Olivia McLennan
T00644569
BIOL3210 - Microbial Community Analysis

Microbial community composition & diversity via amplicon sequencing.

Method:

Amplicon sequencing is an approach that allows researchers to analyze the composition of a genetic region by sequencing targeted polymerase chain reaction (PCR) products (Hugenholtz et al. 1998). For microbial community analysis, 16S rRNA genes, which are ribosomal genes with both variable and highly conserved regions, are often used as phylogenetic markers due to their presence in all microorganisms (Hugenholtz et al. 1998). In some cases, 18S rRNA genes, or Internal Transcribed Spacer (ITS) DNA can be used in lieu, however, 16S rRNA genes are often preferred (Clarridge 2004). For an analysis of microbial community composition and diversity, the RNA is extracted from a sample using an extraction technique effective for a wide range of taxa (Gupta et al. 2019). Then, if analyzing using 16S rRNA genes, a specific region of a 16S rRNA gene is amplified using a primer—a short sequence which binds to a targeted conserved region (Gupta et al. 2019). Once amplified using PCR, the 16S rRNA genes can be sequenced (Gupta et al. 2019). Amplicon sequencing is accomplished using next-generation sequencing (NGS), a high-throughput sequencing technique that determines the order of nucleotides in a targeted region, and then compared to reference sequences available in public databases, such as Green Gene, EBI Metagenomics, or Silva Database (Gupta et al. 2019).

Information Obtained:

In 2007, 16S rRNA genes could be sequenced accurately to the genus level more than 90% of the time, as well as to the species level 65% to 83% of the time; sequencing technology has only improved since (Janda and Abbott 2007). If using NGS, the sequence reads post-PCR, ranging from 75-400 base pairs in length, and can be clustered according to similarity (Hugerth and Andersson 2017). They may then be classified as Operational Taxonomic Units, reducing sequential diversity to biological diversity and allowing for taxonomic classification and comparisons across microbiomics studies (Hugerth and Andersson 2017).

Example:

Disclosing the spectrum of bacteria involved in polymicrobial spontaneous brain abscesses has proved difficult in a clinical setting, due to the aggregative nature of the microbiota

and their ability to avoid detection by conventional culturing methods (Andersen et al. 2022). In 2022, a study by Andersen et al. demonstrated how amplicon-based sequencing can be used to identify and distinguish polymicrobial brain abscesses. In 41 cases of spontaneous brain abscess, 35 cases were bacterial, and amplicon-based sequencing of 16S rRNA genes with culture on selective media supplemented the results of the conventional culture (Andersen et al. 2022). As two-thirds of bacterial infections responsible for brain abscesses are caused by dental and oropharyngeal microbiota, and most infections are polymicrobial, amplicon-based sequencing of 16S rRNA genes with culture on selective media in addition to conventional culture methods could vastly improve treatment options and outcomes (Andersen et al. 2022).

Advantages:

In comparison to traditional culture methods, NGS is much more sensitive. For example, in one study, NGS sampling found 75.70% of the unique bacteria in a sample, while TCM sampling only identified 23.86% of the unique bacteria that NGS found (Gupta and Andersson 2019). Using NGS, the alpha diversity of a sample can be determined, and parallel sequencing of hundreds of samples can be done concurrently within a day (Gupta and Andersson 2019). However, the main advantage of 16S rRNA NGS is that a sample of bacteria can be sequenced regardless of if it can be cultured, providing a tool that circumvents “the great plate anomaly” (Staley and Konopka 1985; Gupta and Andersson 2019).

Disadvantages:

Unfortunately, when using amplicon sequencing, each step, including data collection methods when sampling, sample storage conditions, the amount of material used for extraction, the extraction technique, and the choice of primer, can all lead to biases in the interpretation of microbial community composition and diversity (Hugerth and Andersson 2017; Straub et al. 2020). Furthermore, amplicon sequencing does not provide information on antimicrobial susceptibility, a drawback of its application in the medical field (Gupta and Andersson 2019). Finally, another major limitation of this approach for the analysis of community composition and diversity is that novel sequences cannot be recognized, as classification is commensurable to the underlying databases being used and the sequences they already contain (Konopka 2009; Hugerth and Andersson 2017).

Metagenomics for microbial or viral communities (NOT 16S and 18S rRNA sequencing).

Method:

Metagenomics is a culture-independent analysis which can be done to analyze the genome of a microbial or viral community (Strazzulli et al. 2017). First, metagenomic DNA (mDNA) is extracted from a sample and fragmented using a known restriction enzyme (Strazzulli et al. 2017). To ensure that the mDNA properly represents the community, specific protocols are used to preserve the quality and number of nucleic acids; this improves the sequencing results, and thus the overall metagenomic library (Strazzulli et al. 2017). The fragmented mDNA is then cloned into a vector, transformed into a host bacterium, and screened for the resulting transformants (Handelsman 2004). Next, the transformants are analyzed for phylogenetic markers or other conserved genes, tested for the expression of specific traits, or sequenced (Handelsman 2004). Once filtered for non-target information, analysis results can be matched to various reference databases, allowing for classification (Handelsman 2004; Lee 2019).

Information Obtained:

Metagenomics provides access to the genetic composition of a microbial community and its phylogeny, facilitating studies in the environmental physiology and ecology of microorganisms (Handelsman 2004; Strazzulli et al. 2017). This methodology has revealed genome rearrangements and differences in how microorganisms store elements, resulting in insights about energy and nutrient cycling, as well as gene function and transfer within communities (Handelsman 2004; Strazzulli et al. 2017). Additionally, many novel genes and products have been discovered, elucidating new biological interests (Handelsman 2004; Strazzulli et al. 2017).

Example:

In 2021, the first systematic catalogue of the urban microbial ecosystem, based on 4,728 metagenomic samples from various transit systems worldwide, was presented (Danko et al. 2021). This catalogue was created using a strict methodology, including standard operating procedures for sampling, processing controls, modern DNA extraction and preparation techniques, library preparation using a QIAGEN kit, uncompromising quality controls, sample

identification and analysis methods (Danko et al. 2021). Metagenomics identified a wide range of variations in taxonomy across all cities, with each having its microbial signature, revealing the direct climate and geographic differences (Danko et al. 2021). In addition to the widespread observation of new genes, it was also found that antimicrobial resistance (AMR) genes form distinct clusters (Danko et al. 2021). These results will enable further discoveries of new genes, as well as highlight public health applications and issues, including AMR (Danko et al. 2021).

Advantages:

Being an unbiased, hypothesis-free method, metagenomics provides access to the diversity of the entire genetic composition of a microbial community (Strazzulli et al. 2017; Lee 2019). Metagenomics offers a description that is more whole than current phylogenetic surveys based only on the diversity of the 16S rRNA gene (Strazzulli et al. 2017; Lee 2019). Additionally, metagenomic data can provide unique and novel insights into community dynamics, given that not all genes of interest need to be recognized by sequence analysis (Handelsman 2004).

Disadvantages:

One of the most significant limitations of metagenomics is that many genes are not expressed by domesticated host bacterium selected for cloning, hindering the identification of new genes which encode new or known functions (Handelsman 2004). Another reservation regarding metagenomics is the interpretation of findings (Lee 2019). This is because distinguishing contamination from the target metagenome is made difficult by the sensitivity of NGS and the consistency of reference databases (Lee 2019).

Microbial community transcriptomics (whole-transcriptome shotgun sequencing; RNA-seq).

Method:

Transcriptomics is a powerful tool that enables comprehensive annotation and quantification of bacterial, viral, and other microbial transcripts (Mukherjee and Sudhakara Reddy 2020; Illumina 2022). It can be used on its own or as a supplement to genomics to provide insight into whether genomic DNA is from a viable cell and is actually being expressed under experimental conditions (Bashiardes et al. 2016). To begin, genetic information is sampled, and RNA is extracted and precipitated (Illumina 2022). RNA samples are treated with DNase and may then undergo whole-transcriptome shotgun sequencing or RNA-Seq (Moran 2009; Illumina 2022). If shotgun sequencing is being performed, the RNA is converted to complementary DNA (cDNA) and then sheared or restricted (Weinstock 2012; Illumina 2022). Sequencing is accomplished using random universal primers, and the sequence library is prepared by overlapping the reads (Weinstock 2012; Illumina 2022). If RNA-Seq is being performed, the RNA is converted to cDNA and then sheared or restricted before short sequencing adaptors are added to each molecule (Wang et al. 2009; Illumina 2022). Sequencing is then performed in a high-throughput manner to gain the sequence of each fragment and sort them into exonic reads, junction reads, or polyA tail reads; this method is also quantitative (Wang et al. 2009; Illumina 2022). Sequence reads from either method can be mapped to a reference genome, providing qualitative and quantitative information about the transcript data that can be compared to other datasets (Moran 2009; Illumina 2022).

Information Obtained:

Metatranscriptome methodology renders scientists with a comprehensive set of RNAs encoded by a community's DNA (Moran 2009). Through the mapping of sequence reads from a sequence library or RNA-Seq to a reference genome, transcriptomics allows us to "eavesdrop on microbial ecology," through knowledge of gene expression, and proteins necessary for cell growth, metabolism, transcription, and translation (Moran 2009).

Example:

In 2021, evidence for microbial carbon and nitrogen cycles in deoxygenated seawater was elucidated by using an RNA-Seq methodology (Han et al. 2021). During a seasonal deoxygenation event in the Bohai Sea, samples were collected from two sites whose temperature, pH, and salinity were measured prior (Han et al. 2021). Total sample RNA was extracted, purified, and amplified before being converted to cDNA for paired-end sequencing (Han et al. 2021). Patterns of transcriptional activity were analyzed, and the results demonstrated high expression of genes associated with ammonia oxidation and transportation, as well as carbon fixation (Han et al. 2021). Additionally, photosynthetic, and carbon-fixing transcripts of *Cyanobacteria* genes decreased, whereas transcripts of ribulose bisphosphate carboxylase-encoding genes increased (Han et al. 2021). These results indicate that in a low-oxygen environment, there is an alteration of the microbial groups involved in carbon and nitrogen cycling (Han et al. 2021).

Advantages:

Transcriptomics is amenable to experiments in which gene expression is monitored while an environmental parameter is manipulated; metatranscriptomics is an unparalleled tool for understanding the microbial processes within communities at the transcript level (Moran 2009). Transcriptomics provides a picture of actual gene expression at a given moment, whereas genomics provides gene potential (Moran 2009; Bashiardes et al. 2016). Another notable advantage of shotgun sequencing and RNA-Seq is that universal primers are used to create libraries, eliminating the need to anticipate genes and introduce additional bias (Wang et al. 2009; Weinstock 2012; Moran 2009).

Disadvantages:

Transcriptomics is often unable to capture the entire metatranscriptome for a variety of reasons, e.g., the absence of a polyA tail in non-eukaryotic organisms, making mRNA isolation difficult (Moran 2009; Shakya et al. 2019). In addition, a major disadvantage of shotgun sequencing and RNA-Seq is that universal primers, though not specific, are limited to known sequences, making it difficult to sequence rare transcripts properly (Moran 2009).

Microbial community proteomics.

Method:

Through bridging genetic potential and final phenotype, metaproteomics is a community analysis method that highlights a unique view of the adaptive responses of a microbial community (Wilmes et al. 2015). This analysis method offers the ability to recover a protein complement from a community, then study and characterize it to resolve the components driving the microbial community's function (Wilmes et al. 2015; Wang et al. 2016). Furthermore, proteomic databases, a result of proteomic studies, have facilitated the classification of microbial species by ecological function (Wang et al. 2016). To begin the typical workflow of a proteomic study, a protein extraction protocol based on the specific environment the microbial community resides in is performed (Wang et al. 2016). This is followed by protein separation and identification, which is done in one of two ways (Wang et al. 2016). For a gel-based separation, two-dimensional polyacrylamide gel electrophoresis is performed so that the resolved bands can be exposed to enzymes that will degrade them to peptides to be identified using mass spectroscopy (Wang et al. 2016). Alternatively, a liquid chromatography method coupled with tandem mass spectroscopy could be used to both separate and identify proteins; this methodology is often preferred (Wang et al. 2016). Once identified, proteins can then be interpreted for bioinformatic analysis (Wang et al. 2016).

Information Obtained:

Mass spectroscopy or tandem mass spectroscopy results in mass-to-charge data that can be analyzed through peptide mass fingerprinting, de novo sequencing, or database searching (Wang et al. 2016). These metaproteomic analyses can reflect gene expression regulation, protein synthesis, and microbial physiological activity in various environmental conditions (Wang et al. 2016).

Example:

The health of our gut microbiomes has implications for the pathogenesis of colorectal cancer (CRC). In 2020, Long et al. characterized the gut microbiomes of 14 CRC patients and 14 healthy volunteers using a quantitative metaproteomic method. Stool samples were collected, microbial cells within the samples were enriched, and the proteins were extracted before being

analyzed with liquid chromatography and tandem mass spectroscopy (Long et al. 2020). 341 microbial proteins were found in significantly different abundances between the CRC patients and the healthy volunteers, offering a new diagnostic method for CRC with the 20 most discriminating proteins (Long et al. 2020). As CRC is the fourth most deadly cancer worldwide, the findings this study demonstrated regarding differing abundances of proteins, as well as red meat consumption and oxidative stress, should be noted by oncologists and the general public alike (Long et al. 2020).

Advantage:

In comparison to other "omics" methods, the primary advantage of metaproteomics is the functional information it provides, enabling studies of how microbial communities physiologically respond to environmental conditions (Wilmes and Bond 2006; Wilmes et al. 2015; Wang et al. 2016). This microbial analysis method is an effective tool that can be used to study the characterization of microbial communities in time and space, allowing for the classification of community types (Wang et al. 2016). Another advantage of metaproteomics, like other "omics" methods, is that it is not hindered by the inability to culture samples (Wilmes and Bond 2006). It is thus unbiased, as culturing techniques select for easily culturable organisms and distort our microbial understanding (Wilmes and Bond 2006).

Disadvantages:

Environmental conditions pose a problem for the extraction and maintenance of samples (Wilmes and Bond 2006; Wang et al. 2016). There are neither universal extraction protocols that can be followed nor universal protocols for maintaining the *in situ* environment and thus the *in situ* activity of proteins, which challenges the validity of some metaproteomic studies (Wilmes and Bond 2006; Wang et al. 2016). Lastly, the metaproteome reflects the abundant microbes in microbial populations, but often not rare species, hindering our understanding of the functions of microbial species, and therefore the community as a whole (Wang et al. 2016). It is presumed that only a fraction of the metaproteome (<1%) can be resolved with current methods, given the demanding searches of the robust protein database with the hardware available at this time (Wilmes and Bond 2006; Wang et al. 2016).

Microbial community metabolomics.

Method:

Microbial community metabolomics is a component of systems biology—the analysis of a complex biological system (Hollywood 2006; Tang 2011). Studying the complete set of metabolites within a microorganism or a microbial community, and the outcomes of its interactions may elucidate an unbiased knowledge of an organism or community's state in a specific condition (Tang 2011). Metabolomics is most often performed by nuclear magnetic resonance imaging (NMR) or mass spectroscopy (MS) (Tang 2011). Each method affords researchers the ability to analyze a vast range of metabolites in a single measurement and explicates the metabolites' structures and molecular concentrations (Tang 2011). NMR does not require any physical or chemical treatments, thus analyzed samples can be recovered, whereas MS requires treatments and is frequently coupled to a separation technique, such as a chromatography-based method, but is more sensitive than NMR (Tang 2011). Both technologies are accepted analytical methods for metabolomics studies, and each method has its specific advantages and disadvantages (Tang 2011).

Information Obtained:

Studying a metabolome, the set of metabolites produced by an organism or a community, reflects enzymatic pathways and processes encoded within the genome or metagenome (Tang 2011). Metabolite composition data, in the form of chemical shifts in ppm (NMR) or mass-to-charge ratio (MS) analysis results, reveal how a changing environment may affect the developmental and biological processes of an organism or a community (Tang 2011).

Example:

In 2022, Xia et al. analyzed the quality and formation of metabolic components in fermented koumiss, a traditional low-alcohol dairy beverage. To do so, the milk of 12 mares, 6 black and 6 white was compared using a metabolomic method (Xia et al. 2022). After milk samples were collected in a sterile container, metabolites were extracted using ultrasonication and centrifugation (Xia et al. 2022). Xia et al. then analyzed using ultra-performance liquid chromatography coupled with tandem MS and interpreted using the Kyoto Encyclopedia of Genes and Genome pathway analysis (Xia et al. 2022). In addition to revealing the core

microbiota of each mare, the results demonstrated 59 differential metabolites and 21 differential pathways between the white and black mares (Xia et al. 2022). Xia et al. indicate that further investigation of the microbial metabolic sets would help to better understand the complexities of koumiss and its fermentation process (2022).

Advantages:

A major advantage of metabolomics is that novel pathways with key physiological roles, can be discovered, improving biological knowledge and databases as a whole (Vernocchi et al. 2016). This is because molecular concentrations of metabolites can reveal biochemical functions and processes that are undetectable by other methodologies, filling in the gaps from other traditional "omics" studies and proving groundbreaking for the field of human health (Rochfort 2005; Hollywood 2006; Long 2011). In addition, this approach addresses cellular expression ambiguities from conditional influences and allows researchers to follow phenotypic alterations (Long 2011).

Disadvantages:

A major challenge of metabolomics is the ability to identify and quantify a community's entire set of metabolites, given that time-of-flight MS is preferred when coupling a chromatography-based separation technique, and when using TOF the molecular mass must be lower than 1,000 daltons (Tang 2011; Fiehn 2016). Furthermore, the products of metabolic reactions are both variable and dynamic, which makes simultaneously determining the set of metabolites complicated, given that their compositions and concentrations change depending on environmental conditions (Hollywood 2006; Tang 2011). The results may not directly correlate with a metagenomic analysis performed in tandem (Tang 2011).

Stable Isotope Probing for tracking microbial metabolism through biomarker analysis.

Method:

Stable isotope probing (SIP) links microbial metabolics to metagenomic and phylogenetic information, by allowing researchers to track isotopically labelled biomarkers (Neufeld et al. 2007; Dunford and Neufeld 2010; Uhlik et al. 2012). These biomarkers, associated with communities that assimilate the labelled substrate, are phylogenetically informative, allowing researchers to assign diverse communities specific metabolic functions (Neufeld et al. 2007; Dunford and Neufeld 2010). First, the stable isotope, whether it be ^{13}C, ^{15}N, ^{18}O, or ^{2}H, must be incorporated into the community's cellular components, which is done by way of incubating the community with the label (Uhlik et al. 2012). Next, DNA is isolated and purified directly from the community via cell lysis (Uhlik et al. 2012). The metagenome is then separated based on density ultracentrifugation and retrieved through syringe extraction or fractionation, resulting in unlabeled DNA and stable isotope-labelled DNA (Dunford and Neufeld 2010; Uhlik et al. 2012). Subsequent DNA sequencing can guide additional community analysis approaches (Dunford and Neufeld 2010).

Information Obtained:

SIP allows researchers to identify microbial populations with a defined function (Uhlik et al. 2012). Therefore, SIP is an applicable tool for the identification of active members of a microbial community and associated genes integral to their function. Identification is performed under conditions nearing those *in situ,* without the prerequisite of cultivation (Dunford and Neufeld 2010). This is done by fingerprinting or denaturing gel electrophoresis of the gradient fraction, formed by ultracentrifugation (Dunford and Neufeld 2010).

Example:

The fate of soil organic matter (SOM), such as carbon, is determined by soil microorganisms (Barnett et al. 2021). In 2021, Barnett et al. used multiple DNA-stable isotope probes as substrates to track bacterial assimilation of carbon from different sources, allowing researchers to measure diverse microbial contributions to SOM processing (Barnett et al. 2021). Nine carbon sources were employed in soil microcosms and sampled at 1, 3, 6, 14, 30, and 48 days to measure carbon assimilation dynamics, such as carbon mineralization rates and

cumulative carbon mineralized (Barnett et al. 2021). It was found that uncultivated bacteria are integral to soil carbon cycling, the bioavailability of carbon sources affects phylogenetic conservation, assimilation dynamics define guild structure, and there is a high turnover of bacterial biomass in relation to carbon cycling (Barnett et al. 2021). In total, Barnett et al. determined that 1,286 bacteria were found to incorporate carbon and that their life-history strategies can be linked to carbon assimilation dynamics (Barnett et al. 2021).

Advantages:

SIP is unparalleled in its phylogenetic identification of microbes performing a specific process as well as their ecologically relevant metabolic genes (Dunford and Neufeld 2010; Uhlik et al. 2012). This approach also circumvents the difficulty of detecting rare genes, as it targets subpopulations likely to contain them, ensuring that members of the microbial community found in low abundance are not overlooked, especially when they are integral to a functional process (Uhlik et al. 2012).

Disadvantages:

A major criticism of SIP is the often high concentrations of labelled substrates introduced during the incubation period (Dunford and Neufeld 2010; Uhlik et al. 2012). Although these concentrations may be necessary to achieve adequate yields of labelled DNA for analysis, it is important to use concentrations of the labelled substrate close to that of *in situ* (Dunford and Neufeld 2010; Uhlik et al. 2012). This is because inappropriate concentrations may cause population shifts due to competition (Dunford and Neufeld 2010; Uhlik et al. 2012). An additional challenge of SIP is the requirement to combine a variety of techniques from multiple disciplines; this is an obstacle for some researchers, especially those new to the field of microbial ecology (Uhlik et al. 2012).

Reverse-transcriptase quantitative PCR for monitoring microbial gene expression

Method:

Reverse-transcriptase quantitative PCR (RT-qPCR) is a sensitive tool that enables the detection and quantification of specific microbial transcripts in a community that contains innumerable like-transcripts (Bustin and Nolan 2004). Similar to other transcriptomics methods, such as whole-transcriptome shotgun sequencing and RNA-Seq, RT-qPCR can also be used on its own or as a supplement to genomics (Bashiardes et al. 2016). Again, like the two other metatranscriptomics methods, genetic information is sampled, and RNA is extracted and precipitated (Moran 2009). Again, RNA samples are treated with DNase and reverse transcription is used to convert RNA to cDNA (Moran 2009). The cDNA is then used as a template for sequence detection through amplification with specific primers, targeting conserved regions within genes of interest (Moran 2009). During each cycle of qPCR, the quantity of DNA is measured in real-time by fluorescence-based imaging (Azenta 2022). An RT-qPCR approach is ideal for the rapid quantification of specific target sequences (Azenta 2022). Much can be learned by tracking the expression of marker genes, which provide qualitative and quantitative information about the transcript data when using an RT-qPCR methodology (Moran 2009).

Information Obtained:

Transcriptomics provides researchers with knowledge of gene expression, elucidating which proteins are necessary under various environmental conditions and cellular states (Moran 2009). Transcriptome RT-qPCR methodology renders scientists with qualitative and quantitative knowledge of specific mRNAs encoded by an organism's DNA, or a submetatranscriptome—a subset of all mRNAs encoded by a sample of a community's DNA (Moran 2009). This is through the threshold cycle (C_t) of qPCR, the cycle in which the amplification fluorescence signal exceeded the background fluorescence (Azenta 2022).

Example:

Non-invasive clinical diagnosis of bladder cancer (BCa) is no easy feat (Dubois et al. 2021). In 2021, Dubois et al. designed a set of RNA marker candidates from urine RNA collected from 47 healthy donors and 66 BCa patients and tested them in human bladder cancer cell lines. The markers proved as promising potential tumour markers (Dubois et al. 2021). As

low concentrations of RNA are found in urine, RT-qPCR was a natural assay choice. Transcripts of the TRAM-1 gene locus represented the most promising marker candidate and demonstrated an impressive correlation with the health status of the BCa patients; linear TRAM-1 species consistently showed a minor under-expression in BCa patients, whereas circular TRAM-1 transcripts showed a clear reduction of expression (Dubois et al. 2021). S100A6 was an additional marker candidate that performed well and demonstrated an inverse change of gene expression, like TRAM-1 (Dubois et al. 2021). The work of Dubois et al. strongly indicates that urine RNA markers should be further explored as a tool for non-invasively diagnosing bladder cancer (2021).

Advantages:

Transcriptomics using RT-qPCR provides incredible target-transcript specificity, which is the main advantage of this methodology (Moran 2009; Azenta 2022). This method is highly sensitive and can also, uniquely, be used as a reproducible detection assay (Moran 2009; Azenta 2022). The C_t value for a reaction correlates with the amount of input sequence, and RT-qPCR has a defined limit of detection (Azenta 2022). Like other transcriptomics methods, RT-qPCR is amenable to experiments in which gene expression is monitored while an environmental parameter is manipulated (Moran 2009).

Disadvantages:

There are a few major disadvantages to using RT-qPCR. The first, inherent disadvantage of RT-qPCR is that specific primers must target conserved regions within genes of interest, thus introducing a bias by the need to anticipate genes (Moran 2009). Therefore, by convention, RT-qPCR has no discovery power, given that primer design is limited to known sequences (Azenta 2022). An additional disadvantage of this method is that separate primers are required for different targets (Bustin and Nolan 2004). Unless multiplexing RT-qPCR is being performed, which includes technical challenges of its own such as competition for reagents and potential for unwanted interactions between primers, it is not possible to amplify additional targets at later stages (Henegariu et al. 1997; Bustin and Nolan 2004). This is an issue if the RNA sample is limited (Henegariu et al. 1997; Bustin and Nolan 2004).

Fluorescence *in situ* hybridization (FISH) microscopy.

Method:

Fluorescence *in situ* hybridization (FISH) is a technique used for the identification, localization, and isolation of microbial taxa in their natural context and can be used in tandem with confocal laser scanning microscopy (CLSM) for excellent spatial resolution (Amann et al. 1995; Wagner et al. 2003; Amann and Fuchs 2008; Schimak et al. 2015). FISH provides reliable quantification of populations within complex microbial samples, using the hybridization of fluorescently labelled probes to the rRNA of microbes (Amann and Fuchs 2008; Schimak et al. 2015). The majority of FISH probes target 16S rRNA and are most often between 15 and 30 nucleotides in length, with a fluorescent dye molecule linked to their 5' end (Amann et al. 2001). Typical FISH protocols include fixing and permeabilizing a sample, hybridizing the sample with the oligonucleotide probe, washing the sample to remove the excess probe, and visualizing the labelled sample (Amann et al. 2001). Samples are visualized via epifluorescence microscopy or by fluorescence-activated cell sorting, a specialized type of flow cytometry (Amann et al. 1995; Schimak et al. 2015). As FISH allows for the detection of single cells, it is a suitable tool for phylogenetic studies in microbial ecology (Amann et al. 1995).

Information Obtained:

Through the interpretation of a fluorescently labelled sample using microscopy or flow cytometry data, researchers gain knowledge of microbial identities within a community and thus can infer information about the community's structure (Wagner et al. 2003; Schimak et al. 2015). Additionally, if CLSM is used it will elucidate the spatial rearrangement of microorganisms within the community (Wagner et al. 2003).

Example:

Mediastinitis is a major complication that can occur after cardiac surgery (Spindler et al. 2021). Often, superficial swabbing for mediastinitis does not detect pathogens that are localized in deep tissue, putting patients at risk (Spindler et al. 2021). In 2021, Spindler et al. sampled tissue from 12 patients suffering from deep sternal wound infection (DSWI), intending to map the spatial organization of pathogens and biofilms to suggest alternative DSWI treatments. On average, three wound edge samples were taken per wound and submitted to FISH using a

pan-bacterial probe, as well as a microbiological culture (Spindler et al. 2021). The frequency of DSWI and the time to the occurrence as well as previous operative interventions, complications, and individual risk factors, were documented (Spindler et al. 2021). It was found that microbiological cultures grew for five patients, whereas FISH gave additional diagnostic information in five cases, and confirmed culture results in seven cases (Spindler et al. 2021). This study revealed that FISH confirmed DSWI culture results, and supplemented diagnostic information, in some cases even revealing information that was missed by culturing (Spindler et al. 2021).

Advantages:

A major advantage of FISH is that it can identify and provide information about the spatial resolution and morphology of key microorganisms in a complex microbial community sample, serving as a method that bridges microscopy, microbial culture, biochemical identification and molecular diagnostics (Frickmann et al. 2017). Another advantage of FISH is the rRNA probes (Amann et al. 2001). They are ideal, given rRNA's presence in all living cells, their stability and high copy numbers, and their variable and conserved sequence domains (Amann et al. 2001).

Disadvantages:

Although FISH has its advantages, it also has a few limitations. In terms of typical protocols, not all bacteria and archaea are permeabilized when using standard fixation protocols, and when targeting uncultured prokaryotes, determining the hybridization and washing conditions proves difficult (Wagner et al. 2003). Additionally, quantification—which is done by manual microscopic counting—has proved inaccurate for densely colonized biofilms, which could be the most abundant life form of microorganisms (Wagner et al. 2003). Poor quantification also occurs when studying bacterial cells associated with animal or plant tissues, such as chitin, due to the high number of autofluorescent components (Schimak et al. 2015).

BioOrthogonal Non-Canonical Amino Acid Tagging (BONCAT).

Method:

Cells respond to environmental cues by synthesizing proteins. Previously, identifying new proteins was hindered by all proteins sharing the same pool of amino acids, making them chemically indistinguishable (Dieterich et al. 2006). Bioorthogonal noncanonical amino acid tagging (BONCAT), a method developed for the study of *in situ* microbial activity, allows for the selective identification of newly synthesized proteins (Dieterich et al. 2006; Hatzenpichler and Orphan 2015). A technology based on the cotranslational *in vivo* incorporation of azide groups into newly synthesized proteins, BONCAT is effective in labelling the proteomes of taxonomically and physiologically distinct phyla under changing conditions (Dieterich et al. 2006; Hatzenpichler and Orphan 2015). In a typical BONCAT experiment, a bioorthogonal amino acid, such as L-azidohomoalanine (AHA) or L-homopropargylglycine (HPG), is added to a sample before incubating the solution (Dieterich et al. 2006; Hatzenpichler and Orphan 2015). If the noncanonical amino acid is taken up by a cell and incorporated into a protein, conclusions can be drawn about a cell's activity when fluorescent detection via azide-alkyne click chemistry is performed (Dieterich et al. 2006; Hatzenpichler and Orphan 2015; Saleh et al. 2019). The incorporation of artificial amino acids is unbiased, non-toxic, and does not cause the degradation of proteins (Dieterich et al. 2006; Hatzenpichler and Orphan 2015).

Information Obtained:

BONCAT can label the proteomes of phylogenetically and metabolically diverse cultures without causing changes to protein expression or degradation levels (Hatzenpichler and Orphan 2015). Translational activity, exemplified by the results of AHA and HPG labelling, can be correlated with the ^{15}N-ammonia of specific cells which acts as a proxy for microbial growth (Hatzenpichler and Orphan 2015). In addition, protocols combining BONCAT with fluorescence *in situ* hybridization (FISH) allow researchers to link the identity of a cell with its translational activity using only fluorescence microscopy (Hatzenpichler and Orphan 2015).

Example:

Studies of cystic fibrosis (CF) microbiota may provide insights into the contribution of specific taxa to CF pathogenesis, resulting in an understanding of the polymicrobial basis of the

disease (Valentini et al. 2020). In 2020, Valentini et al. hypothesized that only a subset of microbiota are translationally active in CF patients and tested this hypothesis using BONCAT. To do so, sputum was collected from patients and labelled with AHA before azide-alkyne click chemistry was performed, and the samples were visualized using fluorescence microscopy (Valentini et al. 2020). Samples were also analyzed with fluorescence-activated cell sorting, then sequenced to assign taxonomy (Valentini et al. 2020). It was demonstrated that active bacterial cells vary 6-56% between subjects and that there are heterogenous growth rates in sputum (Valentini et al. 2020). This research may improve future therapies, as it has added to our understanding of CF (Valentini et al. 2020).

Advantages:

BONCAT is useful for monitoring activity at the level of a single cell in complex microbial communities (Dieterich et al. 2006; Steward et al. 2020). It is a convenient molecular approach for analyzing microbial community function, as it avoids the use of radioactive substrates, like [^{35}S]cysteine and [^{35}S]methionine (Dieterich et al. 2006; Steward et al. 2020). BONCAT is also understood to minimally impact protein structure while revealing the selective labelling of newly synthesized proteins and allowing for the immediate validation of candidate proteins without prior knowledge about them (Dieterich et al. 2006; Steward et al. 2020).

Disadvantages:

The primary disadvantage of BONCAT is that post-click chemistry-based enrichment, it is assumed that all proteins are metabolically tagged (Liu et al. 2022). This may not be the case, as there is no step to confirm this assumption, thus altering our understanding of the subproteome of newly synthesized proteins (Liu et al. 2022). Furthermore, on rare occasions, there may be high background fluorescence when visualizing the sample that can be attributed to non-removable click dye; however, this issue can be resolved by using a lower dye concentration or by increasing the number of washing steps (Hatzenpichler and Orphan 2015).

References

Amann, R. I., Ludwig, W., and Schleifer, K. H. 1995. Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. Microbiol. Rev. **59**(1): 143-169. doi:/10.1128/mr.59.1.143-169.1995.

Amann, R. I., Fuchs B. M., and Behrens, S. 2001. The identification of microorganisms by fluorescence *in situ* hybridisation. Curr. Opin. Biotechnol. **12**(3): 231-236. doi:10.1016/S0958-1669(00)00204-4.

Amann, R. I., and Fuchs, B. M. 2008. Single-cell identification in microbial communities by improved fluorescence *in situ* hybridization techniques. Nat. Rev. Microbiol. **6**: 339–348. https://doi:/10.1038/nrmicro1888

Andersen, C., Bergholt, B., Ridderberg, W., and Nørskov-Lauritsen, N. 2022. Culture on selective media and amplicon-based sequencing of 16S rRNA from spontaneous brain abscess—the view from the diagnostic laboratory. Microbiol. Spectr. **10**(2). doi:10.1128/spectrum.02407-21.

Azenta. 2022. NGS, qPCR, or Sanger Sequencing: an assay selection guide [Online]. Available from https://www.azenta.com/blog/ngs-qpcr-or-sanger-sequencing-assay-selection-guide [Accessed November 2022].

Barnett, S. E., Youngblut, N. D., Koechli, C. N., and Buckley, D. H. 2021. Multisubstrate DNA stable isotope probing reveals guild structure of bacteria that mediate soil carbon cycling. PNAS. Nexus. **118**(47): e211529118. doi:10.1073/pnas.2115292118.

Bashiardes, S., Zilberman-Schapira, G., and Elinav, E. 2015. Use of metatranscriptomics in microbiome research. Bioinform. Biol. Insights. **10**: 19-25. doi:10.4137/BBI.S34610.

Bustin, S. A., and Nolan, T. 2004. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. J. Biomol. Tech. **15**(3): 155-166. Available from https://pubmed.ncbi.nlm.nih.gov/15331581/ [Accessed November 2022]. PMID:15331581.

Clarridge, J. E. 2004. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. Clin. Microbiol. Rev. **17**(4), 840–862. doi:10.1128/CMR.17.4.840-862.2004.

Danko, D., Bezdan, D., Afshin, E. E., Ahsanuddin, S., Bhattacharya, C., Butler, D. J., Chng, K. R., Donnellan, D., Hecht, J., Jackson, K., Kuchin, K., Karasikov, M., Lyons, A., Mak, L.,

Meleshko, D., Mustafa, H., Mutai, B., Neches, R. Y., and Zubenko, S. 2021. A global metagenomic map of urban microbiomes and antimicrobial resistance. Cell, **184**(13): 3376-3393. doi:10.1016/j.cell.2021.05.002.

Dieterich, D. C., Link, A. J., Graumann, J., Tirell, D. A., and Schuman, E. M. 2006. Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). PNAS. Nexus. **103**(25): 9482-9487. doi:10.1073/pnas.0601637103.

Dubois, J., Rueger, J., Haubold, B., Kretschmer-Kazemi Far, R., and Sczakiel, G. 2021. Transcriptome analyses of urine RNA reveal tumor markers for human bladder cancer: validated amplicons for RT-qPCR-based detection. Oncotarget, **12**(10): 1011-1023. doi:10.18632/oncotarget.27954.

Dunford, E. A., and Neufeld, J. D. 2009. DNA stable-isotope probing (DNA-SIP). J. Vis. Exp. **42**: e2027. doi:10.3791/2027 PMID:20729803

Fiehn Lab. 2016. Mass resolution and resolving power [Online]. Available from https://fiehnlab.ucdavis.edu/projects/seven-golden-rules/mass-resolution [Accessed November 2022].

Frickmann, H., Zautner, A. E., Moter, A., Kikhney, J., Hagen, R. M., Stender, H., and Poppert, S. 2017. Fluorescence *in situ* hybridization (FISH) in the microbiological diagnostic routine laboratory: a review. Crit. Rev. Microbiol. **43**(3): 263-293. doi:10.3109/1040841X.2016.1169990. PMID:28129707.

Gupta, S., Mortensen, M. S., Schjørring, S., Trivedi, U., Vestergaard, G., Stokholm, J., Bisgaard, H., Krogfelt, K. A., and Sørensen, J. S. 2019. Amplicon sequencing provides more accurate microbiome information in healthy children compared to culturing. Commun. Biol. **2**: 291. doi:10.1038/s42003-019-0540-1.

Han, Y., Zhang, M., Chen, X., Zhai, W., Tan, E., and Tang, K. 2022. Transcriptomic evidences for microbial carbon and nitrogen cycles in the deoxygenated seawaters of Bohai Sea. Environ. Int. **158**: e106889. doi:10.1016/j.envint.2021.106889.

Handelsman, J. 2004. Metagenomics: application of genomics to uncultured microorganisms. Microbiol. Mol. Biol. Rev. **68**(4): 669-685. doi:10.1128/MMBR.68.4.669-685.2004.

Hatzenpichler, R., and Orphan, J. 2015. Detection of protein-synthesizing microorganisms in the environment via bioorthogonal noncanonical amino acid tagging (BONCAT) [Online]. Springer Handbook Protocols: Hydrocarbon and Lipid Microbiology Protocols. Available from http://nebula.wsimg.com/b046306c6f40dd35b12d1819a48c7ce9?AccessKeyId=48B9AC89493A56C67DF6&disposition=0&alloworigin=1 [Accessed November 2022].

Henegariu, O., Heerema, N. A., Dlouhy, S. R., Vance, G. H., and Vogt, P. H. 1997. Multiplex PCR: critical parameters and step-by-step protocol. Biotechniques, **23**(3): 504-11. doi:10.2144/97233rr01. PMID:9298224.

Hollywood, K., Brison, D., and Goodacre, R. 2006. Metabolomics: current technologies and future trends. Proteomics. **6**(17): 4718-4723. doi:10.1002/pmic.200600106.

Hugenholtz, P., Goebel, B. M., and Pace, N. R. 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J. Bacteriol. **180**(18): 4765-4774. doi:10.1128/JB.180.18.4765-4774.

Hugerth, L. W., and Andersson, A. F. 2016. Analysing microbial community composition through amplicon sequencing: from sampling to hypothesis testing. Front. Microbiol. **8**: 1561. doi:10.3389/fmicb.2017.01561.

Illumina. (n.d.). Study microbial transcriptomes and metatranscriptomes with next-generation RNA-Seq [Online]. Available from https://www.illumina.com/areas-of-interest/microbiology/microbial-sequencing-methods/microbial-transcriptomics.html [Accessed November 2022].

Janda, J. M., and Abbott, S. L. 2007. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. J. Clin. Microbiol., **45**(9), 2761-2764. doi:10.1128/JCM.01228-07.

Konopka, A. 2009. What is microbial community ecology? ISME. J. **3**:1223–1230. doi:10.1038/ismej.2009.88.

Lee, A. 2019 Metagenomic next generation sequencing: how does it work and is it coming to your clinical microbiology lab? [Online]. Available from: https://asm.org/Articles/2019/November/Metagenomic-Next-Generation-Sequencing-How-Does-It#:~:text=Metagenomic%20NGS%20(mNGS)%20is%20simply,present%20and%20in%20what%20proportions [Accessed November 2022].

Liu, C., Wong, N., Watanabe, E., Hou, W., Biral, L., DeCastro, J., Mehdipour, M., Aran, K., Conboy, M. J., and Conboy, I. M. 2022. Mechanisms and Minimization of False Discovery of Metabolic Bioorthogonal Noncanonical Amino Acid Proteomics. Rejuvenation Res. **25**(2): 95-109. doi:10.1089/rej.2022.0019.

Long, S., Yang, Y., Shen, C., Wang, Y., Deng, A., Qin, Q., and Qiao, L. 2020. Metaproteomics characterizes human gut microbiome function in colorectal cancer. NPJ. Biofilms. Microbiomes. **6**: 14. doi:10.1038/s41522-020-0123-4.

Moran, M. A. 2009. Metatranscriptomics: eavesdropping on complex microbial communities. Microbe Wash. DC. **4**(7): 329-335. doi:10.1128/MICROBE.4.329.1.

Mukherjee, A., and Reddy, M. S. 2020. Metatranscriptomics: an approach for retrieving novel eukaryotic genes from polluted and related environments. 3 Biotech, **10**(2). doi:10.1007/s13205-020-2057-1. PMID:32030340.

Neufeld, J. D., Dument, M. G., Vohra, J., and Murrel J. C. 2006. Methodological considerations for the use of stable isotope probing in microbial ecology. Microbial. Ecol. **53**: 435-442. doi:10.1007/s00248-006-9125-x.

Neufeld, J. D., Vohra, J., Dumont, M. G., Lueders, T., Manefield, M., Friedrich, M.W., and Murrell, J.C. 2007. DNA stable-isotope probing. Nat Protoc. **2**(4):860-6. doi:10.1038/nprot.2007.109. PMID:17446886.

Rochfort, S. 2005. Metabolomics reviewed: a new "omics" platform technology for systems biology and implications for natural products research. J. Nat. Prod. **68**(12): 1813-1820. doi:10.1021/np050255w.

Saleh, A. M., Jacobson, K. R., Kinzer-Ursem, T. L., and Calve, S. 2019. Dynamics of non-canonical amino acid-labeled intra- and extracellular proteins in the developing mouse. Cell. Mol. Bioeng. **12**(5): 495-509. doi:10.1007/s12195-019-00592-1. PMID:31719929.

Schimak, M. P., Kleiner, M., Wetzel, S., Liebeke, M., Dubilier, N., Fuchs, B. M., and Liu, S.-J. 2015. MiL-FISH: multilabeled oligonucleotides for fluorescence *in situ* hybridization improve visualization of bacterial cells. Microbiology, **82**(1): 62-70. doi:10.1128/AEM.02776-15.

Shakya, M., Lo, C., and Chain, S. G. 2018. Advances and challenges in metatranscriptomic analysis. Front. Genet. **10**. doi:10.3389/fgene.2019.00904. PMID:31608125.

Spindler, N., Moter, A., Wiessner, A., Gradistanac, T., Borger, M., Rodloff, A. C., Langer, S., and Kikhney, J. 2021. Fluorescence *in situ* hybridization (FISH) in the microbiological diagnostic of deep sternal wound infection (DSWI). Infect. Drug. Resist. **14**: 2309-2319. doi:10.2147/IDR.S310139.

Staley, J. T., and Konopka, A. 1985. Measurement of *in situ* activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. Annu. Rev. Microbiol. **39**, 321–346. doi:10.1146/annurev.mi.39.100185.001541.

Steward, K. F., Eilers, B., Tripet, B., Fuchs, A., Dorle, M., Rawle, R., Soriano, B., Balasubramanian, N., Copié, V., Bothner, B., and Hatzenpichler, R. 2020. Metabolic implications of using bioorthogonal non-canonical amino acid tagging (BONCAT) for tracking protein synthesis. Front. Microbiol. **11**: 197. doi:10.3389/fmicb.2020.00197

Straub, D., Blackwell, N., Peltzer, A., Nahnsen, S., and Kleindienst, S. 2019. Interpretations of environmental microbial community studies are biased by the selected 16S rRNA amplicon sequencing pipeline. Front. Microbiol. **11**: e550420. doi:10.3389/fmicb.2020.550420.

Strazzulli, A., Fusco, S., Cobucci-Ponzano, B., Moracci, M., and Contursi, P. 2017. Metagenomics of microbial and viral life in terrestrial geothermal environments. Rev. Environ. Sci. Biotechnol. **16**: 425–454. doi:10.1007/s11157-017-9435-0.

Tang, J. 2011. Microbial Metabolomics. Curr. Genomics. **12**(6): 391-403. doi:10.2174/138920211797248619. PMID:22379393.

Uhlik, O., Leewis, C., Strejcek, M., Musilova, L., Mackova, M., Leigh, M. B., and Macek, T. 2012. Stable isotope probing in the metagenomics era: a bridge towards improved bioremediation. Biotechnol. Adv. **31**(2): 154. doi:10.1016/j.biotechadv.2012.09.003. PMID:23022353

Valentini, T.D., Lucas, S. K., Binder, K. A., Cameron, L. C., Motl, J. A., Dunitz, J. M., and Hunter, R. C. 2020. Bioorthogonal non-canonical amino acid tagging reveals translationally active subpopulations of the cystic fibrosis lung microbiota. Nat. Commun. **11**: 2287. doi:10.1038/s41467-020-16163-2. PMID:32385294.

Vernocchi, P., Del Chierico, F., and Putignani, L. 2015. Gut microbiota profiling: metabolomics based approach to unravel compounds affecting human health. Front. Microbiol. **7**: 1144. doi:10.3389/fmicb.2016.01144.

Wagner, M., Horn, M., and Daims, H. 2003. Fluorescence *in situ* hybridisation for the identification and characterisation of prokaryotes. Curr. Opin. Microbiol. **6**(3): 302-309. doi:10.1016/S1369-5274(03)00054-7.

Wang, Z., Gerstein, M., and Snyder, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. **10**(1): 57. doi:10.1038/nrg2484. PMID:19015660.

Wang, D.-Z., Kong, L.-F., Li, Y.-Y., and Xie, Z.-X. 2016. Environmental microbial community proteomics: status, challenges and perspectives. Int. J. of Mol. Sci. **17**(8): 1275. doi:10.3390/ijms17081275. PMID:27527164.

Weinstock, G. M. 2012. Genomic approaches to studying the human microbiota. Nature, **489**(7415): 250. doi:10.1038/nature11553. PMID:22972298.

Wilmes, P., and Bond, P. L. 2006. Metaproteomics: studying functional gene expression in microbial ecosystems. Curr. Trends Microbiol. **14**(2): 92-97. doi:10.1016/j.tim.2005.12.006.

Wilmes, P., Buschart, A. H., and Bond, P. L. 2015. A decade of metaproteomics: Where we stand and what the future holds. Proteomics, **15**(20): 3409-3417. doi:10.1002/pmic.201500183. PMID:26315987.

Xia, Y., Oyunsuren, E., Yang, Y., and Shuang, Q. 2022. Comparative metabolomics and microbial communities associated network analysis of black and white horse- sourced koumiss. Food Chem. **370**: e130996. doi:10.1016/j.foodchem.2021.130996.

Abstracts

Microbial community composition & diversity via amplicon sequencing:

Culture on Selective Media and Amplicon-Based Sequencing of 16S rRNA from Spontaneous Brain Abscess—the View from the Diagnostic Laboratory (April 11, 2022)

Forty-one stored samples from cases of spontaneous brain abscess were investigated to gain insight into the natural history, causative agents, and relevant laboratory diagnostics of a rare infection. Samples from a larger collection were selected based on retrospective analysis of patient records. All samples were subjected to amplicon sequencing of 16S rRNA gene fragments. Supplementary culture on selected media was performed as suggested by bioinformatics analysis. For three cases, no microorganism was disclosed, while *Toxoplasma gondii*, *Aspergillus fumigatus*, and various bacteria were the cause of 1, 2, and 35 cases, respectively. Bacterial infections were monomicrobial in 20 cases and polymicrobial in 15; the microorganisms of the latter cases were restricted to residents of cavum oris. Amplicon sequencing did not further enhance the importance of the *Streptococcus anginosus* group, which was involved in 17 cases, and the single primer set used may be suboptimal for amplification of *Actinomyces* and *Nocardia*. But, amplicon-based sequencing unquestionably expanded the number of polybacterial infections, with focus on the *Fusobacterium nucleatum* group, *Parvimonas*, and *Porphyromonas*. Culture on selective media confirmed the presence of *F. nucleatum* group bacteria, which attained a prominence in spontaneous brain abscess similar to the *S. anginosus* group. Metagenomics is a powerful tool to disclose the spectrum of agents in polymicrobial infections, but a reliable cutoff value for substantial detection is complex. Commercial media for isolation of *F. nucleatum* group bacteria from mixed infections are available, and these pathogens should be carefully characterized. Isolation of *Parvimonas* and *Porphyromonas* in polymicrobial infections has not been resolved.

Metagenomics for microbial or viral communities (NOT 16S and 18S rRNA sequencing).
A global metagenomic map of urban microbiomes and antimicrobial resistance (June 24, 2021)

We present a global atlas of 4,728 metagenomic samples from mass-transit systems in 60 cities over 3 years, representing the first systematic, worldwide catalog of the urban microbial ecosystem. This atlas provides an annotated, geospatial profile of microbial strains, functional characteristics, antimicrobial resistance (AMR) markers, and genetic elements, including 10,928 viruses, 1,302 bacteria, 2 archaea, and 838,532 CRISPR arrays not found in reference databases. We identified 4,246 known species of urban microorganisms and a consistent set of 31 species found in 97% of samples that were distinct from human commensal organisms. Profiles of AMR genes varied widely in type and density across cities. Cities showed distinct microbial taxonomic signatures that were driven by climate and geographic differences. These results constitute a high-resolution global metagenomic atlas that enables discovery of organisms and genes, highlights potential public health and forensic applications, and provides a culture-independent view of AMR burden in cities.

Microbial community transcriptomics (whole-transcriptome shotgun sequencing; RNA-Seq).

Transcriptomic evidences for microbial carbon and nitrogen cycles in the deoxygenated seawaters of Bohai Sea (October 4, 2021)

Eutrophication-induced water deoxygenation occurs continually in coastal oceans, and alters community structure, metabolic processes, and the energy shunt, resulting in a major threat to the ecological environment. Seasonal deoxygenation events have occurred in the Bohai Sea (China), however, how these affect the functional activity of microorganisms remains unclear. Here, through the use of absolute quantification of 16S rRNA genes amplicon sequencing and metatranscriptomics approaches, we investigated the structure of the microbial community and the patterns of transcriptional activity in deoxygenated seawaters. The dominant phyla were Proteobacteria (average value, 1.4 × 106 copies ml−1), Cyanobacteria (3.7 × 105 copies ml−1), Bacteroidetes (2.7 × 105 copies ml−1), and the ammonia-oxidizing archaea Thaumarchaeota (1.9 × 105 copies ml−1). Among the various environmental factors, dissolved oxygen, pH and temperature displayed the most significant correlation with microbial community composition and functional activity. Metatranscriptomic data showed high transcriptional activity of Thaumarchaeota in the deoxygenated waters, with a significant increase in the expression of core genes representing ammonia oxidation, ammonia transport, and carbon fixation (3-hydroxypropionic acid/4-hydroxybutyric acid cycle) pathways. The transcripts of Cyanobacteria involved in photosynthesis and carbon fixation (Calvin-Benson-Bassham cycle) significantly decreased in low oxygen waters. Meanwhile, the transcripts for the ribulose bisphosphate carboxylase-encoding gene shifted from being assigned to photoautotrophic to chemoautotrophic organisms in surface and bottom waters, respectively. Moreover, the transcription profile indicated that heterotrophs play a critical role in transforming low-molecular-weight dissolved organic nitrogen. Elevated abundances of transcripts related to microbial antioxidant activity corresponded to an enhanced aerobic metabolism of Thaumarchaeota in the low oxygen seawater. In general, our transcriptional evidences showed a population increase of Thaumarchaeota, especially the coastal ecotype of ammonia oxidizers, in low oxygen aquatic environments, and indicated an enhanced contribution of chemolithoautotrophic carbon fixation to carbon flow.

Microbial community proteomics

Metaproteomics characterizes human gut microbiome function in colorectal cancer (March 24, 2020)

Pathogenesis of colorectal cancer (CRC) is associated with alterations in gut microbiome. Previous studies have focused on the changes of taxonomic abundances by metagenomics. Variations of the function of intestinal bacteria in CRC patients compared to healthy crowds remain largely unknown. Here we collected fecal samples from CRC patients and healthy volunteers and characterized their microbiome using quantitative metaproteomic method. We have identified and quantified 91,902 peptides, 30,062 gut microbial protein groups, and 195 genera of microbes. Among the proteins, 341 were found significantly different in abundance between the CRC patients and the healthy volunteers. Microbial proteins related to iron intake/transport; oxidative stress; and DNA replication, recombination, and repair were significantly alternated in abundance as a result of high local concentration of iron and high oxidative stress in the large intestine of CRC patients. Our study shows that metaproteomics can provide functional information on intestinal microflora that is of great value for pathogenesis research, and can help guide clinical diagnosis in the future.

Microbial community metabolomics

Comparative metabolomics and microbial communities associated network analysis of black and white horse- sourced koumiss (February 15, 2022)

The quality and formation of bioactive components in fermented koumiss are based on the complex metabolism of the microbial community. In the present study, changes in the bioactive metabolites and microbial communities in black and white horse-sourced koumiss were evaluated during the fermentation process. 74 and 69 differential metabolites were formed when BLM and WHM were fermentated into koumiss. Lactobacillus and Dekkera grew rapidly and became the dominant genera in the koumiss. Bioactive compounds (e.g., adenine, d-proline) were significantly enhanced after natural fermentation and were positively correlated with Lactobacillus, Dekkera and Acetobacter. The microbial metabolic network showed that Lactobacillus and Dekkera were the functional core microbiota and played significant roles in the formation of bioactive compounds, followed by Acetobacter, Streptococcus and Leuconostoc. The current study results provide new insight into the formation of bioactive components in koumiss, which is useful for directionally isolating functional microorganisms suitable for koumiss fermentation.

Stable Isotope Probing for tracking microbial metabolism through biomarker analysis.
Multisubstrate DNA stable isotope probing revealsguild structure of bacteria that mediate soilcarbon cycling (October 22, 2021)

Soil microorganisms determine the fate of soil organic matter(SOM), and their activities compose a major component of theglobal carbon (C) cycle. We employed a multisubstrate, DNA-stableisotope probing experiment to track bacterial assimilation of Cderived from distinct sources that varied in bioavailability. Thisapproach allowed us to measure microbial contributions to SOMprocessing by measuring the C assimilation dynamics of diversemicroorganisms as they interacted within soil. We identified andtracked 1,286 bacterial taxa that assimilated13C in an agriculturalsoil over a period of 48 d. Overall13C-assimilation dynamics of bac-terial taxa, defined by the source and timing of the13Ctheyassimi-lated, exhibited low phylogenetic conservation. We identifiedbacterial guilds composed of taxa that had similar13C assimilationdynamics. We show that C-source bioavailability explained signifi-cant variation in both C mineralization dynamics and guild struc-ture, and that the growth dynamics of bacterial guilds differedsignificantly in response to C addition. We also demonstrate thatthe guild structure explains significant variation in the biogeo-graphical distribution of bacteria at continental and global scales.These results suggest that an understanding of in situ growthdynamics is essential for understanding microbial contributions tosoil C cycling. We interpret thesefindings in the context of bacteriallife history strategies and their relationship to terrestrial C cycling.

Reverse-transcriptase quantitative PCR for monitoring microbial gene expression

Transcriptome analyses of urine RNA reveal tumor markers for human bladder cancer: validated amplicons for RT-qPCR-based detection (May 11, 2021)

Non-invasive clinical diagnostics of bladder cancer is feasible via a set of chemically distinct molecules including macromolecular tumor markers such as polypeptides and nucleic acids. In terms of tumor-related aberrant gene expression, RNA transcripts are the primary indicator of tumor-specific gene expression as for polypeptides and their metabolic products occur subsequently. Thus, in case of bladder cancer, urine RNA represents an early potentially useful diagnostic marker.

Here we describe a systematic deep transcriptome analysis of representative pools of urine RNA collected from healthy donors versus bladder cancer patients according to established SOPs. This analysis revealed RNA marker candidates reflecting coding sequences, non-coding sequences, and circular RNAs. Next, we designed and validated PCR amplicons for a set of novel marker candidates and tested them in human bladder cancer cell lines. We identified linear and circular transcripts of the S100 Calcium Binding Protein 6 (S100A6) and translocation associated membrane protein 1 (TRAM1) as highly promising potential tumor markers.

This work strongly suggests exploiting urine RNAs as diagnostic markers of bladder cancer and it suggests specific novel markers. Further, this study describes an entry into the tumor-biology of bladder cancer and the development of gene-targeted therapeutic drugs.

Fluorescence *in situ* hybridization (FISH) microscopy.

Fluorescence in situ Hybridization (FISH) in the Microbiological Diagnostic of Deep Sternal Wound Infection (DSWI) (June 21, 2021)

Purpose: Postoperative mediastinitis after cardiac surgery is still a devastating complication. Insufficient microbiological specimens obtained by superficial swabbing may only detect bacteria on the surface, but pathogens that are localized in the deep tissue may be missed. The aim of this study was to analyze deep sternal wound infection (DSWI) samples by conventional microbiological procedures and fluorescence in situ hybridization (FISH) in order to discuss a diagnostic benefit of the culture-independent methods and to map spatial organization of pathogens and microbial biofilms in the wounds.

Methods: Samples from 12 patients were collected and analyzed using classic microbiological culture and FISH in combination with molecular nucleic acid amplification techniques (FISHseq). Frequency of and the time to occurrence of a DSWI was recorded, previous operative interventions, complications, as well as individual risk factors and the microbiologic results were documented.

Results: Tissue samples were taken from 12 patients suffering from DSWI. Classical microbiological culture resulted in the growth of microorganisms in the specimens of five patients (42%), including bacteria and in one case Candida. FISHseq gave additional diagnostic information in five cases (41%) and confirmed culture results in seven cases (59%).

Conclusion: Microbial biofilms are not always present in DSWI wounds, but microorganisms are distributed in a “patchy” pattern in the tissue. Therefore, a deep excision of the wound has to be performed to control the infection. We recommend to analyze at least two wound samples from different locations by culture and in difficult to interpret cases, additional molecular biological analysis by FISHseq.

BioOrthogonal Non-Canonical Amino Acid Tagging (BONCAT).

Bioorthogonal non-canonical amino acid tagging reveals translationally active subpopulations of the cystic fibrosis lung microbiota (May 8, 2020)

Culture-independent studies of cystic fibrosis lung microbiota have provided few mechanistic insights into the polymicrobial basis of disease. Deciphering the specific contributions of individual taxa to CF pathogenesis requires comprehensive understanding of their ecophysiology at the site of infection. We hypothesize that only a subset of CF microbiota are translationally active and that these activities vary between subjects. Here, we apply bioorthogonal non-canonical amino acid tagging (BONCAT) to visualize and quantify bacterial translational activity in expectorated sputum. We report that the percentage of BONCAT-labeled (i.e. active) bacterial cells varies substantially between subjects (6-56%). We use fluorescence-activated cell sorting (FACS) and genomic sequencing to assign taxonomy to BONCAT-labeled cells. While many abundant taxa are indeed active, most bacterial species detected by conventional molecular profiling show a mixed population of both BONCAT-labeled and unlabeled cells, suggesting heterogeneous growth rates in sputum. Differentiating translationally active subpopulations adds to our evolving understanding of CF lung disease and may help guide antibiotic therapies targeting bacteria most likely to be susceptible.