Proteomics: A Revolution in the Evaluation of Gastrointestinal Microflora

Enjoy this white paper by David Quig, PhD.
Proteomics: A Revolution in the Evaluation of Gastrointestinal Microflora

Recent research regarding the gastrointestinal microbiome has irrefutably confirmed the fact that the microbial inhabitants of the gastrointestinal tract, and their astonishing scope of metabolic activities, are at the very core of health and numerous disease processes. It is also clear that clinical microbiology testing should be optimized to address the relative abundance of all bacterial species present in a stool specimen. To that end, culture remains the current standard of practice in high-complexity clinical microbiology testing. A revolutionary methodology has been developed and is employed at Doctor’s Data to markedly expand our ability to identify species and even subspecies of gastrointestinal bacteria and yeast. This article will summarize the pros and cons of current methodologies used to study gastrointestinal flora, with emphasis on the application of a proteomic-based method that optimizes the superior specificity of highly sensitive culture-based comprehensive stool analysis.

Culture-Based Testing

Culture growth-based testing remains the international standard for clinical microbiology testing for stool specimens. Culture enables identification of living, metabolically active microorganisms, as dead bacteria do not contribute to the metabolomic environment of the gut. Further, live, pure bacterial isolates are required for antibiotic (AB) susceptibility testing. Recent research from Boston University indicates that gastrointestinal virus-derived phages harbor numerous antibiotic AB resistance genes that can be transferred to bacteria. However, simply detecting one of the hundreds of AB resistance genes in a bulk stool specimen does not directly indicate that an AB resistance gene is actually present or expressed in a particular microbial species. This requires culture to obtain a pure isolate of the pathogenic species of concern. What’s more, AB resistance genes may be “silent,” or not expressed or activated by specific promoters, and more complicated and genetically diverse mechanisms of antimicrobial resistance remain elusive. The bottom line is that genotyping for functioning AB resistance genes has not been validated for stool specimens for clinical purposes.

Susceptibility to antibiotics must be directly performed on pure, cultured isolates using standardized microbiological techniques according to the Clinical Laboratory Standards Institute’s guidelines, and validated in the specific laboratory. Clinicians should not hesitate to ask the laboratory if such criteria are adhered to.

Figure 1. Growth of beneficial/expected bacteria from Cary-Blair transport media over time.
Misconceptions about Culture

Several exaggerations have been perpetuated regarding the viability and stability of bacterial populations during transport to the laboratory after specimen collection. The recommended transport media for stool specimens, Cary-Blair media, is buffered to prevent shifts in pH and has a very low nutrient content—not “loaded with sugar”—that suspends growth of stool bacteria and inhibits the growth of other bacteria. In addition, the transport media has a low agar content to enhance viability of pathogenic species.

Critics of culture have also suggested that anaerobic bacteria cannot survive sampling and transport to the laboratory. While it is true that some strict anaerobes are not yet able to be cultured, it should be noted that Cary-Blair media contains a reducing agent that markedly lowers the oxygen content and actually favors anaerobes. As shown in Figure 1, stool specimens were collected and held in the transport media for up to 14 days prior to culture under anaerobic conditions at Doctor’s Data, Inc. (DDI). It is clearly evident that all of the levels of beneficial and expected strict anaerobes and facultative anaerobes that DDI reports are stable, with respect to clinical significance, out to 10 days after collection. This is particularly impressive for the strict anaerobes (Bifidobacterium spp. and the Bacteroides fragilis group). Most specimens are received within 48 hours of specimen collection.

Understanding Clostridium Results

DDI reports Clostridium spp. under Expected/Beneficial bacteria because a normal gastrointestinal tract can contain a relatively high amount of the approximate 100 different species of Clostridium, while only 4 are overtly pathogenic. Although Clostridium are labile in an oxygenated environment, we are able to grow the strict anaerobes under our standard anaerobic culture conditions because all Clostridia produce virtually indestructible spores that become vegetative microorganisms when appropriately cultured. Even under standard anaerobic conditions, we occasionally see growth of Clostridium difficile. When this occurs, the isolates are reflexively tested at no charge for the PaLoc gene locus using an FDA-cleared gene amplification platform (LAMP). The PaLoc locus is required for production of toxins A and B, and all toxigenic C. difficile contain the PaLoc gene segment. C. difficile-associated disease occurs almost exclusively when beneficial bacteria are suppressed after exposure to broad-spectrum antibiotics, and only occurs when the bacteria contain the PaLoc gene locus. Therefore, one need not treat C. difficile that do not contain the PaLoc gene. For a complete and discriminating breakdown of Clostridia species, DDI offers a Comprehensive Clostridium Culture test that is optimized specifically for growth of Clostridium species.

Identification of Bacteria and Yeast from Pure Isolates

Identification by Phenotype

In the past, microorganisms were identified manually by phenotypic identification, a vast array of time-consuming biochemical reactions. Eventually, technological advances provided automated systems such as Vitek-2™ to perform phenotypic identification. Such automated identification instruments were considered to be state-of-the-art and markedly reduced manual labor, human error and judgment, and time to identification, and are currently used in most institutional and commercial clinical microbiology laboratories. They are limited, however, with respect to the number of microorganisms that can be confidently identified. If a microorganism is not identifiable with greater than 90% confidence, additional time-consuming off-line testing is required to obtain a computer-assisted, best-fit identification. These
methodologies for identification of microorganisms can take as long as 6 to 24 hours and prolong the time before beginning subsequent antimicrobial susceptibility testing. Despite these inherent limitations, automated instruments are still invaluable for susceptibility testing with commonly prescribed pharmaceutical antimicrobials.

**Identification by Proteomic Analysis**

Proteomics entails the identification and characterization of proteins in cells, tissues, or organisms. This allows identification based on each microorganism’s unique ribosomal protein fingerprint.

Doctor’s Data utilizes a revolutionary state-of-the-science proteomic methodology for rapid, accurate, and reliable identification of thousands of gastrointestinal bacterial and yeast species, and even subspecies. The methodology, matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS), has been recognized by a share of the 2002 Nobel Prize in Chemistry. MALDI-TOF MS has been used to study a host of macromolecules from a wide variety of biological matrices, just as similar methods of detection have been used by NASA on the Mars Curiosity Rover in an effort to determine whether the planet had been hospitable for microbial inhabitation.

MALDI-TOF MS has recently been applied to and established as an invaluable technique in clinical microbiology. In brief, MALDI-TOF MS enables rapid and accurate identification of normal and pathogenic gastrointestinal microorganisms based upon their signature high-abundance proteins.

There are 5 steps in the MALDI-TOF MS identification process: (1) ionization of proteins from a pure isolate; (2) separation of vaporized, ionized proteins that have different masses; (3) detection of the number of the different ionized proteins; (4) collection of data to generate the mass spectrum; and (5) comparison of the unknown spectrum to an extensive database.

In the laboratory, a small portion of a pure isolate from selective subculture is embedded within a saturated liquid matrix on a 96-spot target plate. The matrix solution crystallizes and charges the proteins (primarily +1), then the dried mixture of matrix and proteins is vaporized by nanosecond-duration laser pulses (MALDI) that release the ionized protein molecules that are accelerated in an electric field within a vacuum. Proteins fly through the mass analyzer (TOF) at a rate that is inversely proportional to the mass of the protein. The proteins reach the mass spectrometer over a span of to about 0.01 to 1 millisecond—the timescale to analyze the ions produced by a single laser pulse permits acquisition of data from multiple laser pulses for a specimen. The mass spectrum of the unknown pure isolate is compared to a high-quality reference database of spectra generated from quality-controlled collection strains. The database contains about 7,000 microorganisms and is growing at about 400 more per month, although not all of the microorganisms are derived from the gastrointestinal tract. A goodness of fit is ranked and a log score
threshold is applied for correct identification at the species level (2.300-3.000), secure genus/probable species (2.000-2.299), or possible genus (1.700-1.999). Inclusion in the database requires ≥10 independent reference spectra with a reference score ≥1.9 for a given bacteria or yeast. To date, using MALDI-TOF MS, DDI is able to identify more than 400 bacterial species and 125 yeast species that were not identifiable using the previously employed automated phenotypic instruments used by most clinical microbiology labs. In addition, we have validated secure identification of yeast at the species level at a ranking score of just 1.800, and have been able to identify some bacteria at the subspecies level. Amazingly the MALDI-TOF MS identification process takes less than 1 minute per isolate.

Based on work in her research lab and the plethora of published studies regarding MALDI-TOF MS, Dr. Sandra Richter, Director of Bacteriology in the departments of Clinical Pathology and Molecular Pathology at The Cleveland Clinic, stated that "It is clear that this is a giant step forward in the accuracy and speed of organism identification routinely reported by the lab." (CAP Today. 2013;27[1]).

**Consistent, Validated Results**

Accuracy and speed are important, but if a clinician sends a specimen to several labs will they get the same results? That is a significant issue when laboratories are not using standardized, validated test platforms. The results of an international, multicenter trial to evaluate the reproducibility of a commercially available MALDI-TOF MS platform, the Bruker Biotyper™, were presented at the 19th European Congress of Clinical Microbiology and Infectious Diseases in Helsinki in 2009. Sixty blind-coded bacterial specimens were sent to 8 international laboratories, for a total of 480 specimens. A concordance rate of 98.8% was achieved among the labs. Not only were the samples correctly identified across laboratories, but the reliability of identification was very high, as 97.3% of all log score values were >2. What’s more, the MALDI-TOF MS approach has an impressive capability to differentiate closely related species from a genus such as *Bacillus*. The genus *Bacillus* is one of the largest, most ubiquitous genera of bacteria, with about 65 valid species. The nonpathogenic strains have the highest potential to produce false positive results when analyzed by 16S rRNA sequence analysis due to the high degree of conservation among *Bacillus* species. Sixteen isolates of putative *B. pumilus* spores were isolated from various NASA spacecraft assembly facilities (10-100K clean rooms), and attempts to identify the species included a standard biochemical platform, MALDI-TOF MS, and 16S DNA sequencing. The Bruker Biotyper MALDI-TOF MS protein profiling platform was more accurate than the biochemical platform, more discriminating than the 16S rDNA sequence analysis platform, and complemented the results of genetic methodologies that have greater discriminatory power than 16S rDNA (gyrB sequence analysis and DNA-DNA hybridization). Considering accuracy, interlaboratory reproducibility, cost and the rapidity of identification, the MALDI-TOF MS proteomic approach is best suited for routine comprehensive stool analysis.

**Identification by PCR**

Polymerase chain reaction (PCR) is a relatively old biochemical technology that is used in molecular biology to exponentially amplify select segments of DNA and has been recognized by a share of the 1993 Nobel Prize in Chemistry. After amplification, several means of analysis have been applied in order to identify the preselected target segment of DNA, including real-time-detection PCR (RTD-PCR) and PCR-ELISA. RTD-PCR platforms are able to detect sequence-specific PCR products as they accumulate in real time during the PCR amplification process, and have proven utility for clinical microbiology use in hospital settings for rapid identification (<3 hours) of causative enteric bacteria associated with serious diarrheal diseases. The expeditious nature of the RTD-PCR platform is possible because the bacteria do not have to be cultured. Indeed, recent studies have demonstrated that a 3-hour RTD-PCR platform exhibited 92% (96/104) sensitivity, compared to standard culture and phenotypic identification of *Campylobacter, Salmonella, Yersinia* and *Shigella* species. The investigators acknowledged that, paired with reflexive culture testing, the RTD-PCR platform utilized provides improved care for patients with acute infectious diarrheal disease when time is of the essence. Similarly, RTD-PCR has been applied for the rapid detection
of other enteric pathogens such as *Clostridium difficile*, *Escherichia coli* (EHEC) and *Vibrio* species. Rinttilä and colleagues designed an extensive set of real-time PCR assays targeting a large group of predominant and pathogenic human gut microbial species. They demonstrated that their real-time PCR platform is very sensitive and precise for quantitative detection of human pathogens from fecal samples. Using fecal samples spiked with various amounts of target bacteria, they demonstrated that detection limits could be obtained that were between 6,000 CFU/g for *H. pylori* and 60,000 CFU/g for *Clostridium difficile* and *Campylobacter jejuni*. (J Appl Microbiol. 2004;97:166-77).

16S rDNA PCR is also useful for phylogenetic characterization, or community census, of very abundant microbial communities clustered into enterotypes that appear to be associated with health and specific diseases.

### Clinical Limitations of PCR

PCR-based platforms have significant limitations in the context of clinical medicine, particularly for complementary and integrative medicine (CAM) in a non-acute care setting. CAM-oriented clinicians are interested in global changes in the composition and quantity of gastrointestinal microbiota, even in the absence of enteropathogens. That is because such perturbations can be the cause of changes in colonic metabolism that can be associated with significant symptoms of dysbiosis and various diseases.

Culture has been criticized for the limited number of microorganisms that can be grown, compared to the vast number of microorganisms that have been detected using DNA methodologies. That criticism is used to argue that PCR is superior for evaluation of gastrointestinal microbiota. Of course, culturing all gastrointestinal microorganisms is indeed a daunting, if not impossible, task. However, to put things in perspective, it should be noted that when MALDI-TOF MS is coupled with high-complexity microbiology, Doctor’s Data has the capability to identify at least 1,240 bacteria and yeast organisms—and that does not include speciation of the many *Lactobacilli* or *Bifidobacteria* species. In addition, PCR requires a primer for identification of every genus, let alone species. Therefore, a reliable RTD-PCR based evaluation of microorganisms in a stool specimen that is even 1/10th as comprehensive as that of the MALDI-TOF MS culture approach would be absolutely cost prohibitive, if even possible. Changes in the balance among the prevalent beneficial, expected, and even commensal bacteria and their metabolic activities are as important as the absence of the relatively minute levels of pathogenic bacteria in many symptomatic patients. Some may recall Dr. Leo Galland’s discussions of deficiency dysbiosis. As a clinician, do you want to know as much as possible about the vast array of microorganisms that are present in your patients’ stool specimens, or do you want to simply inquire about a handful of microorganisms? In essence, PCR-based methods are better suited to address the question “Is it there?” rather than “What is there?” The DDI culture-based proteomic approach is analogous to casting a broad net, versus using a fishing pole to obtain a census of the types of fish that inhabit a pond.

The issues of sensitivity and specificity also need to be brought into light in consideration of methods for evaluating gastrointestinal microorganisms. The detection limits of commonly used methods are presented in Table 1.

<table>
<thead>
<tr>
<th>Testing Method</th>
<th>Detection Limit</th>
<th>Linear Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture</td>
<td>10 to 10² CFU/gm</td>
<td>10² to 10⁶ CFU/gm</td>
</tr>
<tr>
<td>Real-time PCR</td>
<td>10³ to 10⁴ CFU/gm</td>
<td>10⁶ to 10⁸ CFU/gm</td>
</tr>
<tr>
<td>PCR ELISA</td>
<td>10³ to 10⁵ CFU/gm</td>
<td>10⁴ to 10⁷ CFU/gm</td>
</tr>
<tr>
<td>PCR</td>
<td>10⁴ to 10⁵ CFU/gm</td>
<td>10⁵ to 10⁷ CFU/gm</td>
</tr>
</tbody>
</table>

CFU = colony forming units, measured per gram of stool

Table 1. Detection limits of various methods for detecting microorganisms. (Crit Care. 2000;4(4):255-61.)
Despite widespread misconception, culture can be 2 to 3 orders of magnitude more sensitive than even RTD-PCR. In fact, culture can reveal the presence of just 1 CFU/g for an enteric pathogen when an enrichment broth is included. PCR-based methods can be less sensitive in part due to a high noise-to-signal ratio obtained when rDNA primers try to find mates among a sea of DNA from a fecal specimen. Further, clinical specificity with respect to microbial identification can be problematic, because the probes used often cannot differentiate closely related microorganisms that have highly conserved sequences of 16S rDNA. PCR-based methods can also be associated with inferior clinical specificity due to use of inappropriate primers, as well as errors introduced during the amplification process.

**Summary**

Doctor’s Data has optimized identification of gastrointestinal microorganisms from cultured stool specimens using MALDI-TOF MS, an accurate, reliable proteomic technique that identifies bacteria and yeast by spectral analysis of microorganisms’ predominant ribosomal proteins. Real-time PCR platforms, albeit less sensitive than culture, have a place in acute-care situations regarding infectious diarrheal disease when time is of the essence. However the real-time PCR approach does not currently lend itself to comprehensive analysis of the vast array of gastrointestinal microorganisms. Next-generation full genetic sequencing techniques in the research setting are moving forward and will further advance our understanding of the role of the gastrointestinal microbiome in health and disease. Currently, culture-based MALDI-TOF MS technology is the most sensitive and discriminating way to identify the greatest number of beneficial, expected, commensal, and pathogenic gastrointestinal microorganisms present in your patients.

**References**


