Challenges of microarray applications for microbial detection and gene expression profiling in food

Yanhong Liu1*, Yan Wang2 and Jie Gang3

1Molecular Characterization of Foodborne Pathogen Research Unit, Eastern Regional Research Center, 600 East Mermaid Lane, Wyndmoor, PA 19038, USA
2Science Research Center, Harbin Institute of Technology, Harbin, Heilongjiang 150001, P. R. China
3Food Engineering Department, College of Life Sciences, Dalian Nationalities University, Dalian, 116000, P. R. China

Abstract

Microarrays are a high throughput technology used in molecular biology. The different types of microarrays have been applied to clinical and environmental microbiology, microbial ecology, human, veterinary, and plant diagnostics. Since multiple genes can be analyzed simultaneously, this technology has been extended to food microbiology for detection and gene expression analyses of food-borne pathogens. Although this technique has multiple applications, microarray technology presently has low sensitivity and also suffers from problems such as reproducibility and reliability. This paper focuses on microarray applications for microbial detection and gene expression profiling of food-borne pathogens including some of the challenges and key issues such as target nucleic acid isolation and selection of target DNA sequences.

Keywords: Microarray; Real-time PCR; Detection; Gene expression; Food

Introduction

Microarray technologies that have been developed since 1995 are very powerful tools to simultaneously analyze a large number of genes or target DNA sequences. Whole-genome sequencing of a number of food-borne pathogens including E. coli, Salmonella, spp. and Listeria, spp., has allowed the construction of microarray chips for use in food microbiology. Microarray applications for food-borne pathogen detection, food quality, and food safety have been extensively reviewed [1-5]. The main applications of microarray technology in foods are: i) analysis of bacterial gene expression, and ii) detection of specific microorganisms. However, there is no review covering microarray applications in both aspects. This review focuses on current applications of microarray technologies (mainly from our laboratory) in both microbial detection and gene expression in food.

The most common microarray application is differential gene expression analyses. The first step in these types of analyses is extraction of RNA from cells. In bacteria, total RNA is usually used as the starting material, although protocols for purification of bacterial mRNA have been developed [6]. Briefly, RNA from two samples (control vs. treated) are reverse transcribed to cDNA and then labeled with different fluorescence dyes either during or after reverse transcription. The most commonly used labels include the fluorophors, Cy3 and Cy5. The different fluorescent labeled cDNA samples (control vs. treated) are mixed and hybridized to a microarray chip containing oligonucleotides or PCR amplified DNA fragments covering the entire genome of a bacterial pathogen. After washing off the non-specific signals, the microarray chip is scanned using a microarray scanner and the fluorescence signals of each spot on the microarray are quantified (Figure 1). After intensive data normalization and statistical analysis, the ratio of the intensities of both dyes (e.g., Cy3/Cy5 and Cy5/Cy3) reflects the relative expression levels for both control and treated samples. The usage of two dyes allows for the quantitative comparison of two samples on a single chip. To avoid dye bias resulting from differences in the incorporation efficiency of the fluorescent dyes, a dye-swamp experiment representing a technical replicate of microarray chip in which the same samples are hybridized but with a reverse labeling is usually performed [7]. Dye swap with true biological replicates achieves the same goals but with greater statistical power. In addition to two channel arrays, one-channel arrays, e.g. Affymatrix and Illumina are also very popular and commonly used in academic and industry setting.

Microarrays can be used to determine whether a specific sequence is present in a sample. The detection microarray chips can cover an entire genome, or a subset of genes from several genomes. However, the most commonly used microarrays for detection of microorganisms contain a limited number of signature sequences in a microorganism and are used for species identification or differentiation among strains. Procedures using microarrays for detection start with purification of genomic DNA from target cells, followed by PCR amplification or whole genome amplification, and then labeling with a fluorescent dye. For this type of analysis, only one fluorescence dye is needed because the assay is designed to detect whether or not a sequence is present in the sample (Figure 2). After hybridization and washing, the microarray chips are scanned using a microarray scanner and the fluorescence signals are quantified.

Microarrays possess considerable advantages and potential for the detection and characterization of pathogenic microorganisms. For example, multiple microorganisms can be detected simultaneously from a sample, and the presence of virulence genes and antibiotic resistance genes carried by a pathogen can also be detected. Various types of microarrays have been developed and evaluated for pathogen detection in food and environmental samples and for biodefense applications [8,9]. Genetic markers used for microarray-based...
microbial detection include but are not limited to housekeeping genes such as 16S-23S rDNA [10], gyrB genes [11], and virulence genes [12]. Microarray techniques have been used to distinguish among different bacterial genera [13], as well as among different strains within the same species such as Listeria monocytogenes [14,15] and Salmonella serovars [16]. Some innovative approaches in detection arrays also include: (a) simultaneous detection of host genes for pathogen identification and virulence/resistance genes for assessment of host potential to develop an illness, (b) horizontal transfer of genetic markers among related and not-related bacterial strains.

Although microarrays have been successfully used for microbial detection in pure-culture studies [17], there are technical difficulties when microarrays are applied to real food samples. Two main issues are discussed in this review: isolation of a sufficient amount of target genomic DNA and the selection of DNA probe sequences for the detection chips.

Isolation of sufficient target genomic DNA from food

A significant challenge in the development of methods for microbial detection in food is the need to detect and characterize a small number of target microorganism(s) in a large sample volume, often with a background flora consisting of similar microorganisms and the presence of interfering food matrix material. Microarrays require a relatively large amount (microgram levels) of DNA, but the fact that food samples usually contain only small numbers of pathogens makes it very difficult to isolate enough genomic DNA. Traditional cultural enrichments of microorganisms in food samples are usually necessary [18].

One way to increase the yield of target DNA is selective amplification of the specific sequences. If the probes on the chip represent a specific locus or a few loci of the gene sequences, target sequences covering the entire locus can be amplified by the PCR. Hybridization of the amplified PCR products to the microarrays will result in greater specificity and sensitivity. For example, the entire O-antigen gene cluster sequences of different E. coli serogroups were amplified by long PCR and labeled before hybridizing the labeled PCR product to microarrays containing oligonucleotides (35-mers) of the specific O-antigen gene clusters [19].

If the probes on the chip represent a number of sequences from different loci, whole genome amplification methods can be used to amplify the target sequences before applying to microarrays. Multiple displacement amplification (MDA) is an isothermal amplification method that uses the bacteriophage Phi29 DNA polymerase and random primers [20]. The strand-displacement synthesis yielded about 20-30 µg of product (>10kb in length) from as few as 1-10 copies of human genomic DNA [21]. MDA can be performed directly on biological samples and has been applied to amplify bacteria in water [22]. Although MDA permits microbial detection in a food that contains low levels of bacterial DNA, it may not increase the specificity of the microarray assays since the whole genome, including the selected target sequences on the chip, is amplified. In addition, some loci may be preferentially amplified by MDA while other loci have low amplification. Recently, a new linear whole genome amplification method (NuGEN Ovation Whole Genome Amplification) that only requires 10 ng of input genomic DNA was developed. The ability to linearly amplify genomic DNA from limited amounts of template DNA offers the possibility for microbial detection in real food matrices.

DNA probe sequences for microbial detection

Microbial detection is based on the ability of a microarray to identify sequences present in the target organism. Typically, microarrays used for detection/characterization of food-borne pathogens contain less than a few hundred probes with well defined target sequences. This is in contrast to gene expression microarrays that often contain the probes representing all genes/loci of bacteria. Since food samples may contain a multitude of non-target genomes, the detection microarrays must contain the sequences that are unique to each microorganism (i.e., signature sequences) with respect to all non-target genomes. Such signature sequences (probes) provide information regarding the identification of the microorganism or its important characteristics. It is a challenge to identify the signature sequences of each microorganism, since a majority of bacterial genome sequences are conserved. The design of the signature sequences (probes) requires identification of unique target sequences through genomic database searches and multiple alignment analysis, and then the design of oligoprobes to represent these sequences on the array. A number of computer programs have been developed for this bioinformatic task.
For bacteria in which the whole genome sequence is not available, gene clusters from each bacterium can be sequenced and compared to identify the sequences that are unique to each bacterium. For example, the O-antigen gene clusters of different E. coli serogroups were sequenced and the OligoArraySelector program [23] was used to select the probe sequences unique to different E. coli serogroups [19].

For microorganisms for which the whole genome sequence has been determined, several software packages are available to identify unique sequences (biomarkers representing each microorganism). For example, a multigenome software package named Tool for Oligonucleotide Fingerprint Identification (TOFI) was developed and has been used to identify DNA fingerprints in Bacillus anthracis, Yersinia pestis, Francisella tularensis and Burkholderia mallei and Burkholderia pseudomallei genomes. Furthermore, the accuracy of these in silico DNA fingerprints was validated by microarray hybridization experiments [24-27]. Targeting these DNA fingerprints greatly increased the specificity of the microarray assays.

A case study: microarray application in E. coli O antigen typing

E. coli serotyping is conventionally performed by agglutination reactions using antisera raised in rabbits against the 179 different O standard references strains. However, traditional serotyping is laborious, time consuming, and often generates equivocal results due to cross-reactions among different serogroups. Furthermore, the antisera for serotyping can only be generated by specialized laboratories with animal facilities. Rapid molecular methods for identifying different E. coli serogroups are needed. The O antigen, which contains many repeats of an oligosaccharide unit (O unit) is present in the outer membrane of Gram-negative bacteria and contributes the major antigenic variability to the cell surface. The genes involved in the biosynthesis of O antigens in E. coli are located in the O-antigen gene cluster and are flanked by the galf and gnd genes on the E. coli chromosome. A number of E. coli O-antigen gene clusters have been sequenced and the genes were annotated [28-32]. Several genes in the clusters, in particular the wex (O antigen flippase) and wzy (O antigen polymerase) genes, show relatively low similarity among different E. coli serogroups, and PCR primers targeting the wex and wzy genes have been used to develop serogroup-specific PCR assays [28-31]. However, since each primer is specific for the respective O-antigen gene cluster genes, PCR assays targeting each of the serogroups need to be performed to identify the E. coli strain. Therefore, an assay for parallel detection of multiple serogroups in a single platform would be useful. To this end, DNA microarrays were developed for rapid identification of different serogroups of E. coli in a single platform. Oligonucleotides, as well as PCR products from genes in the O-antigen gene clusters of E. coli serogroups O7, O104, O111, and O157 were spotted onto glass slides. This was followed by hybridization with labeled long PCR products of the entire O-antigen genes clusters of these serogroups. Results demonstrated that microarrays consisting of either oligonucleotides or PCR products generated specific signals for each serogroup [19]. Since many O antigen gene clusters have been sequenced, targets specific for the sequenced gene clusters can be selected and placed onto the typing array, and there is the potential to expand the DNA array to contain target sequences for all of the E. coli serogroups. This research highlights the importance of microarray technology in pathogen typing.

Gene expression microarray applications in food matrices

The availability of whole genome sequence information for several food-borne pathogens has paved the way for global analyses of microbial adaptation to various environments [33]. For example, microarrays were used to study the responses of L. monocytogenes and B. subtilis to cold stress, and a number of genes that are differentially expressed at low temperatures were identified [34,35]. However, genomic studies of food-borne pathogens in a real food matrix are generally lacking.

There are some crucial challenges to study gene expression of food-borne pathogens in food using microarrays. For example, bacteria usually are present in food in low numbers; therefore, RNA from bacterial cells is not sufficient for target labeling and hybridization for microarrays. Large numbers of bacterial cells must be used to artificially contaminate food to achieve enough RNA. Prior to hybridization, some type of signal amplification is needed. One method, termed whole–community RNA amplification (WCRA) has been developed to provide sufficient amounts of prokaryotic RNAs from environmental samples for microarray analysis [36]. In this approach, a T7 RNA promoter sequence is attached to fusion primers (six to nine random nucleotides), which is then used for reverse transcription of RNA to cDNA. The synthesized cDNA were in turn used for linear RNA amplification with T7 RNA polymerase. About 1200- to 1800-fold amplification was obtained from 10 to 100 ng RNA templates, and very representative detection was obtained with 50 to 100 ng total RNA [36].

Separation of bacteria from food matrices

Another challenge is to separate bacterial pathogens from food matrices. Differential separation and centrifugation are effective ways to separate bacteria from milk. We also developed a system using dialysis tubing containing a high level (10⁹ cfu/ml) of L. monocytogenes to study gene expression in milk. Dialysis tubing with a high molecular weight cutoff (1 million Daltons) was placed in milk. The high cutoff allows the small milk components to travel freely, while keeping the bacteria inside the tubing. Using this system, a sufficient quantity of RNA was isolated from L. monocytogenes and used for microarray experiments to identify the genes expressed in milk [37]. In addition, a filtration step combined with immunomagnetic separation (IMS) has been used to successfully separate Listeria monocytogenes from salmon matrix [38,39]. Bacterial RNA has been isolated successfully from other solid foods such as liver pates and cheese [40,41]. Furthermore, the physiological properties of bacteria can be altered by sampling processing time. In E. coli, rapid changes in levels of specific mRNAs have been observed as a result of changing environmental conditions [42].

Microarrays have been used to study how food-borne pathogens survive and grow in food. For example, microarrays were used to monitor the gene expression profiles of L. monocytogenes strain F2365 in Ultra-High Temperature (UHT) processed skim milk. Total RNA was isolated from strain F2365 in UHT skim milk after 24 hours at 4°C, reverse transcribed, labeled with fluorescent dyes, and hybridized to oligonucleotide (35-mers) microarray chips containing the whole genome of L. monocytogenes strain F2365 [43]. The electrochemically synthesized oligonucleotide microarrays used in this study [37] was CustomArray™ 12K arrays (Combimatrix, Mukilteo, WA). The arrays were constructed to include 35-mer oligonucleotides representing the 2847 L. monocytogenes open reading frames (ORFs). Compared to the spotted microarrays, these in situ custom synthesized microarrays have the advantages of low cost, accuracy, flexibility, and reusability [44]. Genes that were up- (26 genes) and down-regulated (14 genes) in UHT pasteurized skim milk were identified. Furthermore, the gene expression changes determined by microarray assays were confirmed by real-time RT-PCR analyses [37]. The identified genes will be good

Page 3 of 6
Microarrays were used to analyze the transcriptional profiling of bacteria in food because it does not adversely affect the sensory properties or the quality of food. This approach is an alternative to heat and other traditional food preservation methods. Genes encoding energy production and conversion and carbohydrate metabolism were studied during growth of E.coli, and virulence-associated genes were identified simultaneously by the elevated pressure. In addition, an approach was used to identify protein signatures of Bacillus anthracis Sterne strain under different interventions such as heat, pasteurization, and microfiltration. Genes related to spore germination and sporulation have been identified. These genes could help food processors develop effective preservation strategies to better manage pathogens in food. Since food systems are complex with different nutritional components, bacterial pathogens in different foods may grow differently. This study can be extended to evaluate the gene expression profiling of different bacterial pathogens in different food matrices. Eventually, this study will contribute to detailed and fundamental understanding of the mechanisms of bacterial growth and survival in food. Other food matrices used for bacterial gene expression included lettuce leaves and sourdoughs.

In addition to studying food-borne pathogens, microarray approaches have been used to study probiotic microbes in food such as Lactobacillus acidophilus in skim milk. Probiotics are live microorganisms that confer a health benefit to the host when administered in adequate amounts. Probiotics can also control undesirable microorganisms in food animals. Gene expression profiles were studied during growth of Lactobacillus helveticus mutant 2-1 versus a rich culture medium. Moreover, proteomic approaches were used to identify protein signatures of Lactococcus lactis in milk. Lactobacillus sakei genes that were induced during meat fermentation have also been identified. Survival of probiotic strains during storage is a key issue for the food industry. Information from these studies could help in understanding how the probiotic bacteria survive in food and to provide insights on how to improve survival of these bacteria in food.

**Microarray applications in food processing and interventions**

High pressure processing (HPP) has been introduced as an alternative to heat and other traditional food preservation methods because it does not adversely affect the sensory properties or the freshness attributes of certain foods. Several research groups investigated the mechanism of action of high pressure against bacterial cells using microarrays. Different bacterial and different pressure conditions have been used, results are difficult to compare. Microarrays were used to analyze the transcriptional profiling of E. coli grown at 30 and 50 MPa. Both heat and cold stress responses were induced simultaneously by the elevated pressure. In addition, an E. coli mutant, with a deletion in the hup gene encoding a DNA binding protein, exhibited great pressure sensitivity. Sublethal pressure (100 MPa, 15 min) was used to study the transcriptional profiling of E. coli O157:H7. High pressure affected the transcription of many genes including those involved in stress responses, the thiol-disulfide redox system, and Fe-S cluster assembly. Microarrays were used to study the differential gene expression of L. monocytogenes strain S2542 during high hydrostatic pressure (400 MPa and 600 MPa, 5 min). HPP induced increased expression of genes associated with DNA repair mechanisms, transcription and translation protein complexes, the septal ring, the general protein translocase system, flagella assembly and chemotaxis, and lipid and peptidoglycan biosynthetic pathways. On the other hand, HPP suppressed a wide range of genes associated with energy production and conversion and carbohydrate metabolism and virulence-associated genes.

In L. monocytogenes, microarrays have been used to determine the role of the ctsR gene in resistance to high pressure. The hup gene encodes a class III heat shock protein which is the first gene of the four-gene clpC operon. Under normal conditions, the CtsR protein negatively regulates the clpC, clpP and clpE genes by binding specifically to the regulatory regions of these genes. Under stressed conditions, the CtsR protein was degraded by the ClpCP protease. CtsR has been shown to be related to high pressure since several pressure-tolerant mutants contained mutations in this gene. A cts mutant (AK01) of L. monocytogenes Scott A containing a glycine deletion was non-motile and showed greater resistance to high pressure, heat, acid, and H2O2 than the wild type. The L. monocytogenes Scott A cts mutant (2-1) exhibited a 100-fold higher level of variability than the wild type when exposed to 450 MPa. The mutant 2-1 had a deletion in the ctsR gene that resulted in truncation of CtsR of 20 amino acids in CtsR protein. The N-terminal truncation of CtsR lacked the DNA binding and heat sensing domains. Compared to the wild type, the mutant 2-1 showed increased cell length and had no flagella. However, the mechanism of its survival under high pressure is unclear. Microarrays were used to study the gene expression of mutant 2-1 under high pressure. Total RNA was isolated from the pressure-treated (450 Mpa, 3 minutes) ctsR mutant 2-1, reverse transcribed, labeled with fluorescent dyes, and hybridized to commercial oligonucleotide (35-mers) microarray chips representing the whole genome of L. monocytogenes. The gene expression changes determined by microarray assays were confirmed by the real-time RT-PCR analyses. Compared to the pressure-treated L. monocytogenes Scott A wild type, 19 genes were up-regulated (> 2-fold increase) in the ctsR deletion mutant whereas 57 genes were down-regulated (< -2-fold decrease). The up-regulated genes included genes encoding for a transcriptional regulator, ATP-dependent Clp protease, putative accessory gene regulator proteins B and D, transport and binding proteins, phosphotransferase system (PTS) fructose-specific enzyme IIABC components, protein synthesis, and hypothetical proteins. The down-regulated genes included genes that encode for PTS system, mannose-specific IIABC components, transport and binding proteins, flagella synthesis-related proteins, a transcriptional regulator, and hypothetical proteins. The gene expression changes determined by microarray assays were confirmed by the real-time RT-PCR analyses. This study highlights the importance of CtsR in pressure tolerance and will contribute to the appropriate design of safe, effective, and feasible high hydrostatic pressure (HHP) treatments.

In addition to microarrays, other genomic technologies have been used to study food-borne pathogens and food-related bacteria. For example, in vivo expression technology (IVET) was used to identify Lactobacillus sakei genes that were induced during meat fermentation. Real-time PCR assays were used to study the spores of Bacillus anthracis Sterne strain under different interventions such as heat, pasteurization, and microfiltration. Genes related to spore germination and sporulation have been identified. These genes could be used as potential targets for gene knock-out studies to identify the role of the genes. Understanding how bacteria survive in different food systems may help food processors develop effective preservation strategies to better manage pathogens in food. Since food systems are complex with different nutritional components, bacterial pathogens in different foods may grow differently. This study can be extended to evaluate the gene expression profiling of different bacterial pathogens in different food matrices. Eventually, this study will contribute to detailed and fundamental understanding of the mechanisms of bacterial growth and survival in food. Other food matrices used for bacterial gene expression included lettuce leaves and sourdoughs.

---


---

**Figure 3:** Morphological characteristics of exponentially grown cells of L. monocytogenes Scott A wild type (A) and mutant (2-1) with scan electron microscope (SEM). Bars, 1µm. Adapted from Liu et al. [60] with permission.
serve as biomarkers for pasteurization and microfiltration [62]. These approaches can complement microarray assays and provide more information on bacterial survival in food environments.

Conclusions and Future Perspectives

Microarrays hold promise for food safety applications. To make microarrays more practical for use with food samples, we face the challenge of isolating sufficient amounts of bacterial DNA or RNA from food and improving the fluorescent labeling of target DNA or cDNA so that we can reduce the requirement for micrograms of RNA or DNA to nanogram levels. Although amplification of RNA and DNA has been successfully developed [21,22,36] these approaches are mainly limited to pure cultures. Additional work is needed to improve the sensitivity and specificity before these approaches can be used in food samples. Since microarray analysis also suffers from problems associated with reproducibility, reliability, compatibility and standardization of results [63], it is usually confirmed by a complementary platform such as real-time PCR assays. The high costs of reagents (microarray chip and labeling kits) and of instruments (scanner and hybridization chambers) also limit the use of microarrays in food safety. More efforts should be directed toward making detection microarrays affordable so that they can be routinely used in food safety. Another important area in microarray is data analysis. In food industry setup, high-level data analysis and proper interpretation could potentially lead to an identification of targets for interventions.

In gene expression studies, microarrays are limited to microorganisms whose genomes have been fully sequenced and characterized. In addition, microarrays only limit expression studies to the level of RNA, but do not allow distinction between de novo synthesized transcripts and modified transcripts. Given these limitations, microarrays are now being superseded by transcriptomics based on next-generation sequencing technologies [64]. In the future, microarrays should be used together with other emerging approaches such as next-generation sequencing technologies, thus can provide complete pictures of food-borne pathogens in food.

Acknowledgments

The author would like to thank Drs. Pina Fratamico, John B. Luchansky, and James Smith (USDA, Agricultural Research Services, Eastern Regional Research Center, Wyndmoor, PA) for critical reading of the manuscript and helpful discussions.

References

of the Escherichia coli O121 O-antigen gene cluster and detection of enterohemorrhagic E. coli O121 by PCR amplification of the wzx and wzy genes. J Clin Microbiol 41: 3379-3383.


Submit your manuscript and get advantages of OMICS Group submissions

Unique features:

• User-friendly/feasible website-translation of your paper to 50 world’s leading languages
• Audio Version of published paper
• Digital articles to share and explore

Special features:

• 100 Open Access Journals
• 10,000 editorial team
• 21 days rapid review process
• Quality and quick editorial, review and publication processing
• Indexing in PubMed (portal), Scopus, DOAJ, EBSCO, Index Copernicus and Google Scholar etc.
• Sharing Option: Social Networking Enabled
• Authors, Reviewers and Editors rewarded with online Scientific Credits
• Better discount for your subsequent articles

Submit your manuscript at: http://www.editorialmanager.com/jmbt