The Nucleic Acids-Based Genetic Testing for Microorganisms: An Overview

Quanyi Li*

Development Services, BioReliance, USA

*Corresponding Author: Quanyi Li, BioReliance, MilliporeSigma, Rockville, USA.

Received: March 16, 2017; Published: April 01, 2017

Abstract

Many genomes for important microorganisms including unculturable viruses and bacteria have been decoded and the sequences of the genomes are in the public databases. Additionally, many novel microorganisms are discovered using next generation sequencing technologies. With the known genomes and genes for the microorganisms, the nucleic acids-based genetic assays for detection and identification of the microorganisms have been developed. The nucleic acids-based genetic tests have been widely accepted for clinical diagnostics, food safety, environment and biosafety tests for biopharmaceutical products. Although the molecular tests cannot differentiate live or dead microorganisms, the progress in viable PCR assays for microorganisms are very promising. The molecular testing has been used for the microorganism source tracking, monitoring and eventually elimination of microorganisms in a product or the environment. With the advanced knowledge of microorganism genomes, many developed molecular assays and the integration of automation in the biotech laboratories, it is most likely that the nucleic acids-based genetic testing for microorganisms will replace the culture-based testing for microorganisms in near future.

Keywords: Molecular Assay; PCR; Sanger Sequencing; Next Generation Sequencing; Genetic Testing; Microorganisms

Introduction

The culture-based tests for microorganisms are often referred as “golden standards” by microbiologists. The culture methods detect culturable microorganisms, but not unculturable microorganisms [1,2]. Microbiologists faced challenges in detecting slow growing bacteria such as *Mycobacterium* spp. and unculturable viruses such as norovirus for many years. Now, scientists detect those challenging human pathogens using molecular tests [3,4]. Over the past 30 years, microbiology has changed completely and become molecular microbiology era. When an unknown bacteria or virus is isolated and it can be sequenced immediately. With the sequence of the genome, a PCR assay can be developed for source tracking. One of examples was the outbreak of *Shiga*-toxin-producing *Escherichia coli* O104:H4 in 2011 in Germany [52]. The *E. coli* isolated was immediately sequenced with Ion Torrent next-generation sequencing (NGS), then the qPCR assay targeting the unique *Stx2a* allele was developed to detect the outbreak strain [53,54]. Rapidly, the contaminated food samples were found and discarded.

Recently, many new bacteria and viruses have been discovered using nucleic acids-based molecular methods and NGS. Many of the new microorganisms are accepted as new bacteria or virus species, even they remain as unculturable microorganisms. The small ss-DNA virus family Parvoviridae has many new unculturable members including bovine parvovirus 2 and bovine parvovirus 3 [5,6]. The genomes of those bovine parvoviruses were detected in bovine serum products used in biopharmaceutical industries. Presence of the bovine viral nucleic acid sequences in bioproducts caused serious biosafety concern [7].

The sterility testing for microorganisms, *in vitro* cell culture assay with indicator cell lines for adventitious viruses are the standard GMP testing methods, and the molecular tests such as PCR and sequencing are used as additional supporting tests [8,9]. The nucleic acids-based genetic testing with PCR, Sanger DNA sequencing and next generation sequencing has significantly improved the testing quality such as higher sensitivity, better accuracy, and fast turn around time. With the use of advanced technologies and the quality compliance, the molecular tests can be the rapid test of choice for many laboratories [10,11].

1. Nucleic acids-based genetic testing technologies

The most popular molecular technologies used in genetic testing are PCR, DNA sequencing, and next generation sequencing (NGS). All those technologies are based on the principle of DNA replication. PCR amplifies trace amount of DNA or RNA molecules and generates million copies of amplicons. Since the invention of PCR in late 1980, PCR has become the most reliable and sensitive technology in detection of DNA or RNA molecules. Sanger DNA sequencing is used to identify a PCR product. Both PCR and Sanger DNA sequencing are used together for detection and identification of a microorganism. NGS provides entire genome analysis for one microorganism and metagenomics analysis for a community or population of microorganisms [55-57].

The molecular testing analyzes the targeted genes or genomes of the microorganisms with PCR, Sanger DNA sequencing and NGS. The molecular testing delivers rapid tests with high sensitivity and specificity. The molecular testing can be a presumptive or definite test [43,44]. When a molecular testing delivers the results with 100% certainty, the test is definite; and a presumptive test delivers the most likely results with significant confidence although the certainty is not 100%. Thus, molecular genetic testing is the most reliable test compared to culture-based tests.

1.1. PCR and applications in detection of microorganisms

There are many types of PCR including conventional PCR, real-time PCR, quantitative PCR, droplet digital PCR (ddPCR), and reverse transcription-PCR (RT-PCR) [12-15]. The conventional PCR is used for amplicon analysis such as amplicon fragment analysis, amplicon sequencing, nested PCR and multiple amplicons analysis. The conventional PCR is qualitative, or semi-quantitative. When using conventional PCR for detection of microorganisms, PCR product is sequenced and then used for BLAST search GenBank [58-60]. Recently, conventional PCR with degenerative primers was used to develop paramyxoviruses family PCR using capillary electrophoresis of amplicons for accurate sizing. The target gene was the conserved motif of viral polymerase [16]. Clearly, when conventional PCR delivers accurate sizing and sequencing of the targeted genes of microorganisms, it offers an excellent definite assay for detection and identification of broad range of microorganisms.

The real-time PCR is an amplification curve assay, and it uses the threshold of the number of PCR cycles (Ct). In real-time PCR, the amplification of DNA templates is monitored during the cycling of PCR reactions. The quantitation is achieved using the standard curves. The extreme specificity of DNA synthesis with Taq Polymerase makes the real-time PCR as the most popular quantitative assay. However, at the lower detection limit, the Ct value becomes unreliable and causes mis-interpretation [13]. Thus, single target real-time PCR detection of microorganisms is a presumptive assay. The positive test results should be verified by conventional PCR, DNA sequencing, or multiplex real-time PCR with additional targets [44,45]. The ddPCR performs in nanoliter droplets and the amplicons are counted and calculated at the end of PCR. The ddPCR delivers the standard curve independent absolute quantitation. The real-time PCR and ddPCR are the best assays for detection and quantitation of single or few targets of genes [12-15]. The RT-PCR targets the RNA molecule, and the RNA template is converted into cDNA with reverse transcriptase; then the cDNA templates are amplified by PCR. There are many viruses with RNA genomes and the RNA viruses can be detected by RT-PCR [37,45]. Furthermore, the RT-PCR can detect the expression of genes. Since RNA molecules have short life-time in bacteria, the RT-PCR was used to detect viable bacteria, not dead bacteria [80].

Standard PCR detects both viable and nonviable microorganisms. When the DNA from dead cells is damaged, the PCR may only detect DNA derived from viable microorganisms. Thus, viable PCR or vPCR was developed for detection of viable microorganisms. The viable PCR involves treatment of microorganisms in a sample with fluorescent dye which can penetrate damaged cells, but not intact viable cells. The dye then binds to cellular DNA and causes damage of DNA under light exposure. The damaged DNA subsequently cannot be amplified by PCR. By comparing the amplification curves (Ct values) between treated and untreated samples, the viable and nonviable microorganisms in the sample can be detected [17,18]. Alternatively, vPCR is used in detection of viable microorganism after enrichment culture. If a sample does not have viable microorganisms, the Ct values between cultured and uncultured samples would be similar. If a sample has viable microorganisms, the Ct value for cultured sample would be reduced due to the growth of viable microorganisms [19].
1.2. Sanger DNA sequencing and applications in identification of microorganisms

The Sanger DNA was developed by Fred Sanger in 1970. The Sanger sequencing offers the most reliable and accurate sequences [20]. Sanger sequencing can identify bacteria by sequencing 16S rRNA gene using bacteria universal primers and identify fungi by sequencing the ITS regions of rRNA genes using fungal universal primers [21-24]. The genetic identification of bacteria and fungi needs to have pure microorganism isolates for DNA extraction and then amplify 16S RNA gene or ITS region. The PCR products are purified and used for DNA sequencing with sequencing primers. Nowadays, sequencing of the 16S rRNA gene or ITS region is the first experiment for identification of unknown bacteria or fungi.

The DNA sequences can be analyzed using BLAST search against GenBank at National Center for Biotechnology Information (NCBI, www.ncbi.nlm.nih.gov). The similarity of DNA sequences with GenBank entries are reported [61], therefore, the microorganisms are identified based on DNA sequences.

The genetic identification of bacteria and fungal isolates using DNA sequencing of 16S rRNA gene from bacteria and ITS regions form fungi has been commercialized as MicroSeq ID [25]. The MicroSeq ID is the standardized test which can be performed to GMP standards. Furthermore, other genes from microorganisms can be sequenced and used for typing or source tracking of the microorganisms. When a microorganism isolate is obtained, the strain can be determined using multilocus sequence typing (MLST) on PCR amplicons [26,27]. There are several international consortiums for tracking the sources of most important human pathogens using standardized MLST. One of examples, *Legionella pneumophila* sequence-based typing (SBT) is a standard test by EWIGLI (The European Working Group for *Legionella* Infections, www.ewigli.org). Total seven loci including *flaA*, *pile*, *asd*, *mip*, *mpmpS*, * proA* and *neuA* are amplified by PCR and sequenced [28,29], and the variations of the sequences on the loci are called alleles. The allele profiles of the strain compared to the reference sequences are reported to EWIGLI where a database is maintained for international surveillance of the disease and monitoring the outbreak strains.

1.3. NGS and genome sequencing of microorganisms

The NGS can sequence millions of amplicons or small fragments in parallel using nanotechnologies [30,31]. The NGS can analyze a whole genome of microorganism or a population of microorganisms as metagenomics analysis [32,33]. With advanced bioinformatics tools and clouds computing, the NGS data can be processed rapidly with the return of vast number of reads. When a novel virus or adventitious virus was found by NGS universal virus screen [33,34], the presence of virus can be verified by PCR and Sanger sequencing of the amplicons so that the entirety of the assembling genome is confirmed. For identification of isolated virus or bacteriophage, the best approach is using NGS whole genome sequencing [34,35]. Bacteriophage contamination of *E. coli* fermentation for enzymes production is a serious manufacturing problem [91,92]. The bacteriophage genomes can be any type of nucleic acids such as ssDNA, dsDNA, ssRNA, and dsRNA [36]. The type of nucleic acids can be determined using enzyme digestion of the genomic nucleic acids. Then, a DNA or RNA library can be constructed with NGS adaptors and analyzed on either Ion Torrent or Illumina MiSeq. The reads can be assembled using virus or bacteriophage reference genomes.

NGS has been approved for genetic diagnosis by FDA [46,47]. NGS has become a powerful tool for personalized genomics in cancer diagnostics [46,47]. In many cases, Sanger sequencing was used for the confirmation test of genetic mutations detected by NGS [48,49]. The study of microorganisms using metagenomics with NGS has revealed many novel bacteria and viruses [31,50]. The research and early industrial entries for detection of adventitious viruses using NGS showed promising results [51]. Similarly, the novel bacteria and viruses detected by NGS must be confirmed for the entirety of the genomes using Sanger sequencing.

2. Development of molecular assays for detection and identification of microorganisms

There are many PCR assays developed for detection of microorganisms in clinical, food safety, environment and biosafety for biopharmaceutical products [62-65]. An assay development needs an assay concept, then the assay targets, and assay conditions. When an assay is developed, it needs to be validated and then it can be a test for the intended use.
2.1. The concept and targets for PCR assay

The concept comes from the intended use of the assay such as present or absent assay, quantitative assay; to detect a group of pathogens, a species, a strain, and a serogroup. When the intended use of the assay is decided, the identifier sequences are selected as the PCR target regions. The target regions must be unique for the bacteria group, species, or strain. When the PCR assays are for detection of strains, isolates, or serotypes; several target sequences are required. When the PCR targets are identified, the design of PCR primers and probes can be performed using bioinformatics tools. Then, PCR conditions are tested. The PCR positive standards either genomic DNA or RNA from the microorganism or synthesized target DNA sequences are used to evaluate the PCR primers, probes, and PCR conditions.

For qPCR assay for the toxigenic strain of *E. coli* O157:H7, there were several targets required, first target was to identify *E. coli* 157:H7 with *eae*, then with targets for *Shiga* toxin-1 (*stx1*) and *Shiga* toxin-2 (*stx2*); thus, a multiplex qPCR was used [66]. For qPCR assay for MRSA, two target components were used, one was for *Staphylococcus aureus* bacteria using the unique region Sa422, and another one was for the antibiotic resistant gene *meca* [67]. Since *Legionella pneumophila* serogroup 1 causes most human infections, the qPCR assay for serogroup 1 has the clinical interest. The qPCR assay for serogroup 1 was developed using serogroup 1 strains specific targets from LPS cluster genes [68]. For fungi qPCR assays, the ITS sequences were selected to design primers and probes [24,69]. A qPCR assay for *Histoplasma capsulatum* was developed for clinical and environment testing; the human pathogen *Histoplasma capsulatum* cause lung infections through inhalation of fungal spores from air carrying fungal spores [70].

Viruses have diversified genome structures; each family or genus has the signature genome structures and genes. The conserved regions of viral polymerase genes were used to develop herpesvirus family PCR [71], and paramyxoviruses family PCR [16]. For enterovirus, 5' untranslated regions (5'UTR) were used to design the qPCR assay [72]. The viral capsids genes were used as PCR targets for RT-qPCR assays for genogroup I and II noroviruses [73].

2.2. Assay validation

When an assay is developed, the assay specifications need to be validated for the sensitivity, specificity, accuracy, precision, and reproducibility [39,74-76]. The suitability for type of samples or intended uses must be validated as well [74,75]. The validation can be performed by two analysts in single laboratory or in multiple laboratories. The purpose of validation is to prove the assay can be performed repeatedly.

For PCR assays, the primers must be verified electronically for the specificity using BLAST analysis or other bioinformatics tools. The PCR assays need to be performed in the presence of non-specific DNA, typically genomic DNA from bacteria, animal cells and plant tissues. The PCR assays do not amplify no-specific DNA. The sensitivities of the qPCR assay may be compared with another qPCR assay or with a culture method [77,78]. The accuracy needs to be evaluated with spiking of “known” amount of target DNA in various genomic DNA and matrix materials [74,75]. The linearity and slope of a qPCR standard curve are also indicators for accuracy [77,78]. The precision is another important criterion for a qPCR assay and it is expressed statistically using RSD (relative standard deviation) or coefficient of variation (CV) [74,75]. A good PCR assay must have less than 15% of CV [75]. Typically, qPCR has greater variation at lower limit of detection (LOD) and Limit of quantification (LOQ) [74,77]. When using the qPCR assay to test the target in a new matrix or product, a product qualification evaluation must be performed by spiking of known amounts of target [74,75] to prove that the spiked targets can be detected at low concentrations. In addition, an internal control for PCR assay, sample processing control with a microorganism or plasmid DNA to monitor the recovery are also required [74,75]. With a validated PCR assay, the performance of the assay is expected. For GMP testing, the assays must be validated [75].

2.3. Assay and test

A test works with its context, the intended use, the quality standards, the acceptance criteria, and the data interpretation. For example, EPA1615 is an official method developed by US EPA for drinking water test for Enteroviruses [37]. The test utilizes two type of assays, the qPCR assays for detection of viral genomes, and the culture method to access viable or non-viable virus using animal cell culture. The test...
protocol covers the sampling procedure, RNA extraction protocol, qPCR assays and cell culture method, then data recording and analysis, result interpretation and reporting [37]. When molecular assays are used for testing of the microorganisms in clinical samples, food samples, environment samples and biopharmaceutical products, the intended use of the assays must be validated and approved by the regulatory agencies. Then, the molecular assays become regulated tests.

3. Quality systems and regulation requirements

The testing facility, lab and operation must comply with the quality standards known as ISO, cGLP and cGMP. The tests must meet the regulatory requirements including FDA and EPA. The testing laboratory must state clearly about the intended uses such as raw materials, bulk or product lot release tests [88]. The international standard document ISO17025 is one of the most important general requirements for the competence of testing and calibration laboratories. In ISO17025 document, the quality requirements have been clearly explained to cover the entire testing process [79].

In U.S., the microbiological testing including molecular testing is regulated and the corresponding quality standards are applied. The information about the regulatory organizations can be found on their websites [81-90]. In Europe, similar regulatory agencies and ISO standards are compiled for governing and guiding the testing industries. Microbiological tests for food safety are regulated by FDA, the official test methods are from FDA and AOAC validated methods [40]. When a test is developed, then it can be validated by AOAC, after FDA approval of the test, it becomes a FDA approved test method [40]. Similarly, GMP biosafety tests for adventitious microorganisms in biopharmaceutical products are FDA regulated tests, the tests must be validated and approved by FDA [41,42]. Clinical diagnostics are regulated by FDA and the corresponding regulatory agent CLIA [38,39].

Conclusion

The advances of genomics for all living microorganisms including both culturable and unculturable microorganisms have made the genetic testing for them a reality. Many of the unculturable microorganisms are known as molecular species and they can be potential risk and pathogens when they are present in environment, food, biopharmaceutical products. Our understanding about unculturable microorganisms is very limited. With the advanced molecular genetic testing for the microorganisms including both culturable and unculturable microorganisms, we can live in a safer environment with safer food, vaccines and biopharmaceutical products.

Conflict of Interest

The author declares no conflict of interest.

Bibliography


14. Sigma online: PCR Technologies Guide.


Citation: Quanyi Li. "The Nucleic Acids-Based Genetic Testing for Microorganisms: An Overview". *EC Microbiology* 7.3 (2017): 70-79.
The Nucleic Acids-Based Genetic Testing for Microorganisms: An Overview


30. Illumina online: An introduction to next generation sequencing technologies.


The Nucleic Acids-Based Genetic Testing for Microorganisms: An Overview


42. US FDA. "Guidance for industry-characterization and qualification of cell substrates and other biological materials used in the production of viral vaccines for Infectious disease indications" (2010).

43. CDC Online: Guidance for U.S. laboratories testing for Zika virus infection.


47. US FDA: "FDA advances Precision Medicine Initiative by issuing draft guidances on next generation sequencing-based tests" (2016).


The Nucleic Acids-Based Genetic Testing for Microorganisms: An Overview


Citation: Quanyi Li. "The Nucleic Acids-Based Genetic Testing for Microorganisms: An Overview". EC Microbiology 7.3 (2017): 70-79.
75. FDA: Draft guidance for industry bioanalytical method validation (2013).


81. AACC: American Association of Clinical Chemistry for clinical diagnostics.

82. ANSI: American National Standards Institute for food safety and pharmaceutical products.

83. AOAC: AOAC International for food safety and pharmaceutical products.

84. ASTM: ASTM International for environment.

85. CLIA: Clinical Laboratory Improvement Amendments for clinical diagnostics.

86. CLSI: Clinical Laboratory Standards Institute for clinical diagnostics.

87. EPA: Environmental Protection Agency for environment.

88. FDA: Food and Drug Administration for food and pharmaceutical products.

89. ISO: International Organization for Standardization for food safety, environment and pharmaceutical products.

90. USP: United States Pharmacopeial Convention for pharmaceutical products.


Volume 7 Issue 3 April 2017
© All rights are reserved by Quanyi Li.