Review

Rapid sequence-based diagnosis of viral infection

Phenix-Lan Quan, Thomas Briese, Gustavo Palacios, W. Ian Lipkin*

Center for Infection and Immunity, Mailman School of Public Health, Columbia University, New York, NY 10032, United States

Abstract

With globalization of microbial threats and an increasing appreciation for the role of infection in chronic as well as acute diseases, there is burgeoning interest in the development of specific antiviral drugs. Less attention has been focused on the establishment and implementation of rapid viral diagnostic methods, without which it will not be possible to obtain the full benefit of new therapies. Here we review the current status of viral diagnostics and the utility of various sequence-based diagnostic platforms for applications in clinical microbiology, surveillance and pathogen discovery.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Viral diagnostics are becoming increasingly important in clinical medicine and public health. Factors in raising global concern with respect to acute viral diseases include burgeoning international travel and trade, political instability and bioterrorism, climate change and its effects on vector distribution, and the emergence and reemergence of zoonoses. The ability of viruses rapidly to expand their geographic range and appear in unexpected locations is well illustrated by the worldwide spread of the human immunodeficiency virus, the transfer of West Nile virus to the western hemisphere and its subsequent dissemination throughout North and South America, and the recent emergence of chikungunya virus in Europe (Rezza et al., 2007). Unexpected emergences have also occurred in the context of organ transplantation and immunosuppression (Morse, 1995; Fischer et al., 2006; Palacios et al., 2008). There is also a growing appreciation for a potential role for viruses as primary or co-factors in chronic cardiovascular, endocrine, neurodevelopmental and neoplastic disorders (Chang et al., 2007; Perez-Velez et al., 2007). Whereas the absence of effective therapies once made the diagnosis of viral infection primarily an academic exercise, the expanding armamentarium of countermeasures tailored to specific viruses, including small molecules, RNAi, therapeutic antibodies and vaccines, affords new opportunities to significantly reduce morbidity, mortality and health care costs due to viral infections.

The establishment and implementation of platforms that enable early differential diagnosis of infectious diseases will be a key to appreciate the full potency of antiviral therapeutics. Clinical manifestations of infection with individual viral agents are not distinctive, particularly in the early phases of disease. Indeed, even the most seasoned clinicians can find it challenging to differentiate...
between viral and other causes of pneumonia, diarrhea, meningocencephalitis, or hemorrhagic fever. A recent example from our own experience that brought home this point was the workup of samples from an individual who succumbed to acute febrile illness and multi-organ failure during an outbreak of Marburg hemorrhagic fever. Given the clinical presentation and context, the victim was presumed to have viral hemorrhagic fever; however, efforts to find Marburg virus sequences failed and panmicrobiological microarray analysis ultimately implicated *Plasmodium falciparum* (Palacios et al., 2007). Effective antimalarial therapy might have been instituted had the appropriate diagnosis been established at an earlier time point. Accuracy in diagnosis is critical as treatment decisions must frequently balance potential benefit with potential toxicity from antiviral therapy. When toxicity is minimal, but supplies are limited, it is nonetheless imperative that a drug be reserved for those cases for which it will be effective. This issue recently rose to international prominence with the focus on pandemic influenza and recognition that oseltamivir stockpiles might be insufficient; it has been reinforced with the recognition that some influenza strains may be resistant to oseltamivir (Hayden, 2004; Abed et al., 2008).

Early diagnosis and treatment may be critical to the efficacy of antiviral therapy. It is easier to control viral dissemination within a single host and to reduce the potential for transmission to others when viral burden is low. In hepatitis C, for example, early interferon treatment can prevent persistence (Jaekel et al., 2001). This reduces the risk of cirrhosis in the infected individual as well as the risk that she/he will infect others. Another example in which morbidity and mortality are determined by the timing of antiviral intervention is herpes encephalitis (Raschilas et al., 2002). In this instance, rapid suppression of virus replication is particularly important to clinical outcome because the targeted organ, the brain, has only limited potential for repair. Indeed, classic work by Whitley et al. in 1977 demonstrating the efficacy of adenine arabinoside in herpes encephalitis was pivotal in securing support for antiviral research and herpesvirus polymerase chain reaction (PCR) of cerebrospinal fluid has become standard operating procedure in diagnostic evaluation of central nervous system infections (Whitley et al., 1977). Early intervention may also be critical in viral infections in which host responses contribute to pathogenesis. For example, clinical trials of ribavirin revealed no effect when drug was administered after the onset of hantavirus pulmonary syndrome (Chapman et al., 1999; Mertz et al., 2004). However, efforts to find Marburg virus sequences failed and panmicrobiological microarray analysis ultimately implicated *Plasmodium falciparum* (Palacios et al., 2007). Effective antimalarial therapy might have been instituted had the appropriate diagnosis been established at an earlier time point. Accuracy in diagnosis is critical as treatment decisions must frequently balance potential benefit with potential toxicity from antiviral therapy. When toxicity is minimal, but supplies are limited, it is nonetheless imperative that a drug be reserved for those cases for which it will be effective. This issue recently rose to international prominence with the focus on pandemic influenza and recognition that oseltamivir stockpiles might be insufficient; it has been reinforced with the recognition that some influenza strains may be resistant to oseltamivir (Hayden, 2004; Abed et al., 2008).

Multiplex assays can alleviate this problem by allowing investigators to entertain many hypotheses simultaneously. The number of candidates considered is platform-dependent and can range from less than ten with multiplex PCR, to thousands with microarrays, to the entire tree of life with high-throughput pyrosequencing. Although costs and ease of use are improving, only multiplex PCR assays are widely used at the time of writing. In multiplex assays, many genetic targets compete for assay components (e.g., nucleotides, polymerases and dyes), in some instances with variable efficiency. Thus, current multiplex assays are not quantitative, and they also tend to be less sensitive than singleplex assays. However, there is optimism that both challenges will be addressed as platforms evolve.

### 3. Singleplex assays

The most common singleplex assays employed in clinical microbiology and microbial surveillance are polymerase chain reaction assays wherein fluorescent signal is detected as DNA strand replication results in either cleavage or release of a labeled oligonucleotide probe bound to sequence between the forward and reverse primer. Equipment needs are modest, i.e., a thermal cycler, fluorescence reader and laptop computer; thus, fluorescent reporter-dye singleplex assays can be reliably pursued under field conditions, with battery power if necessary. Loop-mediated isothermal amplification (LAMP) is an alternative to PCR that does not require programmable thermal cyclers (Notomi et al., 2000; Hagiwara et al., 2007; Shirato et al., 2007). Although sensitivity is reported to be similar to PCR, LAMP is not quantitative. Products are typically detected in ethidium bromide-stained agarose gels; however, changes in turbidity of the amplification solution may be sufficient; indeed, assays are described where the accumulation of product can be detected by eye (Jayawardena et al., 2007). Nested PCR, where two amplification reactions are pursued sequentially with either one (hemi-nested) or two (fully nested) primers located 3′ with respect to the original primer set, may be more sensitive than fluorescent reporter-dye singleplex assays; however, the potential for contamination is higher because the original reactions must be opened to add reagents for the second, nested, reaction (Casas et al., 1997; Templeton et al., 2004). Thus, nested PCR is challenging even in laboratories with scrupulous experimental hygiene.

### 4. Multiplex PCR assays

Multiplex assays are more difficult to establish because primer sets may differ in optimal reaction conditions (e.g., annealing temperature and magnesium concentration). Additionally, complex primer mixtures are more likely to result in primer–primer interactions that reduce assay sensitivity and/or specificity. To enable
multiplex primer design we developed Greene SC Primer, a software program that automates consensus primer design over a multiple sequence alignment, and allows users to specify primer length, melting temperature and degree of degeneracy (Jabado et al., 2006).

Gel-based multiplex assays, wherein products are distinguished by size, can detect as many as 8–10 distinct targets, albeit with low sensitivity (Casas et al., 1997; Templeton et al., 2004). Fluorescence reporter-based multiplex assays are more sensitive but are limited by the number of fluorescent emission peaks that can be unequivocally separated. At present up to four fluorescent reporter dyes are detected simultaneously. "Sloppy Molecular Beacons" can circumvent this limitation in part by binding to related targets at different melting temperatures (Saunders and Jeffries, 2000); however, as they cannot detect targets that differ by more than a few nucleotides, the improvement does not approach the multiplex capacity of gel-based assays.

Two multiplex platforms have been developed that combine PCR and mass spectroscopy (MS) for sensitive detection of several targets simultaneously. One of them, triangulation identification for genetic evaluation of risks (TIGER) uses matrix-assisted laser desorption/ionization (MALDI) MS to directly measure the molecular weights of PCR products obtained in an experimental sample and to compare them with a database of known or predicted product weights (Van Ert et al., 2004; Hofstadler et al., 2005; Sampath et al., 2007). The other, MassTag PCR, uses atmospheric pressure chemical ionization (APCI) MS to read molecular weight reporter tags attached to PCR primers (Briese et al., 2005; Palacios et al., 2006).

Each system has its strengths. Whereas MALDI-MS (TIGER) is confined to specialized laboratories, APCI MS (MassTag PCR) can be performed on smaller, less expensive, portable instruments. TIGER has the potential to directly indicate candidates for novel variants of known organisms via a divergent product weight but like MassTag PCR, it requires sequencing to characterize a novel sequence or agent.

At the time of writing, syndrome-specific MassTag PCR panels are in use that enable rapid differentiation of viruses, bacteria, fungi and parasites associated with acute respiratory disease, diarrheas, purulent diseases, encephalitis/meningitis and hemorrhagic fevers (Briese et al., 2005; Lamson et al., 2006; Palacios et al., 2006). Employment of these panels to investigate influenza-like illness (ILI) in New York state by MassTag PCR resolved one-third of previously negative samples, and revealed the presence of rhinoviruses in a large proportion of samples, approximately half of which belonged to a previously uncharacterized genetic clade. Studies in Australia, Europe, Asia, Africa and the United States revealed that this novel genetic clade plays a major role not only in ILI but also in asthma and pediatric pneumonia (Arden et al., 2006; Kistler et al., 2007; Lau et al., 2007; Lee et al., 2007; McErlean et al., 2007; Renwick et al., 2007). This and other work confirms the power of multiplex PCR systems for rapidly querying samples for the presence of a wide range of candidate viral and bacterial pathogens that may act alone or in concert. Other multiplex PCR systems have been established that employ flow cytometry to detect amplification products bound to matching oligonucleotides on fluorescent beads (Brunstein and Thomas, 2006; Han et al., 2006; Liet al., 2007). Sensitivity is similar across these platforms. Whether any one of them will become dominant remains to be seen.

5. Viral microarrays

Viral microarrays can be coarsely divided into those that address 10–100 agents and those designed for the detection of thousands of agents including unknowns. Arrays designed to address a limited number of agents, for example respiratory virus resequencing arrays, may employ multiplex consensus PCR (cPCR) to amplify specific genetic targets (Wong et al., 2004; Chiu et al., 2006, 2007; Lin et al., 2007). Sample preparation is not complex and sensitivity in clinical materials is typically 10³ to 10⁴ RNA copies. Although these arrays, which typically employ probes less than 25 nt, may allow speciation of agents they do not truly exploit the utility of microarrays for unbiased microbe detection. Oligonucleotide microarrays can comprise hundreds of thousands of features. Probes of up to 70 nt are not uncommon; thus, unlike PCR, or resequencing arrays, where short primer sequences demand precise complementarity between probe and target, such arrays are less likely to be confounded by minor sequence mismatches. This can be a considerable advantage for the detection of rapidly evolving targets, such as RNA viruses. Additionally, one can incorporate both microbial and host gene targets in high-density arrays. This affords an opportunity both to detect microbes and to assess host responses for signatures consistent with various classes of infectious agents.

The two most familiar larger scope platforms are the Virochip and the GreeneChip (Wang et al., 2003; Palacios et al., 2007). They differ in methods used for probe design, array manufacture, hybridization conditions and bioinformatics analysis of hybridization results. What is common to both is that sample nucleic acids are randomly amplified and labeled. This unbiased amplification is critical to exploiting the full range of probes representing tens of thousands of viral targets. However, because host and microbe sequences compete with similar efficiency for PCR reagents, sensitivity for microbial detection is commonly low (10⁹ to 10¹⁰ copies). It is not surprising therefore, that the most successful applications of this technology have employed samples containing low levels of host nucleic acids, such as cell culture supernatants or clinical samples such as serum, cerebrospinal fluid or urine. Improvements in sensitivity (10³ to 10⁴ copies) have been achieved with more challenging tissue samples by depleting host cell DNA and ribosomal RNA (rRNA) prior to amplification and labeling, and by employing as reporters labeled dendrimers that bind to hybridized sequences to enhance overall signal intensity. Array platforms are described in which hybridization is detected as changes in electrical conductance. These have not yet extended to high-density microarrays but have the potential to further enable improvements in sensitivity.

Viral arrays can facilitate cloning and sequence analysis as well as pathogen identification. Hybridized products typically range from 200 nt to >1000 nt. Because arrays display probes representing several different genomic regions for each virus, one can rapidly recover sequence not only for the hybridized products but also for sequences between those products through use of PCR. The method is simple. The hybridized products are eluted with hot water, and re-amplified using the specific sequence portion of the primers originally employed for random amplification.

6. High-throughput pyrosequencing

The advent of high-throughput sequencing technology affords unique opportunities for pathogen discovery. Unlike cPCR or array methods in which investigators are limited by known sequence information and must make choices regarding the range of pathogens to consider in a given experiment, high-throughput sequencing is unbiased and allows an opportunity to consider the entire tree of life: bacteria, viruses, fungi and parasites. Several systems are in development. We have experience with the pyrosequencing system of 454 Life Sciences; however, the principles for sample preparation and data analysis are broadly applicable across platforms (Margulies et al., 2005). As in viral microarray analyses, elimination of host nucleic acid is critical to boosting pathogen signal toward the threshold for detection. We use the same approach as employed for microarrays. This obviates the potential for detect-
ing DNA genomes of pathogens; however, our reasoning is that an active infection should be associated with transcription. After amplification and sequencing, reads typically range in size from 40 base pairs to 400 base pairs. Raw sequence reads are trimmed to eliminate sequence data.

7. A staged strategy for pathogen detection and discovery

We view the platforms described here as complementary tools for viral diagnostics (Fig. 1). Due to ease of use, sensitivity, capacity for quantitation and low cost, singleplex PCR is ideal in instances where the questions relate to the presence of a single agent or viral burden. Examples include (i) outbreaks of acute infectious disease where decisions are made concerning patient containment or allocation of specific interventions that are in short supply or potentially toxic, or (ii) adjustments of antiviral regimens as in HIV infection. More exploratory, multiplex assays are indicated when singleplex PCR fails to identify an agent, or the clinical presentation and context are not pathognomonic of infection with a single agent. The most time- and cost-effective alternatives for such second-stage analyses are multiplex PCR platforms. Where they fail or the list of candidate agents exceeds 25–30, microarrays are indicated.

Successful diagnosis using multiplex PCR or microarray platforms requires that a pathogen be related to one already known. When agents are novel or sufficiently distant in sequence to related microbes to confound hybridization, it may be necessary to resort to subtractive cloning or high-throughput unbiased sequencing. Multiplex PCR, microarray, subtractive cloning and high-throughput sequencing as used here serve as methods for detection rather than quantitation or detailed characterization. Thus, irrespective of the route that results in identification of a pathogen candidate, subsequent steps include quantitation of pathogen burden in affected hosts and controls, detailed characterization of the pathogen for features that may contribute to virulence or provide clues to provenance, and serology as an index to acute infection and a tool to examine prevalence of infection over time and geography.

8. The future of viral diagnostics

Detection technologies will continue to evolve, allowing faster, more sensitive and less expensive methods for pathogen surveillance and discovery. Multiplex PCR assays are already widely implemented, but microarray technology is less advanced. Improvements now in progress include microfluidic sample processing and direct measurement of conductance changes associated with hybridization. Although we have not addressed proteomics and host-response profiling, it is conceivable that biomarkers could be found that are specific for classes of infectious agents or that provide insights that can guide patient management. Active collaboration between clinicians and laboratorians is the key to success in diagnostics. The most advanced technology will fail if samples are collected without attention to nucleic acid lability, and data will be uninterpretable without accurate information on clinical course and sample provenance. In chronic diseases, in which complex mechanisms such as early exposure and/or genetic susceptibility may contribute to pathogenesis, the most substantive advances in linking viruses to disease are likely to come from investments in prospective serial sample collections and an appreciation that many conditions reflect intersections of genes and the environment in a temporal context.

Acknowledgements

The Center for Infection and Immunity is supported by NIH awards AI070411, HL83850, NS047537 and U54AI5758 (Northeast Biodefense Center-Lipkin).

References


Chiu, C.Y., Alizadeh, A.A., Rouskin, S., Merker, J.D., Yeh, E., Yagi, S., Schnurr, D., Patterson, B.K., Ganem, D., DeRisi, J.L., 2007. Diagnosis of a critical respiratory illness...