REVIEW ARTICLE

WILEY

Impact of host genetics on gut microbiome: Take-home lessons from human and mouse studies

Inbal Cahana | Fuad A. Iraqi

Department of Human Microbiology and Immunology, Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel

Correspondence
Fuad A. Iraqi, Department of Human Microbiology and Immunology, Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv 69978, Israel.
Email: fuadi@tauex.tau.ac.il

Abstract

The intestinal microbiome has emerged as an important component involved in various diseases. Therefore, the interest in understanding the factors shaping its composition is growing. The gut microbiome, often defined as a complex trait, contains diverse components and its properties are determined by a combination of external and internal effects. Although much effort has been invested so far, it is still difficult to evaluate the extent to which human genetics shape the composition of the gut microbiota. However, in mouse studies, where the environmental factors are better controlled, the effect of the genetic background was significant. The purpose of this paper is to provide a current assessment of the role of human host genetics in shaping the gut microbiome composition. Despite the inconsistency of the reported results, it can be estimated that the genetic factor affects a portion of the microbiome. However, this effect is currently lower than the initial estimates, and it is difficult to separate the genetic influence from the environmental effect. Additionally, despite the differences between the microbial composition of humans and mice, results from mouse models can strengthen our knowledge of host genetics underlying the human gut microbial variation.

KEYWORDS
host genetic background, intestinal microbiome in human and mouse, microbial variations and profiles, take-home lessons

1 | HUMAN GUT MICROBIOME

The human body is constantly interacting with various microorganisms that cover its external surfaces and are collectively called the microbiome. The gut microbiome, which contains more than 1000 species of bacteria, is considered the largest and most studied in the field to date.[1,2] Although it was previously thought that the gut microbiome formation begins at birth, recent studies suggest that this initial development occurs at prenatal stages.[3] After birth, especially during the first 3 years of life, the development of this microbiome is accelerated mainly due to environmental exposure. It has been demonstrated that factors, such as nutrition,[4] antibiotics,[5] mode of delivery, and cessation of breast-feeding,[6] have a major impact on its shape in adults.[7] In addition to the dynamics and variability of the microbial communities within the individual, the composition of the microbiome varies significantly between individuals; for example, the species that inhabit the gut, the relative ratios among different bacteria, and the identity of the dominant species.[8-10]

In the past decade, especially after the establishment of the Human Microbiome Project,[10] great progress has been made in

understanding the importance of the gut microbiome in maintaining our health.[11,12] It has been discovered that changes in the gut microbiome are associated with various diseases and health conditions.[13] Increasing evidence of intestinal microbiota involvement in pathogenesis and disease development reveals its potential as a promising therapeutic target for disease management, prevention, and cure.[14] Thus, considerable effort is focused on understanding the factors that shape microbial composition. The gut microbiome is a complex trait affected by multiple factors, including genetics and environment.[15] The typical methods for characterizing the microbiome are either with 16S ribosomal RNA (rRNA) gene sequencing, shotgun, or high-throughput metagenome sequencing techniques.[8] These techniques allow quantification of taxa or gene functions, which can be used as a database for different ecological metrics that characterize diversity in a sample or within a population.[15] The early studies focused on 16S rRNA sequences which are relatively short, often conserved within a species, and generally different between species. Many 16S rRNA sequences have been found, which do not belong to any known cultured species, indicating that there are numerous non-isolated organisms. 16S rRNA sequencing has been developed in response to the need for more rapid and accurate identification of complex microbiome.[16] The 16S rRNA gene "Marker gene"[17] codes for a ribosomal subunit that is widely conserved among bacteria and contains hypervariable regions V1-V9 interspersed among conserved regions of its sequence. These hypervariable regions are unique to each bacterial species, allowing for classification or taxonomy. The conserved regions, on the other hand, allow for the development of universal primers that bind to known sequences shared among most bacteria. The 16S rRNA gene is amplified from total extracted DNA using universal primers to target the conserved regions of the gene, and the resulting PCR products are sequenced to identify the bacterial species present.[18] For the nine hypervariable regions, some regions better characterize bacteria and the choice of which region, and therefore, the appropriate primers to use are an important step in study design.[19] Regardless of the sequencing method, final results are represented in operational taxonomic units (OTUs), which is a sequence identifying an organism usually at the genus or species level. OTUs are based on similarity, in which, the similarity between a pair of sequences is computed as the percentage of sites that agree in a pairwise sequence alignment. A common similarity threshold used is 97%, which was derived from an empirical study that showed most strains had 97% 16S rRNA sequence similarity.[20]

While with the advancement in refinements of DNA amplification, the proliferation of computational power and bioinformatics tools have greatly aided the analysis of DNA sequences recovered from environmental samples, allowing the adaptation of shotgun sequencing to metagenomic samples (known also as whole metagenome shotgun or WMGS sequencing). Shotgun sequencing is a method used for sequencing random DNA strands.[21,22] The approach, used to sequence many cultured microorganisms and the human genome, randomly shears DNA, sequences many short sequences (reads), overlapping ends of different reads to assemble them into a continuous sequence and subsequently reconstruct them into a publically consensus sequence, using sophisticated computer programs. Shotgun metagenomics provides information both about, which organisms are present and what metabolic processes are possible in the community.[23] Because the collection of DNA from an environment is largely uncontrolled, the most abundant organisms in an environmental sample are most highly represented in the resulting sequence data. To achieve the high coverage needed to fully resolve the genomes of under-represented community members, large samples, often prohibitively so, are needed. In contrast, the random nature of shotgun sequencing ensures that many of these organisms, which would otherwise go unnoticed using traditional culturing techniques, will be represented by at least some small sequence segments.[24]

Finally, recently a high-throughput sequencing technique, which does not require cloning the DNA before sequencing, removing one of the main biases and bottlenecks in environmental sampling was developed. The first metagenomic studies conducted using high-throughput sequencing used massively parallel 454 pyrosequencing.[25] Three other technologies commonly applied to environmental sampling are the Ilumina MiSeq and HiSeq, and the Applied Biosystems SOLid systems.[26] These techniques for sequencing DNA generate shorter fragments than Sanger sequencing; 454 pyrosequencing typically produces ~400 bp reads, Illumina MiSeq produces 400-700 bp reads (depending on whether paired-end options are used), and SOLiD produce 25-75 bp reads.[27] Using this technique, it is possible to comprehensively sequence the entire microbiota in a given biological sample and subsequently compare it to a publically consensus sequence.

Using these molecular genetic sequence methods, it will be possible to identify and quantify the different microbiome species and families in a given biological sample. Figure 1 shows a diagram that describes the procedures the assessment of microbial populations in a given biological sample using the three methods including shotgun and high-throughput metagenomics, and 16S-based sequence approaches.

While environmental factors such as diet and lifestyle have been shown to influence the microbiome composition, the role of the host genetics remains unclear.[28] The microbiome structural complexity, the challenge of multiple comparisons, and the large effect of the environment, probably result in variability and inconsistency between the results of recent studies.[28] Therefore, the question of the extent to which human genetics shapes the microbiome composition remains open. In this review, we will present the recent studies and discuss their results in order to provide a current assessment of the role of host genetics in shaping the gut microbiome composition. Additionally, we will review the major insights from mouse studies and discuss how this model can contribute to future human microbiome research.

2 | ASSESSMENT OF THE MICROBIOME HERITABILITY FROM TWIN STUDIES

The term "heritability" describes the inheritance extent of quantitative traits (eg, the abundance of gut bacteria), or in other words,

FIGURE 1 Shotgun and high-throughput metagenomics, and 16S-based sequence approaches for assessing microbial populations in a given biological sample
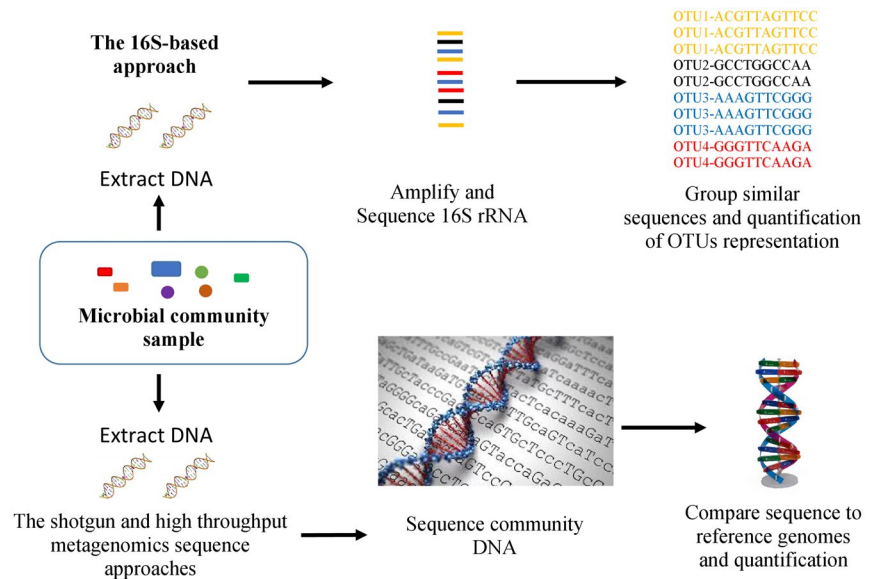
the proportion of variation in a trait that can be explained by the host genetic.[29] In order to investigate and assess the role of genetic background that determines the characteristics of the microbiome, environmental effects should be minimized as much as possible. One of the effective approaches is to compare the differences between groups of monozygotic (MZ) and dizygotic (DZ) twins.[28,30] Assuming that each type of twin experiences the same environmental conditions and that there is a difference in genetic similarity between MZ (100%) and DZ (~50%) twins, the heritability of traits can be estimated. Figure 2 represents a diagram shows the microbial similarity in monozygotic twins, while genetic diversity is observed in dizygotic twins.

Although twin studies are the basic method to investigate whether genetic variation in the host is related to genetic variation in the microbiome, they have not provided the same answers over the years.[8,9] Earlier studies reported that there was no significant difference in gut microbiome composition between MZ and DZ twins.[31,32] In contrast, recent studies have shown opposite results suggesting that the abundances of specific members of the intestinal microbiota are slightly influenced by host genetics.[8,33] While the first two studies[31,32] were based on relatively small cohorts (54 and 87 twin pairs, respectively), the current studies[8,9] significantly increased the sample size (416 and 1,126 twin pairs, respectively) and therefore were able to reveal the associations. By using 16S rRNA gene-based analysis, they identified many heritable taxa.[34,35] The most heritable was the family Christensenellaceae (phylum Firmicutes), which was also the hub of a co-occurrence network that includes other taxa with high heritability. Furthermore, this heritable taxon was enriched in individuals with low BMI, and adding it to an obese-associated microbiome reduced weight gain in transplanted germ-free mice. These observations indicate that heritable microbes can directly contribute to the host phenotype and that the microbial phenotype is additionally influenced by host genetics.[35] In 2016, a follow-up study that included both previous and new data were published.[9] The expanded dataset resulted in heritability estimates with narrower confidence
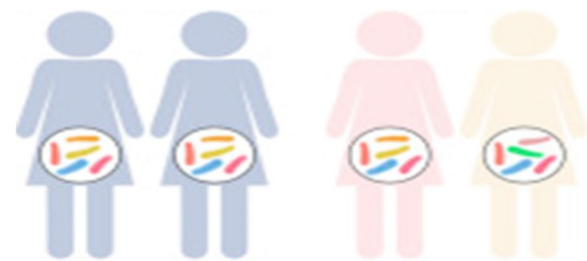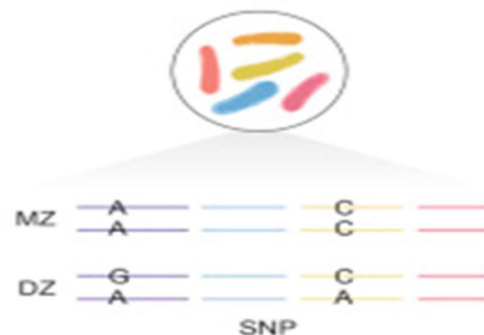

FIGURE 2 Microbial similarity in monozygotic twins, while genetic diversity is observed in dizygotic twins

intervals and revealed additional heritable taxa. They discovered that heritable taxa are temporally stable over long periods, suggesting that their relative abundances are less affected by environmental factors. To further investigate the involvement of host genetics, another study on a subset of Twins-UK participants (250 twins) was performed, using metagenomics shotgun sequencing for the analysis.[36] They showed heritability not only for many microbial taxa but

also for functional modules in the gut microbiome that can relate to the risk of complex diseases. When comparing to previous results that used 16S profiles, most of the microbiome heritability in this study were higher, suggesting that metagenomics analyses can provide greater resolution and power.[36] However, as the twins began to live apart, the microbial similarity between them declined, indicating that the environment overshadows genetics in the gut microbiome design.[36,37] A recent twin study found associations between high heredity microbial taxa and visceral fat accumulation, indicating host genetics as a potential mediator between obese complex phenotype and gut microbiome composition.[38]

Contrary to the findings presented above, several non-twin studies showed a significant bacterial similarity among genetically unrelated individuals who shared a household, but such similarity was not observed across family members that did not share a household.[39-41] These results indicate that the gut microbiome is primarily shaped by environmental factors and that the effect of host genetics is apparently quite modest.

3 | ASSOCIATIONS BETWEEN HOST GENETICS AND GUT MICROBIOME

One of the main approaches to investigate microbiome-host genome associations is microbial genome-wide association studies (mGWAS), which are cohort studies that combine human genotyping or whole-genome sequencing with microbiome analysis (16S rRNA or metagenomics sequencing) from the same individual.[42] Newer and cheaper technologies in recent years have enabled the discovery of several gut microbiomes associated with SNPs. These are related to the innate immune system or metabolism and are located near host genes associated with complex diseases.[9,33,43-49] However, most of the associations were not statistically significant after multiple testing corrections and the reported variants were not repeated in different studies.[28,40] Therefore, these results are limited.

The great complexity of the microbiome structure and the strong environmental effects probably result in variability and inconsistency between the results of recent studies.[28,37] Furthermore, an analysis of 5 previous studies, which included 225 SNPs significantly associated with microbiome parameters, showed almost no overlap between the loci reported in different studies.[40] However, the lactase gene (LCT) is an exception. In fact, the association between LCT locus and the relative abundance of the *Bifidobacterium* (phylum Actinobacteria) is the only consistent finding that has been validated in subsequent cohorts.[40,42,45] The lactase enzyme is responsible for the breakdown of lactose and encoded by the LCT gene, whose variants are associated with lactase persistence and lactose tolerance in adults.[9] *Bifidobacterium* is part of the gut microbiome composition that can utilize lactose, the milk sugar, as an energy source. Thus, its association with the LCT gene may be mediated through environmental factors like diet and lactose consumption that are also influenced by culture and geography.[9,50] The different genotypes of this SNP showed associations with different levels of the *Bifidobacterium*.

To be more specific, the SNP that is typically associated with lactase nonpersistence was also associated with the elevated abundance of *Bifidobacterium*.[34] When combining the dietary information of the subjects, the association was only in lactase non-persistence individuals that consumed lactose.[45] This finding can explain why the phenotype predicted by the genotype (lactose tolerance or intolerance) is not always accurate.[51] In some cases, the microbiome has more influence than the genotype on the observed phenotype.[15]

Studies that used another approach, which defines the entire microbiome composition as one feature (β-diversity) rather than treating taxa separately, managed to find several genetic associations for the overall microbial variation,[52] for example, in VDR gene (vitamin D receptor).[37,49] In addition, it was argued that 42 SNPs can explain over 10% of the β-diversity variance,[49] but since further studies failed to replicate this, the association of any individual SNP with microbiome β-diversity is very limited.[40,53] In contrast, when the associations with environmental factors related to diet and lifestyle were examined, they explain over 20% of the variance in microbiome β-diversity and thereby indicate greater environmental influence than genetic influence.[34,40,53,54]

4 | MAJOR INSIGHTS FROM MOUSE STUDIES

Animal models, in which genetics and environmental factors are better controlled, can be used to further investigate the role of host genetics. Among all nonhuman models, the mouse has been the most common and productive model to investigate the gut microbiome to date.[55,56] Besides high genetic similarity, the mouse shows some similarity to humans at the microbial taxonomic level, thus considered a powerful model for evaluating host-microbiota interactions applicable to humans.[55,57,58] Many mouse studies have demonstrated a consistent and reproducible impact of host genetics on the gut microbiome composition.[57,59-62] For example, the genetic background has shown to be a stronger determinant than gender on the mouse microbiome, since the similarity of microbiota composition within mouse lines was significantly higher than between different lines or within same-sex groups.[62] A recent analysis of F1 hybrids generated from a reciprocal cross between two inbred lines (BALB/c, C57BL/6J) showed a greater genetic effect on the microbial composition than maternal effect or continuous exposure to different microbiota of the other line.[61] These results are compatible with previous studies,[63,64] thus reinforce the hypothesis about the existence of a line-dependent bacterial signature that each line tends to maintain.[61] In addition, several heritable microbial taxa such as *Lactobacillus johnsonii* were identified,[61] in which their relative abundance was affected predominantly by host genetics.[60] It was also discovered that host genetic distance is associated with the overall microbiome composition (β-diversity).[65] By using quantitative trait locus (QTL) mapping methods, studies have managed to find several genetic loci that influence specific microbial taxa or pathways, most related to the immune system and metabolism.[28,59,66-68] A recent

(A) Taxa present in susceptible mice

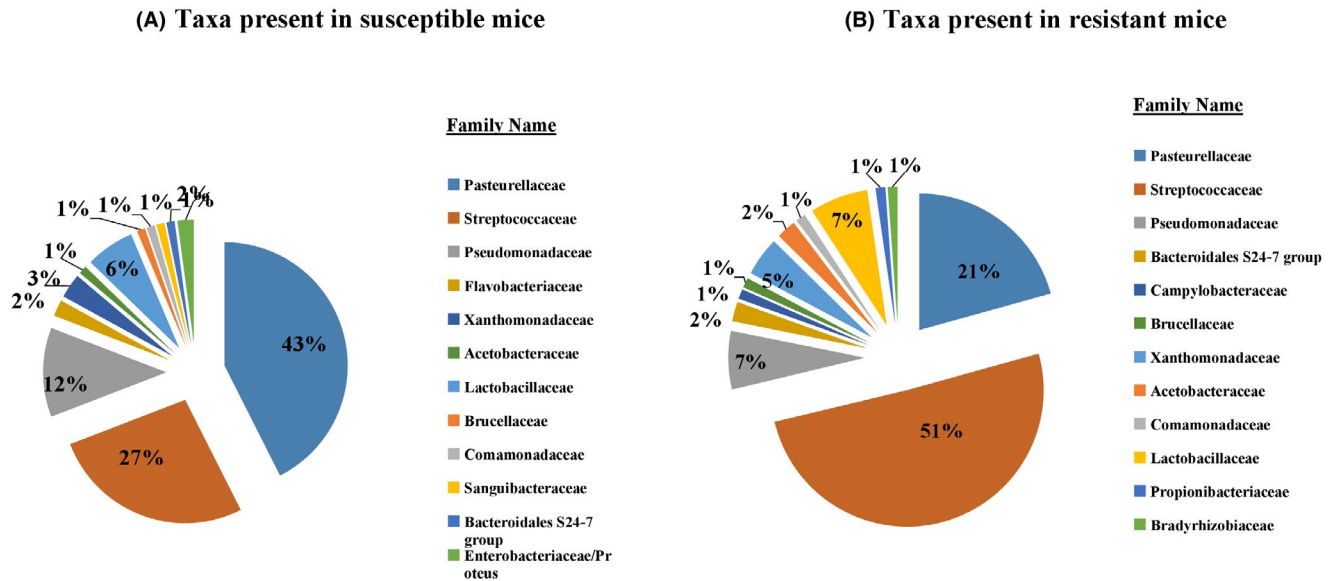**(B) Taxa present in resistant mice**

FIGURE 3 Percentage profiles of the top bacteria families that found in the normal flora in susceptible (3A) and resistant mouse populations (3B) to a specific environmental challenge. While 43% of the Pasteurellaceae taxa was found in the susceptible mouse population (3A), only 21% of this taxa was found in resistant mouse samples (3B) to the same environmental challenge. Streptococcaeae taxa was found 27% and 51% in susceptible (3A) and resistant (3B) mouse populations, respectively, to the same environmental challenge (Unpublished data)

study found a significant overlap between genetic associations discovered in wild mice and humans, in particular genes related to the nervous system, the immune system, and obesity, suggesting a possible mammalian-shared genetic mechanism that affects the gut microbiome composition.[65]

Figure 3 shows an example of percentage profiles of the top bacteria families that are found in the normal flora in susceptible (3A) and resistant mouse populations (3B) to a specific environmental challenge. As shown in Figure 3A, 43% of the Pasteurellaceae taxa was found in the resistant mouse population, while only 21% of this taxa was found in the susceptible mouse population. Interestingly, as shown in Figure 3B, Streptococcaeae taxa were found to be 27% and 51% in susceptible and resistant mouse populations, respectively (Unpublished personal data). Interestingly, B Streptococcaeae taxa was found to be 27% and 51% in susceptible and resistant mouse populations, respectively (Unpublished personal data).

5 | MICROBIOME HERITABILITY

The heritability of the trait in the study population tells us the correlation between the observed phenotype of an individual in that population and the true genetic value of the individual. Since the heritability is usually less than 1.0, this means that an individual with a given observed phenotype can have a true genetic value that varies more or less widely about that phenotype. Calculating and estimating the hereditability rate of the gut microbiome profiles are important in determining the level of the host genetic effect on the microbiome patterns. It is known that as the sample size is increased, the estimated heritability becomes more powerful and accurate.[9]

Nevertheless, in previous studies, the estimated heritability of gut microbiome profiles was found to be relatively low compared to other various complex traits tested in the same research populations.[69-71] While the estimated heritability of the microbial taxa in some of the reported studies was ranged between 0.10 and 0.30, the estimated heritability of other assessed phenotypic traits was much higher. These findings supported the previous statements, that several gut bacteria are heritable; however, the overall heritability tends to be smaller than it was initially estimated and expected.[37,40]

6 | DISCUSSION

The human gut microbiome is increasingly recognized as a key player in maintaining our health. Multiple factors shape its composition; however, the extent to which host genetics shape the microbiome composition remains controversial. One of the classic methods to investigate heritability is to examine differences between MZ and DZ twins.[28] Previous twin studies have shown that the overall microbiome composition[8] and many individual taxa are heritable.[9,36,38,72] Although the studies usually identified the same heritable taxa, the degree of heredity was not the same, probably due to differences in sample size and method of analysis.[36,40] These research characteristics are very important for the significance of heritability results, especially in twin studies on the microbiome. Therefore, larger scale studies in the future may yield additional results that will clarify the extent of the microbiome heritability.

Although we can conclude from the twin studies presented above that host genetics is an integral factor in shaping, at least part, of the gut microbiome composition, it should be kept in mind that

non-twin studies provide evidence for the dominant role of the environment in shaping the gut microbiome.[39-41] It should also be noted that compared to other complex traits tested in twin studies, the heritability of the microbiome components was relatively low.[69-71] While these findings do not contradict the heritability revealed in the twin studies, they do suggest that the existing genetic influence is relatively weak.

Finding associations between genetics and complex traits is a challenging task and when it comes to the gut microbiome, the challenge is sometimes even harder. In recent years, many studies have been published, mainly due to the great technical progress and the decrease in the cost of sequencing.[73] Although much effort has been invested, it is still difficult to deduce a clear trend regarding the role of host genetics in the gut microbiome. This difficulty is caused by the inconsistency and unevenness between the studies' results. Most of the detected SNPs were not repeated in subsequent studies and there was almost no overlap between their reported loci.[28,40] This suggests that most of the SNP-bacteria associations are either weak or population dependent.[40] In addition, a growing number of studies show that environmental factors can explain a greater proportion of the microbiome variability compared to the host genetics.[9,40,53,54] Therefore, the hypothesis that the gut microbiome is predominantly shaped by environmental factors is reinforced.

An exceptional example is a significant association between the LCT gene and the relative abundance of *Bifidobacterium*, which has been consistently repeated across studies.[40,42,45] However, this association may be mediated by lactose consumption.[9,45] This important finding offers a potential mechanism of gene-environment interactions in regulating the gut microbiome variability. It also suggests that more associations may be revealed by looking at the combined effect of genetics and the environment rather than exploring their effects separately.

Despite the inconsistencies in the results, it is important to keep in mind that associations identified in GWAS can be difficult to replicate in other studies because of small effect sizes of single variants, variability between study participants in terms of geography and genetic background, and differences in study design.[33] As presented in twins, larger scale future studies may reveal additional microbiota-associated SNPs but are unlikely to change the overall conclusion that the gut microbiome is mostly shaped by non-genetic factors.[40]

In human studies, large cohort sizes are necessary to uncover the genetic effects on the microbiome since the wide environmental variation among the participants cannot be completely controlled.[15] However, mouse studies can minimize this variation by exposing all mice to the same environmental conditions.[58] Mouse studies, in which the confounding factors are better controlled, suggesting a greater role of host genetics in shaping the gut microbiome than the role presented in humans. Furthermore, from the human and mouse studies discussed in this work, it can be said in general that results obtained in mice were more consistent and replicable, thus presenting a more reliable trend regarding the host genetic background importance. However, these findings raise the question of how representative the mouse model is for human microbiome research. While previous studies have argued that there is a high similarity between mouse and human in the gut microbiome composition, recent analyzes challenge this assumption and suggest far less similarity. Therefore, conclusions must be drawn with caution since the results in mice referring to the intestinal microbiome do not always applicable to humans.[56,74] Nevertheless, a comparison of the associations identified by mGWAS in both humans and wild mice yielded surprising findings.[65] First, the number of genes that overlapped was significantly larger than it was randomly expected, suggesting a possible conserved set of microbiome-associated genes across various mammals. Second, since the overlap host genes were related to the immune system, brain, and obesity, they can serve as strong potential candidates for future research.[65] Despite the dissimilarity of the microbiome system between mice and humans, the mouse model presents important advantages and can provide some insights and future directions for human research.

In conclusion, a number of previous studies have shown that the host genetic background is a fundamental component in determining the composition of the gut microbiome and profile. Furthermore, the composition of the gut microbiota is also known to be modulated by multiple factors, including the age, sex, and diet of the host.[75] Increasing evidence from mouse models has suggested that host genetics also contribute to shaping gut microbiota.[44,59] Recently, Goodrich et al[35] found that a number of gut microbial taxa, including Christensenellaceae, were heritable in humans, which strongly suggests that the human gut microbiota can be influenced by host genetics. At the level of specific locus, several studies have demonstrated that variations in several host genes, such as nucleotide-binding oligomerization domain 2 and fucosyltransferase 2, may contribute to alterations in gut microbial community structure, and consequently, affect Crohn's disease (CD) susceptibility in humans.[76] Therefore, it can be hypothesized that metabolic syndrome (MetS), gut microbial composition, and host genotype are inter-related. However, to our knowledge, no studies to date have investigated which specific gut microbes are responsible for MetS and whether such microbes are related to single nucleotide polymorphisms (SNPs), in particular host genes.

Interestingly, it appears that the effect of the host genetic structure is strong at young ages and seems to be reduced over time, while the environmental impact is getting stronger at older ages, and interperson microbiome variability becomes to be more influenced with environmental factors related to diet, drugs, and anthropometric measurements. Although it is now widely accepted that the gut microbiota plays a crucial role in host metabolism, the current knowledge on the effect of host genetics on specific gut microbes related to MetS status remains limited. In addition, accumulated data have suggested that dysbiosis of the gut microbiota is closely related to the development of obesity and, subsequently, other hallmarks of MetS. These data imply a critical role of the gut microbiota in the development of MetS in humans, and very few studies have investigated the gut microbiota of individuals in relation to MetS. Unfortunately, as of today, there are a limited number of reports that show the association of

host genetic factors with hereditary of gut microbiome profiles, affected by diet, metabolism, and obesity phenotypes. Evaluating the full role of host genetics on its various aspects is challenging and inaccurate since the current research methods are probably not strong enough to detect all genetic effects. Continued microbiome research is essential to understand the comprehensive genetic potential of the host. In addition, comparing results from mice and humans can provide potential candidate genes and guide the focus of future human research.

ACKNOWLEDGMENT

CONFLICT OF INTEREST

None.

ORCID

Fuad A. Iraqi ID https://orcid.org/0000-0001-5525-206X

REFERENCES

1. Lloyd-Price J, Mahurkar A, Rahnavard G. Strains, functions and dynamics in the expanded human microbiome project. *Nature*. 2017;550(7674):61-66.
2. Meisel J, Grice E. The human microbiome. In: Ginsburg G, Willard H, eds. *Genomic and Precision Medicine*, 3rd ed. Boston, MA: Academic Press; 2016:63-77.
3. Hornef M, Penders J. Does a prenatal bacterial microbiota exist? *Mucosal Immunol*. 2017;10(3):598-601.
4. Mills S, Stanton C, Lane JA, Smith GJ, Ross RP. Precision nutrition and the microbiome, part I: current state of the science. *Nutrients*. 2019;11(4):923.
5. Maurice CF, Haiser HJ, Turnbaugh PJ. Xenobiotics shape the physiology and gene expression of the active human gut microbiome. *Cell*. 2013;152(1-2):39-50.
6. Bäckhed F, Roswall J, Peng Y, et al. Dynamics and stabilization of the human gut microbiome during the first year of life. *Cell Host Microbe*. 2015;17(5):690-703.
7. Hasan N, Yang H. Factors affecting the composition of the gut microbiota, and its modulation. *Peer J*. 2019;7:e7502.
8. Goodrich J, Di Rienzi S, Poole A, et al. Conducting a microbiome study. *Cell*. 2014;158(2):250-262.
9. Goodrich JK, Davenport ER, Waters JL, Clark AG, Ley RE. Cross-species comparisons of host genetic associations with the microbiome. *Science*. 2016;352(6285):532-535.
10. Ley RE, Hamady M, Fraser-Liggett CM, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. *Nature*. 2007;449(7164):804-810.
11. Clemente JC, Ursell LK, Parfrey LW, Knight R. The impact of the gut microbiota on human health: an integrative view. *Cell*. 2012;148(6):1258-1270.
12. Lloyd-Price J, Abu-Ali G, Huttenhower C. The healthy human microbiome. *Genome Med*. 2016;8(1):51.
13. Thursby E, Juge N. Introduction to the human gut microbiota. *Biochem J*. 2017;474(11):1823-1836.
14. Durack J, Lynch SV. The gut microbiome: relationships with disease and opportunities for therapy. *J Exp Med*. 2019;216(1):20.
15. Goodrich JK, Davenport ER, Clark AG, Ley RE. The relationship between the human genome and microbiome comes into view. *Annu Rev Genet*. 2017;51(1):413-433.
16. Cook VJ, Turenne CT, Wolfe J, Pauls R, Kabani A. Conventional methods versus 16S ribosomal DNA sequencing for identification of nontuberculous mycobacteria: cost analysis. *J Clin Microbiol*. 2003;41(3):1010-1015.
17. Morgan XC, Huttenhower C. Chapter 12: human microbiome analysis. *PLoS Comput Biol*. 2012;8(12):e10028088.
18. Sergeant MJ, Constantinidou C, Cogan T, Penn CW, Pallen MJ. High-throughput sequencing of 16s rRNA gene amplicons: effects of extraction procedure, primer length and annealing temperature. *PLoS One*. 2012;7(5):1-10.
19. Ghyselinck I, Pfeiffer S, Heylen K, Sessitsch A, De Vos P. the effect of primer choice and short read sequences on the outcome of 16S rRNA gene based diversity studies. *PLoS One*. 2013;8(8):e71360.
20. Gonzalez E, Pitre FE, Brereton NJB. ANCHOR: a 16S rRNA gene amplicon pipeline for microbial analysis of multiple environmental samples. *Environ Microbiol*. 2019;21(7):2440-2468.
21. Wooley JC, Godzik A, Friedberg I. A primer on metagenomics. *PLoS Comput Biol*. 2010;6(2):e1000667.
22. Hugenholtz P, Goebel BM, Pace NR. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol*. 1998;180(18):4765-4774.
23. Segata N, Boernigen D, Tickle TL, Morgan XC, Garrett WS, Huttenhower C. Computational meta'omics for microbial community studies. *Mol Syst Biol*. 2013;9:666.
24. Tyson GW, Chapman J, Hugenholtz P, et al. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature*. 2004;428(6978):37-43.
25. Poinar HN, Schwarz C, Qi J, et al. Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science*. 2006;311(5759):392-394.
26. Rodrigue S, Materna AC, Timberlake SC, et al. Unlocking short read sequencing for metagenomics. *PLoS One*. 2010;5(7):e11840.
27. Schuster SC. Next-generation sequencing transforms today's biology. *Nat Methods*. 2008;5(1):16-18.
28. Kurilshikov A, Wijmenga C, Fu J, Zhernakova A. Host genetics and gut microbiome: challenges and perspectives. *Trends Immunol*. 2017;38(9):633-647.
29. Hill WG. Heritability. In: Brenner S, Miller JH, eds. *Encyclopedia of Genetics*. New York, NY: Academic Press; 2001:921-924.
30. Moayyeri A, Hammond CJ, Hart DJ, Spector TD. The UK adult twin registry (Twins UK resource). *Twin Res Hum Genet*. 2013;16(1):144-149.
31. Turnbaugh PJ, Hamady M, Yatsunenko T, et al. A core gut microbiome in obese and lean twins. *Nature*. 2009;457(7228):480-484.
32. Yatsunenko T, Rey FE, Manary MJ, et al. Human gut microbiome viewed across age and geography. *Nature*. 2012;486(7402):222-227.
33. Goodrich J, Davenport E, Beaumont M, et al. Genetic determinants of the gut microbiome in UK twins. *Cell Host Microbe*. 2016;19(5):731-743.
34. Falony G, Joossens M, Vieira-Silva S, et al. Population-level analysis of gut microbiome variation. *Science*. 2016;352(6285):560-564.
35. Goodrich J, Waters J, Poole A, et al. Human genetics shape the gut microbiome. *Cell*. 2014b;159(4):789-799.
36. Xie H, Guo R, Zhong H, et al. Shotgun metagenomics of 250 adult twins reveals genetic and environmental impacts on the gut microbiome. *Cell Syst*. 2016;21(6):572-584.
37. Weissbrod O, Rothschild D, Barkan E, Segal E. Host genetics and microbiome associations through the lens of genome wide association studies. *Curr Opin Microbiol*. 2018;44:9-19.

38. Le Roy CI, Beaumont M, Jackson MA, et al. Heritable components of the human fecal microbiome are associated with visceral fat. *Gut Microbes*. 2018;9(1):61-67.

39. Abeles SR, Jones MB, Santiago-Rodriguez TM, et al. Microbial diversity in individuals and their household contacts following typical antibiotic courses. *Microbiome*. 2016;4(1):39.

40. Rothschild D, Weissbrod O, Barkan E, et al. Environment dominates over host genetics in shaping human gut microbiota. *Nature*. 2018;555(7695):210-215.

41. Song SJ, Lauber C, Costello EK, et al. Cohabiting family members share microbiota with one another and with their dogs. *eLife*. 2013;2:e00458.

42. Hall AB, Tolonen AC, Xavier RJ. Human genetic variation and the gut microbiome in disease. *Nat Rev Genet*. 2017;18(11):690-699.

43. Blekhman R, Goodrich JK, Huang K, et al. Host genetic variation impacts microbiome composition across human body sites. *Genome Biol*. 2015;16(1):191.

44. Zhang C, Zhang M, Wang S, et al. Interactions between gut microbiota, host genetics and diet relevant to development of metabolic syndromes in mice. *ISME J*. 2009;4:232-241.

45. Bonder MJ, Kurilshikov A, Tigchelaar EF, et al. The effect of host genetics on the gut microbiome. *Nat Genet*. 2016;48(11):1407-1412.

46. Knights D, Silverberg MS, Weersma RK, et al. Complex host genetics influence the microbiome in inflammatory bowel disease. *Genome Med*. 2014;6(12):107.

47. Li D, Achkar J, Haritunians T, et al. A pleiotropic missense variant in SLC39A8 is associated with crohn's disease and human gut microbiome composition. *Gastroenterology*. 2016;151(4):724-732.

48. Turpin W, Espin-Garcia O, Xu W. Association of host genome with intestinal microbial composition in a large healthy cohort. *Nat Genet*. 2016;48(11):1413-1417.

49. Wang J, Thingholm LB, Skiecevičienė 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):1396-1406.

50. Kato K, Ishida S, Tanaka M, et al. Association between functional lactase variants and a high abundance of bifidobacterium in the gut of healthy japanese people. *PLoS One*. 2018;13(10):e0206189.

51. Ranciaro A, Campbell M, Hirbo J, et al. Genetic origins of lactase persistence and the spread of pastoralism in africa. *Am J Hum Genet*. 2014;94(4):496-510.

52. Kolde R, Franzosa EA, Rahnavard G, et al. Host genetic variation and its microbiome interactions within the human microbiome project. *Genome Med*. 2018;10(1):6.

53. Scepanovic P, Hodel F, Mondot S, et al. A comprehensive assessment of demographic, environmental, and host genetic associations with gut microbiome diversity in healthy individuals. *Microbiome*. 2019;7(1):1-15.

54. Zhernakova A, Kurilshikov A, Bonder MJ, et al. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science*. 2016;352(6285):565-569.

55. Kostic AD, Howitt MR, Garrett WS. Exploring host–microbiota interactions in animal models and humans. *Genes Dev*. 2013;27(7):701-718.

56. Nguyen TLA, Vieira-Silva S, Liston A, Raes J. How informative is the mouse for human gut microbiota research? *Dis Model Mech*. 2015;8(1):1-16.

57. Hildebrand F, Nguyen TLA, Brinkman B, et al. Inflammation-associated enterotypes, host genotype, cage and inter-individual effects drive gut microbiota variation in common laboratory mice. *Genome Biol*. 2013;14(1):R4.

58. Spor A, Ley R, Koren O. Unravelling the effects of the environment and host genotype on the gut microbiome. *Nat Rev Microbiol*. 2011;9(4):279-290.

59. 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 USA*. 2010;107(44):18933-18938.

60. Buhnik-Rosenblau K, Danin-Poleg Y, Kashi Y. Predominant effect of host genetics on levels of lactobacillus johnsonii bacteria in the mouse gut. *Appl Environ Microbiol*. 2011;77(18):6531-6538.

61. Korach-Rechtman H, Freilich S, Gerassy-Vainberg S, et al. Murine genetic background has a stronger impact on the composition of the gut microbiota than maternal inoculation or exposure to unlike exogenous microbiota. *Appl Environ Microbiol*. 2019;85(18):826.

62. Kovacs A, Ben-Jacob N, Tayem H, Halperin E, Iraqi FA, Gophna U. Genotype is a stronger determinant than sex of the mouse gut microbiota. *Microb Ecol*. 2011;61(2):423-428.

63. Campbell JH, Foster CM, Vishnivetskaya T, et al. Host genetic and environmental effects on mouse intestinal microbiota. *ISME J*. 2012;6(11):2033-2044.

64. Ericsson AC, Davis JW, Spollen W, et al. Effects of vendor and genetic background on the composition of the fecal microbiota of inbred mice. *PLoS One*. 2015;10(2):e0116704.

65. Suzuki TA, Phifer-Rixey M, Mack KL, et al. Host genetic determinants of the gut microbiota of wild mice. *Mol Ecol*. 2019;28(13):3197-3207.

66. McKnite AM, Perez-Munoz ME, Lu L, et al. Murine gut microbiota is defined by host genetics and modulates variation of metabolic traits. *PLoS One*. 2012;7(6):e39191.

67. Org E, Parks BW, Joo JWJ, et al. Genetic and environmental control of host-gut microbiota interactions. *Genome Res*. 2015;25(10):1558-1569.

68. Wang J, Kalyan S, Steck N, et al. Analysis of intestinal microbiota in hybrid house mice reveals evolutionary divergence in a vertebrate hologenome. *Nat Commun*. 2015;6(1):6440.

69. Davies MN, Verdi S, Burri A, et al. Generalised anxiety disorder – a twin study of genetic architecture, genome-wide association and differential gene expression. *PLoS One*. 2015;10(8):e0134865.

70. Menni C, Mangino M, Zhang F, et al. Heritability analyses show visit-to-visit blood pressure variability reflects different pathological phenotypes in younger and older adults: Evidence from UK twins. *J Hypertens*. 2013;31(12):2356-2361.

71. Polderman JC, Benyamin B, de Leeuw CA, et al. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet*. 2015;47(7):702-709.

72. Davenport ER. Elucidating the role of the host genome in shaping microbiome composition. *Gut Microbes*. 2016;7(2):178-184.

73. Badal VD, Wright D, Katsis Y, et al. Challenges in the construction of knowledge bases for human microbiome-disease associations. *Microbiome*. 2019;7(1):129.

74. Hugenholtz F, de Vos WM. Mouse models for human intestinal microbiota research: a critical evaluation. *Cell Mol Life Sci*. 2018;75(1):149-160.

75. Delzenne NM, Neyrinck AM, Bäckhed F, Cani PD. Targeting gut microbiota in obesity: effects of prebiotics and probiotics. *Nat Rev Endocrinol*. 2011;7:639-646.

76. Rehman A, Sina C, Gavrilova O, et al. Nod2 is essential for temporal development of intestinal microbial communities. *Gut*. 2011;60:1354-1362.