Monitoring MICROBES

In the name of public health, researchers are using microarray technology to identify human pathogens.

BY RANDALL C. WILLIS

IT IS IRONIC that in this era of genomes and proteomes, the smallest organisms, the microbes, still hold humans hostage.

While outbreaks of ebola and epidemics of tuberculosis, malaria, and HIV ravage the developing world, people in the United States are wondering whether their next hamburger will carry *E. coli*, if medical tubing is coated with drug-resistant *Staphylococcus*, and if that unknown letter contains anthrax.

The expanding plethora of multidrug-resistant microbes has the medical community scratching their collective heads. There is grave concern that by the time an infectious pathogen has been identified and its drug-resistance profile determined, it might be too late for the patient. As with any other disease condition—including cancer and heart disease—early diagnosis can be the key that decides between life and death.

Detection and identification

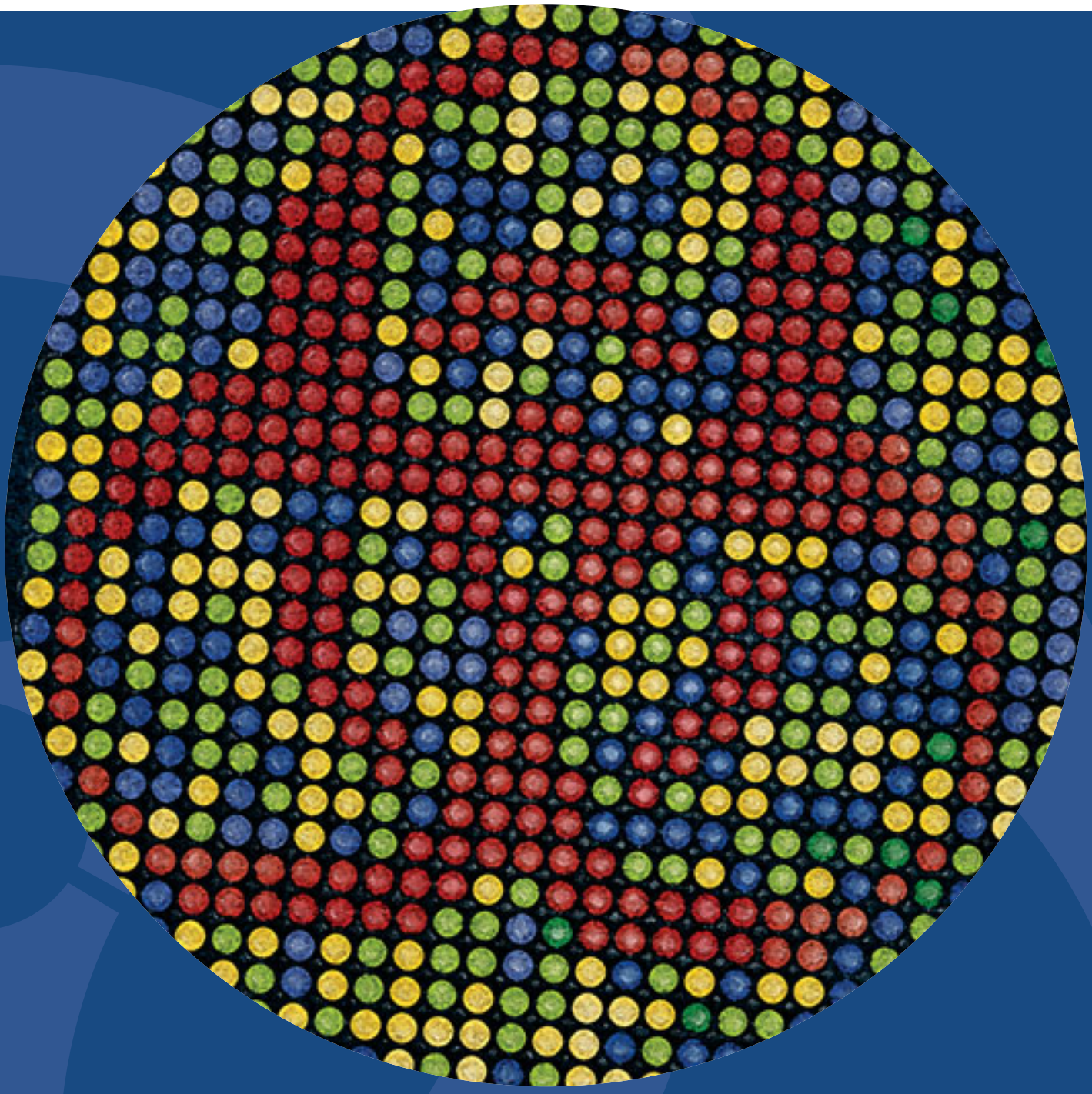
Traditionally, the clinical microbiologist has relied on a variety of culturing and staining methods to identify most bacterial species (*1*). The best known of these stains are the Gram stains. Similarly, microbes such as *Mycobacterium tuberculosis* can be stained and examined using fluorescence microscopy. Parasites can also be identified using a microscope, either directly (e.g., by identifying eggs or cysts) or through staining. Virus identification, however, requires a more indirect route. In this case, the clinician watches how various tissues react to viral infection, typically through cell death or the binding of red blood cells (hemadsorption).

Although these assays have stood the test of time, they are largely gross identifications and cannot narrow down the sample to exact species or strains of the infecting microbe. In addition, some of the assays require lengthy incubations, delaying infection diagnosis. To address this issue, several methods were developed that identify specific molecular features of the pathogen. For example, methods like the enzyme-linked immunosorbent assay (ELISA) rely on the specific interaction of antibodies directed against microbial surface antigens.

Although such assays can be useful in identifying microbial isolates, as well as being rapid and potentially high-throughput, their effectiveness is often compromised because many microbes have found ways of altering their surface antigenicity as a mechanism of avoiding the human immune system. Similarly, other tests that rely on measuring the titer of antibodies in the patient's bloodstream cannot necessarily distinguish between new and old infections. Because these assays rely on indirect or potentially variable factors, methods have been developed that identify microbes by a relatively unchanging factor: their genetic material.

Initial methods relied on the hybridization of small, labeled oligonucleotides to the DNA of infecting microorganisms. But because the titer of these organisms can be quite low, it was necessary to develop an assay that could amplify the incident signal. This was accomplished with the introduction of PCR-based assays. In this case, potentially infected samples are genetically amplified with oligonucleotides that hybridize to opposite ends of a fragment of DNA characteristic of the infecting organism. If a band of predetermined length is found on an agarose gel separation of the PCR reaction, then infection and the infecting species are identified. By varying the sequences of and labels on the oligonucleotides, multiple microbial species and strains can be tested in a single sample well. For example, using the Cepheid (Sunnyvale, CA) Smart Cycler, Canadian researchers recently developed an assay to distinguish between two *Candida* species in blood samples (*2*). This is especially important because *C. albicans* and *C. dubliniensis* are virtually indistinguishable using microbiological methods and require entirely different patient treatment regimens.

But even this method has its limits. The resolving capacity of



most gels is such that it is difficult to determine with any certainty whether one is seeing a band of the appropriate size. Similarly, point mutations within the sequence to which one or more of the oligonucleotides hybridize can lead to false negative readings, as a lack of a PCR band does not necessarily indicate a lack of infection. Finally, because the variation from drug-sensitive to drug-resistant can be as small as a single nucleotide change, even if an appropriate band is detected, the researcher might still need to sequence it to determine which specific isolate of the microbe is in the sample tube.

To address these problems, researchers have taken the PCR-based assays one step further, confirming their microbial identifications by hybridization of the PCR products to strain-specific oligonucleotides bound to a chip.

Microarrays

One side benefit of the various microbial genome projects has been the recent explosion of commercially available genome chips. Ostensibly targeted at the research community looking to determine gene expression profiles during the microbe life cycle, these microarrays can also serve as detection and identification starting points for the clinical setting as hospitals and clinicians begin to identify new methods. Among the chips commercially available are *Helicobacter pylori* arrays from MWG Biotech (*www.mwg-biotech.com*); *M. tuberculosis*, *Plasmodium falciparum*, and *C. albicans* chips from OPERON (*www.operon.com*); *E. coli* arrays from Pan Vera (*www.panvera.com*); *E. coli* and *Pseudomonas aeruginosa* arrays from Affymetrix (*www.affymetrix.com*); and *E. coli* and *M. tuberculosis* arrays from Sigma-Genosys (*www.sigma-genosys.com*). Similarly, companies such as Siebersdorf Research Gmbh (*www.arcs.ac.at/UL/ULB*) and Agilent (*www.chem.agilent.com*) will create custom-made chips, tailored to the organism of interest.

On the academic front, researchers at the St. George's Hospital

Medical School (London) have organized bacterial genome researchers around the world (with funding from the Wellcome Trust) to form a group called BμG@S (Bacterial Microarray Group at St. George's). The group was established to develop whole-genome arrays for 12 pathogens, including *Campylobacter jejuni* (a principal cause of food poisoning), *Haemophilus influenzae*, *Yersinia pestis* (bubonic plague), and *Salmonella typhi*.

The group also offers informatics support to researchers in the form of a database system called BμG@Sbase, which users can access through a Web interface and where they can store data in a searchable, secure manner (*3*).

As group leader Philip Butcher and colleagues explained in a summary of their goals (*4*), "This application will extend an existing microarray group . . . to establish a multicollaborative bacterial microarray facility for the purposes of making available DNA microarray technology to the microbial pathogen research community in the United Kingdom over a period of five years. The timeliness of making whole genome and whole species arrays widely and rapidly available in the postgenomic era is self-evident. Access to and training in high-throughput technology will address both fundamental and applied issues in pathogen biology, pathogenesis, virulence, molecular epidemiology, and public health."

But beyond strictly trying to identify whether the microbe of interest is in the medical sample, clinicians are interested in a more detailed analysis of what subspecies or variant they are going to have to treat. And this is where microarrays come to the fore, as indicated in the following examples of the identification and virulence determination of microbes (see also Table 1).

Virus identification

Recently, Vladimir Chizhikov and colleagues at the Center for Biologics Evaluation and Research (Kensington, MD) and the National Institute of Allergy and Infectious Diseases (Bethesda, MD) developed a microarray system to identify clinically relevant isolates of rotaviruses, the most important agents of severe diarrhea in infants and young children in the developing world (*5*). Traditionally, the classification of rotavirus serotypes (P and G) has been based on the neutralization of antibodies directed against two proteins of the viral outer capsid (VP4 and VP7), but recent efforts have focused on the subtle differences found within the nucleic acid sequences for the genes of these proteins.

Although other groups have relied on PCR-based identification methods, the Maryland researchers analyzed more than 150 VP7 gene sequences from GenBank to identify several genotype-specific oligonucleotides for use in an array format. Also, to ensure the accuracy of their identifications and prevent problems due to spontaneous mutations that often occur in clinical samples, the researchers identified each rotavirus genotype not on the basis of only one oligonucleotide match but rather on the average of nine different ones. Beyond simply confirming the identification results of earlier PCR analyses, the researchers also successfully identified samples that could not be determined by current PCR-based methods. The array-based method also allows for the rapid identification of novel mutations that might not otherwise be detected, an important factor in determining viral mutation rates and molecular evolution.

Bacterium identification

In collaboration with Advanced Array Technologies (Namur, Belgium, *www.aat-array.com*), Sandrine Hamels and colleagues at the Facultés Nôtre Dame de la Paix (Namur), St. Luc University Hospital, and Queen Astrid University Hospital (both in Brussels) developed StaphyChips, a low-density microarray designed to detect various *Staphylococcus* species (*6*). Specifically, they used PCR with degenerate primers to amplify the *femA* gene, a highly conserved staphylococcal peptidoglycan gene. These DNA products were then analyzed on the microarray bearing capture probes specific to the five staph species that are most closely linked to hospitalization-related infections—*S. aureus*, *S. epidermidis*, *S. haemolyticus*, *S. hominis*, and *S. saprophyticus*. In their controlled experiments, the researchers correctly identified each species with no cross-reactions.

Similarly, Jae-Chang Cho and James Tiedje of Michigan State

Table 1

Some recent applications of microarrays to pathogen identification.

Pathogen	Disease	Reference
Viruses		
HIV	AIDS	Hanna, G. L.; et al. *J. Clin. Microbiol.* **2000,** *38,* 2715–2721.
Influenza	Influenza	Li, J.; et al. *J. Clin. Microbiol.* **2001,** *39,* 696–704.
Bacteria		
Escherichia coli	Food poisoning	Chizhikov, V.; et al. *Appl. Environ. Microbiol.* **2001,** *67,* 3258–3263.
Helicobacter pylori	Ulcers	Bjorkholm, B.; et al. *Infect. Immun.* **2001,** *69,* 7832–7838.
Mycobacterium spp.	Tuberculosis	Troesch, A.; et al. *J. Clin. Microbiol.* **1999,** *37,* 49–55.
Salmonella enterica	Food poisoning	Garaizar, J.; et al. *J. Clin. Microbiol.* **2002,** *40,* 2074–2078.
Streptococcus pneumoniae	Strep throat	McCluskey, J.; et al. *Comp. Funct. Genom.* **2002,** *3,* 366–368.
Parasites		
Alexandrium cantenella	Food poisoning	Wilson, W. J.; et al. *Mol. Cell. Probes* **2002,** *16,* 119–127.
Plasmodium falciparum	Malaria	Hayward, R. E.; et al. *Mol. Microbiol.* **2000,** *35,* 6–14.

University (East Lansing) developed a microarray to distinguish between species of *Pseudomonas* bacteria (*7*). Unlike the previous example, however, the Michigan researchers attached PCR-amplified genomic fragments on the slides rather than small oligonucleotides. They then probed this array with labeled fragments of genomic DNA from various bacterial isolates and reference sequences and correctly identified each of the test microbes (Figure 1).

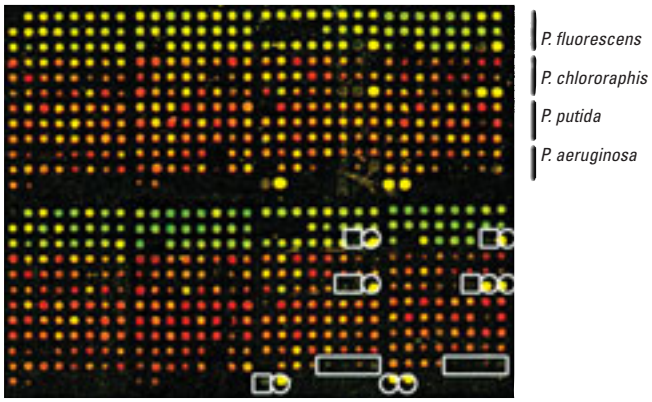


Figure 1. Spot the difference. Using PCR-amplified genomic fragments from various *Pseudomonas spp.*, researchers developed a microarray that can be used in the diagnosis of infection. (Reproduced with permission from Ref. 7.)

Parasite identification

Of course, not all pathogens to be studied will be found in clinical samples alone. In many cases, the microbe also contaminates soil or water. One such example is *Cryptosporidium parvum*, a waterborne parasite that can cause acute diarrhea in otherwise healthy individuals and can be fatal to immunocompromised individuals. Because the concentration of organisms in a water sample can be quite low, PCR has been the method of choice for the identification of *C. parvum*, but as stated earlier, this method can be error-prone and does not tell the full story.

Recently, Timothy Straub and colleagues at the Pacific Northwest National Laboratory (Richland, WA) and the Metropolitan Water District of Southern California (LaVerne) developed an oligonucleotide array for the detection of *Cryptosporidium spp.* that can differentiate between those genotypes known to infect humans and otherwise safe genotypes (*8*).

Specifically, the researchers synthesized a series of oligonucleotides that were homologous to regions of the highly conserved gene *hsp70*, in this case from *Cryptosporidium*, that carried SNPs characteristic of the different isolates. By generating a microarray carrying a series of redundant oligonucleotides that hybridized to seven *hsp70* SNP variants, the researchers ensured that they could derive a diagnostic signature of each species. In testing their system, however, they found that not all of the oligonucleotides bound to the appropriate isolate, indicating that there might be some ambiguity in the sequence information in GenBank, from which they derived their oligonucleotide sequences.

Field work

Each of these examples involved testing samples in the controlled environment of the lab. Although this is appropriate for possible microbial infection or contamination samples that will find their way to a hospital, the incidence of new cases of infection is increasing most dramatically outside the hospital setting, in rural areas of the developing world. To address this problem, Andrei Mirzabekov and colleagues at the Argonne National Laboratory (IL) and the Engelhardt Institute of Molecular Biology (Moscow) developed a portable system for microbial sample preparation and oligonucleotide microarray analysis (*9*).

In the portable system, bacteria are lysed and the extracts are loaded onto a syringe-operated silica microcolumn. After the column is washed, the bound nucleic acids are fragmented and labeled with fluorescent tags. Excess tag is washed away, and the labeled nucleic acids are eluted from the column. The tagged fragments are then hybridized to an array of 20-mer oligonucleotides (in this case, *Bacillis spp.* sequences), and unbound fragments are removed. The bound fluorescent tags are then detected with a CCD camera or a portable microchip imager that uses Polaroid film.

In developing their column-based system, the researchers eliminated steps such as lengthy nucleic acid fractionation methods, including centrifugation and electrophoresis, cutting the whole process down to 3–5 min from 40–60 min. Similarly, the fragmentation and labeling on the column eliminate the need for processes such as PCR. Finally, the portable detection system decreases the costs associated with imaging the microarray to ~$2000 from ~$100,000 for a high-end laser scanner.

The bottom line

Although they have yet to find widespread use in the clinical setting, microarrays are making broad strides toward being viewed as a valuable tool in the diagnosis of infection. Only through continued improvements in the methodology and the efforts of researchers like those involved in BμG@S will the benefits of this technology be realized.

References

(1) Clinical microbiology. In *Microbiology,* 2nd ed.; Prescott, L. M., Harley, J. P., Klein, D. A., Eds.; Wm. C. Brown Publishers: Dubuque, IA, 1993; pp 672–697.
(2) Trépanier, H.; et al. One-hour Detection of *Candida albicans* and *Candida dubliniensis* in Blood Samples Using the Smart Cycler. Presented at the 101st General Meeting of the American Society of Microbiology, Orlando, FL, May 2001; www.smartcycler.com/pdfs/candidas_poster.pdf.
(3) Witney, A. A.; Hinds, J. *Comp. Funct. Genom.* **2002,** *3,* 369–371.
(4) Butcher, P. D.; et al. A multi-collaborative microbial pathogen microarray facility at St. George's Hospital Medical School; www.sghms.ac.uk/depts/medmicro/bugs/BUGSMission/WT_Summary.htm.
(5) Chizhikov, V.; et al. *J. Clin. Microbiol.* **2002,** *40,* 2398–2407.
(6) Hamels, S.; et al. *Biotechniques* **2001,** *31,* 1364–1372.
(7) Cho, J.-C.; Tiedje, J. M. *Appl. Environ. Microbiol.* **2001,** *67,* 3677–3682.
(8) Straub, T. A.; et al. *Appl. Environ. Microbiol.* **2002,** *68,* 1817–1826.
(9) Bavykin, S. G.; et al. *Appl. Environ. Microbiol.* **2001,** *67,* 922–928.

Randall C. Willis is a senior associate editor of *Modern Drug Discovery*. Send your comments or questions about this article to mdd@acs.org or to the Editorial Office address on page 3. ■

KEY TERMS: cell biology, clinical, genomics, high throughput, imaging, informatics, proteomics, screening, technique