SAGE-Hindawi Access to Research
Journal of Pathogens
Volume 2011, Article ID 310135, 11 pages
doi:10.4061/2011/310135

Review Article

Recent Advances in Molecular Technologies and Their Application in Pathogen Detection in Foods with Particular Reference to *Yersinia*

Jin Gui[1] and Isha R. Patel[2]

[1] *College of Management and Technology, Walden University, 155 Fifth Avenue South, Minneapolis, MN 55401, USA*
[2] *Division of Molecular Biology, Office of Applied Research and Safety Assessment, Center for Food Safety and Applied Nutrition, U.S. Food and Drug Administration, 8301 Muirkirk Road, MOD 1 Facility, Laurel, MD 20708, USA*

Correspondence should be addressed to Jin Gui, jin.gui@waldenu.edu

Received 22 June 2011; Accepted 8 September 2011

Academic Editor: Latiful Bari

Copyright © 2011 J. Gui and I. R. Patel. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Yersinia enterocolitica is an important zoonotic pathogen that can cause yersiniosis in humans and animals. Food has been suggested to be the main source of yersiniosis. It is critical for the researchers to be able to detect *Yersinia* or any other foodborne pathogen with increased sensitivity and specificity, as well as in real-time, in the case of a foodborne disease outbreak. Conventional detection methods are known to be labor intensive, time consuming, or expensive. On the other hand, more sensitive molecular-based detection methods like next generation sequencing, microarray, and many others are capable of providing faster results. DNA testing is now possible on a single molecule, and high-throughput analysis allows multiple detection reactions to be performed at once, thus allowing a range of characteristics to be rapidly and simultaneously determined. Despite better detection efficiencies, results derived using molecular biology methods can be affected by the various food matrixes. With the improvements in sample preparation, data analysis, and testing procedures, molecular detection techniques will likely continue to simplify and increase the speed of detection while simultaneously improving the sensitivity and specificity for tracking pathogens in food matrices.

1. Introduction

The genus *Yersinia* mainly includes animal pathogens, but animals can transmit disease to humans through direct or indirect contact [1]. Symptoms of illness can include diarrhea, vomiting, abdominal pain, and fever. There are three species within the genus *Yersinia* that are pathogenic for humans: *Yersinia enterocolitica*, *Yersinia pseudotuberculosis*, and *Yersinia pestis*. All these species have evolved with different clinical symptoms. *Y. enterocolitica* infections have been observed all over the world, but appear to be more common in Europe, especially in some Scandinavian regions, with much lower rates in the United States [2]. Food has often been suggested to be the main source of yersiniosis. Enteropathogenic *Yersinia*, *Y. enterocolitica,* and *Y. pseudotuberculosis,* entering the human body in contaminated food invade the M cells of the Peyer's patches [3]. The process and its effect on the host cell are driven by a large array of virulence factors that are deployed under genetic and environmental regulation. *Y. enterocolitica* can be categorized by biotype. Biotype 1A strain is considered as nonpathogenic, while 1B strain is considered as high-pathogenic, and biotypes 2, 3, 4, and 5 strains are considered as low-to-moderate pathogenic. The pathogenic phenotype can be differentiated due to the virulence-associated genes identified in these strains.

In the event of foodborne disease outbreaks, rapid identification of foodborne pathogens rely on the speed and simplicity of the detection method, which are critical for early detection and quick response [4]. The new advancement of high-throughput OMICS technologies provides scientists with the means to identify the agent and attribute it to a specific source of pathogenic *Yersinia* in food systems [5].

2. Current Advances in Detection Methods

One of the most challenging issues in food safety is the detection of foodborne pathogens. Since the infectious dose of many pathogens is as low as a few cells or particles [6], the sensitivity of the diagnostic tool becomes essential. In fact, the detection of pathogens in nonprocessed or minimally processed foods is not easy. Such foods are not sterile; the native microflora in such foods can mask the presence of a pathogen by interfering with isolation [7]. Thus, more sensitive and reliable detection methods have been developed in accordance with the advancement of molecular and biochemical technologies.

Isolation of *Y. enterocolitica* from clinical, food, and environmental samples can be challenging primarily due to the difficulty of growing *Y. enterocolitica in vitro* [8]. Traditional culture-dependent methods have several limitations, such as long incubation steps, lack of identification between species, and lack of discrimination between pathogenic and nonpathogenic strains [8, 9].

Numerous molecular techniques have emerged, that offer the advantage of speed along with specific and sensitive detection [10, 11]. Due to the relative simplicity, rapidity, reliability, and sensitivity, DNA-based detection technology plays an important role and provides detection methods in the form of next-generation sequencing [12], microarray [13], fluorescent in situ hybridization (FISH) [14], polymerase chain reaction (PCR) [15], molecular beacon technology [16], and many others. DNA testing is now possible on a single molecule, and high-throughput analysis allows thousands of detection reactions to be performed at once, thus allowing a range of characteristics to be rapidly and simultaneously determined. Some of the current molecular detection methods not only can be performed in the laboratory or clinical settings but also can be run at the observation site, such as on the farm or in the field, in the form of "all-in-one" kits [17, 18].

2.1. Genome Sequence. The release of the complete genome sequence of *Y. enterocolitica* strain 8081 provided important insights into the pathology of this bacterium [19]. There are 18 completed and over 160 incomplete *Yersinia* strains past and ongoing *Yersinia* genome sequencing projects (http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi/) including *Y. pestis* strain CO92 [20] and *Y. pseudotuberculosis* strain IP 31758 at J. Craig Venter Institute/The Institute for Genomic Research [21]. These sequencing projects will enable the study of the evolution of the pathogenic changes in each species as they have adapted to new environmental surroundings. The information gathered from the genome sequences of the three major pathogenic *Yersinia* species will allow the development of a cross-species microarray for pathogenic *Yersinia* and will lead to invaluable insights into how the enteropathogens are adapted to their lifestyle.

Recently, Fuchs and coworkers took advantage of a whole-genome shotgun sequencing approach to assemble, annotate, and analyze the sequence of strain W22703 of *Y. enterocolitica* [22]. Their research study provided valuable information on the strategies utilized by *Y. enterocolitica* to cope with its environment. Wang et al. [23] sequenced the complete genome of *Y. enterocolitica* strain 3/O:9 and strain 8081 (1B/O:8); the comparison of the genome sequences of these two strains indicated that these two strains' different pathogenicity may have been a result of completely separate evolutionary events. Recent efforts by Batzilla et al. [24] to compare the complete genome of *Y. enterocolitica palearctic* serobiotype O:3/4 to the available genome of *Y. enterocolitica* ssp. *enterocolitica* 8081 O:8/1B indicated that gene loss and acquisition during evolution through mobile genetic elements could be the contributing factor to differentiate pathogenic bacteria from apathogenic bacteria of the same species. *Y. enterocolitica* is a heterogeneous bacterial species with a complex life cycle encompassing aquatic and biological environments. Further genome sequencing and analysis will help us to learn more about the evolution of *Y. enterocolitica* strains and provide the necessary information for the development of molecular-based detection methods for *Yersinia* in food systems.

Rouillard and Gulari developed a pangenomic oligonucletide microarray probe set database called OligoArrayDb [25]. OligoArrayDb was designed for most of the sequenced genomes that are not covered by commercial catalog arrays. Based on their algorithm of analysis, the *Y. enterocolitica* strain 8081 genome, a total of 4137 transcripts and containing 11821 oligonucletides, were chosen to represent the *Y. enterocolitica* strain 8081 transcriptome. Among these oloigonucleotides, 11251 are considered to be fully specific to their targets. This microarray probe set can be accessed through the website at http://berry.engin.umich.edu/oligoarraydb/index.html.

2.2. Microarray Analysis. The dominant application of microarrays has been in measuring gene expression in different biological conditions [26–28]. Other important microarray applications include comparative genomic hybridization [29], chromatin immunoprecipitation [30], mutation detection [31], genotyping [32], and array-mediated localized cell transfection [33, 34]. Microarray technology involves the placement of user-defined oligonucleotide probes in specific locations on a solid matrix such as glass or filters. The concept behind all microarrays is the precise placement of DNA fragments at high density on the solid support, so that they can act as molecular detectors. There are many variations of this method based on the solid matrix used and more importantly, the different types of DNA fragments on the array, including cDNA, oligonucleotides, and genomic fragments. Currently, there are three main types of microarrays: filter arrays [35], spotted glass slide arrays [36], and in situ synthesized oligonucluetide arrays [37] available for research purposes.

Following the hybridization of target DNA sequences to probes on the solid matrix, fluorescence-based detection can be used to monitor binding signal and be recorded. Along with the rapid development of microarray technologies, there has been an unprecedented amassing of data collected by academic institutes, as well as industrial organizations. Software applications can be used to conduct data analysis and greatly facilitate the data analysis

process. There are many open-source, public-domain, and commercial solutions for data storage, analysis, management, and exportation. Most of the applications are being updated frequently to keep current with the new demands from research. Several applications have been released that integrate data acquisition, processing, analysis, and exportation [25, 38]. The commercial GeneSifter (http://www.geospiza.com/Products/AnalysisEdition.shtml), the academic GenMAPP (http://genmapp.org/), and the open-source BASE (http://base.thep.lu.se/) aim to provide the functionalities for data analysis. Some software applications also provide comprehensive solutions for image analysis and data extraction. Most recent software applications for microarray data analysis are listed in Table 1.

Microarray methods provide an effective way of distinguishing between nonspecific and target product formation following PCR amplification of target DNA sequences from the samples. Amplification methods have been used previously in combination with microarray technology for the detection of *Y. pestis*. Huang et al. [39] were able to specifically detect *Y. pestis* from *Y. enterocolitica* and *Y. pseudotuberculosis* using a microarray method combined with PCR amplification. Myers and coworkers [40] developed a microarray chip combined with PCR amplification for detection and characterization of four virulence genes (*virF*, *ail*, *yst*, and *blaA*) in *Y. enterocolitica.* They were able to identify *Y. enterocolitica* from adulterated pasteurized whole milk using this approach. Ikeda et al. [41] were able to detect three foodborne bacteria: *Salmonella enterica* serovar Enteritidis, *Y. enterocolitica,* and *Bacillus cereus* in fresh vegetables using a DNA microarray method. Kim et al. [42] used comparative genomics to select 70-mer ologonucleotide probes specific for 11 major foodborne pathogens for use in microarray analysis. All of these studies have demonstrated that genome sequencing and DNA microarray analysis have a powerful application in detection of pathogenic *Yersinia* in food systems.

2.3. Immunoassay. Antibodies have been used for many years to type bacterial isolates serologically [43–45]. The development of the enzyme-linked immunosorbent assay (ELISA) introduced highly sensitive tests for specific targets with great reliability. Key advantages of ELISA are its ease of use, flexibility, and low cost. The highly specific nature of antibodies, especially monoclonal antibody (MAbs), and the simplicity and versatility of antigen-antibody reactions have facilitated the design of a variety of assays, and they comprise the largest group of molecular biological methods being used in foodborne pathogen detection [46–48].

Yersinia pestis is antigenically homogenous, but *Y. enterocolitica* and *Y. pseudotuberculosis* have multiple O and H antigens [49]. ELISA kits for detection of *Y. enterocolitica* are commercially available for the detection of the O antigen; for example, Mabs anti-O:3 and -O:9 can be purchased from LifeSpan BioSciences for research purposes.

Other methods for evaluating immunological binding events include fluorescence-based microscopy and surface plasma resonance. A commonly used field-portable immunoassay is the lateral flow disposable membrane technology. This technology is designed for threshold or qualitative testing. Advantages of this format include low-cost, portability, room-temperature stability and no need for specialized equipment and only minimal user training is required [50].

Multiplexing format immunoassays, suitable for the simultaneous evaluation of multiple targets in a sample, can be developed to increase the analytical productivity and drastically reduce analysis costs and sample and reagent consumption. For the low-multiplexing assay without automation, quantitative PCR, ELISA, or Western blotting allow multiple targets to be measured simultaneously and quantitatively. For the high multiplexing OMIC technologies, microarrays, SELDI, and LC/MS allow measurement of several hundred potential targets, but the output is essentially qualitative. There are two main multiplex immunoassay formats currently being applied widely in research: (1) protein attached microarrays [51, 52] and (2) bead-based microarrays [53, 54]. Magliulo et al. [55] developed a simple, multiplexed sandwich chemiluminescent enzyme immunoassay for the simultaneous detection of four of the major foodborne pathogens: *Escherichia coli* O157:H7, *Y. enterocolitica, Salmonella* Typhimurium, and *Listeria monocytogenes*. The accuracy and precision of this method were comparable to those achievable with the conventional culturing methodology yet detection was completed significantly faster than in traditional practices.

Protein microarray is a novel technology for quickly detecting and identifying proteins [56]. A protein detecting microarray comprises many different affinity reagents arrayed at high spatial density on a solid support. Each agent captures its target protein from a complex mixture, and the captured proteins are subsequently identified. For routine detection purposes, there is substantial benefit to be gained from using protein microarray technology. In principle, thousands of proteins can be spotted on a single slide, enabling one to interrogate simultaneously the presence of many different proteins with minimal sample consumption. Furthermore, hundreds of copies of an array can be manufactured, enabling the same proteins to be probed repeatedly with many different molecules from different samples. Rucker and coworkers have successfully developed antibody-based microarray techniques for the multiplexed detection of cholera toxin β-subunit, diphtheria toxin, anthrax lethal factor, and protective antigen, *Staphyloccus aureus* enterotoxin B, and tetanus toxin C fragment from spiked samples [57]. Li et al. used a protein microarray spotting with 149 *Y. pestis* proteins to profile antibody responses to a *Y. pestis* live vaccine [58]. With the continuing innovation for this technology, some limitations need to be addressed, as well. For protein detection microarrays, the cross-reactivity of affinity reagents need to be assessed and reduced. For a protein function microarray, the purity and integrity of the proteins need to be determined.

Immunoassays have an important role in the diagnosis and monitoring of diseases in routine-based pathological laboratories. However, immunoassay sensitivity and potential cross-reactivity should be carefully considered in comparing detection methods. Nucleic-acid-based technology

Table 1: Current software applications for microarray data analysis.

Software	Application	Provider	Platform	Web link
Array Designer	Primer design for microarray construction	Premier Biosoft International	Windows Linux	http://www.premierbiosoft.com/dnamicroarray/index.html
ArrayMiner	Analysis tool for microarray gene expression data	Optimal Design	Mac OS Windows	http://www.optimaldesign.com/ArrayMiner/ArrayMiner.htm
ArrayTrack	Database solution for managing, analyzing, and interpreting microarray gene expression data	National Center for Toxicological Research U.S. Food and Drug Administration	Web-based	http://www.fda.gov/ScienceResearch/BioinformaticsTools/Arraytrack/default.htm
ArrayVision	Automated analysis of macro- and microarrays	GE Healthcare	Windows	http://www.gelifesciences.com/aptrix/upp01077.nsf/Content/Products?OpenDocument&ParentId=957136
BAMarray	Detecting differentially expressed genes from microarray data using Bayesian analysis	Case Western Reserve University	Mac OS Windows Linux	http://www.bamarray.com/default.htm
BASE	Database solution for the massive amounts of data generated by microarray analysis	Lund University	Web-based	http://base.thep.lu.se/
Cluster	Perform a variety of types of cluster analysis and other types of processing on large microarray datasets	University of Tokyo	Mac OS Windows Linux/Unix	http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm
GenePattern	Gene expression analysis tools	Broad Institute, MIT	Web-based	http://www.broadinstitute.org/cancer/software/genepattern/desc/expression.html
GeneSifter	Tools for exploring the statistically significant interplay of the data with factors of biological relevance to understand the expression pattern in microarray data.	Geospiza Inc.	Web-based	http://www.geospiza.com/Products/AnalysisEdition.shtml
GenMAPP	Tools for visualizing data from gene expression experiments in the context of biological pathways.	Gladstone Institute, University of California at San Francisco	Windows	http://genmapp.org/
GenMaths XT	Analysis of high density microarrays and gene chips	Applied Maths	Windows	http://www.applied-maths.com/genemaths/genemaths.htm
Genowiz	A comprehensive multi platform software for microarray data analysis	Ocimum Biosolutions	Mac OS Windows Linux/Unix	http://www3.ocimumbio.com/data-analysis-insights/analytical-tools/genowiz/
Microarray tools	Including: a Comparative Genomic Hybridization (CGH) and expression microarray data analysis, data management and export system	J. Craig Venter Institute	Windows Linux/Unix	http://www.jcvi.org/cms/research/software/#c622/

Table 1: Continued.

Software	Application	Provider	Platform	Web link
Partek Genomics Suite	Statistical analysis and data mining tools to facilitate powerful and intuitive exploratory data analysis	Partek Incorporated	Windows Linux/Unix	http://www.partek.com/partekgs/
TreeArrange and Treeps	Software for displaying and manipulating hierarchical clustered data	University of Waterloo	Windows Linux/Unix	http://monod.uwaterloo.ca/downloads/treearrange/
waviCGH	For the analysis and visualization of array-CGH data	Spanish National Cancer Center, Bioinformatics Unit	Web-based	http://wavi.bioinfo.cnio.es/

may be a suitable alternative for a range of molecular targets traditionally detected by immunoassays [59].

2.4. Next-Generation Sequencing. DNA sequencing is one of the most important molecular tools in any life sciences field [12, 60]. Over the past 30 years, there has been more than a millionfold improvement in the rate of sequence generation with the progression from radio-labeled products using slab gels to fluorescent products and capillary electrophoresis to next-generation sequencing technologies [60]. According to Stratton, in the future, the cost of sequencing may drop greatly where, for example, the costs of sequencing whole cancer genomes can drop to US$1000. Routine sequencing in a clinical, diagnostic setting will then become feasible [60].

Next Generation Sequencing (NGS) technology has been adopted as a sequencing tool for quite some time [61–63]. This sequencing technology has the following features: massively paralleled sequencing without electrophoresis, samples need to be prepared and amplified, and extensive usage of computer resources. NGS can be categorized into (1) microelectrophoretic methods, (2) sequencing by hybridization, (3) real-time observation of single molecules, and (4) cyclic array sequencing [64].

There are significant differences between conventional sequencing technologies and NGS platforms in terms of sequencing chemistry, application, and cost [64, 65]. The comparison of major NGS technologies and conventional sequencing technologies is summarized in Table 2. The applications of conventional sequencing using the Sanger approach are suitable for small-scale sequencing within the kilobase to megabase range [66, 67]. The requirements of a Sanger sequencing approach include major costs such as robotic support of reagents, processing of multiple samples in either 96- or 384-well formats, and regular maintenance of capillary-based sequencers. NGS has fewer infrastructure requirements than the Sanger sequencing approach. Among the NGS platforms, there are important differences that may result in advantages with respect to specific applications (Table 2). Some applications may be more tolerant of short read lengths than others. The accuracy, as well as the specific error distributions of individual technologies, may also be relevant [68–71].

The diversity and advancement of NGS technology pose challenges for bioinformaticists to address, such as the issues of alignment, assembly, sequence scoring, data storage, and data release. Two major computational approaches are performed with NGS reads, assembly and alignment. The assembly approach is performed when no reference genome exists for the DNA sequenced, such as in the case of a genetically uncharacterized pathogen. Assembly algorithms take sequence reads, align overlapping sections, and generate longer length contigs, which serve as the scaffold for genome assembly, and subsequent alignments [72–74]. Alignment process is used to determine the best match between sequence reads and the reference sequence. To accommodate the large number of reads generated by NGS, a number of new alignment algorithms have been developed. These algorithms share the characteristic that alignment is performed in a multistep or heuristic approach in which the first phase consists of converting either the sequence reads or the reference sequence into an index of shorter length sequences, which are given read identifiers [75–77]. Postalignment, programs generate key information including the number of aligned reads, a list of sequence variants relative to the reference, and the percentage of reads containing the variant. A variety of software applications have been developed using these algorithms and are being widely utilized by researchers. Some of the popular tools are listed in Table 3.

Some of the key applications for NGS include (1) whole genome *de novo* sequencing and single nucleotide polymorphism (SNP) discovery [63, 68, 78], (2) mapping of structural rearrangements and transformation events [79], (3) expressed sequence tags (ESTs) or serial analysis of gene expression [80], (4) transcriptome assembly for gene discovery and transcription profiling [81], (5) large-scale analysis of DNA methylation [82], (6) genome-wide mapping of DNA-protein interactions [83], (7) confirmatory sequencing in gene cloning [84], and (8) genome-map-based cloning [85].

Cummings and coworkers [86] used the SOLiD system (Applied Biosystems, Calif) to conduct parallel microbial whole genome typing to detect strain-specific polymorphism in *Bacillus anthracis* and *Y. pestis*. Their research results

Table 2: Comparison of major next generation DNA sequencing technologies and conventional sequencing.

Platform	Application	Sequencing chemistry	Read length (bases)	Throughput per run (Gb)	Read per run (million)	Throughput per 24 hr (Gb)	Raw accuracy Range (%)	Cost Pe Mb ($)
ABI 3730	(1) Complement *de novo* assemblies for high-quality assembly of complex genomes; (2) Custom sequencing (3) Targeted resequencing for polymorphism discovery and genotyping	Sanger Dideoxy	800	0.00008	0.000096	0.00064	99.0 to 99.999	4000
ABI SOLID 5500	(1) Whole genome SNP discovery; (2) Transcriptome assembly and expression profiling; (3) Whole methylome resequencing	Sequencing by ligation	60 × 2	310	5167	45	99.0 to 99.9	0.05
Illumina HiSeq	(1) Whole genome SNP discovery; (2) Transcriptome assembly and expression profiling; (3) Whole methylome resequencing; (4) Bacterial and megaplasmid *de novo* assembly	Sequencing by synthesis	100 × 2	600	6000	75	96.2 to 99.7	0.02
Life Technologies Ion Torrent	(1) Whole methylome resequencing; (2) Bacterial and megaplasmids *de novo* assembly; (3) Sequencing quality control; (4) Sequencing requirement lower complexity	pH meter	200	0.2	1	2.4	>99.0	0.5
Roche 454	(1) *De novo* assemblies of complex genomes; (2) Metagenomics; (3) Analysis of large structural variations	Pyrosequencing	600	0.8	1	0.5	96.0 to 97.0	8

suggested the possibility of using NGS technology during a forensic or epidemiological investigation facilitating high-resolution strain tracking. Morelli et al. [87] utilized both conventional sequencing and NGS technologies to identify patterns of global phylogenetic diversity through the comparison of 17 whole genomes of *Y. pestis* isolates from global sources. Chen et al. used NGS technology to obtain and compare sequencing data from 3 pathogenic and 8 nonpathogenic members of the *Yersinia* genus [88]. They identified 100 regions within the genome of *Y. enterocolitica* that represented potential candidates for the design of nucleotide sequence-based assays for detection of the pathogen.

NGS has fundamentally impacted various fields of biological research, including food safety. This technology can be transitioned into the clinical diagnostic area. Similar to the development of microarray technology, the challenges will shift from mastering this technology to the question

Table 3: Software applications for NGS analysis.

Software	Categories	Sequencing file format compatibility	Created by	Operating platform	Web link
ABySS	Assembly	FASTA FASTQ QSEQ SAM BAM	Jared Simpson et al. Michael Smith Genome Sciences Centre	Mac OS Linux POSIX	http://www.bcgsc.ca/platform/bioinfo/software/abyss/
Edena	Assembly	FASTQ	David Hernandez University of Geneva Hospitals	Windows Linux	http://www.genomic.ch/edena.php/
Exonerate	Alignment	FASTA	Guy Slater and Ewan Birney European Bioinformatics Institute	Windows Linux Unix	http://www.ebi.ac.uk/~guy/exonerate/
Maq	Alignment	FASTA FASTQ Illumina Bustard & Gerald Illumina ELAND	Heng Li	Windows Linx	http://maq.sourceforge.net/
Mosaik	Alignment	FASTA FASTQ Illumina Bustard & Gerald SRF	Michael Stromberg and Gabor Marth Boston College	Mac OS Windows Linux	http://code.google.com/p/mosaik-aligner/
Phrap/ Cross_match/ Swat	Alignment	FASTA	Phil Green, Brent Ewing and David Gordon University of Washington	Mac OS Windows Linux	http://www.phrap.org/phredphrapconsed.html
PyroBayes	Base Caller	SFF	Aaron Quinlan et al. Boston College	Linux	http://bioinformatics.bc.edu/marthlab/PyroBayes/
SHARCGS	Assembly	Illumina Bustard & Gerald Illumina ELAND	Juliane Dohm et al. Max Planck Institute	Linux	http://sharcgs.molgen.mpg.de/
SHRiMP	Alignment	FASTA FASTQ SAM Illumina Bustard & Gerald	Michael Brudno and Stephen Rumble University of Toronto	Mac OS Linux	http://compbio.cs.toronto.edu/shrimp/
SOAP	Alignment Burrows-Wheeler	Illumina Bustard & Gerald Illumina ELAND	Ruiqing Li et al. Beijing Genomics Institute	Unix	http://soap.genomics.org.cn/
SSAHA2	Alignment Smith-Waterman	FASTA FASTQ SAM Illumina Bustard & Gerald	The Wellcome Trust Sanger Institute	Mac OS Linux	http://www.sanger.ac.uk/resources/software/ssaha2/
SSAKE	Assembly	FASTA	Rene Warren et al. Michael Smith Genome Sciences Centre	Linux	http://www.bcgsc.ca/platform/bioinfo/software/ssake/

Table 3: Continued.

Software	Categories	Sequencing file format compatibility	Created by	Operating platform	Web link
VCAKE	Assembly k-mer extension	FASTA	William Jeck et al.	Mac OS Linux	http://vcake.sourceforge.net/
Velvet	Assembly	FASTA FASTQ Illumina Bustard & Gerald Illumina ELAND	Daniel Zerbino et al.	Mac OS Linux Cygwin	http://www.ebi.ac.uk/~zerbino/velvet/

of how best to extract meaningful biological or clinical information from the large amount of data generated by this technology.

3. Summary

Food has often been suggested to be the main source of yersiniosis. Current methods to detect foodborne pathogens rely traditionally on culture media to select and propagate viable cells in foods. However, the isolation rates of pathogenic *Y. enterocolitica* have been low, which may be due to the limited sensitivity of the culture methods. The new advancement of the current technologies will provide cheaper, more accurate, and faster methods to identify pathogenic *Yersinia* in food systems during a food-related pathogenic crisis.

Despite better detection efficiencies, results derived using molecular biology methods can be affected by the various food matrices, the presence of normal bacterial flora, and interferences by some of the food ingredients. It still remains a challenge to develop methods that are rapid, sensitive, and specific in detection of foodborne pathogens. With the improvements in sample preparation, data analysis, and testing procedures, molecular detection techniques will likely continue to simplify and increase the speed of detection while simultaneously improving the sensitivity and specificity for tracking pathogens in food matrices.

The molecular-based detection methods discussed, above all, have advantages and limitations. Even use of the same detection method such as real-time PCR approach, different target genes used for the assay can limit the detection sensitivity. The detection range can vary from single colony forming unit (CFU) per ml to 10^3 CFU/mL. Similarly, the lateral flow stripe requires a relatively high concentration of target organisms between 10^7 CFU/mL to 10^{10} CFU/mL. Due to the limitations of individual detection methods, the combination with other techniques should be used for verification to ensure adequate specificity and sensitivity of the detection results. Combining with other methods also enhances the performance of individual assays. Owing to the complex variables in food analysis, most molecular-based methods for detecting foodborne pathogens are used for screening purposes, where the positive results need to be confirmed by cultural methods.

Acknowledgments

The authors thank Drs. Vijay Juneja, Xianghe Yan, and Robert W Li (U.S. Department of Agriculture, Agricultural Research Service) for their valuable comments and input.

References

[1] B. Swaminathan, M. C. Harmon, and I. J. Mehlman, "A review: *Yersinia enterocolitia*," *Journal of Applied Bacteriology*, vol. 52, no. 2, pp. 151–183, 1982.

[2] E. J. Bottone, "*Yersinia enterocolitica*: overview and epidemiologic correlates," *Microbes and Infection*, vol. 1, no. 4, pp. 323–333, 1999.

[3] C. Pujol and J. B. Bliska, "Turning Yersinia pathogenesis outside in: subversion of macrophage function by intracellular yersiniae," *Clinical Immunology*, vol. 114, no. 3, pp. 216–226, 2005.

[4] D. M. Hunter and D. V. Lim, "Rapid detection and identification of bacterial pathogens by using an ATP bioluminescence immunoassay," *Journal of Food Protection*, vol. 73, no. 4, pp. 739–746, 2010.

[5] D.-H. Kim, B.-K. Lee, Y.-D. Kim, S.-K. Rhee, and Y.-C. Kim, "Detection of representative enteropathogenic bacteria, Vibrio spp., pathogenic *Escherichia coli*, *Salmonella spp.*, *Shigella spp.*, and *Yersinia enterocolitica*, using a virulence factor gene-based oligonucleotide microarray," *Journal of Microbiology*, vol. 48, no. 5, pp. 682–688, 2010.

[6] E. C. D. Todd, J. D. Greig, C. A. Bartleson, and B. S. Michaels, "Outbreaks where food workers have been implicated in the spread of foodborne disease. Part 4. Infective doses and pathogen carriage," *Journal of Food Protection*, vol. 71, no. 11, pp. 2339–2373, 2008.

[7] E. R. Rocha, C. D. Herren, D. J. Smalley, and C. J. Smith, "The complex oxidative stress response of *Bacteroides fragilis*: the role of OxyR in control of gene expression," *Anaerobe*, vol. 9, no. 4, pp. 165–173, 2003.

[8] A. S. Waage, T. Vardund, V. Lund, and G. Kapperud, "Detection of low numbers of pathogenic *Yersinia enterocolitica* in environmental water and sewage samples by nested polymerase chain reaction," *Journal of Applied Microbiology*, vol. 87, no. 6, pp. 814–821, 1999.

[9] V. Thibodeau, E. H. Frost, S. Chénier, and S. Quessy, "Presence of *Yersinia enterocolitica* in tissues of orally-inoculated pigs and the tonsils and feces of pigs at slaughter," *Canadian Journal of Veterinary Research*, vol. 63, no. 2, pp. 96–100, 1999.

[10] S. C. A. Chen and D. P. Kontoyiannis, "New molecular and surrogate biomarker-based tests in the diagnosis of bacterial

and fungal infection in febrile neutropenic patients," *Current Opinion in Infectious Diseases*, vol. 23, no. 6, pp. 567–577, 2010.

[11] R. Girones, M. A. Ferrús, J. L. Alonso et al., "Molecular detection of pathogens in water—the pros and cons of molecular techniques," *Water Research*, vol. 44, no. 15, pp. 4325–4339, 2010.

[12] J. Shendure and H. Ji, "Next-generation DNA sequencing," *Nature Biotechnology*, vol. 26, no. 10, pp. 1135–1145, 2008.

[13] S. M. Yoo and S. Y. Lee, "Diagnosis of pathogens using DNA microarray," *Recent Patents on Biotechnology*, vol. 2, no. 2, pp. 124–129, 2008.

[14] L. Cerqueira, N. F. Azevedo, C. Almeida, T. Jardim, C. W. Keevil, and M. J. Vieira, "DNA mimics for the rapid identification of microorganisms by fluorescence in situ hybridization (FISH)," *International Journal of Molecular Sciences*, vol. 9, no. 10, pp. 1944–1960, 2008.

[15] H. P. Dwivedi and L. A. Jaykus, "Detection of pathogens in foods: the current state-of-art and future directions," *Critical Review in Microbiology*, vol. 37, no. 1, pp. 40–63, 2011.

[16] A. Tsourkas and G. Bao, "Shedding light on health and disease using molecular beacons," *Brief Funct Genomic Proteomic*, vol. 1, no. 4, pp. 372–384, 2003.

[17] J. J. Adamczyk, L. C. Adams, and D. D. Hardee, "Field efficacy and seasonal expression profiles for terminal leaves of single and double *Bacillus thuringiensis* toxin cotton genotypes," *Journal of Economic Entomology*, vol. 94, no. 6, pp. 1589–1593, 2001.

[18] Z. Li, Y. Wang, J. Wang, Z. Tang, J. G. Pounds, and Y. Lin, "Rapid and sensitive detection of protein biomarker using a portable fluorescence biosensor based on quantum dots and a lateral flow test strip," *Analytical Chemistry*, vol. 82, no. 16, pp. 7008–7014, 2010.

[19] N. R. Thomson, S. Howard, B. W. Wren et al., "The complete genome sequence and comparative genome analysis of the high pathogenicity *Yersinia enterocolitica* strain 8081," *PLoS Genetics*, vol. 2, no. 12, pp. 2039–2051, 2006.

[20] J. Parkhill, B. W. Wren, N. R. Thomson et al., "Genome sequence of *Yersinia pestis*, the causative agent of plague," *Nature*, vol. 413, no. 6855, pp. 523–527, 2001.

[21] M. Eppinger, M. J. Rosovitz, W. F. Fricke et al., "The complete genome sequence of *Yersinia pseudotuberculosis* IP31758, the causative agent of Far East scarlet-like fever," *PLoS Genetics*, vol. 3, no. 8, pp. 1508–1523, 2007.

[22] T. M. Fuchs, K. Brandt, M. Starke, and T. Rattei, "Shotgun sequencing of *Yersinia enterocolitica* strain W22703 (biotype 2, serotype O:9): genomic evidence for oscillation between invertebrates and mammals," *BMC Genomics*, vol. 12, article 168, pp. 1–12, 2011.

[23] X. Wang, Y. Li, H. Jing et al., "Complete genome sequence of a *Yersinia enterocolitica* "old world" (3/O:9) strain and comparison with the "new world" (1B/O:8) strain," *Journal of Clinical Microbiology*, vol. 49, no. 4, pp. 1251–1259, 2011.

[24] J. Batzilla, U. Antonenka, D. Höper, J. Heesemann, and A. Rakin, "*Yersinia enterocolitica* palearctica serobiotype O:3/4—a successful group of emerging zoonotic pathogens," *BMC Genomics*, vol. 12, pp. 1–37, 2011.

[25] J. M. Rouillard and E. Gulari, "OligoArrayDb: pangenomic oligonucleotide microarray probe sets database," *Nucleic Acids Research*, vol. 37, no. 1, pp. D938–D941, 2009.

[26] S. B. Plaisier, R. Taschereau, J. A. Wong, and T. G. Graeber, "Rank-rank hypergeometric overlap: identification of statistically significant overlap between gene-expression signatures," *Nucleic Acids Research*, vol. 38, no. 17, pp. e169–e169, 2010.

[27] H. He, H. Zhang, X. Wang et al., "Development of a versatile, target-oriented tiling microarray assay for measuring allele-specific gene expression," *Genomics*, vol. 96, no. 5, pp. 308–315, 2010.

[28] M. N. Miller, B. W. Okaty, S. Kato, and S. B. Nelson, "Activity-dependent changes in the firing properties of neocortical fast-spiking interneurons in the absence of large changes in gene expression," *Developmental Neurobiology*, vol. 71, no. 1, pp. 62–70, 2011.

[29] R. Schwarz, B. Joseph, G. Gerlach et al., "Evaluation of one- and two-color gene expression arrays for microbial comparative genome hybridization analyses in routine applications," *Journal of Clinical Microbiology*, vol. 48, no. 9, pp. 3105–3110, 2010.

[30] S. Kim, J. Hu, Y. Oh et al., "Combining ChIP-chip and expression profiling to model the MoCRZ1 mediated circuit for Ca/calcineurin signaling in the rice blast fungus," *PLoS Pathogens*, vol. 6, no. 5, Article ID e1000909, 2010.

[31] M. Gauthier, B. Bonnaud, M. Arsac et al., "Microarray for hepatitis B virus genotyping and detection of 994 mutations along the genome," *Journal of Clinical Microbiology*, vol. 48, no. 11, pp. 4207–4215, 2010.

[32] B. Quiñones, M. S. Swimley, A. W. Taylor, and E. D. Dawson, "Identification of *Escherichia coli* O157 by using a novel colorimetric detection method with DNA microarrays," *Foodborne Pathogens and Disease*, vol. 8, no. 6, pp. 705–711, 2011.

[33] R. Z. Wu, S. N. Bailey, and D. M. Sabatini, "Cell-biological applications of transfected-cell microarrays," *Trends in Cell Biology*, vol. 12, no. 10, pp. 485–488, 2002.

[34] R. Natrajan, S. E. Little, J. S. Reis-Filho et al., "Amplification and overexpression of CACNA1E correlates with relapse in favorable histology Wilms' tumors," *Clinical Cancer Research*, vol. 12, no. 24, pp. 7284–7293, 2006.

[35] P. Simpson, C. Jones, A. Mackay, and S. R. Lakhani, "Gene expression analysis using filter cDNA microarrays," *Methods in molecular medicine.*, vol. 120, pp. 415–424, 2006.

[36] T. Hart, A. Zhao, A. Garg, S. Bolusani, and E. M. Marcotte, "Human cell chips: adapting DNA microarray spotting technology to cell-based imaging assays," *PLoS One*, vol. 4, no. 10, Article ID e7088, pp. 1–7, 2009.

[37] H. Lee, B. D. O'Connor, B. Merriman et al., "Improving the efficiency of genomic loci capture using oligonucleotide arrays for high throughput resequencing," *BMC Genomics*, vol. 10, article 646, pp. 1–12, 2009.

[38] L. H. Saal, C. Troein, J. Vallon-Christersson, S. Gruvberger, A. Borg, and C. Peterson, "BioArray Software Environment (BASE): a platform for comprehensive management and analysis of microarray data," *Genome Biology*, vol. 3, no. 8, Article ID SOFTWARE0003, pp. 1–6, 2002.

[39] H. Huang, W. L. Ma, X. Q. Dong, B. Zhang, Q. Wu, and W. L. Zheng, "DNA microarray for the detection of Yersinia pesits," *Di Yi Jun Yi Da Xue Xue Bao*, vol. 24, no. 1, pp. 47–49, 2004.

[40] K. M. Myers, J. Gaba, and S. F. Al-Khaldi, "Molecular identification of *Yersinia enterocolitica* isolated from pasteurized whole milk using DNA microarray chip hybridization," *Molecular and Cellular Probes*, vol. 20, no. 2, pp. 71–80, 2006.

[41] M. Ikeda, N. Yamaguchi, K. Tani, and M. Nasu, "Detection of food poisoning bacteria in fresh vegetables using DNA microarray," *Journal of Health Science*, vol. 52, no. 1, pp. 36–42, 2006.

[42] H. J. Kim, S. H. Park, T. H. Lee, B. H. Nahm, Y. R. Kim, and H. Y. Kim, "Microarray detection of food-borne pathogens using

specific probes prepared by comparative genomics," *Biosensors and Bioelectronics*, vol. 24, no. 2, pp. 238–246, 2008.

[43] M. Boye, A. A. Feenstra, C. Tegtmeier, L. O. Andresen, S. R. Rasmussen, and V. Bille-Hansen, "Detection of Streptococcus suis by *in situ* hybridization, indirect immunofluorescence, and peroxidase-antiperoxidase assays in formalin-fixed, paraffin-embedded tissue sections from pigs," *Journal of Veterinary Diagnostic Investigation*, vol. 12, no. 3, pp. 224–232, 2000.

[44] U. Gasanov, D. Hughes, and P. M. Hansbro, "Methods for the isolation and identification of *Listeria spp.* and *Listeria monocytogenes*: a review," *FEMS Microbiology Reviews*, vol. 29, no. 5, pp. 851–875, 2005.

[45] F. Dziva, A. P. Muhairwa, M. Bisgaard, and H. Christensen, "Diagnostic and typing options for investigating diseases associated with Pasteurella multocida," *Veterinary Microbiology*, vol. 128, no. 1-2, pp. 1–22, 2008.

[46] A. G. Gehring, P. L. Irwin, S. A. Reed et al., "Enzyme-linked immunomagnetic chemiluminescent detection of *Escherichia coli* O157:H7," *Journal of Immunological Methods*, vol. 293, no. 1-2, pp. 97–106, 2004.

[47] V. M. Bohaychuk, G. E. Gensler, R. K. King, J. T. Wu, and L. M. McMullen, "Evaluation of detection methods for screening meat and poultry products for the presence of foodborne pathogens," *Journal of Food Protection*, vol. 68, no. 12, pp. 2637–2647, 2005.

[48] L. M. Clotilde, C. Bernard IV, G. L. Hartman, D. K. Lau, and J. M. Carter, "Microbead-based immunoassay for simultaneous detection of Shiga toxins and isolation of *Escherichia coli* O157 in foods," *Journal of Food Protection*, vol. 74, no. 3, pp. 373–379, 2011.

[49] A. Roggenkamp, K. Ruckdeschel, L. Leitritz, R. Schmitt, and J. Heesemann, "Deletion of amino acids 29 to 81 in adhesion protein YadA of *Yersinia enterocolitica* serotype O:8 results in selective abrogation of adherence to neutrophils," *Infection & Immunity*, vol. 64, no. 7, pp. 2506–2514, 1996.

[50] B. Ngom, Y. Guo, X. Wang, and D. Bi, "Development and application of lateral flow test strip technology for detection of infectious agents and chemical contaminants: a review," *Analytical & Bioanalytical Chemistry*, vol. 397, no. 3, pp. 1113–1135, 2010.

[51] M. J. Taussig and U. Landegren, "Progress in antibody arrays," *Drug Discovery Today*, vol. 2, no. 4, pp. 169–176, 2003.

[52] C. Wingren and C. A. K. Borrebaeck, "Progress in miniaturization of protein arrays-a step closer to high-density nanoarrays," *Drug Discovery Today*, vol. 12, no. 19-20, pp. 813–819, 2007.

[53] K. L. Kellar and K. G. Oliver, "Multiplexed microsphere assays for protein and DNA binding reactions," *Methods in Cell Biology*, vol. 2004, no. 75, pp. 409–429, 2004.

[54] S. Derveaux, B. G. Stubbe, K. Braeckmans et al., "Synergism between particle-based multiplexing and microfluidics technologies may bring diagnostics closer to the patient," *Analytical & Bioanalytical Chemistry*, vol. 391, no. 7, pp. 2453–2467, 2008.

[55] M. Magliulo, P. Simoni, M. Guardigli et al., "A rapid multiplexed chemiluminescent immunoassay for the detection of *Escherichia coli* O157:H7, *Yersinia enterocolitica*, *Salmonella typhimurium*, and *Listeria monocytogenes* pathogen bacteria," *Journal of Agricultural and Food Chemistry*, vol. 55, no. 13, pp. 4933–4939, 2007.

[56] A. Sreekumar, M. K. Nyati, S. Varambally et al., "Profiling of cancer cells using protein microarrays: discovery of novel radiation-regulated proteins," *Cancer Research*, vol. 61, no. 20, pp. 7585–7593, 2001.

[57] V. C. Rucker, K. L. Havenstrite, and A. E. Herr, "Antibody microarrays for native toxin detection," *Analytical Biochemistry*, vol. 339, no. 2, pp. 262–270, 2005.

[58] B. Li, L. Jiang, Q. Song et al., "Protein microarray for profiling antibody responses to *Yersinia pestis* live vaccine," *Infection & Immunity*, vol. 73, no. 6, pp. 3734–3739, 2005.

[59] J. Hoorfar, D. L. Baggesen, and P. H. Porting, "A PCR-based strategy for simple and rapid identification of rough presumptive *Salmonella* isolates," *Journal of Microbiological Methods*, vol. 35, no. 1, pp. 77–84, 1999.

[60] M. R. Stratton, P. J. Campbell, and P. A. Futreal, "The cancer genome," *Nature*, vol. 458, no. 7239, pp. 719–724, 2009.

[61] S. Brenner, M. Johnson, J. Bridgham et al., "Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays," *Nature Biotechnology*, vol. 18, no. 6, pp. 630–634, 2000.

[62] T. D. Harris, P. R. Buzby, H. Babcock et al., "Single-molecule DNA sequencing of a viral genome," *Science*, vol. 320, no. 5872, pp. 106–109, 2008.

[63] W. Brockman, P. Alvarez, S. Young et al., "Quality scores and SNP detection in sequencing-by-synthesis systems," *Genome Research*, vol. 18, no. 5, pp. 763–770, 2008.

[64] J. A. Shendure, G. J. Porreca, and G. M. Church, "Overview of DNA sequencing strategies," *Current Protocols in Molecular Biology*, chapter 7: unit 7.1, no. 81, pp. 7.1.1–7.1.11, 2008.

[65] X. Zhou, L. Ren, Q. Meng et al., "The next-generation sequencing technology and application," *Protein & Cell*, vol. 1, no. 6, pp. 520–536, 2010.

[66] F. Sanger, "Sequences, sequences, and sequences," *Annual Review of Biochemistry*, vol. 57, pp. 1–28, 1988.

[67] K. V. Voelkerding, S. A. Dames, and J. D. Durtschi, "Next-generation sequencing:from basic research to diagnostics," *Clinical Chemistry*, vol. 55, no. 4, pp. 641–658, 2009.

[68] J. C. Dohm, C. Lottaz, T. Borodina, and H. Himmelbauer, "SHARCGS, a fast and highly accurate short-read assembly algorithm for *de novo* genomic sequencing," *Genome Research*, vol. 17, no. 11, pp. 1697–1706, 2007.

[69] M. J. Chaisson and P. A. Pevzner, "Short read fragment assembly of bacterial genomes," *Genome Research*, vol. 18, no. 2, pp. 324–330, 2008.

[70] D. Hernandez, P. François, L. Farinelli, M. Østerås, and J. Schrenzel, "*De novo* bacterial genome sequencing: millions of very short reads assembled on a desktop computer," *Genome Research*, vol. 18, no. 5, pp. 802–809, 2008.

[71] A. D. Smith, Z. Xuan, and M. Q. Zhang, "Using quality scores and longer reads improves accuracy of Solexa read mapping," *BMC Bioinformatics*, vol. 9, article 128, pp. 1–8, 2008.

[72] R. L. Warren, G. G. Sutton, S. J. M. Jones, and R. A. Holt, "Assembling millions of short DNA sequences using SSAKE," *Bioinformatics*, vol. 23, no. 4, pp. 500–501, 2007.

[73] A. V. Zimin, D. R. Smith, G. Sutton, and J. A. Yorke, "Assembly reconciliation," *Bioinformatics*, vol. 24, no. 1, pp. 42–45, 2008.

[74] J. R. Miller, A. L. Delcher, S. Koren et al., "Aggressive assembly of pyrosequencing reads with mates," *Bioinformatics*, vol. 24, no. 24, pp. 2818–2824, 2008.

[75] K. Prüfer, U. Stenzel, M. Dannemann, R. E. Green, M. Lachmann, and J. Kelso, "PatMaN: rapid alignment of short sequences to large databases," *Bioinformatics*, vol. 24, no. 13, pp. 1530–1531, 2008.

[76] B. D. Ondov, A. Varadarajan, K. D. Passalacqua, and N. H. Bergman, "Efficient mapping of Applied Biosystems SOLiD sequence data to a reference genome for functional genomic

applications," *Bioinformatics*, vol. 24, no. 23, pp. 2776–2777, 2008.

[77] D. Campagna, A. Albiero, A. Bilardi et al., "PASS: a program to align short sequences," *Bioinformatics*, vol. 25, no. 7, pp. 967–968, 2009.

[78] J. Butler, I. MacCallum, M. Kleber et al., "ALLPATHS: *de novo* assembly of whole-genome shotgun microreads," *Genome Research*, vol. 18, no. 5, pp. 810–820, 2008.

[79] N. Navin and J. Hicks, "Future medical applications of single-cell sequencing in cancer," *Genome Medicine*, vol. 3, no. 5, pp. 1–12, 2011.

[80] A. C. Fierro, F. Vandenbussche, K. Engelen, Y. Van de Peer, and K. Marchal, "Meta analysis of gene expression data within and across species," *Current Genomics*, vol. 9, no. 8, pp. 525–534, 2008.

[81] Y. Surget-Groba and J. I. Montoya-Burgos, "Optimization of *de novo* transcriptome assembly from next-generation sequencing data," *Genome Research*, vol. 20, no. 10, pp. 1432–1440, 2010.

[82] F. B. Rahmatpanah, S. Carstens, S. I. Hooshmand et al., "Large-scale analysis of DNA methylation in chronic lymphocytic leukemia," *Epigenomics*, vol. 1, no. 1, pp. 39–61, 2009.

[83] D. S. Johnson, A. Mortazavi, R. M. Myers, and B. Wold, "Genome-wide mapping of in vivo protein-DNA interactions," *Science*, vol. 316, no. 5830, pp. 1497–1502, 2007.

[84] O. Wurtzel, M. Dori-Bachash, S. Pietrokovski, E. Jurkevitch, and R. Sorek, "Mutation detection with next-generation resequencing through a mediator genome," *PLoS One*, vol. 5, no. 12, Article ID e15628, 2010.

[85] J. M. Kidd, Z. Cheng, T. Graves, B. Fulton, R. K. Wilson, and E. E. Eichler, "Haplotype sorting using human fosmid clone end-sequence pairs," *Genome Research*, vol. 18, no. 12, pp. 2016–2023, 2008.

[86] C. A. Cummings, C. A. Chung, R. Fang et al., "Accurate, rapid and high-throughput detection of strain-specific polymorphisms in *Bacillus anthracis* and *Yersinia pestis* by next-generation sequencing," *Investigative Genetics*, vol. 1, no. 1, article 5, 2010.

[87] G. Morelli, Y. Song, C. J. Mazzoni et al., "*Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity," *Nature Genetics*, vol. 42, no. 12, pp. 1140–1143, 2010.

[88] P. E. Chen, C. Cook, A. C. Stewart et al., "Genomic characterization of the *Yersinia* genus," *Genome Biology*, vol. 11, no. 1, article R1, 2010.