FOOD BIOLOGICAL CONTAMINANTS

DNA Microarray Technology Used for Studying Foodborne Pathogens and Microbial Habitats: Minireview

SUFIAN F. AL-KHALDI
U.S. Food and Drug Administration, CFSAN, Division of Microbiological Studies, HFS-517, 5100 Paint Branch Pkwy, College Park, MD 20740-3855
SCOTT A. MARTIN
University of Georgia, Department of Animal and Dairy Science, 312 Animal and Dairy Science Complex, Athens, GA 30602-2771
AVRAHAM RASOOLY
U.S. Food and Drug Administration, CFSAN, Division of Microbiological Studies, HFS-517, 5100 Paint Branch Pkwy, College Park, MD 20740-3855
JEFF D. EVANS
U.S. Department of Agriculture, Richard B. Russell Agricultural Research Center, PO Box 5677, Athens, GA 30604

Microarray analysis is an emerging technology that has the potential to become a leading trend in bacterial identification in food and feed improvement. The technology uses fluorescent-labeled probes amplified from bacterial samples that are then hybridized to thousands of DNA sequences immobilized on chemically modified glass slides. The whole gene or open reading frame(s) is represented by a polymerase chain reaction fragment of double-strand DNA, approximately 1000 base pair (bp) or 20–70 bp single-strand oligonucleotides. The technology can be used to identify bacteria and to study gene expression in complex microbial populations, such as those found in food and gastrointestinal tracts. Data generated by microarray analysis can be potentially used to improve the safety of our food supply as well as ensure the efficiency of animal feed conversion to human food, e.g., in meat and milk production by ruminants. This minireview addresses the use of microarray technology in bacterial identification and gene expression in different microbial systems and in habitats containing mixed populations of bacteria.

During the past 10 years, the complete nucleotide base sequence of 37 microbial genomes has been published, and it is estimated that there are 128 microbial genome sequencing projects currently under way (http://www.tigr.org/tdb/mdb/mdbcomplete.html; 1, 2). Sequencing those different microbial genomes, as well as human genomes and plants (e.g., *Arabidopsis*), has generated large amounts of data. Fast-capacity screening technologies, such as DNA microarray hybridization, can incorporate these data in complex investigations to compare gene expression levels and to test cells and study genes that work together when activated and repressed by stress or environmental stimuli (networking genes).

An example of a DNA microarray hybridization strategy is shown in Figure 1. Publicly available sequences of bacterial genomes are used to synthesize 20–70 single-strand nucleotide target sequences representing different genes, which are then spotted on glass slides. Two primers are designed by sequence alignment of conserved regions of gene families, and fluorescent-labeled DNA probes are amplified and hybridized to the target sequence, which has been spotted on the area of the glass slide, conventionally called the "chip." This strategy is used mainly to identify bacterial genes (3).

A second approach to study gene expression relies on generating a chip spotted with 500–5000 base pairs (bp) of polymerase chain reaction (PCR)-amplified fragments or synthesized single-strand oligonucleotides representing genes. Fluorescent-labeled DNA probes containing complementary DNA (c-DNA) are amplified by reverse transcriptase PCR (RT-PCR) from RNA isolated from bacterial strain samples. Subsequently, the Flurolink Cy5 d(C/U)TP (control) and Flurolink Cy3 d(C/U)TP (sample) probes are hybridized to the DNA on the glass chip by using automated hybridization stations. The relative signal from each spot is quantitated with confocal LASER beam scanners, and the pattern of expression is detected and confirmed.

Microarray technology is useful as a tool to solve complex problems in many microbial habitats (4), with potential appli-

Received August 14, 2001. Accepted by AH January 18, 2002.
Corresponding author's e-mail: Sufian.Al-Khaldi@cfsan.fda.gov.

Downloaded from https://academic.oup.com/jaoac/article/85/4/906/5656851 by guest on 21 August 2022

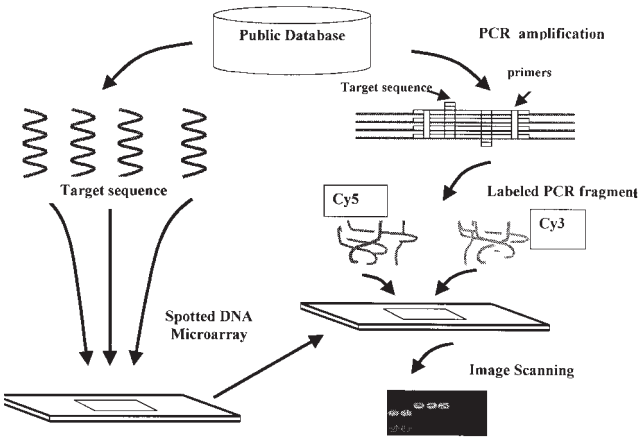

Figure 1. Steps involved in preparing a gene chip by using DNA microarray technology.

cation for improving food and feed safety in humans and animals, respectively.

Characterization of Bacterial Pathogens Contaminating Food

Pathogenic microorganisms are identified routinely by using PCR, surrogate biochemical, and immunological markers. Many PCR assays have been developed for the characterization and identification of microbial pathogens, such as *Escherichia coli* O157:H7, a major food pathogen (5). Common target genes for PCR amplification include the conserved regions of the *slt*-I and -II toxin structural genes (6) and the *eaeA* (intimin) gene (7), whose product mediates the adherence of O157:H7 to host cells. The *slt*-I and -II and *eaeA* genes in *E. coli* are not unique for serotype O157:H7 (7–14); more specific target genes have also been used, including *rfbE*, which encodes for an enzyme involved in biosynthesis of the O157 antigen (15, 16), and *fliC*, which encodes the H7 antigenic flagellin. Combining all of these virulence factors on one DNA microarray chip may allow for simultaneous analysis and eliminate the need for more time-consuming individual gene-specific PCR amplifications.

Another approach for the identification of food pathogens is the use of ribosomal DNA (rDNA) sequences that are available publicly from the GeneBank database. Microarray-based rDNA identification of bacterial contaminants allows fast and reliable detection and identification of contaminating pathogens, thus improving food safety.

Microbial Target Gene Amplification for DNA Microarray Analysis

One of the main potential uses of DNA microarray analysis is the simultaneous characterization and identification of multiple food bacterial contamination. To perform this task, multiple potential genes, such as virulence- and antibiotic-resistant genes have to be amplified simultaneously by using a DNA template. Until now, a common approach for amplification of multiple DNA sequences has been multiplex PCR, where multiple DNA sequences are amplified simultaneously by using more than one set of primers.

A multiplex PCR method has been developed, which directly detects genes that are involved in biosynthesis of O157 and H7 antigens, along with the *slt*-I and -II toxins and the outer membrane protein intimin (17). Although this approach is effective for analysis of bovine feces for *E. coli* O157:H7 (17), the number of different amplifications is limited because of primer–primer interactions when multiple sets of primers are present in the same reaction mixture (18). Other limitations include nonspecific amplification that results in multiple bands and problems with identification of the PCR product because of DNA fragment size similarity. DNA microarray analysis of multiplex PCR products may overcome these problems because it is based on DNA hybridization. Thus, the amplified PCR product must match the control sequence to be considered positive rather than relying on a gross size estimate for identification. In addition, it is quite possible to amplify the whole genome during DNA labeling by using random hexamers (19). This solves both problems of nonspecific amplification and DNA fragment size analysis.

In a recent study (3), the following 6 genes were labeled by fluorescent dyes by using multiplex PCR and were hybridized to 25 bp gene-specific targets on a microchip: *eaeA* encodes for outer membrane, *slt*-I and -II for shiga-like toxins, *fliC* for H7 antigen, *rfbE* for an enzyme involved in the biosynthesis of the O157 antigen, and *ipaH* for invasion plasmid antigen. The chip could uncover virulence factor genes in 15 *Salmonella*, *Shigella*, and *E. coli* pure culture strains. Use of the bacterial chip was a relatively fast, flexible, and reliable way to characterize virulence genes compared with multiplex PCR.

Research has been performed to detect the 16S ribosomal RNA (rRNA) of microarray bacterial chips by using RNA isolated from bacterial environmental samples (20). A short single-strand 20 bp oligonucleotide sequence representing 16S rRNA from *Geobacter chapellei* and *Desulfovibrio desulfuricans* was arrayed on a chip. The RNA was incubated with a labeled (biotin) probe, which hybridizes to an RNA sequence located near the hybridization site of the target sequence (Figure 2). The whole mixture was then hybridized to the microarray chip, and the biotin-labeled RNA could hybridize to a different target sequence on the chip. The detection sensitivity level was 0.5 μg total RNA isolated from the samples containing *G. chapellei* and *D. desulfuricans* environmental bacteria, and this represented approximately 7.5×10^6 cells (20).

In another study, random genome fragments arrayed on a chip were used to identify different bacterial species (19). Bacterial genomes were fragmented and cloned randomly, 1 kb random fragments were spotted on a glass slide chip, and the DNA isolated from pure bacterial cultures was labeled with Cy3 fluorescent dye by using PCR DNA labeling using

Downloaded from https://academic.oup.com/jaoac/article/85/4/906/5656851 by guest on 21 August 2022

random priming. In this random priming with PCR, the primers were 6 dNTP (NNNNNN) oligonucleotides, each N representing adenine/cytosine/guanine/thymine. In the PCR reaction, the primers were annealed to the DNA template at random locations, and primer extension was performed with DNA polymerase by using fluorescent dye-labeled precursor dNTP with Cy3. The labeled DNA was mixed with another different genomic DNA labeled with Cy5 (to detect ratio of the background), and the mixture was then hybridized to the chip. The strategy enabled identification of different bacterial species of *Pseudomonas* (19).

Gene Expression of Pathogenic Bacteria

Numerous virulence factors, as well as the nonvirulence genes of pathogenic bacteria, are good candidates for high-throughput gene expression experiments. Virulence factor genes (adhesions, invasins, and toxins) can easily be identified because of the availability of very advanced databases. However, genes that might help to increase the adaptation and quick recovery of pathogenic bacteria in stressful conditions (fitness genes) may be difficult to identify because of the lack of a detectable phenotype or appropriate assay. DNA microarray technology may be used to explore the relationships between virulence genes and fitness genes. For example, is the expression of certain virulence genes accompanied by the expression of other fitness genes? Answering this question will help to determine whether there is an interaction between the 2 types of genes.

Evolutionary changes of bacterial genomes due to the acquisition of new genes could compromise the effectiveness of antibiotic sensitivity screening as well as large-scale genome sequencing. In a recent study in which genomic DNA from 2 strains of *Helicobacter pylori* were sequenced, it was found that 206 coding regions were different, indicating 7% evolutionary divergence over time (21). Similarly, 12 specific deletions of a *Mycobacterium tuberculosis* vaccine strain were different from several *M. tuberculosis* strains, which helps to explain the limited effectiveness of the BCG (Bacille-Calmette-Guerin) vaccine in some people (22, 23). To avoid using complete genome sequencing to examine lateral gene transfer between different bacterial strains, a method

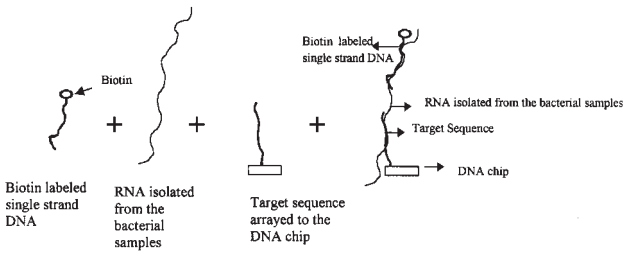

Figure 2. Steps involved in preparing the labeled Cy3 RNA probe with biotin before chip hybridization.

was developed for directly identifying gene transfer in *Pseudomonas aeruginosa* by using 2-dimensional DNA fragment electrophoresis (24). This method involved pooling DNA from 2 strains of *P. aeruginosa* and digesting the DNA with selected restriction enzymes; the DNA fragments were separated by their molecular weight with polyacrylamide gel electrophoresis. Each fragment size was then separated according to sequence composition by using denaturing-urea gradient gel electrophoresis. The fragments were transferred to nitrocellulose membrane, and Southern hybridization was conducted with 2 probes: the probes for strain 1 (Cy3) and strain 2 labeled with Cy5. This method enabled detection of single-copy gene insertions (2.5 kb) responsible for gentamicin resistance.

In conclusion, genome sequencing and gene polymorphism, including lateral gene transfer studies and building virulence gene chips, might provide a better way of discovering functions of different networking genes in different bacteria.

Gene Expression of Ruminal Bacteria

Another potential application for microarray technology is to study rumen microbes. The content of the rumen is a diverse and unique microbial ecosystem composed of bacteria, protozoa, and fungi. Foodstuffs entering the rumen are fermented by these microorganisms to volatile fatty acids, methane, ammonia, lactic acid, and heat (25, 26). Ingestion of large amounts of cereal concentrates provides the substrate that encourages rapid proliferation of ruminal bacteria, such as *Streptococcus bovis*, which produce large amounts of lactic acid (27), and causes a condition known as ruminal acidosis (27). Increased numbers of *S. bovis* are often associated with low ruminal pH (< 6.0) because lactic acid is a stronger acid (pK 3.9) than the volatile fatty acids (pK 4.8 for acetic acid), and this causes dysfunction of the rumen, which in some cases results in death (27). Conservative estimates of the annual impact on U.S. feedlot economy associated with acidosis range between $60 and $100 million (28). Economic losses associated with ruminal acidosis in dairy cattle are not available but are most likely quite significant. These monetary losses provide justification for conducting research to better understand lactate metabolism within the rumen.

For decades, microbiologists and nutritionists have attempted to alter ruminal fermentation to decrease some of the losses associated with feedstuff digestion. With the complexity of the ruminal microbial ecosystem and general lack of information about the genetics of important ruminal bacteria, improvements in the efficiency of feedstuff fermentation have been limited over the past 10+ years, with most research focusing on cellulolytic ruminal bacteria. An approach to overcoming the limitations associated with studying one type of bacterium at a time is the use of DNA microarray technology (21, 29, 30).

The genome of many ruminal bacteria remains to be sequenced. However, it is possible to construct a gene microarray library representing the mixed population of environmental organisms. Shotgun DNA microarray has been ap-

Downloaded from https://academic.oup.com/jaoac/article/85/4/906/5656851 by guest on 21 August 2022

plied in evaluating gene expression in *Plasmodium falciparum* malaria by using 3648 random clones. The success rate of identifying the functions of genes (clones) was at least one order-of-magnitude above all previous studies combined (31). The same concept can be used to examine gene expression by ruminal bacteria obtained from animals fed different diets. Random *E. coli* shotgun genomic libraries of ruminal bacteria could be generated, and 1–2 kb fragments from the *E. coli* libraries could be amplified by high-throughput PCR technique. The PCR fragments could be arrayed on a gene chip representing the mixed bacterial flora of the rumen (Figure 3), and the rumen gene chip could then be used to monitor bacterial gene expression under different environmental conditions. Labeled cDNA isolated from ruminal bacterial RNA under different environmental conditions could be hybridized to the chip. For example, adding the organic acid malate to the diet has been shown to reduce ruminal acidosis (28, 32–34). Therefore, it is quite possible that the expressed genes on a ruminal bacterial gene chip, altered by malate treatment, could be identified, selected, and sequenced. Further data analysis could reveal the uniqueness of the DNA sequences, as well as the involvement of different bacterial genes in acidosis. Important bacterial genes involved in fiber digestion and protein degradation within the rumen could also be identified.

Although the rumen is believed to be a reservoir for many pathogenic bacteria, different research studies have shown that well-fed animals appear less likely to harbor pathogenic *E. coli* (35, 36). Diez-Gonzalez et al. (37) suggested that switching animal feed from grain (high starch) to hay (cellulose) decreased both the number and acid resistance of *E. coli*, indicating that current feed practice that depends heavily on grain feeding might contribute to increased cases of human food poisoning with *E. coli* O157:H7. Gene expression and identification have been studied by using a DNA microarray chip for *E. coli* caused by contamination of beef. Measuring the correlation between the transcript level of tryptophan genes under a variety of conditions (38) showed that only 3 operons, *trp*, *mtr*, and *aroH*, represent the core of a highly responsive *trp* repressor regulon, thus confirming regulatory patterns established in previous studies (38). Another study (39) used a microarray chip that contained an array of 25 bp single-strand target sequences of 4 virulence genes (intimin, shiga-like toxins I and II, and hemolysin A genes) of *E. coli* O157:H7. When the chip was hybridized with biotin-labeled DNA of the different genes, it was 32-fold more sensitive in identifying genes than was PCR gel electrophoresis.

Limitation of the Technology

Microarray technology, like any new technology, is not trouble-free. Some of the main problems associated with it are high background interference (noise), differences in efficiency of DNA labeling between the 2 fluorescent dyes, and instability of RNA isolated from different bacterial systems. In addition, the high cost of automated hybridization, variability in results of the same experiment performed under similar conditions, and different statistical methods of data analysis

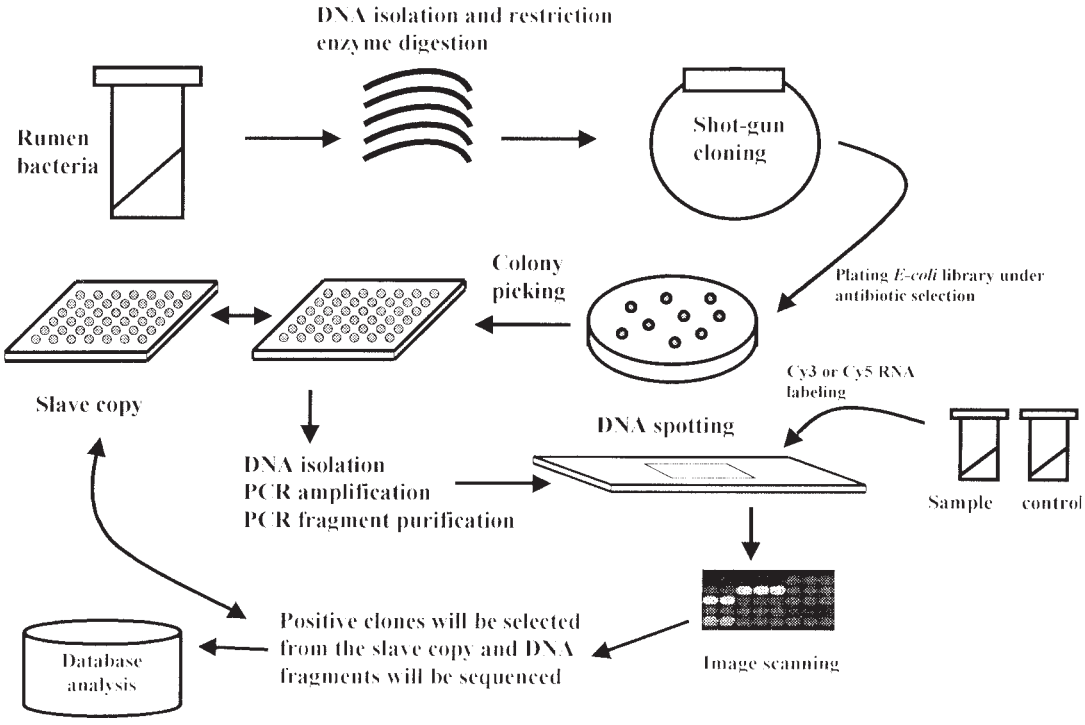

Figure 3. Example of applying DNA microarray technology to the rumen.

Downloaded from https://academic.oup.com/jaoac/article/85/4/906/5656851 by guest on 21 August 2022

decrease the accessibility of the technology. These problems have to be addressed to make the technology useful and the analysis reproducible.

Conclusions

Microarray is a useful technology for the analysis not only of gene expression but also gene identification. Using 20–70 bp single-strand oligonucleotides to represent a gene is more economical than whole-gene PCR amplification. Development of DNA microarrays will make it possible to rapidly detect and identify any of numerous pathogens and to rapidly assess the presence of various virulence factors or antibiotic-resistance genes. Furthermore, this method has the potential to be automated, which would facilitate its use in food processing facilities, medical laboratories, and food safety monitoring agencies.

The ability to create a first generation gene chip containing a random gene library representing a microbial habitat of interest, such as the rumen, is feasible. The creation of a shotgun library for ruminal bacteria similar to the gene chip created for *P. falciparum* (31) represents an efficient strategy when genome sequences are not available. When the role of gene interaction in different bacterial species is explored, the creation of a more comprehensive microarray chip would be valuable for studying the complex interactions between bacterial species, thereby helping to elucidate the global interaction between genes, that is, the big picture.

References

(1) Weinstock, G.M. (2000) *Emerg. Infect. Dis.* **6**, 496–504

(2) Cummings, C.A., & Relman, D.A. (2000) *Emerg. Infect. Dis.* **6**, 513–525

(3) Chizhikov, V., Rasooly, A., Chumakov, K., & Levy, D.D. (2001) *Appl. Environ. Microbiol.* **67**, 3258–3263

(4) Harrington, C.A., Rosenow, C., & Retief, J. (2000) *Curr. Opin. Microbiol.* **3**, 285–291

(5) Vilgalys, R., & Hester, M. (1990) *J. Bacteriol.* **172**, 4238–4246

(6) Nagano, I., Kunishima, M., Itoh, Y., Wu, Z., & Takahashi, Y. (1998) *Microbiol. Immunol.* **42**, 371–376

(7) Yu, J., & Kaper, J.B. (1992) *Mol. Microbiol.* **6**, 411–417

(8) Gannon, V.P., D'Souza, S., Graham, T., & King, R.K. (1997) *Adv. Exp. Med. Biol.* **412**, 81–82

(9) Gannon, V.P., King, R.K., Kim, J.Y., & Thomas, E.J. (1992) *Appl. Environ. Microbiol.* **58**, 3809–3815

(10) Gooding, C.M., & Choudary, P.V. (1999) *Mol. Cell. Probes* **13**, 341–347

(11) Gooding, C.M., & Choudary, P.V. (1997) *J. Dairy Res.* **64**, 87–93

(12) Paton, A.W., & Paton, J.C. (1998) *J. Clin. Microbiol.* **36**, 598–602

(13) Paton, J.C., & Paton, A.W. (1998) *Clin. Microbiol. Rev.* **11**, 450–479

(14) Nikolich, M.P., Hong, G., Shoemaker, N.B., & Salyers, A.A. (1994) *Appl. Environ. Microbiol.* **60**, 3255–3260

(15) Desmarchelier, P.M., Bilge, S.S., Fegan, N., Mills, L., Vary, J.C., & Tarr, P.I. (1998) *J. Clin. Microbiol.* **36**, 1801–1804

(16) Bilge, S.S., Vary, J.C., Dowell, S.F., & Tarr, P.I. (1996) *Infect. Immun.* **64**, 4795–4801

(17) Hu, Y., Zhang, Q., & Meitzler, J.C. (1999) *J. Appl. Microbiol.* **87**, 867–876

(18) Lin, Z., Cui, X., & Li, H. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 2582–2587

(19) Cho, J.C., & Tiedje, J.M. (2001) *Appl. Environ. Microbiol.* **67**, 3677–3682

(20) Small, J., Call, D.R., Brockman, F.J., Straub, T.M., & Chandler, D.P. (2001) *Appl. Environ. Microbiol.* **67**, 4708–4716

(21) Alm, R.A., Ling, L.S., Moir, D.T., King, B.L., Brown, E.D., Doig, P.C., Smith, D.R., Noonan, B., Guild, B.C., deJonge, B.L., Carmel, G., Tummino, P.J., Caruso, A., Uria-Nickelsen, M., Mills, D.M., Ives, C., Gibson, R., Merberg, D., Mills, S.D., Jiang, Q., Taylor, D.E., Vovis, G.F., & Trust, T.J. (1999) *Nature* (*London*) **397**, 176–180

(22) Field, D., Hood, D., & Moxon, R. (1999) *Curr. Opin. Genet. Dev.* **9**, 700–703

(23) Behr, M.A., Wilson, M.A., Gill, W.P., Salamon, H., Schoolnik, G.K., Rane, S., & Small, P.M. (1999) *Science* **284**, 1520–1523

(24) Malloff, C.A., Fernandez, R.C., & Lam, W.L. (2001) *J. Mol. Biol.* **312**, 1–5

(25) Evans, J.D., & Martin, S.A. (2000) *Curr. Microbiol.* **41**, 336–340

(26) Amgarten, M., Schatzmann, H.J., & Wuthrich, A. (1981) *J. Vet. Pharmacol. Ther.* **4**, 241–248

(27) Slyter, L.L. (1976) *J. Anim. Sci.* **43**, 910–929

(28) Martin, S.A., Sullivan, H.M., & Evans, J.D. (2000) *J. Dairy Sci.* **83**, 2574–2579

(29) Brooker, J.D., Thomson, A.M., & Ward, H. (1992) *Australas. Biotechnol.* **2**, 288–291

(30) Doblhoff-Dier, O., Bachmayer, H., Bennett, A., Brunius, G., Cantley, M., Collins, C., & Collard, J.-M. (2000) *Trends Biotechnol.* **18**, 141–146

(31) Hayward, R.E., Derisi, J.L., Alfadhli, S., Kaslow, D.C., Brown, P.O., & Rathod, P.K. (2000) *Mol. Microbiol.* **35**, 6–14

(32) Martin, S.A., Bertrand, J.A., Sauls, B., & Hill, G.M. (2000) *J. Dairy Sci.* **83**, 308–312

(33) Martin, S.A., Streeter, M.N., Nisbet, D.J., Hill, G.M., & Williams, S.E. (1999) *J. Anim. Sci.* **77**, 1008–1015

(34) Martin, S.A., & Streeter, M.N. (1995) *J. Anim. Sci.* **73**, 2141–2145

(35) Rasmussen, M.A., Cray, W.C., Jr, Casey, T.A., & Whipp, S.C. (1993) *FEMS Microbiol. Lett.* **114**, 79–84

(36) Cray, W.C., Jr, Casey, T.A., Bosworth, B.T., & Rasmussen, M.A. (1998) *Appl. Environ. Microbiol.* **64**, 1975–1979

(37) Diez-Gonzalez, F., Callaway, T.R., Kizoulis, M.G., & Russell, J.B. (1998) *Science* **281**, 1666–1668

(38) Khodursky, A.B., Peter, B.J., Cozzarelli, N.R., Botstein, D., Brown, P.O., & Yanofsky, C. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 12170–12175

(39) Call, D.R., Brockman, F.J., & Chandler, D.P. (2001) *Int. J. Food. Microbiol.* **67**, 71–80

Downloaded from https://academic.oup.com/jaoac/article/85/4/906/5656851 by guest on 21 August 2022