

ABSTRACT: The human intestinal microbiome is a microbial ecosystem that expresses as many as 100 times more genes than the human host, thereby constituting an important component of the human *holobiome*, which contributes to multiple health and disease processes. As most commensal species are difficult or impossible to culture, genomic characterization of microbiome composition and function, under various environmental conditions, comprises a central tool in understanding its roles in health and disease. The first decade of microbiome research was mainly characterized by usage of DNA sequencing-based 16S rDNA and shotgun metagenome sequencing, allowing for the elucidation of microbial composition and genome structure. Technological advances in RNA-seq have recently provided us with an ability to gain insight into the genes that are actively expressed in complex bacterial communities, enabling the elucidation of the functional changes that dictate the microbiome functions at given contexts, its interactions with the host, and functional alterations that accompany the conversion of a *healthy* microbiome toward a disease-driving configuration. Here, we highlight some of the key metatranscriptomics strategies that are implemented to determine microbiota gene expression and its regulation and discuss the advantages and potential challenges associated with these approaches.

KEYWORDS: gut, microbiome, bacteria, RNA-Seq, next generation sequencing, transcriptome

CITATION: Bashiardes et al. Use of Metatranscriptomics in Microbiome Research. *Bioinformatics and Biology Insights* 2016;10 19–25 doi: 10.4137/BBI.S34610.

TYPE: Review

RECEIVED: January 16, 2016. **RESUBMITTED:** March 02, 2016. **ACCEPTED FOR PUBLICATION:** March 07, 2016.

ACADEMIC EDITOR: J. T. Efrid, Associate Editor

PEER REVIEW: Four peer reviewers contributed to the peer review report. Reviewers' reports totaled 760 words, excluding any confidential comments to the academic editor.

FUNDING: EE is supported by Yael and Rami Ungar, Israel; the Abisch Frenkel Foundation for the Promotion of Life Sciences; the Gurwin Family Fund for Scientific Research; the Leona M. and Harry B. Helmsley Charitable Trust; the Crown Endowment Fund for Immunological Research; the estate of Jack Gitlitz; the estate of Lydia Hershkovich; the Benozio Endowment Fund for the Advancement of Science; the Adelis Foundation; John L. and Vera Schwartz, Pacific Palisades; Alan Markovitz, Canada; Cynthia Adelson, Canada; Centre National de la Recherche Scientifique; the estate of Samuel and Alwyn J. Weber; Mr. and Mrs. Donald L. Schwarz, Sherman Oaks; grants funded by the European Research Council; the Kenneth Rainin Foundation; the German-Israel Bination foundation; the Israel Science Foundation; the Minerva Foundation; the Rising Tide Foundation; and the Alon Foundation scholar award. EE is the incumbent of the Rina Gudinski Career Development Chair. The authors confirm that the funder had no influence over the study design, content of the article, or selection of this journal.

COMPETING INTERESTS: Authors disclose no potential conflicts of interest.

CORRESPONDENCE: eran.elinav@weizmann.ac.il

COPYRIGHT: © the authors, publisher and licensee Libertas Academica Limited. This is an open-access article distributed under the terms of the Creative Commons CC-BY-NC 3.0 License.

Paper subject to independent expert blind peer review. All editorial decisions made by independent academic editor. Upon submission manuscript was subject to anti-plagiarism scanning. Prior to publication all authors have given signed confirmation of agreement to article publication and compliance with all applicable ethical and legal requirements, including the accuracy of author and contributor information, disclosure of competing interests and funding sources, compliance with ethical requirements relating to human and animal study participants, and compliance with any copyright requirements of third parties. This journal is a member of the Committee on Publication Ethics (COPE).

Published by Libertas Academica. Learn more about this journal.

The Microbiome

Microbiome research has greatly benefited from the technological breakthroughs that enabled the genomic characterization of the human genome less than two decades ago. Soon after the deciphering of the human genome, attention has been shifting to the enormous genomic prokaryotic gene pool within the human body, which far exceeds that of the human eukaryotic genome, but whose contribution to human physiology remained elusive. The Nobel laureate Joshua Lederberg first termed *microbiome* as the combination of commensal, symbiotic, and pathogenic microorganisms that colonize the human body.¹ This microbial ecosystem is composed of bacteria, fungi, and viruses, with the predominant focus of study today being on the bacterial component of this community. In a tour de force effort, David Relman's group first characterized the intestinal microbiome bacterial composition, extraordinarily without the use of next-generation sequencing but rather by the use of Sanger sequencing.² Work by Gordon enabled an understanding of the factors affecting the bacterial community structure and the role this compositional structure may play in mammalian physiology and risk of disease.^{4,5}

The Human Microbiome Project and the European-based MetaHit project followed with a large-scale (and multicenter) effort to comprehensively characterize these microorganisms that are found on and in our bodies and to further determine their various roles.^{6,7}

Collectively, these pioneering works led to the discovery that our microbiome consists of vast numbers of cells, with the latest estimates indicating approximately equal numbers between microbiota cells and our own cells,⁸ and express as many as 100 times more genes when compared to the human eukaryotic gene pool. A great deal of heterogeneity was shown to exist between individuals in their microbiome composition.⁶ While the basis to this heterogeneity is not entirely comprehended, human microbiome structure and stability is estimated to be influenced by a multifactorial array of host genetics and immune and environmental factors.^{6,9} Later, it was realized that healthy individuals harbor a *core microbiota* colonizing and characterizing various body habitats. For example, the gut microbiome comprises more than 1,000 species of bacteria, but the most common phyla are Firmicutes, Bacteroidetes, Actinobacteria, and Proteobacteria.¹⁰ Similarly, the



female urogenital tract has its own microbiota profile with more than 150 species with a core microbiota of Firmicutes, Bacteroidetes, and Actinobacteria.¹¹ In this review, the term core microbiota has been used to describe commonly observed bacterial phyla. It is worth noting that this term has been used by Gordon in a landmark study on obese and lean twins,¹² to describe identifiable bacterial metabolic gene networks rather than characterized phyla.

The microbiota communities, of which the gut microbiome is the very best studied, play important multifactorial roles in human physiology. They are important in controlling pathogen colonization¹³ and in immune system development,¹⁴ and they help us in our digestion by hydrolyzing the compounds in our diet, which could not be broken down through enzyme production,¹⁵ and in the production of vitamins, such as vitamins B12, B5, and K.¹⁶ Changes in microbiota populations, termed dysbiosis, have been associated with a number of human conditions, such as inflammatory bowel disease,¹⁷ obesity,¹² and nonalcoholic fatty liver disease.¹⁸ Furthermore, dysbiosis has been shown to occur as a result of pathogen infections, such as Human Immunodeficiency Virus (HIV)¹⁹ and influenza virus.²⁰ Concise general reviews of microbiome functions and associations with disease are comprehensively described elsewhere.^{21–24}

Characterization of the Microbiome

Characterizing the microbiota in terms of their taxonomy and phylogeny has been carried out in a large number of studies by sequencing of the 16S ribosomal RNA subunit gene.^{25–27} This gene contains regions that are conserved throughout bacterial species and hypervariable regions that are unique for specific genera, which are targeted for sequencing and used for taxonomic characterization. The sequenced variable regions are then clustered into Operational Taxonomic Units, providing invaluable information on their taxonomic characterization.^{28–30}

Whole genome shotgun sequencing followed by metagenomic analysis adds a more detailed layer of information to the taxonomical characterization of a sample, by generating information on the gene composition of the bacteria present.³¹ This information can in turn be used to discover new genes and to formulate putative functional pathways and modules, thus providing insight into functional and genetic microbiome variability.³² Metagenomic analysis is carried out on genomic DNA isolated from the environment under study, but it does not distinguish whether this genomic DNA comes from cells that are viable or not or whether the predicted genes are actually expressed and under what conditions.³³ In addition, other *-omics* approaches, namely metaproteomics and metabolomics, are increasingly being implemented and are contributing to the understanding of microbiota community function. Metaproteomics provides an image into the entire protein complement of the microbiota communities studied under different conditions and an insight into the genes expressed and the key metabolic activities characterizing

the bacterial communities.^{34,35} Metabolomics provides information on the secreted or modulated metabolite composition of the microbiota community, thereby enabling an understanding of the functional dynamics influencing community and host interactions.^{36–38} Together, these nongenomic-based approaches may complement metagenomics and metatranscriptomics data and altogether add to our understanding of the complex pathways at play in the dynamic environment of microbiota communities.

Characterization of the Metatranscriptome

Recent advances in sequencing technologies that have revolutionized metagenomic analyses have also advanced the approaches aimed at studying and understanding gene expression on a global scale. Understanding the critical roles that host gene expression plays on a cellular or tissue level has come a long way from the elegantly described differential display approach in the early days³⁹ to a global *transcriptome* approach through the use of microarrays.⁴⁰ In recent years, the advent of massively parallel sequencing and RNA-seq has provided new and exciting opportunities in the area of transcriptome analysis, providing insight and dynamic range previously unimaginable.⁴¹ In addition to the insights gained from elucidation of host gene expression profile, we are now in a position to also study the gene expression of complex bacterial communities at a given environment (be it in the gut, oral cavity, or respiratory tract) that includes the gene expression profile of bacteria that can or cannot be cultured. Data from metatranscriptome analyses, thus, complement metagenomics data by elucidating accurately which of the genes that were annotated in the metagenomic analysis are transcribed and to what extent,⁴² thereby enabling to demonstrate the functions from a potential repertoire of bacteria that are actually in use at a given context. From such functional data, active metabolic pathways can be identified in the bacterial communities and can be associated to particular environmental conditions.⁴³ Therefore, metatranscriptomics offer a more informative perspective compared with metagenomics, as it can reveal details about populations that are transcriptionally active and not just identify the genetic content of bacterial populations as shown in metagenomic analysis. This is of utmost importance, as elusive, and clinically important differences in distinct groups of active bacteria often occur between human individuals.^{42,44}

Isolation and Processing of Microbiome mRNA

Typically, a metatranscriptome experiment of the microbiome involves isolation of total RNA from bacteria colonizing the area of interest (eg, gut, skin, and oral cavity). In eukaryotes, mRNA can be selected by synthesizing cDNA using oligo-d(T) primers, therefore taking advantage of the poly-A tail characterizing mRNA species. Prokaryotic mRNA makes up only 1%–5% of total RNA species, with the majority being 16S and 23S rRNAs as well as tRNAs.⁴⁵ However, in contrast to eukaryotic mRNA, prokaryotic mRNA lacks a poly-A



tail, making its selection during cDNA synthesis inapplicable. Therefore, various approaches have been developed and implemented to address this issue.^{46–48} Removal of rRNA with the use of probes targeting specific rRNA regions that are attached to magnetic beads represents an attractive option. The process involves annealing of probes to target sequences (rRNA) followed by their removal with the use of a magnet.⁴⁶ As with all methods involving RNA manipulation, the challenge of avoiding degradation by contaminating ribonucleases is also presented in this approach. Maintaining the commonly implemented laboratory practices for these types of sensitive protocols to avoid introduction of contaminating ribonucleases is important, and the incorporation of RNase inhibitors into the procedure can represent an effective protection strategy. What remains is an enriched population of other mRNAs that are representative of transcriptionally active genes. For massively parallel sequence analysis, these RNAs are fractionated, cDNA is synthesized, and adapters are ligated to the cDNA ends (following end repair) generating a library that is amplified and then sequenced. Sequence reads are mapped to reference genomes, and the expressed genes are identified based on the sequence reads covering these regions.⁴⁹

Computational Analysis of Metatranscriptomics Data

A typical metatranscriptome dataset contains many millions of sequenced mRNA molecules, termed RNA-seq reads. Moreover, as metatranscriptome experiments are consistently increasing in size and number, automated, efficient, high-throughput analyses are essential to infer the biological meaning from these datasets.^{33,50} Several comprehensive analysis suites (eg, HUMAnN⁵¹ and MG-RAST⁵²) have been developed over the past few years, are extensively used, and provide an end-to-end solution. These are applied alongside the combinations of specialized bioinformatic tools (eg, BOWTIE⁵³ and GEM⁵⁴ for mapping, Trimmomatic⁵⁵ for quality filtering, and CuffDiff⁵⁶ for differential gene expression) to achieve the same overall goal of inferring the gene expression levels and changes in gene expression levels, from the raw sequenced mRNA reads. A few analytic steps are essential in this process and are, therefore, present uniformly exist in all metatranscriptome analyses. These steps consist of the filtering of non-mRNA reads, and as well as the host reads, filtering and trimming low-quality reads and nucleotides (similar to the quality control process in high-throughput metagenomic analysis), identifying the open-reading frames, mapping the reads to a reference database, normalization, and calculations of the gene expression levels along with other summarizing statistics.⁴¹

An analytic step that is optional is the assembly of the reads into contigs, which can be executed after the initial filtering. If executed, the assembly step is followed by mapping the contigs to reference genomes, when these are available. While an assembly step is challenging computationally and requires higher quality experimental sequencing data,

it holds the potential to uncover information regarding the gene expressions that is not attainable without it, such as the relation between adjacent genes and the start and stop sites. Experimentally, to enable the assembly, deeper sequencing is required, and therefore, commonly only highly abundant regions can be assembled from a larger set of reads.⁵⁷ An assembly step is essential in cases in which a reference genome and subsequent gene annotations are not available. This is less common in the context of the gut microbiota but is relevant in RNA-seq of nonmodel organisms. In the event that a reference genome is not available, the annotations of the sequenced transcripts are usually obtained by sequence similarity to sequenced and annotated proteins. In other words, the assembled transcripts are aligned against large annotated protein databases with software, such as Blast2GO,⁵⁸ and if highly similar proteins are found, a similar biological function is usually inferred. Some suites for full transcriptome reconstruction have been developed and are based on extensive computational techniques, usually relying on graph-theoretic concepts (eg, Trinity).⁵⁹

Another important issue in the analysis and inference of biological information from metatranscriptomics data is combining the analysis of the RNA-seq data and the whole DNA data, ie metagenomics. Analyzing these two types of data simultaneously for a sample enables us to conclude the actual expressed genes vs the potentially existing genes.⁴³ Regardless of the existence of the assembly step, at the end of the RNA-seq analysis and the postnormalization process, a summary of the data is converted into relative gene expression values and can then be further analyzed similar to the statistical analyses seen in 16S and metagenomic sequencing (eg, gene expression level within a sample, richness within samples, and similarity between samples).

Utilization of Metatranscriptomics in Health and Disease

Assessment of microbial activity. Identifying functionally active bacteria within a mixed bacterial microbiome may highlight the disease-driving bacteria within a generally inactive microbial pool. Several strategies of determining transcriptionally active bacteria have been described.^{33,60} Gosalbes et al.³³ utilized the presence of 16S rRNA transcript as a way to determine the phylogenetic structure of active bacteria in the gastrointestinal tract (finding the phyla Firmicutes as predominantly active followed by Bacteroidetes). In healthy individuals, characterization of mRNA revealed activation of pathways involved in carbohydrate metabolism, cell component synthesis, and energy production.

The value of characterizing microbiota in a combined metagenomics–metatranscriptomics approach is highlighted by the effect of commensal bacteria on xenobiotics. As discussed earlier, gut microbiota plays an important role in metabolizing carbohydrates and proteins by producing and secreting an array of enzymes. These metabolic activities can potentially



affect xenobiotic stability when the xenobiotics are substrates of these enzymes. There are over 40 xenobiotics (that are or have been on the market) that are affected by the gut microbiota without the underlining mechanisms being fully understood.⁶¹ This effect is best shown in a tragic turn of events where an antiviral drug marketed under the generic name sorivudine was converted by the gut microbiota into a compound that inhibited the metabolism of the anticancer drug 5-fluorouracil causing its accumulation leading to toxicity. Within a short period of time, 18 patients who were prescribed a combination of sorivudine and 5-fluorouracil died.⁶² Although it is established that gut bacteria are responsible for the metabolism of many drugs, the exact bacteria involved and the molecular pathways implicated are frequently unknown. Recent studies are beginning to unravel the microbiota metabolic processes that influence drug metabolism. An elegant study by Maurice et al.⁶⁰ implemented a flow cytometry approach to isolate active bacterial populations from the gut that were then characterized by 16S rRNA gene sequencing and metatranscriptomics to determine the gene expression profiles in response to xenobiotics. This study showed that there are distinct sets of active bacteria in the gut (composed mainly of *Firmicutes*) and that over half the gut microbiota showed a high level of metabolic activity. They also showed that approximately one-third of the gut microbiota were damaged cells (not active but gene composition can be detected by metagenomics analysis). Exposure to xenobiotics was seen to have an effect on the active gut microbiome's structure as well as on its gene expression profile.⁶⁰ These types of studies show that moving beyond determining the phylogenetic community profile and delving into understanding the active gut microbiome metabolic activity involved in xenobiotic metabolism and resistance will have important repercussions in understanding interpatient variations in drug efficacy and toxicity.

Another important example can be seen in the case of the cardiac drug digoxin that can be inactivated by gut microbiome metabolism. Transcriptional profiling revealed that specific strains of the gut bacteria, *Eggerthella lenta*, have a cytochrome-encoding operon that is upregulated by digoxin and is predictive of the cardiac drug inactivation.⁶¹ Using gnotobiotic mice, it was shown that increasing dietary protein could significantly reduce digoxin microbial metabolism and result in increased concentrations of the drug present in serum.⁶¹ Overall, this highlights the importance of taking into consideration the gut microbiome with respect to drug efficacy.

Assessment of microbiome-immune interactions. The effects of microbiome on the mucosal immune system are considered pivotal in affecting host physiology. Studies focusing on toll-like receptor 5 (TLR5) knockout (KO) mice are an interesting example of the use of metatranscriptomics to complement metagenomics and 16S rRNA characterization of such microbiota-immune interactions. TLR5 is expressed in the intestinal mucosa and recognizes flagellin, the principal

component of the bacterial flagella. Mice lacking TLR5 were shown to develop metabolic syndrome⁶³ and colitis⁶⁴ and were characterized by dysbiosis of the gut microbiota.^{63,65} Another impairment noted in TLR5-deficient mice relates to the maintenance of barrier function. Indeed, the mucosal innate immune system plays diverse roles in microbial containment through mechanisms, such as regulation of protective mucus production⁶⁶ and secretion of high concentrations of IgA enabling coating of *locally invasive* commensal bacteria, thereby dampening inflammatory responses.⁶⁷ In the absence of innate recognition of flagellin, there is a reduction in the concentration of anti-flagellin antibodies that play a role in containing the microbiota, which can have an overall effect on microbiota and mucosal barrier functionality.⁶⁸ Gut dysbiosis in mice that lack TLR5 (and display metabolic syndrome) was characterized through a 16S rRNA sequencing approach.^{63,65} Although dysbiosis was indicated by bacterial phylogenetic analysis of stool samples, the underlying mechanism affecting and taking place within the bacterial community in the absence of TLR5 was only enabled using a metatranscriptomics approach. Metagenomics analysis of TLR5 KO mice showed that there was no significant difference in the presence of microbiota functional genes between these mice and wild-type mice.⁶⁹ In contrast, metatranscriptomics analysis revealed an upregulation in flagella motility-associated genes of commensal microbiota in the TLR5 KO mice compared to wild-type mice.⁶⁹ In this model, TLR5 flagellin recognition brings about anti-flagellin antibody production that leads to a downregulation of flagellar motility genes from a variety of bacteria, keeping the microbiota contained. Absence of TLR5 has, as a consequence, a reduction in anti-flagellin antibody production, causing upregulation of bacterial flagellar motility genes and, therefore, increasing bacterial motility in the gastrointestinal tract environment, resulting in a breach of the mucosal barrier.⁶⁹

Studying microbiome antisense RNA. Although traditional metatranscriptome analysis involves characterizing the mRNA transcripts under specific environmental conditions, and from this data determining metabolic pathways that are activated, the bacterial transcriptome represents a high level of complexity.⁷⁰ As RNA-seq methods matured allowing for strand-specific libraries to be constructed, it was revealed that the bacterial transcriptome encodes a surprisingly large number of *cis*-encoded RNAs denoted as antisense transcripts that are transcribed from the opposite strand of DNA-encoding genes.⁷¹ Through base pair interaction, these have the potential to interact with the sense transcript acting as a regulator of mRNA.⁷¹ This interaction has been shown experimentally, for example, in the case of the cyanobacteria *Synechocystis* sp.⁷² and the bacterium *Listeria monocytogenes*.⁷³ In fact, it has been proposed that these antisense RNAs may not only target their corresponding sense RNA but also mRNAs in other locations.⁷⁴ The numbers of antisense RNAs identified vary between bacterial species. For example, *Escherichia coli* was reported to have antisense



transcripts for 22% of its genes⁷⁵ whereas *Staphylococcus aureus* was reported to have antisense transcripts for just over 1% of genes.⁷⁶ Antisense transcripts are involved in a wide variety of regulatory roles, such as modulation of transcription regulators and toxic protein synthesis repression.⁷¹ It appears that antisense RNAs play an important role in prokaryotic gene regulation that can further enrich our understanding of underlying mechanisms of microbiota gene expression and its regulation. Although many studies on prokaryotic antisense RNAs were predominantly carried out on single-cultured strains, implementing such analyses on a *global* microbiota scale can complement traditional metatranscriptomics analysis of mRNAs and provide a greater depth of understanding into the regulatory mechanisms affecting these expression profiles. An example of such a global approach can be seen in the analysis of antisense RNAs in the human gut microbiota that showed a significant and dynamic presence.⁷⁷ There was variation in the transcription of antisense RNA between individuals (the number of strain-specific antisense RNAs) as well as between individuals in the genes that had antisense RNA. Common genes with antisense RNA transcripts included transposases, defense mechanism genes, as well as bacterial house-keeping genes.⁷⁷

Studying microbiome small noncoding RNAs. The bacterial transcriptome includes small noncoding RNAs (sRNAs) that are generally between 50 and 500 bp in size and are involved in gene regulation.⁷⁰ They do so by interacting, through base pairing, with the 5'-Untranslated Region (UTR) of target mRNA sequences regulating the translation or stability of the transcript.⁷⁰ They are involved in regulating important processes in bacteria, such as iron metabolism,⁷⁸ virulence,⁷³ and quorum sensing,⁷⁹ and are important as they allow for rapid adaptation to changing environments.⁸⁰ sRNAs have been identified in multiple bacterial species.⁸¹ The advent of next-generation sequencing methods has accelerated their identification for various bacterial species, such as *Salmonella*⁸¹ and *Bacillus subtilis*.⁸² Next-generation sequencing methods have also presented the opportunity to study bacterial sRNAs on a community level. For example, metatranscriptomics analysis of bacteria in the ocean at various depths has shown the potential role of sRNAs in niche adaptation.⁸³ Metatranscriptomics of the human active gut microbiota identified a number of sRNAs, although their role in the gut microbial community was not elucidated.³³ Analysis of sRNAs has generally been limited to single bacterial species, and metatranscriptomics approaches of community-based analysis are limited. An interesting example of following a global metatranscriptomics approach to analyze sRNA in a complex community was carried out in the oral cavity.⁸⁰ In this study, interestingly, community mRNA expression profiles were combined with sRNA profiles to understand the active metabolic pathways and the underlying regulatory mechanisms influencing oral microbiome dynamics in periodontitis. Periodontitis is a biofilm-induced inflammatory disease that affects the tissues

surrounding the teeth (periodontium) and is characterized by dysbiosis of the bacterial community that is considered to be causative in the disease.⁸⁴ The microbiome of the oral cavity represents a complex microbiota community that is exposed to a frequent alteration in environment brought about by ingested food that requires rapid microbiota adaptation to these fluctuating environmental conditions. This is enabled through adaptation of metabolic processes to these environmental fluctuations, and sRNAs play an important regulatory role in enabling this.⁸⁰ Metatranscriptomics approaches have been implemented on subgingival biofilms in an effort to understand the regulatory mechanisms potentially facilitating dysbiosis that is characteristic of periodontitis.⁸⁵ Community mRNA expression profiles identified specific metabolic signatures defining disease progression and interestingly indicate that not only known periodontal pathogens but also bacteria, which are not normally associated with the disease, show an increase in virulence leading to disease progression.⁸⁵ Associating microbial community sRNA expression profiles with mRNA expression data showed that sRNAs are potentially involved in regulating the oral microbiome community metabolic activities and, in this way, the transition of the oral microbiome from commensal to dysbiotic.⁸⁰

Limitations and Challenges of Metatranscriptomics Analysis

Several challenges associated with metatranscriptome analysis merit mentioning. The isolation of high-quality RNA samples from some biological samples (such as feces) can be a difficult if not daunting task. Experimental strategies have been developed to tackle some of these issues⁴²; nevertheless, significant challenges do remain. The potential of host RNA contamination in the sample that can occur to various degrees depending on the sample (eg, contamination is high in biopsy samples) can prove to be problematic. In these cases, rRNA from the host cannot be removed by following a strategy of annealing probes to target bacterial rRNA sequences followed by their removal with the use of a magnet, and they remain as contaminants that can increase the overall processing costs and complicate downstream analysis of data. Another issue to consider is that mRNA has a short half-life and thus it may be hard to detect rapid/short-lived responses to environmental stimuli.⁴³ Furthermore, the presence of mRNA is not always synonymous with the presence of protein (or protein activity for that matter). As such, pipelines integrating metagenomics, metatranscriptomics, metabolomics, and metaproteomics datasets may potentially enable to gain a *holistic* view of microbiome composition and function at multiple layers. Finally, at present, multiple metagenomic analysis methods may at times produce variable results, even if identical databases are used in the analysis. Thus, standardization of RNA isolation, processing, sequencing, and analysis is warranted to enable further dissemination of metatranscriptomics methods and their integration into microbiome research.



On a final note, traditionally, large-scale expression studies using methods, such as microarrays and serial analysis of gene expression, have been accompanied by validation of results by an independent technique, Quantitative Polymerase Chain Reaction (qPCR) being considered as the *gold standard*.^{86,87} Now that large-scale expression approaches have shifted toward utilizing Next Generation Sequencing (NGS) approaches, namely RNA-seq, validation of discoveries using this technology with the use of qPCR should not be overlooked and can provide added value to the observed expression patterns.⁸⁸

Conclusion

Metatranscriptomics holds great potential to uncover biological information that may be otherwise obscured by other genomic methodologies. It provides an accurate snapshot, at a given moment in time and under specific conditions, of the actual gene expression profile rather than its potential, as inferred from DNA-based shotgun metagenomic sequencing. As such, deciphering microbiome metatranscriptomics may better enable the elucidation of functional changes that dictates the microbiome functions at given contexts, its interactions with the host, and functional alteration that accompany the conversion of a *healthy* microbiome toward a disease-driving configuration. Moreover, metatranscriptomics may open a window into discovering the regulatory mechanisms orchestrating observed gene expressions, thereby uncovering how host–microbe and microbe–microbe interactions regulate microbiome activity. The integration of various microbiome analysis approaches can each contribute a single piece toward completing a large and complex puzzle. Taking a global integrated approach of 16S rRNA characterization, shotgun metagenomics, metatranscriptomics, metaproteomics, and metabolomics may merit careful consideration in cases where budget constraints and sample availability are not prohibitive.

While metatranscriptomics microbiome analysis holds promise in enhancing our understanding of the complex community behavior of the microbiome, several challenges need to be met in order to enhance the reproducibility and general applicability of metatranscriptome analysis. Despite these challenges, metatranscriptomics analysis of the microbiome may be of great value in moving from a descriptive microbiome facet to a deeper understanding of causality in microbial contribution to homeostasis and disease susceptibility. As such, integration of metatranscriptomics into microbiome research may enable to gain better understanding of its diverse roles in mammalian physiology and integrate these data into the clinical world.

Acknowledgments

We thank the members of the Elinav Lab for fruitful discussions. We apologize to authors whose relevant work was not included in this review owing to space constraints.

Author Contributions

Wrote the first draft of the manuscript: SB, GZS, EE. Agree with the manuscript results and conclusions: SB, GZS, EE. Jointly developed the structure and arguments for the paper: SB, GZS, EE. Made critical revisions and approved the final version: SB, GZS, EE. All authors reviewed and approved of the final manuscript.

REFERENCES

- Lederberg J, McCray AT. 'Ome Sweet 'Omics – a genealogical treasury of words. *Scientist*. 2001;15:8.
- Eckburg PB, Bik EM, Bernstein CN, et al. Diversity of the human intestinal microbial flora. *Science*. 2005;308(5728):1635–8.
- Rawls JF, Mahowald MA, Ley RE, Gordon JI. Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. *Cell*. 2006;127(2):423–33.
- Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature*. 2006;444(7122):1027–31.
- Bäckhed F, Manchester JK, Semenkovich CF, Gordon JI. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proc Natl Acad Sci U S A*. 2007;104(3):979–84.
- Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature*. 2012;486(7402):207–14.
- Human Microbiome Jumpstart Reference Strains Consortium, Nelson KE, Weinstock GM, et al. A catalog of reference genomes from the human microbiome. *Science*. 2010;328(5981):994–9.
- Sender R, Fuchs S, Milo R. Are we really vastly outnumbered? Revisiting the ratio of bacterial to host cells in humans. *Cell*. 2016;164(3):337–40.
- Goodrich JK, Waters JL, Poole AC, et al. Human genetics shape the gut microbiome. *Cell*. 2014;159(4):789–99.
- Arumugam M, Raes J, Pelletier E, et al. Enterotypes of the human gut microbiome. *Nature*. 2011;473(7346):174–80.
- González A, Vázquez-Baeza Y, Knight R. SnapShot: the human microbiome. *Cell*. 2014;158(3):690–0.e1.
- Turnbaugh PJ, Hamady M, Yatsunenko T, et al. A core gut microbiome in obese and lean twins. *Nature*. 2009;457(7228):480–4.
- Kamada N, Chen GY, Inohara N, Núñez G. Control of pathogens and pathobionts by the gut microbiota. *Nat Immunol*. 2013;14(7):685–90.
- Yamamoto M, Yamaguchi R, Munakata K, et al. A microarray analysis of gnotobiotic mice indicating that microbial exposure during the neonatal period plays an essential role in immune system development. *BMC Genomics*. 2012;13:335.
- Flint HJ, Scott KP, Duncan SH, Louis P, Forano E. Microbial degradation of complex carbohydrates in the gut. *Gut Microbes*. 2014;3(4):289–306.
- Kau AL, Ahern PP, Griffin NW, Goodman AL, Gordon JI. Human nutrition, the gut microbiome and the immune system. *Nature*. 2011;474(7351):327–36.
- Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci U S A*. 2007;104(34):13780–5.
- Jiang W, Wu N, Wang X, et al. Dysbiosis gut microbiota associated with inflammation and impaired mucosal immune function in intestine of humans with non-alcoholic fatty liver disease. *Sci Rep*. 2015;5:8096.
- Lozupone CA, Rhodes ME, Neff CP, Fontenot AP, Campbell TB, Palmer BE. HIV-induced alteration in gut microbiota: driving factors, consequences, and effects of antiretroviral therapy. *Gut Microbes*. 2014;5(4):562–70.
- Wang J, Li F, Wei H, Lian Z-X, Sun R, Tian Z. Respiratory influenza virus infection induces intestinal immune injury via microbiota-mediated Th17 cell-dependent inflammation. *J Exp Med*. 2014;211(12):2397–410.
- Levy M, Thaiss CA, Elinav E. Metagenomic cross-talk: the regulatory interplay between immunogenomics and the microbiome. *Genome Med*. 2015;7(1):120.
- Yu Y-N, Fang J-Y. Gut microbiota and colorectal cancer. *Gastrointest Tumors*. 2015;2(1):26–32.
- McKenney PT, Pamer EG. From hype to hope: the gut microbiota in enteric infectious disease. *Cell*. 2015;163(6):1326–32.
- Shapiro H, Thaiss CA, Levy M, Elinav E. The cross talk between microbiota and the immune system: metabolites take center stage. *Curr Opin Immunol*. 2014;30:54–62.
- David LA, Maurice CF, Carmody RN, et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature*. 2014;505(7484):559–63.
- Zeevi D, Korem T, Zmora N, et al. Personalized nutrition by prediction of glycemic responses. *Cell*. 2015;163(5):1079–94.
- David LA, Materna AC, Friedman J, et al. Host lifestyle affects human microbiota on daily timescales. *Genome Biol*. 2014;15(7):R89.



28. Morgan XC, Huttenhower C. Chapter 12: human microbiome analysis. *PLoS Comput Biol*. 2012;8(12):e1002808.
29. Liu Z, DeSantis TZ, Andersen GL, Knight R. Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. *Nucleic Acids Res*. 2008;36(18):e120.
30. Kuczynski J, Lauber CL, Walters WA, et al. Experimental and analytical tools for studying the human microbiome. *Nat Rev Genet*. 2012;13(1):47–58.
31. Wooley JC, Godzik A, Friedberg I. A primer on metagenomics. Bourne PE, ed. *PLoS Comput Biol*. 2010;6(2):e1000667.
32. Qin J, Li R, Raes J, et al; MetaHIT Consortium. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*. 2010;464(7285):59–65.
33. Gosalbes MJ, Durbán A, Pignatelli M, et al. Metatranscriptomic approach to analyze the functional human gut microbiota. *PLoS One*. 2011;6(3):e17447.
34. Verberkmoes NC, Russell AL, Shah M, et al. Shotgun metaproteomics of the human distal gut microbiota. *ISME J*. 2009;3(2):179–89.
35. Erickson AR, Cantarel BL, Lamendella R, et al. Integrated metagenomics/metaproteomics reveals human host-microbiota signatures of Crohn's disease. *PLoS One*. 2012;7(11):e49138.
36. Marcobal A, Kashyap PC, Nelson TA, et al. A metabolomic view of how the human gut microbiota impacts the host metabolome using humanized and gnotobiotic mice. *ISME J*. 2013;7(10):1933–43.
37. Vázquez-Castellanos JF, Serrano-Villar S, Latorre A, et al. Altered metabolism of gut microbiota contributes to chronic immune activation in HIV-infected individuals. *Mucosal Immunol*. 2015;8(4):760–72.
38. Ursell LK, Haider HJ, Van Treuren W, et al. The intestinal metabolome: an intersection between microbiota and host. *Gastroenterology*. 2014;146(6):1470–6.
39. Liang P, Pardee AB. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science*. 1992;257(5072):967–71.
40. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*. 1995;270(5235):467–70.
41. Wang Z, Gerstein M, Snyder M. RNA-seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 2009;10(1):57–63.
42. Franzosa EA, Morgan XC, Segata N, et al. Relating the metatranscriptome and metagenome of the human gut. *Proc Natl Acad Sci U S A*. 2014;111(22):E2329–38.
43. Shirley B, Alejandra V-L, Fernanda C-G, et al. Combining metagenomics, metatranscriptomics and viromics to explore novel microbial interactions: towards a systems-level understanding of human microbiomes. *Comput Struct Biotechnol J*. 2015;13:390–401.
44. Benítez-Páez A, Belda-Ferre P, Simón-Soro A, Mira A. Microbiota diversity and gene expression dynamics in human oral biofilms. *BMC Genomics*. 2014;15:311.
45. Peano C, Pietrelli A, Consolandi C, et al. An efficient rRNA removal method for RNA sequencing in GC-rich bacteria. *Microb Inform Exp*. 2013;3(1):1.
46. Sultan M, Amstislavskiy V, Risch T, et al. Influence of RNA extraction methods and library selection schemes on RNA-seq data. *BMC Genomics*. 2014;15:675.
47. Wendisch VF, Zimmer DP, Khodursky A, Peter B, Cozzarelli N, Kustu S. Isolation of *Escherichia coli* mRNA and comparison of expression using mRNA and total RNA on DNA microarrays. *Anal Biochem*. 2001;290(2):205–13.
48. Sharma CM, Hoffmann S, Darfeuille F, et al. The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature*. 2010;464(7286):250–5.
49. Giannoukos G, Ciulla DM, Huang K, et al. Efficient and robust RNA-seq process for cultured bacteria and complex community transcriptomes. *Genome Biol*. 2012;13(3):R23.
50. Korf I. Genomics: the state of the art in RNA-seq analysis. *Nat Methods*. 2013;10(12):1165–6.
51. Abubucker S, Segata N, Goll J, et al. Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput Biol*. 2012;8(6):e1002358.
52. Glass EM, Wilkening J, Wilke A, Antonopoulos D, Meyer F. Using the metagenomics RAST server (MG-RAST) for analyzing shotgun metagenomes. *Cold Spring Harb Protoc*. 2010;2010(1):db.rot5368.
53. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–9.
54. Marco-Sola S, Sammeth M, Guigó R, Ribeca P. The GEM mapper: fast, accurate and versatile alignment by filtration. *Nat Methods*. 2012;9(12):1185–8.
55. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20.
56. Ghosh S, Chan C-KK. Analysis of RNA-seq data using TopHat and Cufflinks. *Methods Mol Biol*. 2016;1374:339–61.
57. Morgan XC, Huttenhower C. Meta'omic analytic techniques for studying the intestinal microbiome. *Gastroenterology*. 2014;146(6):1437–48.e1.
58. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005;21(18):3674–6.
59. Grabherr MG, Haas BJ, Yassour M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29(7):644–52.
60. Maurice CF, Haider HJ, Turnbaugh PJ. Xenobiotics shape the physiology and gene expression of the active human gut microbiome. *Cell*. 2013;152(1–2):39–50.
61. Haider HJ, Turnbaugh PJ. Is it time for a metagenomic basis of therapeutics? *Science*. 2012;336(6086):1253–5.
62. Sousa T, Paterson R, Moore V, Carlsson A, Abrahamsson B, Basit AW. The gastrointestinal microbiota as a site for the biotransformation of drugs. *Int J Pharm*. 2008;363(1–2):1–25.
63. Vijay-Kumar M, Aitken JD, Carvalho FA, et al. Metabolic syndrome and altered gut microbiota in mice lacking toll-like receptor 5. *Science*. 2010;328(5975):228–31.
64. Vijay-Kumar M, Sanders CJ, Taylor RT, et al. Deletion of TLR5 results in spontaneous colitis in mice. *J Clin Invest*. 2007;117(12):3909–21.
65. Carvalho FA, Koren O, Goodrich JK, et al. Transient inability to manage proteobacteria promotes chronic gut inflammation in TLR5-deficient mice. *Cell Host Microbe*. 2012;12(2):139–52.
66. Slack E, Hapfelmeier S, Stecher B, et al. Innate and adaptive immunity cooperate flexibly to maintain host-microbiota mutualism. *Science*. 2009;325(5940):617–20.
67. Turner JR. Intestinal mucosal barrier function in health and disease. *Nat Rev Immunol*. 2009;9(11):799–809.
68. Sanders CJ, Yu Y, Moore DA, Williams IR, Gewirtz AT. Humoral immune response to flagellin requires T cells and activation of innate immunity. *J Immunol*. 2006;177(5):2810–8.
69. Cullender TC, Chassaing B, Janzon A, et al. Innate and adaptive immunity interact to quench microbiome flagellar motility in the gut. *Cell Host Microbe*. 2013;14(5):571–81.
70. Stazic D, Voß B. The complexity of bacterial transcriptomes. *J Biotechnol*. 2015. DOI: 10.1016/j.jbiotec.2015.09.041.
71. Thomason MK, Storz G. Bacterial antisense RNAs: how many are there, and what are they doing? *Annu Rev Genet*. 2010;44(1):167–88.
72. Georg J, Voss B, Scholz I, Mitschke J, Wilde A, Hess WR. Evidence for a major role of antisense RNAs in cyanobacterial gene regulation. *Mol Syst Biol*. 2009;5:305.
73. Toledo-Arana A, Repoila F, Cossart P. Small noncoding RNAs controlling pathogenesis. *Curr Opin Microbiol*. 2007;10(2):182–8.
74. Mutalik VK, Qi L, Guimaraes JC, Lucks JB, Arkin AP. Rationally designed families of orthogonal RNA regulators of translation. *Nat Chem Biol*. 2012;8(5):447–54.
75. Dornenburg JE, Devita AM, Palumbo MJ, Wade JT. Widespread antisense transcription in *Escherichia coli*. *MBio*. 2010;1(1):e24–10.
76. Beaume M, Hernandez D, Farinelli L, et al. Cartography of methicillin-resistant *S. aureus* transcripts: detection, orientation and temporal expression during growth phase and stress conditions. *PLoS One*. 2010;5(5):e10725.
77. Bao G, Wang M, Doak TG, Ye Y. Strand-specific community RNA-seq reveals prevalent and dynamic antisense transcription in human gut microbiota. *Front Microbiol*. 2015;6:896.
78. Massé E, Salvail H, Desnoyers G, Arguin M. Small RNAs controlling iron metabolism. *Curr Opin Microbiol*. 2007;10(2):140–5.
79. Lenz DH, Mok KC, Lilley BN, Kulkarni RV, Wingreen NS, Bassler BL. The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell*. 2004;118(1):69–82.
80. Duran-Pinedo AE, Yost S, Frias-Lopez J. Small RNA transcriptome of the oral microbiome during periodontitis progression. *Appl Environ Microbiol*. 2015;81(19):6688–99.
81. Livny J, Waldor MK. Identification of small RNAs in diverse bacterial species. *Curr Opin Microbiol*. 2007;10(2):96–101.
82. Irnov I, Sharma CM, Vogel J, Winkler WC. Identification of regulatory RNAs in *Bacillus subtilis*. *Nucleic Acids Res*. 2010;38(19):6637–51.
83. Shi Y, Tyson GW, DeLong EF. Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column. *Nature*. 2009;459(7244):266–9.
84. Hajishengallis G. The inflammophilic character of the periodontitis-associated microbiota. *Mol Oral Microbiol*. 2014;29(6):248–57.
85. Yost S, Duran-Pinedo AE, Teles R, Krishnan K, Frias-Lopez J. Functional signatures of oral dysbiosis during periodontitis progression revealed by microbial metatranscriptome analysis. *Genome Med*. 2015;7(1):27.
86. Jeanty C, Longrois D, Mertes P-M, Wagner DR, Devaux Y. An optimized protocol for microarray validation by quantitative PCR using amplified amino allyl labeled RNA. *BMC Genomics*. 2010;11:542.
87. Felix RS, Colleoni GWB, Caballero OL, et al. SAGE analysis highlights the importance of p53cvs, ddx5, mapkapk2 and ranbp2 to multiple myeloma tumorigenesis. *Cancer Lett*. 2009;278(1):41–8.
88. Fang Z, Cui X. Design and validation issues in RNA-seq experiments. *Brief Bioinform*. 2011;12(3):280–7.