Capturing dynamic shifts in the microbiome with metatranscriptomics

Diversigen leverages Illumina technology to gain a deeper understanding of the functional activity in complex microbial communities

TONYA WARD, PHD,
SR DIRECTOR, BIOINFORMATICS
& LABORATORY OPERATIONS,
DIVERSIGEN

DR. EMILY HOLLISTER, PHD,
DIRECTOR, R&D COMPUTATIONAL
BIOLOGY, DIVERSIGEN

The microbiome is a dynamic collection of microbes that have the potential to impact host physiology in health and disease. Commonly used DNA-based metagenomic approaches do not reflect the functional activity of the microbiome accurately. Metatranscriptomics, leverages next-generation sequencing technology (NGS) to profile all the microbial transcripts within a complex sample, providing an unbiased view into which individual microbes are active, what they are doing, and how their function impacts the rest of the microbial community and host physiology.

Diversigen and sister company DNA Genotek have developed optimized workflows for microbiome sample collection, extraction, sequencing, and data analytics for high-performance metatranscriptomics applications with a goal to accelerate microbiome research. We spoke with Dr Tonya Ward and Dr Emily Hollister from Diversigen and Dr Brice Le François from DNA Genotek about the evolution of metatranscriptomics, key considerations for designing metatranscriptomics studies, and the future of metatranscriptomics research.

Q: Can you tell me about Diversigen?

Tonya Ward (TW): Diversigen provides innovative and scalable services, mostly centered around profiling all of the microbes that live in and around us. Our mission is to accelerate scientific discovery and innovation by leveraging multiomic expertise and cutting-edge microbiome science. Our goal is making the best methods available to researchers to help drive their biological breakthroughs. These methods include advances to amplicon sequencing, shotgun sequencing, metatranscriptomics, and the

DR. BRICE LE FRANÇOIS, PHD,
SR R&D MANAGER, DNA GENOTEK

"Metatranscriptomics...seeks to characterize gene expression within a microbial community and provide insight into active versus potential function in the community."

ability to integrate multiple 'omic' data types with our comprehensive annotations and custom analysis. Because the microbiome field is ever evolving, our focus is on innovating in the realm of microbial profiling to help provide our customers with cutting-edge results.

Q: What is metatranscriptomics and how is it different from metagenomics?

Emily Hollister (EH): Broadly, metagenomics is the study of genetic material from all microorganisms present within an environment or ecosystem. It can tell you which organisms are present and give insights into their potential gene-encoded function. Metatranscriptomics, which has its roots in single-organism transcriptomic studies, seeks to characterize gene expression within a microbial community and provide insight into active versus potential function in the community. Although metagenomic-based understanding of potential function has led to important insights regarding microbial community dynamics and host–microbe interactions, there are also examples in the literature demonstrating that what you might predict based on the DNA composition of a microbial community does not necessarily reflect the function of that community. The choice between metagenomics and metatranscriptomics ultimately boils down to whether you want to understand what is present in a microbial community versus what is active within that community.

Q: What are the key considerations for planning metatranscriptomics studies?

Brice LeFrançois (BLF): Bacteria are, by design, extremely adaptable. They have the ability to change their transcriptome within seconds in response to subtle changes in their environment. In order to rapidly modify their gene expression, bacteria use enzymes associated with ribosomes to quickly degrade messenger RNA (mRNA) that is being translated. To make sure that the sample collected accurately represents what the bacteria are doing *in vivo*, it is critical that the bacteria are captured in the state they're in at the point of collection. Sample processing and storage are key sources of bias. For example, mRNA tends to be fairly easily degraded, so loss of some mRNA transcripts during sample handling means that they are not captured during library preparation, ultimately leading to skewed gene expression profiles. Furthermore, recent research in the microbiome field has shown that RNA extraction is the single most important factor in getting accurate transcriptomic profiles. Inefficient cell lysis can cause important groups of cells that may be active and contribute to the overall transcriptome to be missed altogether.

EH: From an analytics perspective, there are also some important questions to consider. First, how well do you understand your microbial community? Do you have reasonable reference genomes available to support your annotation? Are you going to perform read mapping or assembly? How are you going to do your downstream differential expression analysis? The answers to each of these questions will shape the ways in which you design and carry out your studies.

Q: How can researchers optimize their sample collection, RNA extraction, and library preparation processes?

BLF: Optimizing sample collection depends on the sample type. For example, swabs are typically used to collect vaginal or skin microbiome samples, whereas fecal samples are collected directly using our OMNIgene™•GUT collection kit. It is critical that the sample is stabilized right away to prevent any changes in the metatranscriptomic profile. The best approach is to use a buffer that lyses the bacteria and prevents them from reacting to changes in their environment as part of the collection process. A caveat is that bacterial lysis also releases RNA, leaving it vulnerable to degradation by nucleases present in the samples. So, a strong stabilization chemistry that mitigates the impact of nucleases is necessary to keep RNA intact for downstream sequencing and analysis. Choosing an RNA extraction kit that minimizes the potential for RNA degradation during sample processing is also extremely important. Effective bacterial ribosomal RNA depletion is another key factor to consider for any transcriptomic study.

"Choosing an RNA extraction kit that minimizes the potential for RNA degradation during sample processing is also extremely important. Effective bacterial ribosomal RNA depletion is another key factor to consider for any transcriptomic study."

Q: Why is ribosomal RNA depletion necessary for metatranscriptomics studies?

BLF: When total RNA is extracted from a microbiome sample, the vast majority, 98% or even more, is ribosomal RNA. This type of RNA is not particularly useful as it does not provide any information about functional status or metabolic activity. For instance, some cells are metabolically inactive but still contain thousands of copies of ribosomal RNA. This is why eliminating ribosomal RNA is critical to reveal the true metabolic activity in complex microbial samples.

EH: The presence of abundant ribosomal RNA in metatranscriptomics samples means that, in the absence of effective ribosomal RNA depletion, researchers have to sequence very deeply to access the most biologically relevant transcripts from complex microbial communities.

"When total RNA is extracted from a microbiome sample, the vast majority, 98% or even more, is ribosomal RNA... eliminating ribosomal RNA is critical to reveal the true metabolic activity in complex microbial samples."

Q: Can you tell me about your experience with codeveloping the Illumina Ribo-Zero™ Plus Microbiome Depletion Kit?

BLF: When we started working on metatranscriptomics, we found that commercially available ribosomal RNA depletion kits performed adequately for pure cultures or simple microbial communities. However, for fecal samples, which comprise hundreds of microbial species, they failed to effectively deplete the ribosomal RNA fraction. Around the same time, Illumina was working on the beta version of a new ribosomal RNA depletion kit. We tested that kit and saw promising performance, though depletion was definitely not enough. A majority of reads were still mapping to ribosomal RNA. So, we engaged with the Illumina research team to codevelop a solution that could deplete unwanted ribosomal RNA in complex microbiome samples that can contain hundreds of bacterial species. This collaboration led to the development of the Illumina Ribo-Zero Plus Microbiome Depletion Kit. We are seeing excellent performance and high pan-microbial ribosomal RNA depletion efficiency across all the samples we have tested, which is great.

"In the absence of effective ribosomal RNA depletion, researchers have to sequence very deeply to access the most biologically relevant transcripts."

Q: What are the critical steps in the data analysis workflows for metatranscriptomics studies?

EH: Our data analysis pipeline starts with quality control, including the trimming of adapters and removal of low-quality reads. We also filter out host transcripts, which we are not necessarily interested in because our focus is on the microbes. Host content varies tremendously based on sample type. Even with effective chemical ribosomal RNA depletion, some ribosomal reads do still make it through to sequencing. So, the next step is to filter out the unwanted ribosomal RNA reads. From there, we start a process of looking at both taxonomy–that is, the composition of microbial community–and function, using curated microbial databases. This allows us to understand expressed genes and the microbes that are likely to be expressing them. These outputs can be used to analyze differential expression depending on study type and research goals. Whether it's longitudinal data, or a case control study, our data analysis pipelines can reveal the key factors driving the differences in complex microbial communities and how their metabolic function is impacted as a result of these differences.

"We engaged with the Illumina research team to co-develop a solution that could deplete unwanted ribosomal RNA in complex microbiome samples that can contain hundreds of bacterial species."

For Research Use Only. Not for use in diagnostic procedures.

Q: What insights can researchers gain from metatranscriptomics data sets?

EH: Metatranscriptomics can answer several important questions about the microbiome. First, what microbes are active in the community and what are they doing? Are they reacting in a specific way to an intervention, be that medication, diet, or the way in which the samples were handled? If you sample across individuals, do their microbiomes behave in similar ways or are there highly individualized responses? Another interesting application of metatranscriptomics is profiling viruses. Because metatranscriptomics looks at RNA, it can be a useful tool for exploring viruses, particularly RNA viruses that might not be captured by metagenomics studies.

"Because metatranscriptomics looks at RNA, it can be a useful tool for exploring viruses, particularly RNA viruses that might not be captured by metagenomics studies."

Q: What is the future of metatranscriptomics in microbome research?

BLF: Currently, metatranscriptomics is not being widely used because it is a relatively new field and researchers don't fully know what to expect with metatranscriptomics data. Metatranscriptomics has the potential to uncover insights into the true functional activity of the microbiome at the time the sample is collected. Another fantastic application of metatranscriptomic studies is host–microbiome interactions. With metatranscriptomics, we have the opportunity to study the interplay between the host and the microbiome because it doesn't just reveal what microbes are present, but also which microbes are active in the community, and how they may influence host biology. For instance, acne is caused by a skin commensal that everybody has, but whether a person develops acne depends upon what these bacteria are doing and how the host immune system responds to them. So, studying microbial gene expression opens a new window onto the human microbiome and how it plays into our health. The field is at a turning point, and I am excited to see how metatranscriptomics evolves in the next few years.

"Anywhere microbes are present, we have the opportunity to learn more about how they're functioning on an individual level, thereby helping to drive the overall ecosystem function."

EH: An exciting space to watch as metatranscriptomics develops will involve studying mechanisms. Many predictions can be made based on the microbiome composition in a given sample, but with metatranscriptomics we can drill down and understand how microbial function drives biology, impacting the host and the broader ecosystem. The promise of this area of research is not limited to human health. Anywhere microbes are present, we have the opportunity to learn more about how they're functioning on an individual level, thereby helping to drive the overall ecosystem function.

TW: Metatranscriptomics has developed from single-organism transcriptomics. In the areas of human health, transcriptomics has enabled a lot of discovery, as opposed to genomics alone. Similarly for microbiome studies, having a robust metatranscriptomics methodology can empower researchers to accelerate discovery in the microbiome space. This, in turn, will help to accelerate getting a diagnostic end-point potential out of the microbiome and guide future therapeutic discoveries as well. I believe that the widespread application of metatranscriptomics has the potential to unlock a lot of new aspects of microbial function in complex communities.

"... having a robust metatranscriptomics methodology can empower researchers to accelerate discovery in the microbiome space."

Learn more

Ribo-Zero Plus Microbiome, illumina.com/products/by-type/accessory-products/ribo-zero-plus-microbiome-depletion

Illumina Stranded Total RNA Prep with Ribo-Zero Plus Microbiome, illumina.com/products/by-type/sequencing-kits/library-prep-kits/stranded-total-rna-prep

Diversigen, diversigen.com

DNA Genotek, dnagenotek.com

illumina®

1.800.809.4566 toll-free (US) | +1.858.202.4566 tel | techsupport@illumina.com | www.illumina.com

© 2022 Illumina, Inc. All rights reserved. All trademarks are the property of Illumina, Inc. or their respective owners. For specific trademark information, see www.illumina.com/company/legal.html.
M-GL-00917 v1.0

OMNIgene™•GUT is a trademark of DNA Genotek, Inc.