing marker genes (for example, 16S rRNA) that poorly represent the phenotypic diversity of communities shaped by HGT<sup>31</sup>. In the following sections, we discuss whether, despite these concerns, other microbiome traits are heritable.

### Alpha diversity

Alpha diversity describes local diversity, or the diversity within a sample (or habitat)<sup>40</sup>. As is the case with beta diversity, there are multiple ways to characterize alpha diversity.

Perhaps the most widely known measure of alpha diversity is Shannon's diversity index<sup>41</sup>, which estimates the uncertainty in a sampling process by weighing the informativeness of elements in a series (for community and microbial ecologists, the elements are species) on the basis of their observed frequencies. This entropy is then exponentiated to estimate the effective number of species within a community, or true alpha diversity<sup>42</sup>. The widespread use of Shannon's index in ecology stems from the fact that it is biased toward neither common nor rare species. Despite this intuitive appeal, however, the phylogenetic markers that are used in most studies—sections of bacterial 16S rRNA<sup>43</sup> and the ITS regions of fungi<sup>44</sup>—vary widely in gene copy number across species. This means that Shannon's index is inappropriate unless corrective steps are taken<sup>24</sup> (for example, dividing each species' abundance by its predicted number of marker gene copies) that are difficult to apply to many taxa.

Many researchers instead focus on species richness, the number of taxa within a sample, to understand alpha diversity. Species richness does not take abundance into account, and it can be investigated without rarefaction by using long-established statistical models (for example, negative binomial or quasi-Poisson generalized linear models) that allow one to adjust

for differences in the number of sequences among samples by using “offsets” to model the exposure rate. This simplicity, combined with its low statistical burden, has made richness a popular phenotype.

As an example, bacterial richness in the maize rhizosphere<sup>38</sup> is shaped by genetic differences among accessions ( $R^2 = 19\%$ ), as are the number of bacterial (heritability [ $H^2$ ] = 0.57) and fungal ( $H^2 = 0.47$ ) taxa in the root microbiome of *A. thaliana*<sup>18</sup>. Of course, bacteria and fungi co-occur in many microbiomes, which suggests that they should be combined in analyses when appropriate. In the case of the root microbiome of *A. thaliana*, combining bacterial and fungal richness data enabled Bergelson and colleagues<sup>18</sup> to fine-map *CELLULOSE1* and other promising candidate genes implicated in immunity. The top candidate SNPs fall nearby the disease resistance gene *NPRI*<sup>45</sup>, which was not mapped in either of the marginal analyses.

Another measure of alpha diversity is Faith’s phylogenetic diversity (PD)<sup>46</sup>, which uses phylogenetic trees generated with molecular or cladistic data to estimate diversity. Like species richness, Faith’s PD does not take “abundance” into account and thus avoids the marker gene copy-number problem. Faith’s PD has become a popular measure in conservation biology, where geographic regions with higher phylogenetic diversity are considered to be more diverse (and thus more valuable) than communities of closely related species. In the case of microbiome studies, analyses of Faith’s PD have revealed that the gut bacterial communities of both wild mouse<sup>47</sup> and humans<sup>48</sup> are heritable (for example, human  $H^2 = 0.37$ ). Curiously, GWASs of the wild mouse<sup>47</sup> and human gut microbiome<sup>11</sup> have pinpointed the same candidate gene: *CSMD1*.

### Individual taxa

Characterizing microbiome beta and alpha diversity has greatly improved our understanding of the overall processes that shape microbial communities. Although microbes interact in these communities, many researchers are nevertheless interested in the presence/absence and abundances of individual taxa and whether individual taxa are also shaped by host genes.

As an example in plants, Walters and colleagues<sup>13</sup> recently identified 143 heritable operational taxonomic units (OTUs) in the rhizosphere of maize. Overall, the broad-sense  $H^2$  of these taxa were low ( $15 < H^2 < 25\%$ ) compared with taxa in the phyllosphere of maize<sup>49</sup>, which raises the question of whether these two habitats are differentially affected by environmental effects and host genes. Support for this hypothesis comes from work in *Boechera*<sup>50</sup>, which similarly suggests that leaf microbiomes are more affected by host genes than root microbiomes.

In the case of the human gut microbiome, Goodrich and colleagues<sup>3,12</sup> have identified several heritable bacterial groups. A member of the most highly heritable ( $H^2 = 0.42$ ) taxon, the Christensenellaceae (specifically, *Christensenella minuta*), has been shown to reduce weight gain in transplantation experiments with germfree recipient mice<sup>3</sup>. Moreover, network analyses

have revealed that the Christensenellaceae are the key group (based on node “degree”, or the number of network connections) in a sub-network of several heritable bacterial groups within the gut microbiome. This suggests that host genes may indirectly shape microbial communities through more direct interactions with key taxa (that is, “hub points”) in microbial community networks.

Genotype-by-environment interactions (GxEs) are widely believed to shape host microbiome traits, but the prevalence of GxEs remains unclear. The best-characterized example involves the lactase enzyme, which is encoded by the *LCT* locus. Polymorphisms at the human *LCT* locus determine the ability of adult humans to digest milk. These variants, however, not only are linked to lactase persistence but also with the abundance of the genus *Bifidobacterium* in the human gut microbiome. Strains of *Bifidobacterium* that can metabolize lactose are found in (significantly) increasing abundance with the increasing consumption of dairy products in the microbiomes of individuals homozygous for the haplotype associated with hypolactasia (lactose intolerance) in Europeans; this is strong evidence of a GxE<sup>10</sup>.

Most microbial communities contain a large number of taxa, and the number increases with increasing sample size as rare taxa are discovered. Although this poses challenges for beta diversity (see above), studies of individual taxa often focus on the most heavily sequenced taxa (for example, the top 100 species) to improve speed and reduce the burden of multiple testing. So it is perhaps surprising that recent studies of bacteria have sought to examine both the heritability of individual taxa and (separately) the same data binned at increasingly higher taxonomic levels. As discussed above, prokaryotes experience high rates of HGT and gene loss, which suggests that their physiology is not conserved enough to justify the increased (statistical) burden of testing taxa at both low (for example, genus) and high (for example, phylum) taxonomic levels.

### Microbial metabolism

As described above, amplicon-based analyses of beta diversity, alpha diversity, and the abundances of individual taxa are hampered in varying degrees by both HGT and biases introduced during PCR and sequencing<sup>25,26</sup>. One solution is to directly sequence microbial genes or RNA by using metagenome shotgun sequencing or transcriptomics. The key advantage of these two approaches is that the metabolic potential of the microbiome is directly measured rather than predicted from marker gene sequences. For studies focused on understanding community assembly processes, these untargeted sequencing approaches also have the potential to characterize species membership in a less biased and PCR-free manner while enabling the identification of traditionally overlooked groups (oomycetes, viruses, Cyanobacteria, and so on).

Among the disadvantages of metagenomes and metatranscriptomes is that they are expensive to generate, which is perhaps why they are rarely characterized outside of human and crop research. Moreover, host sequences may be enriched in some

habitats (for example, 51,52), resulting in contamination that further increase the costs of sequencing. The studies that have been performed so far have shown that microbial metabolism varies widely among species grown in the same environment, suggesting that hosts do affect microbial gene expression in some manner. As an example, enzymes involved in nitrate reduction are more highly expressed in the wheat than cucumber microbiome of plants grown in a randomized experiment<sup>53</sup>. In the cucumber microbiome, enzymes associated with sulfur assimilation and pectin degradation are upregulated relative to wheat. As noted above, pectin, which is generally more abundant in the cell walls of dicots than monocots<sup>54</sup>, is one of the polysaccharides that regulate host cell-wall integrity.

It is unclear how heritable individual microbial enzymatic categories and pathways are, but the limited number of studies that have explored the relationship between microbial metabolism and host variation have identified promising associations, although the abundances of taxa and microbiome functions appear to be influenced by different host genes<sup>10</sup>. This is consistent with an emerging theme from microbial research: the microbial composition of samples is a poor proxy of microbial gene expression within the same samples and this is due in part to HGT<sup>55</sup>. For example, the inner leaves of bromeliads form water traps that in turn form aquatic ecosystems. By sequencing the bacteria and the microbial genes within these aquatic islands, Louca and colleagues demonstrated that, despite hosting highly diverse bacterial communities, the genes expressed within these communities are remarkably similar<sup>56</sup>. If a microbial community experiences high rates of HGT, it will be far more informative to characterize microbial metabolism than to sequence marker genes that are, in effect, decoupled from the genomes that they are intended to represent.

### Future research perspectives

Research in the past decade has illustrated a close relationship between a host's microbiome and health<sup>3,5</sup>, which naturally leads to questions about how these communities form and whether host genes play a role. In this review, we summarize recent research focused on this topic while discussing whether

the choice of microbiome traits affects our understanding of these communities.

What is the most meaningful way to summarize the microbiome? The most common approach is to characterize the overall similarity of microbiomes by using beta-diversity measures. These studies often find that microbiomes vary dramatically among conspecifics that live in different areas<sup>13,50,57</sup>, which leads to the reasonable conclusion that spatially varying environmental factors differentially shape these communities. However, in most studies, it is unclear whether environmental factors are confounded with ecological processes, such as ecological drift and dispersal limitation. Whatever the research focus, if everything is *not* everywhere, approaches based on beta-diversity will be underpowered when samples from distant geographic regions are under consideration.





In comparison, analyses of individual taxa<sup>13,48,49</sup> and species richness<sup>18,38</sup> have detected (modestly more) heritable components of the microbiome<sup>18,49</sup>. Despite these promising results, however, it should be noted that analyses of alpha diversity and individual OTUs share many of the same problems as beta diversity, including whether microbiomes distributed across the environment experience the same exposure rate (that is, due to dispersal limitation). Confounding factors are regularly included as covariates to study human diseases, and the strategies used in epidemiology<sup>58</sup> probably deserve more attention in the study of host microbiota<sup>11,59</sup>.

A major limitation of current host microbiome research is the fact that microbes are often shaped by HGT but characterized with marker genes (for example, 16S rRNA) that fail to distinguish among genetically divergent strains<sup>60</sup>. As we move forward in this research, it will be critical to determine whether metagenomic or metatranscriptomic studies, which are unaffected by HGT, can uncover additional components of heritability.

### Acknowledgments

We thank Jana Mittelstrass (University of Zurich) for helpful discussions and comments.

### References

1.  Mendes R, Kruijt M, de Bruijn I, *et al.*: **Deciphering the rhizosphere microbiome for disease-suppressive bacteria.** *Science*. 2011; **332**(6033): 1097–100.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [F1000 Recommendation](#)
2. Panke-Buisse K, Poole AC, Goodrich JK, *et al.*: **Selection on soil microbiomes reveals reproducible impacts on plant function.** *ISME J*. 2015; **9**(4): 980–9.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
3.  Goodrich JK, Waters JL, Poole AC, *et al.*: **Human genetics shape the gut microbiome.** *Cell*. 2014; **159**(4): 789–99.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
4. Chevalier C, Stojanović O, Colin DJ, *et al.*: **Gut Microbiota Orchestrates Energy Homeostasis during Cold.** *Cell*. 2015; **163**(6): 1360–74.  
[PubMed Abstract](#) | [Publisher Full Text](#)
5.  Erny D, Hrabě de Angelis AL, Jaitin D, *et al.*: **Host microbiota constantly control maturation and function of microglia in the CNS.** *Nat Neurosci*. 2015; **18**(7): 965–77.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
6. Micallef SA, Shiaris MP, Colón-Carmona A: **Influence of *Arabidopsis thaliana* accessions on rhizobacterial communities and natural variation in root exudates.** *J Exp Bot*. 2009; **60**(6): 1729–42.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
7.  Benson AK, Kelly SA, Legge R, *et al.*: **Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors.** *Proc Natl Acad Sci U S A*. 2010; **107**(44): 18933–8.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)



8. Bodenhausen N, Bortfeld-Miller M, Ackermann M, *et al.*: **A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota.** *PLoS Genet.* 2014; **10**(4): e1004283.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
9. Lebeis SL, Paredes SH, Lundberg DS, *et al.*: **PLANT MICROBIOME. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa.** *Science.* 2015; **349**(6250): 860–4.  
[PubMed Abstract](#) | [Publisher Full Text](#)
10. Bonder MJ, Kurilshikov A, Tigchelaar EF, *et al.*: **The effect of host genetics on the gut microbiome.** *Nat Genet.* 2016; **48**(11): 1407–12.  
[PubMed Abstract](#) | [Publisher Full Text](#)
11. Wang J, Thingholm LB, Skieceviciene J, *et al.*: **Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota.** *Nat Genet.* 2016; **48**(11): 1396–406.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
12. Goodrich JK, Davenport ER, Beaumont M, *et al.*: **Genetic Determinants of the Gut Microbiome in UK Twins.** *Cell Host Microbe.* 2016; **19**(5): 731–43.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. **F** Walters WA, Jin Z, Youngblut N, *et al.*: **Large-scale replicated field study of maize rhizosphere identifies heritable microbes.** *Proc Natl Acad Sci U S A.* 2018; **115**(28): 7368–73.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
14. **F** Bulgarelli D, Rott M, Schlaeppi K, *et al.*: **Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota.** *Nature.* 2012; **488**(7409): 91–5.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [F1000 Recommendation](#)
15. **F** Lundberg DS, Lebeis SL, Paredes SH, *et al.*: **Defining the core *Arabidopsis thaliana* root microbiome.** *Nature.* 2012; **488**(7409): 86–90.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
16. **F** Rothschild D, Weissbrod O, Barkan E, *et al.*: **Environment dominates over host genetics in shaping human gut microbiota.** *Nature.* 2018; **555**(7695): 210–5.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [F1000 Recommendation](#)
17. Srinivas G, Möller S, Wang J, *et al.*: **Genome-wide mapping of gene-microbiota interactions in susceptibility to autoimmune skin blistering.** *Nat Commun.* 2013; **4**: 2462.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
18. Bergelson J, Mittelstrass J, Horton MW: **Characterizing both bacteria and fungi improves understanding of the *Arabidopsis* root microbiome.** *Sci Rep.* 2019; **9**(1): 24.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
19. Bray JR, Curtis JT: **An Ordination of the Upland Forest Communities of Southern Wisconsin.** *Ecol Monogr.* 1957; **27**(4): 325–49.  
[Publisher Full Text](#)
20. Lozupone C, Knight R: **UniFrac: a new phylogenetic method for comparing microbial communities.** *Appl Environ Microbiol.* 2005; **71**(12): 8228–35.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
21. **F** Legendre P, De Cáceres M: **Beta diversity as the variance of community data: dissimilarity coefficients and partitioning.** *Ecol Lett.* 2013; **16**(8): 951–63.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [F1000 Recommendation](#)
22. **F** Martino C, Morton JT, Marotz CA, *et al.*: **A Novel Sparse Compositional Technique Reveals Microbial Perturbations.** *mSystems.* 2019; **4**(1): pii: e00016-19.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
23. Wong RG, Wu JR, Gloor GB: **Expanding the UniFrac Toolbox.** *PLoS One.* 2016; **11**(9): e0161196.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
24. Kembel SW, Wu M, Eisen JA, *et al.*: **Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance.** *PLoS Comput Biol.* 2012; **8**(10): e1002743.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
25. **F** Wen C, Wu L, Qin Y, *et al.*: **Evaluation of the reproducibility of amplicon sequencing with Illumina MiSeq platform.** *PLoS One.* 2017; **12**(4): e0176716.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
26. Edgar RC: **UNBIAS: An attempt to correct abundance bias in 16S sequencing, with limited success.** *bioRxiv.* 2017.  
[Publisher Full Text](#)
27. Goodall DW: **Objective methods for the classification of vegetation. III. An essay in the use of factor analysis.** *Aust J Bot.* 1954; **2**(3): 304–324.  
[Publisher Full Text](#)
28. Jaccard P: **Contribution au problème de l'immigration post-glaciaire de la flore alpine.** *Bull de la Soc Vaudoise des Sci Nat.* 1900; **36**(136): 87–130.  
[Publisher Full Text](#)
29. Ochman H, Lawrence JG, Groisman EA: **Lateral gene transfer and the nature of bacterial innovation.** *Nature.* 2000; **405**(6784): 299–304.  
[PubMed Abstract](#) | [Publisher Full Text](#)
30. **F** Smillie CS, Smith MB, Friedman J, *et al.*: **Ecology drives a global network of gene exchange connecting the human microbiome.** *Nature.* 2011; **480**(7376): 241–4.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [F1000 Recommendation](#)
31. Stackebrandt E, Goebel BM: **Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology.** *Int J Syst Evol Microbiol.* 1994; **44**(4): 846–9.  
[Publisher Full Text](#)
32. Ritpitakphong U, Falquet L, Vimoltust A, *et al.*: **The microbiome of the leaf surface of *Arabidopsis* protects against a fungal pathogen.** *New Phytol.* 2016; **210**(3): 1033–43.  
[PubMed Abstract](#) | [Publisher Full Text](#)
33. Redford AJ, Bowers RM, Knight R, *et al.*: **The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves.** *Environ Microbiol.* 2010; **12**(11): 2885–93.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
34. Horton MW, Bodenhausen N, Beilsmith K, *et al.*: **Genome-wide association study of *Arabidopsis thaliana* leaf microbial community.** *Nat Commun.* 2014; **5**: 5320.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
35. Jacobs MA, Alwood A, Thaipisuttikul I, *et al.*: **Comprehensive transposon mutant library of *Pseudomonas aeruginosa*.** *Proc Natl Acad Sci U S A.* 2011; **100**(24): 14339–44.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
36. **F** Blekhan R, Goodrich JK, Huang K, *et al.*: **Host genetic variation impacts microbiome composition across human body sites.** *Genome Biol.* 2015; **16**: 191.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
37. **F** Wang J, Kurilshikov A, Radjabzadeh D, *et al.*: **Meta-analysis of human genome-microbiome association studies: the MiBioGen consortium initiative.** *Microbiome.* 2018; **6**(1): 101.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
38. Peiffer JA, Spor A, Koren O, *et al.*: **Diversity and heritability of the maize rhizosphere microbiome under field conditions.** *Proc Natl Acad Sci U S A.* 2013; **110**(16): 6548–53.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
39. Turpin W, Espin-Garcia O, Xu W, *et al.*: **Association of host genome with intestinal microbial composition in a large healthy cohort.** *Nat Genet.* 2016; **48**(11): 1413–7.  
[PubMed Abstract](#) | [Publisher Full Text](#)
40. Whittaker RH: **Vegetation of the Siskiyou Mountains, Oregon and California.** *Ecol Monogr.* 1960; **30**(3): 279–338.  
[Publisher Full Text](#)
41. Shannon CE: **A Mathematical Theory of Communication.** *Bell Labs Tech J.* 1948; **27**(4): 623–56.  
[Publisher Full Text](#)
42. Jost L: **Entropy and diversity.** *Oikos.* 2006; **113**(2): 363–75.  
[Publisher Full Text](#)
43. Větrovský T, Baldrian P: **The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses.** *PLoS One.* 2013; **8**(2): e57923.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
44. **F** Schoch CL, Seifert KA, Huhndorf S, *et al.*: **Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for *Fungi*.** *Proc Natl Acad Sci U S A.* 2012; **109**(16): 6241–6.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
45. Cao H, Glazebrook J, Clarke JD, *et al.*: **The *Arabidopsis* *NPR1* gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats.** *Cell.* 1997; **88**(1): 57–63.  
[PubMed Abstract](#) | [Publisher Full Text](#)
46. Faith DP: **Conservation evaluation and phylogenetic diversity.** *Biol Conserv.* 1992; **61**: 1–10.  
[Publisher Full Text](#)
47. **F** Suzuki TA, Phifer-Rixey M, Mack KL, *et al.*: **Host genetic determinants of the gut microbiota of wild mice.** *Mol Ecol.* 2019; **28**(13): 3197–207.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
48. Goodrich JK, Davenport ER, Waters JL, *et al.*: **Cross-species comparisons of host genetic associations with the microbiome.** *Science.* 2016; **352**(6285): 532–5.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
49. **F** Wallace JG, Kremling KA, Kovar LL, *et al.*: **Quantitative Genetics of the Maize Leaf Microbiome.** *Phytozomes J.* 2018; **2**(4): 208–24.  
[Publisher Full Text](#) | [F1000 Recommendation](#)
50. Wagner MR, Lundberg DS, Del Rio TG, *et al.*: **Host genotype and age shape the leaf and root microbiomes of a wild perennial plant.** *Nat Commun.* 2016; **7**: 12151.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
51. Jiang Y, Xiong X, Danska J, *et al.*: **Metatranscriptomic analysis of diverse microbial communities reveals core metabolic pathways and microbiome-specific functionality.** *Microbiome.* 2016; **4**: 2.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
52. Karasov TL, Almario J, Friedemann C, *et al.*: ***Arabidopsis thaliana* and**

- Pseudomonas* Pathogens Exhibit Stable Associations over Evolutionary Timescales.** *Cell Host Microbe*. 2018; **24**(1): 168–179.e4.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
53. **F** Ofek-Lalzar M, Sela N, Goldman-Voronov M, *et al.*: **Niche and host-associated functional signatures of the root surface microbiome.** *Nat Commun*. 2014; **5**: 4950.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [F1000 Recommendation](#)
54. McNeil M, Darvill AG, Fry SC, *et al.*: **Structure and function of the primary cell walls of plants.** *Annu Rev Biochem*. 1984; **53**: 625–63.  
[PubMed Abstract](#) | [Publisher Full Text](#)
55. Burke C, Steinberg P, Rusch D, *et al.*: **Bacterial community assembly based on functional genes rather than species.** *Proc Natl Acad Sci U S A*. 2011; **108**(34): 14288–93.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
56. **F** Louca S, Jacques SMS, Pires APF, *et al.*: **High taxonomic variability despite stable functional structure across microbial communities.** *Nat Ecol Evol*. 2016; 1(1): 15.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [F1000 Recommendation](#)
57. **F** Yatsunenko T, Rey FE, Manary MJ, *et al.*: **Human gut microbiome viewed across age and geography.** *Nature*. 2012; **486**(7402): 222–7.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
58. Schardl CL, Leuchtman A, Spiering MJ: **Symbioses of grasses with seedborne fungal endophytes.** *Annu Rev Plant Biol*. 2004; **55**: 315–40.  
[PubMed Abstract](#) | [Publisher Full Text](#)
59. **F** Sanna S, van Zuydam NR, Mahajan A, *et al.*: **Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases.** *Nat Genet*. 2019; **51**(4): 600–5.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
60. **F** Jaspers E, Overmann J: **Ecological significance of microdiversity: identical 16S rRNA gene sequences can be found in bacteria with highly divergent genomes and ecophysiologicals.** *Appl Environ Microbiol*. 2004; **70**(8): 4831–9.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)



# Open Peer Review

Current Peer Review Status:  

---

## Editorial Note on the Review Process

F1000 Faculty Reviews are written by members of the prestigious F1000 Faculty. They are commissioned and are peer reviewed before publication to ensure that the final, published version is comprehensive and accessible. The reviewers who approved the final version are listed with their names and affiliations.

---

## The reviewers who approved this article are:

### Version 1

- Rachael E. Antwis**  
School of Science, Engineering and Environment, University of Salford, Salford, UK  
**Competing Interests:** No competing interests were disclosed.
- Maggie R. Wagner**  
University of Kansas, Lawrence, KS, 66045, USA  
**Competing Interests:** No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)

F1000Research