Harnessing CRISPR Effectors for Infectious Disease Diagnostics

Roby P. Bhattacharyya,†§ Sri Gowtham Thakku,†§ and Deborah T. Hung*,†∥⊥,‡

†Infectious Disease and Microbiome Program, Broad Institute of Harvard and MIT, 415 Main Street, Cambridge, Massachusetts 02142, United States
‡Division of Infectious Diseases and Department of Medicine, Massachusetts General Hospital, 55 Fruit Street, Boston, Massachusetts 02114, United States
§Program in Health Sciences and Technology, Harvard Medical School and Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, United States
∥Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114, United States
⊥Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115, United States

ABSTRACT: Nucleic acid detection is an important method for pathogen identification but can be expensive, have variable sensitivity and specificity, and require substantial infrastructure. Two new methods capitalize on unexpected in vitro properties of clustered regularly interspaced short palindromic repeats (CRISPR) effectors, turning activated nucleases into intrinsic amplifiers of a specific nucleic-acid binding event. These effectors are coupled with a variety of reporters and used in tandem with existing isothermal amplification methods to produce sensitive, sequence-specific pathogen identification in multiple field-deployable formats. While still in their infancy, these modular CRISPR-based methods have the potential to transform pathogen identification and other aspects of infectious disease diagnostics.

Nucleic acid amplification and detection has come to play a central role in modern infectious disease diagnostics, enhancing sensitivity, specificity, and breadth. Initially, molecular methods played a critical role in expanding the scope of detectable pathogens to those that are not amenable to traditional culture-based methods. More recently, their potential to obviate the complex infrastructure required for culture and to expedite diagnosis has greatly broadened interest in such methods in both the developed and developing worlds. Considerable progress has been made to try to meet the technological requirements of these assays, such as the need and capacity for thermal cycling and the cold chain required for enzymes and nucleic acids for PCR. Despite this progress, however, current approaches clearly do not yet function as a panacea for infectious disease diagnostics. The ideal pathogen detection technology platform—accurate, sensitive, cheap, portable, and functional across a range of pathogens and clinical settings, from inpatient to outpatient and from the developed to the developing world—remains elusive. New approaches are very much still needed to achieve this ambitious goal. Recently, new methods capitalizing on the clustered regularly interspaced short palindromic repeats (CRISPR) effectors, also known as CRISPR-associated (Cas) proteins, have been reported to have intriguing potential to impact the infectious disease diagnostic field.

CRISPR effectors represent a diverse group of bacterial and archaeal proteins that appear to have evolved to play a role in bacterial adaptive immunity. Central to their role in immune defense is the ability to recognize specific foreign nucleic acid sequences (such as those from invading viruses) and to subsequently neutralize them through cleavage. Unlike restriction endonucleases, target recognition by Cas nucleases is mediated in a modular way by base pairing to a nucleic acid “guide.” This modularity has allowed Cas proteins to be reprogrammed to target nearly any desired sequence, simply through design of a guide sequence that hybridizes with the desired target.

One class of CRISPR effectors, defined by single, multi-domain proteins, has formed the core of a remarkable array of biotechnological innovations in recent years, led by gene editing using Cas9 as a novel therapeutic approach. The exquisite specificity of Cas9 has also been applied to diagnostics with the demonstration of its ability to enable detection of single nucleotide polymorphisms (SNPs) to enhance the resolution of viral diagnosis. More recently, newer additions to the CRISPR technology toolbox, the Cas12 and Cas13 protein families, have enabled alternative strategies for diagnostics. These protein families display unexpected properties that have allowed two groups to demonstrate sensitive and specific nucleic acid detection modules that can be used for a wide range of diagnostic applications, including for infectious diseases. These new methods are termed SHERLOCK (“specific high-sensitivity enzynzymatic reporter unlocking”), which utilizes a protein from the Cas13 family (Cas13a), and DETECTR (“DNA endonuclease targeted CRISPR trans reporter”), which utilizes a protein from the Cas12 family (Cas12a). Cas12a and Cas13a have “collateral” cleavage activity, which is triggered by a specific binding event in vitro and serves as an intrinsic amplifier for the initial binding event. This facilitates sensitive and specific detection of the nucleic acid sequences.

Received: July 12, 2018
TWO NOVEL DETECTION MODULES BASED ON CRISPR-ASSOCIATED NUCLEASES

The molecular machinery behind the two detection modules, SHERLOCK and DETECTR, are conceptually similar (Figure 1), though they utilize different effectors, Cas13a and Cas12a.

![Diagram of molecular machinery](image)

Figure 1. Schematic depiction of a CRISPR effector nuclease exploited for nucleic acid detection in either SHERLOCK (Cas13a) or DETECTR (Cas12a). In the absence of its nucleic acid target, the Cas nuclease is inactive. Upon binding of its guide crRNA to a cognate target (RNA for Cas13a, ssDNA or dsDNA for Cas12a), the nuclease is activated, leading to not only target cleavage but also catalytic cleavage of off-target nucleic acids (RNA for Cas13a, ssDNA for Cas12a). This “collateral” nuclease activity is turned into an amplified signal by providing reporter probes with a fluorophore (orange) linked to a quencher (black) by a short oligonucleotide (gray). Upon cleavage of the reporter by the activated nuclease, the fluorophore is liberated from the quencher and thus fluoresces bright. In both SHERLOCK and DETECTR, target abundance may be enhanced by isothermal preamplification using RPA with or without in vitro transcription and/or reverse transcription.

respectively. The specificity for each method is encoded in the selective recognition of a target nucleic acid sequence by a guide crRNA that can be replaced in a modular fashion. After nuclease activation by specific binding of a guide crRNA to its cognate target, Cas13a and Cas12a both catalytically cleave collateral nucleic acids in vitro. Although its biological significance remains speculative, this unexpected collateral nuclease activity appears to be an intrinsic property of each effector protein, perhaps serving to induce dormancy or even cell death in a bacterial cell overwhelmed by viral targets.4,5,6

When this activity is leveraged to collaterally cleave reporter oligonucleotides to liberate a fluorophore from a quencher, detection of the initial specific binding event can be achieved using conventional fluorometry. Sensitivity is enhanced in both methods by isothermal preamplification of the target using recombinase-polymerase amplification (RPA), which allows these methods to detect targets in the low attomolar range.5,16 By designing target-specific RPA primers that flank, but are distinct from, the recognition site for the crRNA, this amplification step provides an added layer of specificity beyond that of the effector binding event alone, in much the same way as a nested two-step PCR reaction or TaqMan qPCR.

SHERLOCK uses as its effector Cas13a (initially termed C2c2), which binds and cleaves RNA targets through its guide crRNA.12 SHERLOCK was first engineered to detect the presence of nucleic acids to identify and characterize pathogens and tumor DNA.5 More recently, SHERLOCKv2 adds several additional features including multiplexing, quantitation, increased sensitivity, and a simpler readout.13 Multiplexing was achieved through identification of Cas effectors with orthogonal sequence preferences for off-target cleavage, enabling design of mutually exclusive reporter probes, each with a different fluorophore, that could report on up to four different target sequences in the same reaction. Quantitation involved finding nonsaturating conditions for the preamplification and detection reactions. Sensitivity was enhanced independent of target preamplification by designing a cascade system in which a Cas13a reporter, when cleaved, would serve as an activator for a separate CRISPR effector, Csm6, which in turn can catalytically cleave reporters when activated. Finally, the SHERLOCK assay has taken two different approaches to enhancing portability. First, all of the reagents involved in the original SHERLOCK assay were shown to withstand lyophilization on paper before being reconstituted, allowing for the possibility of field deployment.14 Second, an alternative reporter probe was described in SHERLOCKv2 that allows the assay to be adapted to a lateral flow device with a visual readout, similar to enzyme-linked immunosorbent assays (ELISAs) used in antigen detection.17 This lateral flow assay format is not yet compatible with multiplexing or quantitation but provides a rapid, visual binary readout.

In contrast to SHERLOCK, DETECTR harnesses Cas12a, which binds to DNA targets through its guide crRNA. Cas12a is also capable of target-recognition-triggered collateral cleavage, in this case of unrelated ssDNA upon binding to dsDNA.16 Cas12 proteins from several, though not all, species exhibited this off-target ssDNA “shredding” upon activation. In the DETECTR assay, activation of Cas12a results in liberation of quenched fluorophores from reporter probes with remarkable efficiency in vitro, with over 1000 turnovers per second from a maximally activated Cas12a molecule. The measured $k_{cat}/K_m$ approached the theoretical diffusion limit, suggesting that this collateral nuclease activity observed in vitro is not likely to be incidental but rather may have been evolutionarily optimized. To date, DETECTR has used fluorescence as a readout.

APPLICATIONS TO INFECTIOUS DISEASE DIAGNOSTICS

The SHERLOCK assay has been demonstrated to recognize both viruses and bacteria. It was used to detect Dengue, Zika, West Nile, and Yellow Fever viruses, all single-stranded RNA viruses of the flavivirus genus, as well as HIV, from RNA extracted either from cell culture or from patient serum or urine samples.15,17,18 It was also used to detect RNA genes from different bacterial species for identification and the carbapenemase genes KPC and NDM in order to identify the presence of genotypic causes of resistance to the carbapenem antibiotics,16 an important antibiotic class of last resort.20–23 A subsequent publication describes an upstream nucleic acid stabilization and viral inactivation sample preparation step, called HUDSON (heating unextracted diagnostic samples to obliterate nucleases), that pairs directly with SHERLOCK and bypasses the need for RNA extraction, further simplifying the infrastructure requirements for assay deployment.18

DETECTR, meanwhile, was initially demonstrated in type-specific detection of the double-stranded DNA human papillomavirus (HPV) from both cell culture and patient anal swabs, with good reported concordance with conventional PCR studies.16 The assay targeted a variable region that allowed identification of the two types of HPV most associated with malignancy, HPV16 and HPV18. Though Cas12a and Cas13a recognize DNA and RNA, respectively, the demonstrated ability to combine reverse transcription, RPA, and in vitro transcription assays in a single workflow without...
intervening purification steps means that, in theory, either method should be able to detect either type of target; the ability to detect both RNA and DNA targets has been demonstrated with SHERLOCK.$^{15}$

These CRISPR-based diagnostic methods, to date, are well-suited to detect specific genetic targets as a proxy for the presence or absence of a pathogen. With proper calibration, perhaps this can be extended to estimating pathogen abundance. Targeted detection can also aim to identify other genes of clinical relevance, such as antibiotic resistance genes, as illustrated for SHERLOCK by the detection of carbapenemase genes. In some cases, these assays may even be tunable to perform single nucleotide polymorphism (SNP) detection, when a critical SNP is known in advance, by carefully designing a crRNA to target the region containing the SNP in a manner that favors selective binding of one variant over another. The potential of this approach was demonstrated with SHERLOCK for detection of a specific Zika virus mutation linked to microcephaly$^{18}$ and also for detection of several key cancer-associated mutations.$^{15}$ The readout for SNP detection was not binary, however, but relied on relative quantification of binding to crRNAs targeting the wild-type and the SNP, which requires either more sophisticated, quantitative detection measurements or perhaps a carefully calibrated lateral flow readout.$^{17,18}$

### COMPARISON TO EXISTING TECHNOLOGIES

Compared with more mature pathogen detection methods such as culture, antigen or antibody detection, PCR and related amplification-based detection, and even unbiased sequencing, CRISPR-based nucleic acid detection technologies are new and relatively untested beyond the proofs of concept discussed above. However, their conceptual strengths and weaknesses relative to these other methods are worth considering further, within the context of the key desired characteristics of a pathogen detection assay. These characteristics include range or breadth of pathogens detected, sensitivity, specificity, speed, reproducibility, robustness to variation in both target sequence and real-world experimental conditions, iterability (i.e., barriers to introducing new targets for detection), multiplexing capacity, portability, ease of use, infrastructure requirements, and cost (Table 1).

The conceptually most direct comparators to these CRISPR-based detection systems are methods based on nucleic acid amplification, e.g., PCR. Like amplicon-based methods, CRISPR-based tests require foreknowledge of the pathogen being sought. DETECTR and SHERLOCK distinguish themselves from PCR-based methods, however, in that they couple an initial amplification step with a subsequent detection step that both further amplifies the signal and provides an additional layer of specificity based on target recognition by the guide crRNA. These advantages could be matched in PCR methods only by the addition of a second, nested PCR reaction, a step that requires amplicon isolation between the two steps, a sample manipulation that is not necessary in the CRISPR systems. In addition, these CRISPR-based assays function at more accessible temperatures than PCR, widening the possible scope of deployment conditions. While some amplification-based assays, including the RPA step used in preamplification in these assays, can function under similar conditions, they often suffer from relatively low specificity related to low-temperature primer annealing; these CRISPR effectors have evolved to augment the specificity of nucleic acid recognition at physiological temperatures.

One drawback to both CRISPR-based methods and PCR is the challenge of multiplexing simultaneous, independent detection of different pathogens. Particularly in cases where a pathogen is unknown and of low abundance, a single assay that could detect a range of pathogens is much preferred to the need to split the sample for parallel reactions. Even though some progress is being made, complex multiplexed PCR is challenging and remains an active area of research.$^{16}$ At the same time, SHERLOCKv2 currently offers some degree of multiplexing (4-plex); however, this is quite limited in scope to date and requires a detector that can distinguish four separate fluorophores. While preamplification might in theory facilitate sample splitting for spatial multiplexing, this only helps if that preamplification step can itself be multiplexed across the full range of potential pathogens.

Mismatch tolerance is another area of potential difference between these two assays, with important implications. High tolerance for mismatches increases robustness to natural variation in target sequence, particularly in mutation-prone pathogens like RNA viruses. PCR assays are uniquely susceptible to mismatches at the 3’ end of either primer that impair primer extension and thus amplification. Hybridization such as that required for CRISPR detection should be more tolerant to such mismatches. However, the RPA preamplification step will be as dependent as PCR on successful priming, and while SHERLOCKv2’s unique potential for coupling of Cas13a to Csm6 enables polymerization-free amplification, this comes at the cost of the additional layer of specificity for the intended target provided by RPA priming, since the Csm6-

### Table 1. Comparisons of CRISPR-Based Nucleic Acid Detection Methods with Current Pathogen Detection Methods

<table>
<thead>
<tr>
<th>feature</th>
<th>culture</th>
<th>antibody detection</th>
<th>antigen detection</th>
<th>PCR</th>
<th>isothermal amplification$^{\text{a}}$</th>
<th>amplicon sequencing</th>
<th>sequencing</th>
<th>CRISPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>breadth of pathogens</td>
<td>limited</td>
<td>limited</td>
<td>limited</td>
<td>broad</td>
<td>broad</td>
<td>most broad (unbiased)</td>
<td>broad</td>
<td></td>
</tr>
<tr>
<td>sensitivity</td>
<td>variable</td>
<td>variable</td>
<td>low</td>
<td>high</td>
<td>medium</td>
<td>high</td>
<td>medium</td>
<td>can be high</td>
</tr>
<tr>
<td>specificity</td>
<td>high</td>
<td>low</td>
<td>case-specific</td>
<td>case-specific</td>
<td>case-specific</td>
<td>high</td>
<td>high</td>
<td>can be high</td>
</tr>
<tr>
<td>speed</td>
<td>days, weeks</td>
<td>retrospective minutes</td>
<td>hours</td>
<td>minutes, hours</td>
<td>hours</td>
<td>days, hours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>portability</td>
<td>low</td>
<td>high</td>
<td>low</td>
<td>high</td>
<td>low</td>
<td>low</td>
<td>low</td>
<td>can be high</td>
</tr>
<tr>
<td>ease of use</td>
<td>moderate</td>
<td>easy</td>
<td>easy</td>
<td>moderate</td>
<td>moderate</td>
<td>hard</td>
<td>hardest</td>
<td>easy–moderate</td>
</tr>
<tr>
<td>cost</td>
<td>low</td>
<td>low</td>
<td>low</td>
<td>moderate</td>
<td>high</td>
<td>high</td>
<td>can be low</td>
<td></td>
</tr>
</tbody>
</table>

$^{\text{a}}$Examples of isothermal amplification techniques: LAMP = loop-mediated isothermal amplification; NASBA = nucleic acid sequence-based amplification; RCA = rolling circle amplification; RPA = recombinase-polymerase amplification.
mediated amplification process requires only the initial Cas13a activation. Conversely, because hybridization events can be regulated with nucleotide specificity, crRNA probes can be carefully designed to distinguish between strains of the same viral or bacterial species or between clinical isolates with critical SNPs.\textsuperscript{15–18} The same distinction can in principle be made with PCR, as primers can be designed to be sensitive to known polymorphisms; these concepts are implemented, for instance, in rifampin resistance detection by the TB diagnostic GeneXpert.\textsuperscript{18} However, the separation between preamplification and hybridization-based detection can perhaps be more specifically exploited in the new CRISPR-based assays, as the preamplification step could be targeted to conserved regions flanking the SNP of interest, and the relative plasticity of the guide-target hybridization can be exploited for SNP detection, at least in theory. One would therefore expect a trade-off between mismatch tolerance and specificity, depending on the details of assay design and implementation. In initial reports, thus far, these CRISPR-based assays appear to be quite robust and specific, but these characteristics will be challenged as they are more widely deployed against a range of real-world positive and negative samples.

Finally, another promising distinguishing feature of these CRISPR-based diagnostics to date is the apparent robustness of the enzymes to storage and implementation conditions, which has been reported for Cas13a.\textsuperscript{15} The demonstrated potential for independence of cold-chain transport and storage suggested by lyophilization tolerance,\textsuperscript{15} for compatible and relatively simple upstream sample processing,\textsuperscript{18} for isothermal functionality at accessible temperatures, and for adaptation to a simple lateral-flow device\textsuperscript{17} together offer unique, potentially transformative possibilities for deployment.

\section*{CONCLUSIONS}

Excitement around the development of CRISPR/Cas-based diagnostics is based on the critical need for transformative infectious disease diagnostics that can be deployed in numerous different clinical settings and against a diverse set of pathogens. Such tests have the potential for application to both clinical care and surveillance for pathogen outbreaks. They are a novel, alternative approach that deserve full investigation of their potential, as much of the impact of these assays will depend on the details of implementation. Moving beyond proof of principle studies to larger numbers of relevant samples under real-world conditions, it will be exciting to see the degree to which these new methods live up to the tantalizing promise of being robust, point-of-care, field deployable tests in diverse environments, with remarkable sensitivity, specificity, speed, and cost. As the true performance characteristics and thresholds for detection are determined and the technical specifications for a deployable test are optimized, the transformative possibilities for this diagnostic approach and its ability to meet the current demand for revolutionizing infectious disease diagnostics will become more and more clear.

\section*{AUTHOR INFORMATION}

\section*{Corresponding Author}
*E-mail: hungg@molvio.mgh.harvard.edu.

\section*{ORCID}
Deborah T. Hung: 0000-0003-4262-0673

\section*{Notes}

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The authors declare the following competing financial interest(s): D.T.H. and R.P.B. are co-authors on the SHERLOCK paper (ref 15) and are named co-inventors on patent applications related to SHERLOCK.

\section*{ACKNOWLEDGMENTS}

This publication was supported in part by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award 1K08AI119157-03 (R.P.B.).

\section*{REFERENCES}


(14) Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Eesletzibichler, P., Vols, S. E., Joung, J., van der


