Nutrition Research Reviews (2003), **16**, 21–35 DOI: 10.1079/NRR200358

Genomics of food-borne bacterial pathogens

Jerry M. Wells* and Marjon H. J. Bennik

Institute of Food Research, Norwich Research Park, Norwich NR4 7UA, UK

Despite continued efforts to prevent and control food-borne illness it remains a major cause of morbidity and mortality throughout the world. The problem is made worse by the continuous threat from emerging pathogens that can evolve to adapt to the different environments resulting from ongoing changes in farming or food production. The present review discusses the impact of genomics and post-genomic technologies on research in the area of food-borne bacterial pathogens. Genomics research is moving at a fast pace and these are exciting times for microbial research. The genome sequences of approximately ninety bacterial genomes have recently been completed and genome sequences are already available for several food-borne pathogens and closely related species. Comparative genomics is providing new insights into mechanisms of bacterial evolution and has helped in determining virulence factors of pathogens. Genomics has also provided tools such as DNA microarrays that can be used to examine the genetic composition and whole genome expression profiles of bacterial strains by hybridisation of fluorescently labelled DNA. This is helping to identify genes associated with particular phenotypes such as virulence and host preference, and to identify genes in uncharacterised genomes of closely related organisms. Microarrays are also being developed for the detection of food-borne pathogens and investigation of the evolutionary relationship between different species of bacteria. The review concludes with a discussion of the use of functional genomics tools to investigate bacterial responses to environmental stresses and also host–pathogen interactions. These research areas will be valuable in designing future strategies for controlling food-borne pathogens.

Food-borne pathogens: DNA microarrays: Genomics: Host–microbe interactions: Stress response

Introduction

It has been estimated that more than 200 infectious diseases can be transmitted through food (Bryan, 1982). A recent study in the USA estimated that of the 76 million illnesses and 323 000 hospitalisations caused each year by food-borne pathogens, 67 % are caused by viruses, 30 % by bacteria and 3 % by parasites (Mead *et al*. 1999). In the USA the most important bacterial causes of food-borne infection are *Salmonella* and *Campylobacter* (shown in Fig. 1), which account for 26 % of all illnesses and 17 % of hospitalisations. In terms of mortality rates the major bacterial food-borne pathogens are *Salmonella* (non-typhoidal), *Listeria monocytogenes*, *Campylobacter*, and enterohaemorrhagic *Escherichia coli* (Mead *et al*. 1999). These pathogens are also prominent causes of food-borne illness throughout Europe but their relative prevalence varies at the national level (Tirado & Schmidt, 2001; Adak *et al*. 2002). An overview of food-borne bacterial illnesses in the USA as reported by Mead *et al*. (1999) is given in Table 1. Decades of research and surveillance have almost eliminated some diseases due to technological developments and better methods of detection and risk assessment (Lund *et al*. 2000). However, new trends in farming practices, the development of new products and production methods, and the globalisation of markets are presenting new challenges for food safety (Kaferstein & Abdussalam, 1999).

Food-borne illness can manifest as an intoxication from the consumption of food with preformed toxins and can occur even if viable pathogens are no longer present when heating is not required or if the temperature is not sufficient to inactivate the toxin. The major pathogens implicated in this form of illness are *Bacillus cereus*, *Clostridium botulinum* and *Staphylococcus aureus* (Table 1) that produce

Abbreviations: G+C, guanine plus cytosine; MALDI, matrix-assisted laser desorption and ionisation; LOS, lipo-oligosaccharide; PCR, polymerisation chain reaction; PI, pathogenicity islands; *S. Typhimurium*, *Salmonella enterica* serovar Typhimurium; SPI, *Salmonella enterica* serovar Typhimurium pathogenicity islands.

***Corresponding author:** Dr Jerry Wells, fax +44 1603 207732, email jerry.wells@bbsrc.ac.uk

emetic toxin, botulinum toxin, and enterotoxins, respectively (Peck, 1997; Balaban & Rasooly, 2000; Kotiranta *et al*. 2000). *B. cereus* and *Cl. perfringens* can also produce toxins in the gut leading to diarrhoeal disease (Granum, 1990; Kotiranta *et al*. 2000). Other food-borne pathogens such as species of *Salmonella*, *Campylobacter*, *Shigella* and *Yersinia* infect the intestine of human and animal hosts and cause gastroenteritis through cell invasion and damage to the epithelium and/or the production of one or more toxins and virulence factors. A detailed overview of food-borne pathogens is found in Lund *et al*. (2000). In many industrialised countries *Campylobacter* and *Salmonella* account for the majority of bacteria-related food-borne illness and poultry and poultry products have been reported to be the major sources of infection. Certain species of *Campylobacter* and *Salmonella* can colonise a variety of animal hosts that form a zoonotic reservoir for the spread of these pathogens in the environment.

Approximately 80 % of *Campylobacter* infections are caused by two species of the genus, namely, *Campylobacter*

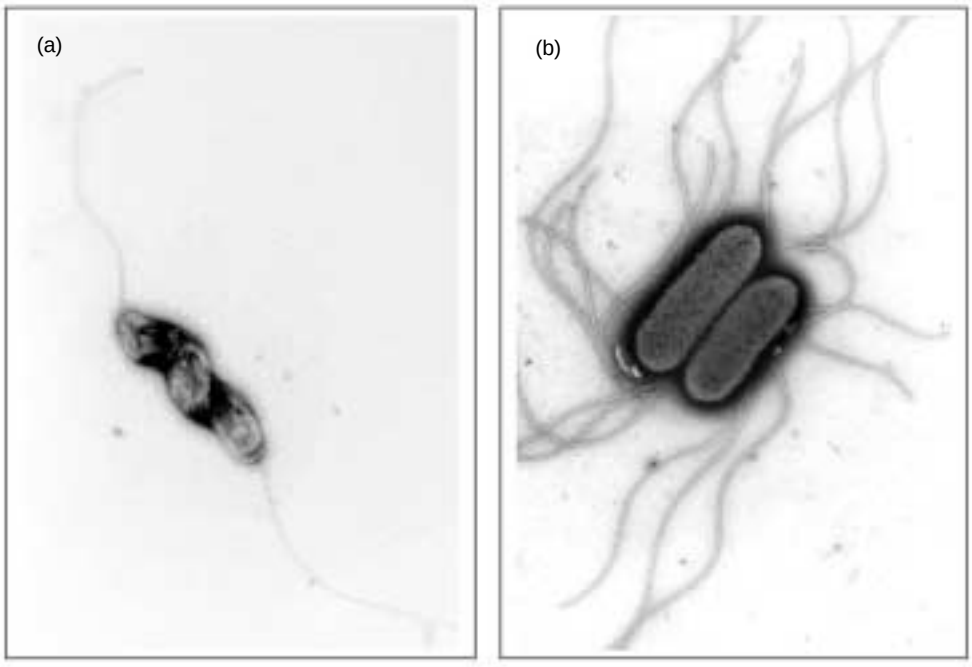

Fig. 1. *Campylobacter jejuni* (a) and *Salmonella enterica* serovar Typhimurium (b) are the leading causes of food-borne bacterial illness. (Images courtesy of Mary Parker, Institute of Food Research).

Table 1. Estimated illnesses, frequency of food-borne transmission, hospitalisation rate and case fatality rates for known bacterial pathogens in the USA

Bacteria	Estimated cases	Percentage food-borne transmission	Hospitalisation rate	Case-fatality rate
Bacillus cereus	27 360	100	0·006	0·0000
Brucella spp.	1 554	50	0·550	0·0500
Campylobacter spp.	2 453 926	80	0·102	0·0010
Clostridium botulinum	58	100	0·800	0·0769
Clostridium perfringens	248 520	100	0·003	0·0005
Escherichia coli O157:H7	73 480	85	0·295	0·0083
E. coli, non-O157 STEC	36 740	85	0·295	0·0083
E. coli, enterotoxigenic	79 420	70	0·005	0·0001
E. coli, other diarrhoeogenic	79 420	30	0·005	0·0001
Listeria monocytogenes	2 518	99	0·922	0·2000
Salmonella enterica serovar Typhimurium	824	80	0·750	0·0040
Salmonella, non-typhoidal	1 412 498	95	0·221	0·0078
Shigella spp.	448 240	20	0·139	0·0016
Staphylococcus food poisoning	185 060	100	0·180	0·0002
Streptococcus, food-borne	50 920	100	0·133	0·0000
Vibrio cholerae, toxigenic	54	90	0·340	0·0060
V. vulnificus	94	50	0·910	0·3900
Vibrio, other	7 880	65	0·126	0·0250
Yersinia enterocolitica	96 368	90	0·242	0·00050
Total	5 204 934			

* Data taken from Mead *et al*. (1999).

https://doi.org/10.1079/NRR200358 Published online by Cambridge University Press

jejuni and *C. coli* with *C. jejuni* being more frequently associated with disease in humans. A variety of other species, including *C. upsaliensis* and *C. lari* also infect humans but their incidence remains unclear (Mishu *et al*. 1993). Symptoms from infection with *C. jejuni* can vary from very mild diarrhoea to profuse bloody diarrhoea with mucosal damage and inflammation especially in the ileum and jejunum (Wassenaar & Blaser, 1999). Invasion of the intestinal epithelium and the production of one or more toxins are thought to play a major role in pathogenesis of disease caused by *Campylobacter*. The infections are self-limiting and bacteraemia is rarely tested for, or reported in clinical cases. In a rare number of cases, infection with *Campylobacter* is associated with the peripheral neuropathies known as Guillain-Barré syndrome and Miller Fisher (Nachamkin *et al*. 1998). Lipo-oligosaccharide (LOS) antigens present on certain serotypes of *C. jejuni* mimic gangliosides found on human nerve tissue and are thought to provoke autoimmune reactions that give rise to neuropathology.

The genus *Salmonella* comprises two species: *Salmonella enterica* comprising some 2000 serovars and *S. bongori*, a species that does not cause disease in mammals. *S. enterica* serovar Typhimurium (further referred to as *S. Typhimurium*) can evoke a localised inflammatory response in the gastrointestinal tract of man and many other animals. Some other serovars of *S. enterica* cause systemic infections and typhoid fever in man (Chalker & Blaser, 1988; Rabsch *et al*. 2001).

Shigella spp. are the causative agents of bacillary dysentery and invade the intestinal epithelium causing inflammation and ulceration of the mucosa. Cases of food-borne shigellosis are much lower than that of salmonellosis or campylobacteriosis but it remains an important cause of food-borne illness. The main species involved are *Shigella sonnei* (70 % of cases) and *Shigella flexneri* (25 % of cases) and the foods frequently implicated as sources of infection are potato salad or salad dishes containing fish, raw vegetables, chicken and seafood (Kotloff *et al*. 1999).

Yersinia enterocolitica is another invasive enteric pathogen causing diarrhoeal illness and pseudo-appendicitis that can be transmitted through food. Cases are rare in the USA but more common in Northern Europe and most frequently associated with undercooked pork (Bottone, 1999).

The intestinal tract is also the main portal of entry for *Listeria monocytogenes*, the causative agent of food-borne listeriosis, a severe disease associated with gastroenteritis, meningitis, meningoencephalitis, septicaemia, abortion and perinatal infections (Schlech, 2001). *L. monocytogenes* is able to tolerate extremes of pH and salt, and can grow at temperatures as low as –0·4°C (Walker *et al*. 1990). This organism can be encountered in certain unpasteurised foods, for example, unpasteurised milk and associated dairy products. After ingestion of contaminated food *Listeria* disseminates from the gastrointestinal tract to phagocytes and can eventually end up in the central nervous system. In pregnant women *Listeria* also disseminates to the placental tissues causing abortion and perinatal infections (Vazquez-Boland *et al*. 2001).

E. coli is probably the most famous member of the food-poisoning enteric bacteria due to its extensive use of non-pathogenic *E. coli* K12 derivatives as a vector host system in research laboratories throughout the world. Wild-type strains of *E. coli* are able to survive in a wide variety of environments including the intestinal tract of man and other animals. *E. coli* are in fact predominantly non-pathogenic but certain clonal types such as serotype O157 and other enterohaemorrhagic *E. coli* (EHEC) have acquired genes from other enterobacteriacae and contain mutations that allow them to propagate through the food chain and cause disease in man (Strockbine *et al*. 1986). Ruminants such as cattle and sheep appear to be the major source of enterohaemorrhagic *E. coli* infections.

At the time of writing the present review the genomes of several of the above-mentioned pathogens had been completely sequenced or were in progress. Genomics is the term used to describe genome sequencing and the use of computational approaches to provide information directly from genome-scale analysis of DNA and protein sequences. The increasing availability of genome sequences has spurred on the development and application of post-genomics tools such as microarrays and proteomics for the genome-scale analysis of gene transcription, protein expression and gene function. The aim of the present review is to discuss the contribution of genomic and post-genomic research to our scientific understanding of food-borne bacterial pathogens, and its potential impact on future research.

Genome sequencing and comparative genomics of food-borne pathogens

An estimated ninety microbial genomes have been completely sequenced including the food-borne pathogens *C. jejuni*, *S. enterica* serovar Enteritidis, *S. Typhimurium*, *E. coli* O157:H7, *L. monocytogenes*, and *Shigella flexneri*. Another 130 microbial genome sequencing projects are currently in progress and these include *C. botulinum*, *B. cereus* and a strain of *C. perfringens* associated with food-borne illness. An up-to-date overview of the status of these genome projects with links to the organisations responsible for the sequencing can be found on the web page of The Institute of Genomic Research at www.tigr.org and at the web page of the National Centre for Biotechnology Information at www.ncbi.nih.gov. Typically the first stage of a genome sequencing project involves the sequencing of clones from a random shotgun library until a six- to ten-fold coverage of the genome is achieved. Not all sequences will be represented in the library of clones and the genome sequence will comprise a number of contigs that are ultimately assembled into the finished sequence using one or more strategies for gap closure (Fraser & Fleischmann, 1997; Frangeul *et al*. 1999). Initially, computer tools are used to assemble a contiguous sequence from the raw data and identify the open reading frames. Homology-based algorithms are then used to try and predict the function of the proteins. Further clues can be gained from examining the gene order since bacterial genes are often arranged in functional groups or operons that may be co-transcribed and co-regulated.

The first genome of a food-borne bacterium to be sequenced was that of *C. jejuni* and it led to the discovery

https://doi.org/10.1079/NRR200358 Published online by Cambridge University Press

of important new aspects of the biology of this organism (Parkhill *et al*. 2000). Historically, the heat-stable antigens used for serotyping of *C. jejuni* were assumed to be lipopolysaccharides as in the case of *E. coli* and *Salmonella* (Penner *et al*. 1983) although it had been suggested that Penner serotyping antibody might recognise a capsular polysaccharide antigen (Chart *et al*. 1996). A 42·6 kb cluster of putative capsule biosynthetic genes was discovered in the genome of *C. jejuni* flanked by genes with similarity to the type 2 capsule transport genes found in *E. coli*. The transport genes were conserved among different strains, whereas many of the putative glycosyltransferases and polysaccharide biosynthetic enzymes are not found in strains of a different Penner serotype, lending support to the idea that the capsule antigen is indeed an important determinant of the serotyping systems for *Campylobacter* (Dorrell *et al*. 2001; B Pearson, J Wright, C Pin, K I'Anson, T Humphrey and J Wells, unpublished results). Furthermore, Karlyshev *et al*. (2000) found that four out of six different serotypes lacking the transport gene *kpsM* became untypable using Penner serotyping.

Another striking discovery made from the sequencing of the *C. jejuni* genome was the identification of twenty-five hypervariable genes containing homopolymer pG:C tracts of the contingency type described previously (Moxon *et al*. 1994; van Belkum *et al*. 1998). Variation in the length of the homopolymer tracts was observed among different clones selected for genome sequencing, indicating that the hypervariable genes have potential to undergo a high rate of slipped-strand mispairing (Parkhill *et al*. 2000). Many of the hypervariable genes are found in loci involved in the production or modification of surface structures such as the flagellum, LOS and capsule, suggesting that they play a role in antigenic variation. Hypervariable genes have not been found in the genome of *E. coli* or *Salmonella* food-borne pathogens although some generally conserved chromosomal regions are more divergent than others (for example, fimbrial biosynthetic operons), perhaps as a consequence of positive selection in the host environment. However, increased mutation rates at some loci or differential retention of paralogous genes during strain evolution may also account for the presence of these hypervariable genes.

Recent evidence suggests that plasmids may play a more important role in virulence of *C. jejuni* than once believed (Alm *et al*. 1999; Bacon *et al*. 2000; B Pearson, personal communication) and thus more effort should be given to the sequencing of plasmids from well-characterised strains of *Campylobacter*. Plasmids play important roles in virulence and pathogenesis of disease caused by other enteric pathogens. For example, studies on pathogenic enterohaemorrhagic *E. coli* and enteropathogenic *E. coli* have identified a number of important virulence factors such as adhesions and proteases on plasmids. A body of evidence supports the hypothesis that *Shigella* species, including the food-borne pathogen *Shigella flexneri*, have evolved from a non-pathogenic *E. coli* ancestor and that the main virulence factors, such as the enterotoxin and factors required for invasion and intracellular spread were acquired by horizontal transfer of a large plasmid (Ochman & Groisman, 1995; Ochman *et al*. 2000; Pupo *et al*. 2000). Most strains of *S. Typhimurium* contain a virulence plasmid of about 90 kb and in the LT2 strain the 94 kb plasmid pSLT was sequenced in addition to the genome (McClelland *et al*. 2001).

The sequencing of bacterial genomes is also providing insights into bacterial evolution and the mechanisms involved in gene transfer that lead to the acquisition of new genetic traits in different species and genus. The recent genome sequence of the important enterohaemorrhagic food-borne pathogen *E. coli* O157:H7 unexpectedly revealed that this strain contained approximately 25 % more DNA that the non-pathogenic laboratory strain *E. coli* K12 (Hayashi *et al*. 2001; Perna *et al*. 2001). *E. coli* O157:H7 contains 1387 new genes in clusters up to 88 kb in length, designated O-islands, to indicate that they are O157 serotype-derived. Many gene clusters were also found to be specific to *E. coli* K12 (Blattner *et al*. 1997) and are arranged in so-called 'K clusters' totalling 0·53 Mb. The O-clusters of genes found in the sequenced strain of *E. coli* O157:H7 encode candidate virulence factors, alternative metabolic capacities, several prophages and other new functions. Clusters of genes or islands that are important in virulence and have an average base composition different from the bulk of the genome are commonly referred to as pathogenicity islands (PI) and they have been found in a variety of Gram-positive and Gram-negative bacteria. In many cases PI are flanked by sequences associated with DNA transfer such as bacteriophage attachment sites, insertion sequence elements, integrase-like genes and plasmid origins of replication, indicating that they have been spread among members of the bacterial kingdom by horizontal transfer, especially via plasmids and bacteriophages (Hacker *et al*. 1997). In *E. coli* the atypical base composition of the O and K islands mentioned earlier (p. 24) indicates that they most probably arose through horizontal transfer from a donor species with a different DNA base composition, but only a subset of these islands are associated with elements likely to be autonomously mobile. Thus it is possible that many of the sequences associated with mobile genetic elements such as direct repeats and insertion sequence elements have been deleted and lost from the genome of *E. coli*.

In the food-borne pathogen *S. Typhimurium*, five pathogenicity islands (SPI I–V) were discovered before the genome sequence became available (McClelland *et al*. 2001). SPI-I and SPI-II play a role in invasion and survival in macrophages, respectively, and have been found in all *S. Typhimurium* strains tested. The five major PI are absent from *E. coli* suggesting that these were acquired by all serovars of *Salmonella*. It has been speculated that these SPI elements were transmitted to *Salmonella* from a common ancestor soon after its divergence from *E. coli* more than 100 million years ago (Groisman & Ochman, 1996; Lee, 1996). The genome sequence of *S. Typhimurium* strain LT2 revealed more than sixty-two gene clusters containing four or more genes called 'islands' that are unique to certain species. Some of these 'islands' are typical PI containing putative virulence factors and sequence elements associated with their transfer from other hosts (McClelland *et al*. 2001). The genome of the Gram-positive food-borne pathogen *L. monocytogenes* and the non-pathogenic species

L. innocua have both been sequenced; 270 *L. monocytogenes* (2·94 Mb) and 149 *L. innocua* (3·01 Mb) specific genes were identified (Glaser *et al*. 2001). The two *Listeria* genomes have a conserved co-linear organisation with the species-specific genes distributed in multiple islands (between 1 to 23 kb) on the chromosome in situations similar to the comparison of *E. coli* strain O157:H7. Approximately half of these *Listeria* species-specific islands had an average base composition different from the bulk of the genome, suggesting that they had been acquired by horizontal transfer. However, this is not true for all the islands and thus some islands may have been acquired by horizontal gene transfer from an organism with a similar guanine plus cytosine (G+C) composition or have become adapted over a long period of time. This seems to be the case for the previously studied virulence locus of *L. monocytogenes*. Analysis of the flanking regions around the virulence gene cluster in *Listeria* species and *B. subtilis* suggested that it was probably acquired by an ancestral species and then lost in *L. innocua* (Glaser *et al*. 2001).

The high number of gene clusters and islands discovered in *E. coli*, *Salmonella* and *Listeria* indicates that lateral gene transfer of DNA between different species and even among strains of the same species occurs more frequently than was once believed. Undoubtedly this has been a driving force in the adaptability and evolution of these pathogens.

In contrast to *E. coli*, *Salmonella* and *Listeria*, the genome sequence of the food-borne pathogen *C. jejuni* contained very few repeat sequences and no prophages, insertion sequence elements or plasmid origins of replication despite the fact that bacteriophages and plasmids have been reported in *Campylobacter* (Lee *et al*. 1994; Sails *et al*. 1998). The implication is that mobile elements have played a less important role in the evolution of *C. jejuni*, perhaps because this organism is naturally competent for transformation with genomic DNA. Additionally, there are no PI or other gene clusters in the *C. jejuni* genome that have a markedly different base composition to that of the bulk of the genome. The only regions that do show a slightly lower than average G+C composition are the capsule and LOS biosynthetic loci, suggesting that some of these were acquired from an organism with an overall lower G+C composition that occupies the same niches as *Campylobacter*.

Initially, microbial genomics focused on the sequencing of genomes of medically and industrially important species from diverse parts of the evolutionary tree. This approach has provided insights into the major evolutionary relationships between these microbes. As more genome sequences become available attention is turning to the comparative genomics of closely related organisms as this provides insights into the evolutionary events occurring over a shorter time scale (Ochman & Jones, 2000; Ochman *et al*. 2000). The genome sequence data for multiple *Salmonella* serovars is now available and recent genomic studies have focused on a comparison of these closely related genomes (Edwards *et al*. 2002). Approximately 500–600 kb was found to be specific to each serovar and these unique sequences are distributed throughout the genome in regions ranging from 1 to 50 kb in length. Many of these unique regions also have a lower G+C content compared with the average G+C content of the genome suggesting that they were acquired from an ancestral organism with lower G+C content than *Salmonella* that shared the same ecological niches. The ultimate aim of these studies is to understand how the acquisition of genetic traits and the loss of ancestral traits relate to the serovar-specific differences in host preference, disease outcomes and survival in new environments such as host tissues. The loss of ancestral traits through negative selection has been called pathoadaptive mutation (Sokurenko *et al*. 1999) and good examples of this are found in evolutionary studies on pathogenic *Shigella* spp. causing food-borne illness. Lysine decarboxylase blocks the activity of *Shigella* enterotoxins and the four pathogenic species of *Shigella* all contain pathoadaptive deletions resulting in loss of the lysine decarboxylase gene (cad locus) (Nataro *et al*. 1995; Maurelli *et al*. 1998). Different insertion sequences, prophages and insertions of other gene loci have displaced the cad region in the four species of *Shigella* indicating that these convergent evolutionary events occurred independently in each species (Day *et al*. 2001).

In the near future, genome sequences of different strains of the same species of various food-borne pathogens will become available enabling the effect of sequence divergence to be investigated with regard to the individual virulence and other biochemical attributes of a strain. Comparative genomics does not, however, guarantee that the factors influencing the characteristics of individual strains will be readily identified without additional research, as even point mutations in non-coding regulatory sequences or in specific genes could have profound effects on virulence and other strain properties. Although many discoveries have been made through genome sequencing, large gaps in our understanding of the function of genes and their biological significance remain. For example, the genome sequence of *C. jejuni* did not reveal any new virulence factors that might be involved in colonisation or invasion of epithelial cells or cytotoxicity despite the fact that relatively little was known about the virulence of this organism. There is much that could be done to investigate genes of unknown function using the traditional genetic, biochemical and physiological approaches. Some caution should also be exercised in the computational assignment of function based on homology, as proteins with similar sequences can indeed have different functions (Brenner, 1999).

Whole genome expression profiling using microarrays

Arrays of nucleic acids have been used in research for many years but it is only recently that it has been possible to synthesise or deposit nucleic acids at densities of more than 250 000 oligonucleotide probes or 10 000 different cDNA per square centimetre (Bowtell, 1999; Lipshutz *et al*. 1999). The miniaturisation of nucleic acid arrays on glass or silicon surfaces and use of fluorescence for detection enables highly parallel analyses with performance specifications that could not be achieved with the earlier nylon membrane technologies (Fodor *et al*. 1991; Schena *et al*. 1995; Shalon *et al*. 1996; Graves, 1999). There are two

https://doi.org/10.1079/NRR200358 Published online by Cambridge University Press

possible approaches for making microarrays. The first was pioneered by Affymetrix and utilises a computational approach involving photolithography and miniature masks to chemically synthesise unique oligonucleotides onto a silicon surface (Fodor *et al*. 1991). The second method involves the robotic deposition of DNA onto glass slides and was developed by the Stanford Laboratory (Schena *et al*. 1995). These two main approaches are now in common use but this field continues to move rapidly and new technical approaches are constantly appearing in the literature (Lipshutz *et al*. 1999; Singh-Gasson *et al*. 1999; Brenner *et al*. 2000; Okamoto *et al*. 2000; Steemers *et al*. 2000; Taton *et al*. 2000; Westin *et al*. 2000; Blohm & Guiseppi-Elie, 2001). Information on general and technical aspects of microarraying is also available from recent reviews (Cheung *et al*. 1999; Graves, 1999; Lipshutz *et al*. 1999; Niemeyer & Blohm, 1999; DeRisi, 2000; Lee *et al*. 2000; Meldrum, 2000; Schuchhardt *et al*. 2000).

The most widely used application of microarrays is for the analysis of gene transcription at the level of the whole genome. The seminal papers on expression profiling come from studies on the yeast *Saccharomyces cerevisiae* and *E. coli* K12 (Chu *et al*. 1998; Spellman *et al*. 1998; Richmond *et al*. 1999; Wei *et al*. 2001).

To date there are only a few publications specifically dealing with microarray expression profiling of food-borne pathogens. Acetate-induced acid tolerance and autoinducer II-induced gene expression have been studied in *E. coli* O157:H7 (Arnold *et al*. 2001; Sperandio *et al*. 2001) and intracellular gene expression in *S. Typhimurium* following infection of macrophages (discussed later (p. 30) in host–microbe interactions; Eriksson *et al*. 2003). However, there have been several microarray studies in non-pathogenic strains of *E. coli* that serve as relevant models for certain aspects of the biology of the food-borne *E. coli* pathogens (see section on stress response; pp. 28–29). Most microarray expression profiling experiments generate long lists of genes that respond directly or indirectly to the experimental variable tested. The complexity of the transcriptional response arises as a consequence of the specific regulatory network or interconnected networks that are triggered by the experimental variable but also from any indirect effects such as growth rate-related perturbations in gene expression. Thus, additional physiological or biochemical and genetic experiments are often required to corroborate the results and fully interpret the data. One useful approach that can be employed to help dissect the regulatory pathways involved in a response to a stimulus is to study transcriptional changes over time so that temporally co-regulated genes can be classified separately. Non-specific effects in a regulatory pathway can also be minimised by ensuring that the stimulus applied (for example, oxidative stress) has minimal effects on bacterial growth rate but nevertheless induces a stress response (Pomposiello *et al*. 2001). One faces similar issues in the microarray analysis of genetic mutants, particularly if they are in regulatory genes. This problem was solved in studies on yeast by generating a ‘compendium’ database of expression profiles from 300 mutants and isogenic wild-type strains grown under sixty-three different physiological conditions or treatments (Hughes *et al*. 2000). Clustering analysis was then used to classify transcripts on the basis of co-regulation and to distinguish direct from indirect effects. For example, treatment of yeast cells with the sterol biosynthesis inhibitor itraconazole resulted in hundreds of changes in gene expression compared with untreated cells, and seemed to involve diverse pathways such as starvation response, yeast mating, and cell surface stress response (Bammert & Fostel, 2000). However, when these changes were analysed as a group under different physiological conditions it became clear that the primary effect of itraconazole treatment was the induction of the sterol biosynthetic genes and that other changes in gene expression were non-specific (Hughes *et al*. 2000). In the former study a gene of unknown function was found to be co-regulated with the genes involved in sterol biosynthesis and was subsequently shown to have a function in this pathway (Hughes *et al*. 2000). It is now generally accepted that the classification of genes by expression profiling is a useful approach to predict the general function of unknown genes. Once groups of regulated transcripts have been identified there are several search tools that can be used to try and identify common upstream or downstream sequence elements that could serve as DNA binding sites for a regulatory factor (Roth *et al*. 1998; Tavazoie *et al*. 1999). For illustrations of the discovery of new regulatory factors and circuits using this approach, see Spellman *et al*. (1998) and Zhu *et al*. (2000). Another useful genetic approach that can be used to determine gene members of a regulon is to examine the transcription profile after experimentally inducing expression of the regulator (Arfin *et al*. 2000). Despite the issues and limitations of microarrays discussed earlier (p. 26), microarray expression profiling has been highly successful in revealing the trends in microbial gene expression that reflect the organisation of operons and regulons.

In the future it is hoped that raw microarray data on bacteria and other organisms will become available through centralised databases so that comparative studies can be carried out in different laboratories as new computational tools are developed (for a review, see Rhodius *et al*. 2002). This will potentially increase the value of the individual studies being carried out as they can be combined and re-analysed in different ways in the future. This highlights the need for microarray standards and guidelines such as those developed by the Microarray Gene Expression Data group (www.mged.org/Workgroups/MIAME/miame_checklist.html). Proper experimental replication and statistical analysis are also crucial aspects of microarray work (for reviews, see Kim *et al*. 2002; Hatfield *et al*. 2003).

Microarray-based comparative genomics

DNA microarrays have an important application in comparative genomics as they can be used to compare whole genomes of different bacterial isolates for the presence or absence of genes. The limitations of this approach are that point mutations, small deletions and gene rearrangements, and novel genes not present on the array will not be detected. Additionally, intergenic regions containing promoter elements and non-translated RNA are not usually included on DNA spotted microarrays and thus are not included in the analysis. Nevertheless, this approach can

https://doi.org/10.1079/NRR200358 Published online by Cambridge University Press

provide important information on the genetic diversity of isolates in natural populations and, for example, could include strains of different virulence or strains adapted to particular environmental conditions.

In a recent study, eleven strains of *C. jejuni* were compared using a low cost DNA microarray comprising selected polymerisation chain reaction (PCR) products from the genome sequencing project (Dorrell *et al*. 2001). Although the microarray had limitations and only 34 % of the PCR products were gene-specific, it proved to be suitable for performing genotyping where the data generated are qualitative rather than quantitative. In this study 21 % of the genes were deemed to be absent or highly divergent in one or more strains tested. Many of the strain variable genes were found in the loci responsible for the synthesis of surface structures such as the capsule, LOS and flagella (Dorrell *et al*. 2001).

These findings are similar to recent results generated using an open reading frame (ORF)-specific DNA microarray where 16·3 % of genes present in the sequenced strain (NCTC 11168) were absent or highly divergent from the eighteen other *C. jejuni* strains tested (Pearson *et al*. 2003). In this study seven highly variable regions of the genome (plasticity regions) were identified that contained genes for hypothetical proteins and transporters that might be linked to adaptation in different environments. In addition many regions contained genes associated with the expression and modification of surface structures such as membrane proteins, flagellum, capsule and LOS.

These initial comparative microarray studies on *Campylobacter* have looked at overall diversity of a collection of strains and there has been no attempt to correlate the virulence properties of a particular strain with its genetic composition. This would be a valuable thing to do on a larger set of well-characterised strains especially as studies on *Campylobacter* invasion have shown striking differences among different strains in respect of their invasiveness, rates of translocation across cell monolayers and the effect of biochemical inhibitors of invasion *in vitro* (Hu & Kopecko, 1999). It is not yet known to what extent these observations might reflect differences in genetic factors contributing to cell invasion and translocation. A good example of the power of such an approach is illustrated by the microarray comparison between two *Helicobacter pylori* strains that produced different levels of gastritis in a gerbil gastritis model. It was found that the strain producing less severe gastritis lacked a PI containing the cag toxin (Israel *et al*. 2001).

Genomic comparisons between non-pathogenic and pathogenic strains of the same species can also be particularly informative as it can reveal putative virulence factors for further investigation. Additionally, interspecies microarray hybridisations can rapidly identify thousands of genes in uncharacterised genomes as described for the various different serovars of *S. enterica* including those causing food-borne illness (McClelland *et al*. 2001). In a recent study, the presence of *S. Typhimurium* LT2 gene homologues was assessed by microarray analysis in twenty-two other *Salmonella* serovars including avirulent *S. bongori*, a potential progenitor of the *S. enterica* serovars (Porwollik *et al*. 2002). The phylogenetic tree compiled on the basis of LT2 gene homologues was largely consistent with the previous studies based on multilocus enzyme electrophoresis and DNA sequencing of several loci. Most strikingly, the data indicated that many gene clusters were acquired on multiple independent occasions between the different *Salmonella* lineages. Similar studies using mixed-genome microarrays of *Listeria* showed that many genes have diverged between lineages of *L. monocytogenes* (Call *et al*. 2003).

Microarray detection methods

The rapid detection and isolation of food-borne pathogens is crucial for the monitoring of outbreaks of illness and for epidemiological surveillance. Numerous detection methods have been developed for epidemiological investigation of food-borne illnesses or surveillance of food products, many of which rely on PCR-based assays of species-specific DNA or 16S rRNA sequences (Woese, 1987; Martin & Trimmers, 1997). Immunological and mass spectroscopic methods have also been employed for this purpose but are generally more expensive or cumbersome than PCR methods (Ezzell *et al*. 1990; Fox *et al*. 1990). The use of microarrays represents the latest development in detection technology. In a recent study it was demonstrated that glass slide microarrays containing oligonucleotide probes to four virulent loci from *E. coli* O157:H7 was 32-fold more sensitive than gel electrophoresis in detecting PCR products. Furthermore, the array was capable of detecting amplification products from a single cell (approx 1 fg DNA) and 55 colony-forming units of *E. coli* O157:H7/ml in chicken rinsate following immunomagnetic capture and PCR (Call *et al*. 2001). Similar approaches are being developed for the detection and species identification of *Listeria* and other important pathogens (Soini & Musser, 2001; Volokhov *et al*. 2002). A portable system for the preparation of labelled RNA or DNA from 1 ml volumes of bacterial cell culture and glass slide microarray analysis has recently been described (Bavykin *et al*. 2001). Starting with a bacterial culture the whole procedure takes 50 min to acquire a microarray image but direct detection of food-borne pathogens is not possible unless the methods can be adapted for detection of PCR products.

In a recent study it was shown that DNA microarrays can be used to quantify specific target genes in DNA samples using a fluorescently labelled DNA reference for standardisation (Cho & Tiedje, 2002). However the detection limit for the target gene was approximately 10 pg of DNA indicating that direct detection of less than about 10^5 bacteria per sample (for example, *E. coli* O157) is not possible without enrichment for the organisms of interest. The Cho & Tiedje (2002) study also showed that the detection of specific genes in amounts above the detection limit was still problematic in environmental samples due to the presence of other sources of DNA.

A new approach for the identification of bacterial species has been proposed that takes advantage of microarray technology (Cho & Tiedje, 2001). The method relies on whole genomic DNA–DNA hybridisation as the criterion for determining a bacterial species (Wayne *et al*. 1987). Random genome fragments (for example, > 100 fragments

https://doi.org/10.1079/NRR200358 Published online by Cambridge University Press

of approximately 1 kb) are spotted onto glass slide microarrays and hybridised to purified Cy dye-labelled DNA from test strains. The approach was evaluated with twelve well-characterised *Pseudomonas* strains and cluster analysis of the hybridisation profiles revealed the expected taxonomic relationships between the strains at the species level. Given the spot capacity of existing microarrays the technique could easily be used with around 1000 reference strains and thus cover the full taxonomic range of either Gram-negative or Gram-positive bacteria.

Oligonucleotide microarrays have also been exploited for determining bacterial species. By analysing hybridisation of fluorescently labelled DNA to oligonucleotides covering all possible sequence variants in a region of interest it is possible to determine the sequence of the target DNA. This approach has been used to sequence 16S RNA gene sequences and simultaneously identify different species of *Mycobacteria* (Gingeras *et al*. 1998).

Proteomics

Proteomics is the large-scale analysis of proteins and their post-translational modifications (Washburn & Yates, 2000). Protein preparations are first separated by one- or two-dimensional gel electrophoresis in polyacrylamide gels and then identified using MS. The mass spectrometric identification relies on the analysis of peptides generated by digestion of the protein of interest with a sequence-specific protease such as trypsin. Typically the protein spots are stained and excised from the gel and then digested with trypsin before elution and analysis with matrix-assisted laser desorption and ionisation (MALDI), which results in a peptide mass fingerprint of the protein that can be searched against the predicted spectra for proteins in the databases (Jensen *et al*. 1997; Berndt *et al*. 1999). Advances in automation have made the high throughput identification of proteins possible and the availability of a genome sequence greatly facilitates the identification of the proteins by MS. It is also possible to gain sequence information for some of the peptides by their initial separation using MS and then subsequent ionisation of the individual peptides into fragments detected by a tandem mass spectrometer unit. MALDI Quadrupole time-of-flight (Q-tof) spectrometers combine a MALDI ion source and a sensitive tandem mass spectrometer unit that can fragment the individual peptides (Shevchenko *et al*. 2000). In addition to the above methods, peptides from the most abundant proteins in complex mixtures of proteins can be quantified by differential labelling of two samples with stable isotopes and analysis of the mixture by MS (Oda *et al*. 1999). For example, microbes can be grown in the presence of medium containing ^{14}N or ^{15}N. Recently, a similar approach was described based on the differential labelling of cysteines with a non-radioactive isotope in protein mixtures before mixing and quantification by MS (Gygi *et al*. 1999).

Proteomics can therefore be of considerable value in identifying novel proteins and proteins with altered levels of expression under different conditions or in mutant strains of bacteria. Silver staining has become a popular method of staining proteins in gels because it is more sensitive and compatible with MS. However, the newer fluorescent protein stains such as Sypro Ruby are gaining popularity for visualisation and quantitative measurements of protein expression because they are highly sensitive and have a greater dynamic range of detection than Silver or Coomassie staining (Leimgruber *et al*. 2002).

Proteomics is complementary to microarray analysis of transcription because it can measure changes in protein expression that are the end result of transcriptional and translational regulatory processes as well as post-translational modifications. It is already known that protein abundance does not always correlate with mRNA abundance and that post-translational modifications of proteins are not apparent from genomic sequence or mRNA expression data. However, phosphorylation, glycosylation and protein processing can be extremely important factors influencing the activity, stability and localisation of proteins in both prokaryotic and eukaryotic cells. Proteomics can also be used to identify subsets of the proteome as reported for the cell envelope proteins of *S. Typhimurium* (Qi *et al*. 1996). Protein expression analysis will also be useful in the analysis of secreted proteins found in the supernatant fractions of microbial cultures and for the elucidation of their biological functions (Antelmann *et al*. 2001). In addition, immunoblotting methods can be combined with two-dimensional gel electrophoresis to identify microbial antigens recognised by the host during infection that might be exploited for the development of vaccines and anti-infective strategies (Jungblut, 2001).

Genomics and stress response of food-borne bacterial pathogens

Food-borne bacterial pathogens are often ubiquitous in nature or present in large animal reservoirs. Their capacity to survive stresses in the external environment and, in the case of zoonotic pathogens, within animal hosts are important determinants in their transmission to foods. To inactivate pathogenic organisms in foods or control their outgrowth during storage, food processing and preservation techniques are employed that impose various physiological stresses such as heat, acid, high osmolarity, various preservatives and relatively novel techniques, for example, modified atmosphere storage or high pressure treatment (Smelt, 1998; Leistner, 2000). An understanding of how bacteria cope with stressful conditions is therefore potentially useful in designing new intervention strategies and control methods for managing food safety.

Most bacteria have a general stress response that provides multiple resistances. It has been shown for several Gram-negative bacteria including *Salmonella* and *E. coli* that the general stress response is mediated by the alternative sigma factor RpoS (σ^S) (Fang *et al*. 1992; Loewen *et al*. 1998; Hengge-Aronis, 2000, 2002). The RpoS regulon has been best studied in the model organism *E. coli*; σ^S directs the RNA polymerase to promoters of over seventy genes and can mediate resistance to oxidative stress, near-u.v. irradiation, potentially lethal heat shocks, hyperosmolarity, low pH, and ethanol (for more details, see Hengge-Aronis, 2000). Often, a response to one stress simultaneously mediates increased resistance to other stresses (cross-resistance). Although the general functions

https://doi.org/10.1079/NRR200358 Published online by Cambridge University Press

of RpoS seem to be similar in enteric bacteria, Ibanez-Ruiz *et al*. (2000) demonstrated that the composition of the regulon shows differences between closely related species such as *Salmonella* and *E. coli*. The genome sequences of various food-borne *E. coli* and *Salmonella* isolates have recently been identified or are currently in progress, and will probably reveal new RpoS-regulated genes, including genes involved in virulence. In Gram-positive bacteria, the alternative sigma factor σ^B plays a similar important role in regulation of the general stress response (Voelker *et al*. 1999; Price, 2000; Hecker & Voelker, 2001). DNA microarray and proteomic studies are now enabling rapid identification of genes that are under the control of σ^B (Gertz *et al*. 2000; Petersohn *et al*. 2001).

Exposure of bacteria to a certain stress can also trigger specific adaptive responses that subsequently confer increased resistance to a higher level of the same stress that would otherwise be lethal. Genomics is having a major impact on the understanding and analysis of such regulatory systems. DNA microarray expression profiling of genetic mutants and wild-type strains has been used to investigate the genes regulated by specific regulators involved in the oxidative stress response of *E. coli*, such as OxyR (Zheng *et al*. 2001) and SoxS (Pomposiello *et al*. 2001). Furthermore, studies have been undertaken to determine genes controlled by pleiotropic regulators such as H-NS (Hommais *et al*. 2001) and the leucine responsive protein in *E. coli* (Hung *et al*. 2002). DNA microarray analysis has also proven to be useful in determining genes controlled by two-component regulatory systems involved in sporulation in *Bacillus* (Fawcett *et al*. 2000), and signal transduction at low temperatures (Kobayashi *et al*. 2001). Transcription profiling and proteomics approaches have furthermore been used to investigate cell responses to a variety of stress conditions relevant to food processing and food safety; for example, heat stress (Helmann *et al*. 2001; Riehle *et al*. 2001; Periago *et al*. 2002), acid stress (Blankenhorn *et al*. 1999; Arnold *et al*. 2001; Oh *et al*. 2002), oxidative stress (Pomposiello *et al*. 2001; Zheng *et al*. 2001) and osmotic stress (Duche *et al*. 2002). Increased knowledge about the stress responses of food-borne bacteria can possibly predict for cross-resistances to various preservation treatments, and help in the rational design of combined treatments that target different cellular mechanisms.

Upon entry of food-borne bacteria in the gastrointestinal tract of humans, the organisms are exposed to host defences such as highly acidic conditions in the stomach and high bile salt concentrations in the duodenum. To cause infection at the right place at the right time, invasive bacteria such as *L. monocytogenes*, *Salmonella*, *E. coli* and *Campylobacter* regulate gene expression of their 'weaponry' tightly, often in response to their environment (for example, temperature, Fe availability, oxidative burst and other factors, many of which are still unknown). These pathogens usually express specific virulence factors and toxins that mediate infection. In *L. monocytogenes*, it is becoming clear that its ability to cause disease is also influenced by the expression of stress response genes. Various reports have recently demonstrated that altered expression of the ATP-dependent ClpP protease and one of its subunits, ClpC, both belonging to the class III heat shock proteins, can lead to significantly attenuated virulence of *L. monocytogenes* (Rouquette *et al*. 1996; Nair *et al*. 1999; Gaillot *et al*. 2000; Karatzas *et al*. 2002). Typical stress response genes may therefore also play an important role in infection.

Analysis of complete genomes has clearly shown that considerable diversity, and thereby also the ability to adapt to certain environments, can result from acquisition and deletion of genes (see also the section on genomics; pp. 23–25). In addition, more subtle changes can lead to profoundly different phenotypes. Mechanisms responsible for generating phenotypic variation often involve changes in repetitive DNA elements by strand slipping of DNA polymerases (for example, surface structure synthesis in *C. jejuni*; Parkhill *et al*. 2000) or by recombination of repeat regions (for example, the *fim* switch in *E. coli*; Blomfield, 2001). Such mechanisms can generate a subset of cells within a population with intrinsically different properties that may allow for the survival of only a subset of cells during stress imposed by the environment or the host (Bayliss *et al*. 2001; Blomfield, 2001; Hallet, 2001). An important functional role of such repeats in survival is suggested by the findings that the genome of *E. coli* contains an over-representation of repeats in stress response genes (Rocha *et al*. 2002) while, for example, in *C. jejuni* many of the genes with repeats are found in loci involved in the synthesis of surface structures (Parkhill *et al*. 2000).

With the completion of the genome sequences of various food-borne pathogens, rapid progress can be expected in defining the cellular response of different food-borne pathogenic bacteria. Knowledge of bacterial stress and survival mechanisms may not only help the design of new intervention strategies and control methods for known pathogens but may also be valuable in combating emerging pathogens in new niches in the food chain.

Genomics and host–microbe interactions

As mentioned earlier (p. 29), the majority of the food-borne illnesses of bacterial origin are caused by pathogens that enter the gastrointestinal tract of man and overcome various host barriers (for example, gastric juice and bile salts). Different pathogens cause different types of illnesses, resulting from toxin production in the gut lumen and/or invasion of the gut tissue, sometimes leading to systemic infections. The analysis of bacterial transcription in the environment of the host is a promising approach to understanding bacterial virulence. Furthermore, the availability of sequences for human and animal genomes will allow the response of the host to be investigated during infection.

DNA microarray transcription profiling holds exceptional promise for the analysis of host–microbe interactions although investigation of bacterial gene expression in the environment of the host is currently technically challenging. Rapid procedures are needed to ensure that changes in gene transcription are not altered during the isolation of sufficient numbers of bacteria from the infected animal. An alternative approach is to rapidly fix the tissue or luminal samples before mRNA is prepared for expression analysis. However, sensitivity of detection of bacterial gene expression will probably be a problem if the sample contains sub-

https://doi.org/10.1079/NRR200358 Published online by Cambridge University Press

stantial amounts of mRNA from other prokaryotes or from the host cells and tissues. The use of customised primers for the labelling of pathogen-specific mRNA and PCR-based methods for linear amplification of cDNA may help increase the sensitivity of detection (Talaat *et al*. 2000). Alternatively, pathogen gene expression can be investigated in environments that mimic aspects of the host environment or in cell and organ culture models (Mekalanos, 1992; Cotter & Miller, 1998). Recently Eriksson *et al*. (2003) reported on the complete transcription profile of intracellular *S. Typhimurium* following infection of macrophages. Four hundred and eight genes of unknown function, several well-characterised virulence factors and SOS genes were altered in expression during infection of macrophages. The *Salmonella* response also provided some new insights into the possible environment of the host cell vacuole. Genome microarrays also promise to accelerate the discovery of events occurring in host cells as a consequence of host–microbe interactions (Rosenberger *et al*. 2001). Here the uses of different cultured cell lines will probably be highly informative in interpreting the complex response of different host cell types and tissues. The first examples of using this approach to study the host response have been with human cell lines infected with cytomegalovirus or HIV-1 (Zhu *et al*. 1998; Corbeil *et al*. 1999; Geiss *et al*. 2000). Studies with cultured macrophages and *Salmonella* showed that the expression of pro-inflammatory cytokines, chemokines, signalling molecules and transcriptional activators as well as several other genes not previously known to be regulated in response to infection were increased in expression (Eckmann *et al*. 2000; Rosenberger *et al*. 2000). Interestingly, activation of the macrophages modified the response to infection with *Salmonella*. In similar studies on the transcriptional response of human promyelocytic cells to *L. monocytogenes* infection, several pro-inflammatory factors identified in the experiments with *Salmonella* were also increased in expression. There were also a number of notable differences; for example, in the induction of apoptosis-promoting genes by *Salmonella* and anti-apoptotic genes by *L. monocytogenes*. It is tempting to speculate that this might reflect differences in the bacterial interactions with the host cells but the use of different cell lines and methods prevents firm conclusions being drawn from this comparison.

Several other recent papers have described the host transcriptional response to non-food-related bacterial pathogens and have provided new insights into the complex interaction between host and microbe (Belcher *et al*. 2000; Ichikawa *et al*. 2000; Coombes & Mahony, 2001; Diehn & Relman, 2001). Transcriptional analysis of intestinal tissue in gnotobiotic mice colonised with *Bacteroides thetaiotaomicro*, a prevalent commensal of the gastrointestinal tract, revealed changes in gene expression associated with several important intestinal functions (Hooper *et al*. 2001). In the same study a comparison of transcriptional responses to other gut commensal bacteria revealed both common and bacterial species-specific responses of the host and indicated the importance of host–commensal interactions in modulating normal intestinal functions.

In the not too distant future, microarray expression studies will probably generate a large body of data on host–pathogen interactions and enormous efforts using traditional life science approaches will be required to confirm new hypotheses. Nevertheless, advances in this field have exciting prospects for novel medical applications and therapeutics.

Conclusions

The complete genomes of several food-borne pathogens are now known, and powerful genetic and bioinformatics tools have been developed to investigate the biological significance of the information contained within the genome and the extrachromosomal elements. Genome sequences of several closely related species and of different strains of a species of some important bacterial pathogens are becoming available at a rapid pace. Genome comparisons between closely related pathogenic and non-pathogenic species or less virulent strain variants could identify DNA elements that contribute to the different virulence potential of individual strains.

Genome sequence analysis and comparative genome sequence analysis has also highlighted the extent of genetic diversity among many food-borne bacterial pathogens and the important role that horizontal gene transfer has played in generating the genetic heterogeneity that exists in bacterial species. This research is providing fascinating insights into microbial evolution and the mechanisms leading to the evolution of pathogenic bacteria. A challenge for the future will be to understand the biological significance of this heterogeneity in the context of the complex interplay between the host and microbe and to predict the potential for emergence of new pathogens in different ecological niches. Here DNA microarrays offer the possibility to investigate this diversity in bacterial populations and to correlate the presence or absence of specific DNA sequences with phenotypic properties. Once virulence factors have been identified that correlate with infection or different disease outcomes they could be potential diagnostic markers for the comprehensive surveillance and monitoring of food-borne pathogens in the food chain. If multiplex PCR methods for amplifying diagnostic sequences are combined with detection on microarrays, more sensitive and informative tools might be expected to be seen for the future surveillance and typing of food-borne pathogens in food, livestock and the environment.

Proteomics and microarray technologies allow a genome-wide analysis of gene and protein expression under different experimental conditions and this is having a major impact on the study of differentially regulated genes during adaptation to environmental stress and under conditions that are intended to simulate *in vivo* conditions. The transcriptional responses observed in response to change in the environmental condition are often complex and pleiotrophic but the use of short time courses and conditions that result in the same growth rates under both conditions can help elucidate primary from secondary effects on gene transcription. Expression profiling is providing new information about the gene composition of regulons and differentially regulated genes even if the conditions have been previously well studied. Additionally the classification of genes by expression profiling patterns is a useful

https://doi.org/10.1079/NRR200358 Published online by Cambridge University Press

approach to predict the general function of unknown genes. The genomics era is an exciting time for research on microbial pathogens and tremendous progress in research has already been made in the field of food-borne pathogens especially for those food-borne pathogens that have not been so well studied previously.

Acknowledgements

J. M. W. and M. H. J. B. are grateful for financial support from the BBSRC Core Strategic Grant. J. M. W. is grateful for financial support for post-genomic research in his laboratory from The European Bacterial Proteomics (EBP) Network (EC contract QLK2-CT-2000–01536), TCS-TARGETS (EC contract QLK2-CT-2000-00543) and BAS antimicrobials (EC contract QLK3-CT-2001-00277). M. H. J. B. and J. M. W. would also like to thank Drs Bruce Pearson, Francis Mulholland, Kathryn Holmes, Carmen Pin, Willem de Vos, Tjakko Abee, Andreas Karatzas and Trinad Chakraborty for helpful discussions and Thérèse Hall for help with preparation of the manuscript.

References

Adak GK, Long SM & O'Brien SJ (2002) Trends in the indigenous foodborne disease and deaths, England and Wales: 1992–2000. *Gut* **51**, 832–841.

Arfin SM, Long AD, Ito ET, Tolleri L, Riehle MM, Paegle EES & Hatfield GW (2000) Global gene expression profiling in *Escherichia coli* K12. The effects of integration host factor. *Journal of Biological Chemistry* **275**, 29672–29684.

Alm RA, Ling LS, Moir DT, King BL, Brown ED, Doig PC, Smith DR, Noonan B, Guild BC, deJonge BL, Carmel G, Tummino PJ *et al*. (1999) Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* **397**, 176–180.

Antelmann H, Tjalsma H, Voigt B, Ohlmeier S, Bron S, Maarten van Dijl J & Hecker M (2001) A proteomic view on genome-based signal peptide predictions. *Genome Research* **11**, 1484–1502.

Arnold CN, McElhanon J, Lee A, Leonhart R & Siegele DA (2001) Global analysis of *Escherichia coli* gene expression during the acetate-induced acid tolerance response. *Journal of Bacteriology* **183**, 2178–2186.

Bacon DJ, Alm RA, Burr DH, Hu L, Kopecko DJ, Ewing CP, Trust TJ & Guerry P (2000) Involvement of a plasmid in virulence of *Campylobacter jejuni* 81–176. *Infection and Immunity* **68**, 4384–4390.

Balaban N & Rasooly A (2000) Staphylococcal enterotoxins. *International Journal of Food Microbiology* **61**, 1–10.

Bammert GF & Fostel JM (2000) Genome-wide expression patterns in *Saccharomyces cerevisiae*: a comparison of drug treatments and genetic alterations affecting biosynthesis of ergosterol. *Antimicrobial Agents and Chermotherapy* **44**, 1255–1265.

Bavykin SG, Akowski JP, Zakhariev VM, Barsky VE, Perov AN & Mirzabekov AD (2001) Portable system for microbial sample preparation and oligonucleotide microarray analysis. *Applied and Environmental Microbiology* **67**, 922–928.

Bayliss CD, Field D & Moxon ER (2001) The simple sequence contingency loci of *Haemophilus influenzae* and *Neisseria meningitidis*. *Journal of Clinical Investigation* **107**, 657–662.

Belcher CE, Drenkow J, Kehoe B, Gingeras TR, McNamara N, Lemjabbar H, Basbaum C & Relman DA (2000) From the cover: the transcriptional responses of respiratory epithelial cells to *Bordetella pertussis* reveal host defensive and pathogen counter-defensive strategies. *Proceedings of the National Academy of Sciences USA* **97**, 13847–13852.

Berndt P, Hobohm U & Langen H (1999) Reliable automatic protein identification from matrix-assisted laser desorption/ionization mass spectrometric peptide fingerprints. *Electrophoresis* **20**, 3521–3526.

Blankenhorn D, Phillips J & Slonczewski JL (1999) Acid- and base-induced proteins during aerobic and anaerobic growth of *Escherichia coli* revealed by two-dimensional gel electrophoresis. *Journal of Bacteriology* **181**, 2209–2216.

Blattner FR, Plunkett G III, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW *et al*. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1474.

Blohm DH & Guiseppi-Elie A (2001) New developments in microarray technology. *Current Opinion in Biotechnology* **12**, 41–47.

Blomfield IC (2001) The regulation of pap and type 1 fimbriation in *Escherichia coli*. *Advances in Microbial Physiology* **45**, 1–49.

Bottone EJ (1999) *Yersinia enterocolitica*: overview and epidemiologic correlates. *Microbes and Infection* **1**, 323–333.

Bowtell DD (1999) Options available – from start to finish – for obtaining expression data by microarray. *Nature Genetics* **21**, 25–32.

Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M, Roth R, George D *et al*. (2000) Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nature Biotechnology* **18**, 630–634.

Brenner SE (1999) Errors in genome annotation. *Trends in Genetics* **15**, 132–133.

Bryan FL (1982) *Diseases Transmitted by Foods – A Classification and Summary*. Atlanta, GA: Centers for Disease Control.

Call DR, Brockman FJ & Chandler DP (2001) Detecting and genotyping *Escherichia coli* O157:H7 using multiplexed PCR and nucleic acid microarrays. *International Journal of Food Microbiology* **67**, 71–80.

Call DR, Borucki MK & Besser TE (2003) Mixed-genome microarrays reveal multiple serotype and lineage-specific differences among strains of *Listeria monocytogenes*. *Journal of Clinical Microbiology* **41**, 632–639.

Chalker RB & Blaser MJ (1988) A review of human salmonellosis: III. Magnitude of Salmonella infection in the United States. *Review of Infectious Diseases* **10**, 111–124.

Chart H, Frost JA, Oza A, Thwaites R, Gillanders S & Rowe B (1996) Heat-stable serotyping antigens expressed by strains of *Campylobacter jejuni* are probably capsular and not long-chain lipopolysaccharide. *Journal of Applied Bacteriology* **81**, 635–640.

Cheung VG, Morley M, Aguilar F, Massimi A, Kucherlapati R & Childs G (1999) Making and reading microarrays. *Nature Genetics Review* **21**, 15–19.

Cho JC & Tiedje JM (2001) Bacterial species determination from DNA-DNA hybridization by using genome fragments and DNA microarrays. *Applied and Environmental Microbiology* **67**, 3677–3682.

Cho JC & Tiedje JM (2002) Quantitative detection of microbial genes by using DNA microarrays. *Applied and Environmental Microbiology* **68**, 1425–1430.

Chu S, DeRisi J, Eisen M, Mulholland J, Botstein D, Brown PO & Herskowitz I (1998) The transcriptional program of sporulation in budding yeast. *Science* **282**, 699–705.

Coombes BK & Mahony JB (2001) cDNA array analysis of altered gene expression in human endothelial cells in response

to *Chlamydia pneumoniae* infection. *Infection and Immunity* **69**, 1420–1427.

Corbeil J, Sheeter D, Rought S, Du P, Ferguson M, Masys DR, Welsh JB, Fink JL, Sasik R, Huang D, Drenkow J, Richman DD & Gingeras T (1999) Magnitude and specificity of temporal gene expression during HIV-1 infection of a CD4+ T cell. *Nature Genetics* **23**, 39–40.

Cotter PA & Miller JF (1998) *In vivo* and *ex vivo* regulation of bacterial virulence gene expression. *Current Opinion in Microbiology* **1**, 17–26.

Day WAJ, Fernandez RE & Maurelli AT (2001) Pathoadaptive mutations that enhance virulence: genetic organization of the cadA regions of *Shigella* spp. *Infection and Immunity* **69**, 7471–7480.

DeRisi J (2000) Overview of nucleic acid arrays. In *Current Protocols in Molecular Biology*, pp. 22.1.1–22.3.6 [FA Ausubel, editor]. New York, NY: John Wiley & Sons.

Diehn M & Relman DA (2001) Comparing functional genomic datasets: lessons from DNA microarray analyses of host-pathogen interactions *Current Opinion in Microbiology* **4**, 95–101.

Dorrell N, Mangan JA, Laing KG, Hinds J, Linton D, Al-Ghusein H, Barrell BG, Parkhill J, Stoker NG, Karlyshev AV, Butcher PD & Wren BW (2001) Whole genome comparison of *Campylobacter jejuni* human isolates using a low-cost microarray reveals extensive genetic diversity. *Genome Research* **11**, 1706–1715.

Duche O, Tremoulet F, Glaser P & Labadie J (2002) Salt stress proteins induced in *Listeria monocytogenes. Applied and Environmental Microbiology* **68**, 1491–1498.

Eckmann L, Smith JR, Housley MP, Dwinell MB & Kagnoff MF (2000) Analysis by high density cDNA arrays of altered gene expression in human intestinal epithelial cells in response to infection with the invasive enteric bacteria salmonella. *Journal of Biological Chemistry* **275**, 14084–14094.

Edwards AE, Olsen GJ & Maloy SR (2002) Comparative genomics of closely related salmonellae. *Trends in Microbiology* **10**, 94–99.

Eriksson S, Lucchini S, Thompson A, Rhen M & Hinton JC (2003) Unravelling the biology of macrophage infection by gene expression profiling of intracellular *Salmonella enterica*. *Molecular Microbiology* **47**, 103–118.

Ezzell JW Jr, Abshire TG, Little SF, Lidgerding BC & Brown C (1990) Identification of *Bacillus anthracis* by using monoclonal antibody to cell wall galactose-N-acetylglucosamine polysaccharide. *Journal of Clinical Microbiology* **28**, 223–231.

Fang FC, Libby SJ, Buchmeier NA, Loewen PC, Switala J, Harwood J & Guiney DG (1992) The alternative sigma factor KatF (RpoS) regulates *Salmonella* virulence. *Proceedings of the National Academy of Sciences USA* **96**, 7502–7507.

Fawcett P, Eichenberger P, Losick R & Youngman P (2000) The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. *Proceedings of the National Academy of Sciences USA* **97**, 8063–8068.

Fodor SP, Read JL, Pirrung MC, Stryer L, Lu AT & Solas D (1991) Light-directed spatially addressable parallel chemical synthesis *Science* **251**, 767–773.

Fox AJ, Gilbart J & Morgan S (1990) Analytical microbiology: a perspective. In *Analytical Microbiology Methods: Chromatography and Mass Spectrometry*, pp. 1–17 [A Fox, L Larsson, SL Morgan and G Odham, editors]. New York, NY: Plenum Press.

Frangeul L, Nelson KE, Buchrieser C, Danchin A, Glaser P & Kunst F (1999) Cloning and assembly strategies in microbial genome projects. *Microbiology* **145**, 2625–2634.

Fraser CM & Fleischmann RD (1997) Strategies for whole microbial genome sequencing and analysis. *Electrophoresis* **18**, 1207–1216.

Gaillot O, Pellegrini E, Bregenholt S, Nair S & Berche P (2000) The ClpP serine protease is essential for the intracellular parasitism and virulence of *Listeria monocytogenes*. *Molecular Microbiology* **35**, 1286–1294.

Geiss GK, Bumgarner RE, An MC, Agy MB, van 't Wout AB, Hammersmark E, Carter VS, Upchurch D, Mullins JI & Katze MG (2000) Large-scale monitoring of host cell gene expression during HIV-1 infection using cDNA microarrays. *Virology* **266**, 8–16.

Gertz S, Engelmann S, Schmid R, Ziebandt AK, Tischer K, Scharf C, Hacker J & Hecker M (2000) Characterization of the sigma(B) regulon in *Staphylococcus aureus*. *Journal of Bacteriology* **182**, 6983–6991.

Gingeras TR, Ghandour G, Wang E, Berno A, Small PM, Drobniewski F, Alland D, Desmond E, Holodniy M & Drenkow J (1998) Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic *Mycobacterium* DNA arrays. *Genome Research* **8**, 435–448.

Glaser P, Frangeul L, Buchrieser C, Rusniok C, Amend A, Baquero F, Berche P, Bloecker H, Brandt P, Chakraborty T, Charbit A, Chetouani F *et al*. (2001) Comparative genomics of *Listeria* species. *Science* **294**, 849–852.

Granum PE (1990) *Clostridium perfringens* toxins involved in food poisoning. *International Journal of Food Microbiology* **10**, 101–111.

Graves DJ (1999) Powerful tools for genetic analysis come of age. *Trends in Biotechnology* **17**, 127–134.

Groisman EA & Ochman H (1996) Pathogenicity islands: bacterial evolution in quantum leaps. *Cell* **87**, 791–794.

Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH & Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnology* **17**, 994–999.

Hacker J, Blum-Oehler G, Muhldorfer I & Tschape H (1997) Pathogenicity islands of virulent bacteria: structure function and impact on microbial evolution. *Molecular Microbiology Reviews* **23**, 1089–1097.

Hallet B (2001) Playing Dr Jekyll and Mr Hyde: combined mechanisms of phase variation in bacteria. *Current Opinion in Microbiology* **4**, 570–581.

Hatfield GW, Hung SP & Baldi P (2003) Differential analysis of DNA microarray gene expression data. *Molecular Microbiology* **47**, 871–877.

Hayashi T, Makino K, Ohnishi M, Kurokawa K, Ishii K, Yokoyama K, Han C-G, Ohtsubo E, Nakayama K, Murata T, Tanaka M, Tobe T *et al*. (2001) Complete genome sequence of *Enterohemorrhagic Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Research* **8**, 11–22.

Hecker M & Voelker U (2001) General stress response of *Bacillus subtilis* and other bacteria. *Advances in Microbial Physiology* **44**, 35–91.

Helmann JD, Wu MF, Kobel PA, Gamo FJ, Wilson M, Morshedi MM, Navre M & Paddon C (2001) Global transcriptional response of *Bacillus subtilis* to heat shock. *Journal of Bacteriology* **183**, 7318–7328.

Hengge-Aronis R (2000) The general stress response in *Escherichia coli*. In *Bacterial Stress Responses*, pp. 161–178 [G Storz and R Hengge-Aaronis, editors]. Washington, DC: ASM Press.

Hengge-Aronis R (2002) Signal transduction and regulatory mechanisms involved in control of the σ^S (RpoS) subunit of RNA polymerase. *Microbiology and Molecular Biology Reviews* **66**, 373–395.

Hommais F, Krin E, Laurent-Winter C, Soutourina O, Malpertuy A, Le Caer JP, Danchin A & Bertin P (2001) Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Molecular Microbiology* **40**, 20–36.

Hooper LV, Wong MH, Thelin A, Hansson L, Falk PG & Gordon JI (2001) Molecular analysis of commensal host-microbial relationships in the intestine. *Science* **291**, 881–884.

Hu L & Kopecko DJ (1999) *Campylobacter jejuni* 81–176 associates with microtubules and dynein during invasion of human intestinal cells. *Infection and Immunity* **67**, 4171–4182.

Hughes TR, Marton MJ, Jones AR, Roberts CJ, Stoughton R, Armour CD, Bennett HA, Coffey E, Dai H, He YD, Kidd MJ, King AM *et al*. (2000) Functional discovery via a compendium of expression profiles. *Cell* **102**, 109–126.

Hung SP, Baldi P & Hatfield GW (2002) Global gene expression profiling in *Escherichia coli* K12. The effects of leucine-responsive regulatory protein. *Journal of Biological Chemistry* **277**, 40309–40323.

Ibanez-Ruiz M, Robbe-Saule V, Hermant D, Labrude S & Norel F (2000) Identification of RpoS (σ^S)-regulated genes in *Salmonella enterica* Serovar Typhimurium. *Journal of Bacteriology* **182**, 5749–5756.

Ichikawa JK, Norris A, Bangera MG, Geiss GK, van 't Wout AB, Bumgarner RE & Lory S (2000) Interaction of *Pseudomonas aeruginosa* with epithelial cells: identification of differentially regulated genes by expression microarray analysis of human cDNAs. *Proceedings of the National Academy of Sciences USA* **97**, 9659–9664.

Israel DA, Salama N, Arnold CN, Moss SF, Ando T, Wirth HP, Tham KT, Camorlinga M, Blaser MJ, Falkow S & Peek RM Jr (2001) *Helicobacter pylori* strain-specific differences in genetic content, identified by microarray, influence host inflammatory responses. *Journal of Clinical Investigation* **107**, 611–620.

Jensen ON, Mortensen P, Vorm O & Mann M (1997) Automation of matrix-assisted laser desorption/ionization mass spectrometry using fuzzy logic feedback control. *Analytical Chemistry* **69**, 1706–1714.

Jungblut PR (2001) Proteome analysis of bacterial pathogens. *Microbes and Infection* **3**, 831–840.

Kaferstein F & Abdussalam M (1999) Food safety in the 21st century. *Bulletin of the World Health Organization* **77**, 347–351.

Karatzas KAG, Wouters JA, Gahan CGM, Hill C, Abee T & Bennik MHJ (2002) A single amino acid deletion in the glycine-rich region of Ctsr effects mobility, piezotolerance, and virulence of *Listeria monocytogenes* Scott A. In *Listeria Monocytogenes* – Inactivation by Essential Oils and High Pressure and Contribution of Genetic Variation to Stress Resistance, pp. 54–62. PhD Thesis. Wageningen, The Netherlands: Wageningen University.

Karlyshev AV, Linton D, Gregson NA, Lastovica AJ & Wren BW (2000) Genetic and biochemical evidence of a *Campylobacter jejuni* capsular polysaccharide that accounts for Penner serotype specificity. *Molecular Microbiology* **35**, 529–541.

Kim CC, Joyce EA, Chan K & Falkow S (2002) Improved analytical methods for microarray-based genome composition analysis. *Genome Biology* **3**, 1–16.

Kobayashi K, Ogura M, Yamaguchi H, Yoshida K, Ogasawara N, Tanaka T & Fujita Y (2001) Comprehensive DNA microarray analysis of *Bacillus subtilis* two-component regulatory systems. *Journal of Bacteriology* **183**, 7365–7370.

Kotiranta A, Lounatmaa K & Haapasalo M (2000) Epidemiology and pathogenesis of *Bacillus cereus* infections. *Microbes and Infection* **2**, 189–198.

Kotloff KL, Winickoff JP, Ivanoff B, Clemens JD, Swerdlow DL, Sansonetti PJ, Adak GK & Levine MM (1999) Global burden of shigella infections: implications for vaccine development and implementation of control strategies. *Bulletin of the World Health Organization* **77**, 651–656.

Lee CA (1996) Pathogenicity islands and the evolution of bacterial pathogens. *Infectious Agents and Disease* **5**, 1–7.

Lee CY, Tai CL, Lin SC & Chen YT (1994) Occurrence of plasmids and tetracycline resistance among *Campylobacter jejuni* and *Campylobacter coli* isolated from whole market chickens and clinical samples. *International Journal of Food Microbiology* **24**, 161–170.

Lee ML, Kuo FC, Whitmore GA & Sklar J (2000) Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. *Proceedings of the National Academy of Sciences USA* **97**, 9834–9839.

Leimgruber RM, Malone JP, Radabaugh MR, LaPorte ML, Violand BN & Monahan JB (2002) Development of improved cell lysis, solubilization and imaging approaches for proteomic analyses. *Proteomics* **2**, 135–144.

Leistner L (2000) Basic aspects of food preservation by hurdle technology. *International Journal of Food Microbiology* **55**, 181–186.

Lipshutz RJ, Fodor SP, Gingeras TR & Lockhart DJ (1999) High density synthetic oligonucleotide arrays. *Nature Genetics Review* **21**, Suppl., 20–24.

Loewen PC, Hu B, Strutinsky J & Sparling R (1998) Regulation in the rpoS regulon of *Escherichia coli*. *Canadian Journal of Food Microbiology* **44**, 707–717.

Lund BM, Baird-Parker TC & Gould GW (2000) *The Microbiological Safety and Quality of Food*, vol. 2. Frederick, MD: Aspen Publishers.

McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, Hou S, Layman D *et al*. (2001) Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* **413**, 852–856.

Martin GP & Trimmers E (1997) PCR and its modifications for the detection of infectious diseases. In *Nucleic Acid Amplification Technologies: Application to Disease Diagnosis*, pp. 79–99 [H Lee, S Morse and O Olsvik, editors]. Natick, MA: Eaton Publishing.

Maurelli AT, Fernandez RE, Bloch CA, Rode CK & Fasano A (1998) 'Black holes' and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp and enteroinvasive *Escherichia coli*. *Proceedings of the National Academy of Sciences USA* **95**, 3943–3948.

Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, Griffin PM & Tauxe RV (1999) Food-related illness and death in the United States. *Emerging and Infectious Diseases* **5**, 607–625.

Mekalanos JJ (1992) Environmental signals controlling expression of virulence determinants in bacteria. *Journal of Bacteriology* **174**, 1–7.

Meldrum D (2000) Automation for genomics part two: sequencers microarrays and future trends. *Genome Research* **10**, 1288–1303.

Mishu B, Ilyas AA, Koski CL, Vriesendorp F, Cook SD, Mithen FA & Blaser MJ (1993) Serologic evidence of previous *Campylobacter jejuni* infection in patients with the Guillain-Barre syndrome. *Annals of Internal Medicine* **118**, 947–953.

Moxon ER, Rainey PB, Nowak MA & Lenski RE (1994) Adaptive evolution of highly mutable loci in pathogenic bacteria. *Current Biology* **4**, 24–33.

Nachamkin I, Allos BM & Ho T (1998) *Campylobacter* species and Guillain-Barre syndrome. *Clinical Microbiological Reviews* **11**, 555–567.

Nair S, Fréhel C, Nguyen L, Escuyer V & Berche P (1999) ClpE, a novel member of the HSP100 family, is involved in cell division and virulence of *Listeria monocytogenes*. *Molecular Microbiology* **31**, 185–196.

Nataro JP, Seriwatana J, Fasano A, Maneval DR, Guers LD, Noriega F, Dubovsky F, Levine MM & Morris JG Jr (1995)

Identification and cloning of a novel plasmid-encoded enterotoxin of enteroinvasive *Escherichia coli* and *Shigella* strains. *Infection and Immunity* **63**, 4721–4728.

Niemeyer CM & Blohm D (1999) DNA microarrays. *Angewandte Chemie (International Edition in English)* **38**, 2865–2869.

Ochman H & Groisman EA (1995) The evolution of invasion by enteric bacteria. *Canadian Journal of Microbiology* **41**, 555–561.

Ochman H & Jones IB (2000) Evolutionary dynamics of full genome content in *Escherichia coli*. *EMBO Journal* **19**, 6637–6643.

Ochman H, Lawrence JG & Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**, 299–304.

Oda Y, Huang K, Cross FR, Cowburn D & Chait BT (1999) Accurate quantitation of protein expression and site-specific phosphorylation. *Proceedings of the National Academy of Sciences USA* **96**, 6591–6596.

Oh MK, Rohlin L, Kao KC & Liao JC (2002) Global expression profiling of acetate-grown *Escherichia coli*. *Journal of Biological Chemistry* **277**, 13175–13183.

Okamoto T, Suzuki T & Yamamoto N (2000) Microarray fabrication with covalent attachment of DNA using bubble jet technology. *Nature Biotechnology* **18**, 438–441.

Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, Basham D, Chillingworth T, Davies RM, Feltwell T, Holroyd S, Jagels K, Karlyshev AV *et al*. (2000) The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* **403**, 665–668.

Peck MW (1997) Clostridium botulinum and the safety of refrigerated processed foods of extended durability. *Trends in Food Science and Technology* **8**, 186–192.

Penner JL, Hennessy JN & Congi RV (1983) Serotyping of *Campylobacter jejuni* and *Campylobacter coli* on the basis of thermostable antigens. *European Journal of Clinical Microbiology* **2**, 378–383.

Periago PM, Van Schaik W, Abee T & Wouters JA (2002) Identification of proteins involved in the heat stress response of *Bacillus cereus* ATCC 14579. *Applied and Environmental Microbiology* **68**, 3486–3495.

Perna NT, Plunkett G III, Burland V, Mau B, Glasner JD, Rose DJ, Mayhew GF, Evans PS, Gregor J, Kirkpatrick HA, Pósfai G, Hackett J *et al*. (2001) Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**, 529–533.

Petersohn A, Brigulla M, Haas S, Hoheisel JD, Volker U & Hecker M (2001) Global analysis of the general stress response of *Bacillus subtilis*. *Journal of Bacteriology* **183**, 5617–5631.

Pomposiello PJ, Bennik MHJ & Demple B (2001) Genome-wide transcriptional profiling of the *Escherichia coli* responses to superoxide stress and sodium salicylate. *Journal of Bacteriology* **183**, 3890–3902.

Porwollik S, Wong RM & McClelland M (2002) Evolutionary genomics of Salmonella: gene acquisitions revealed by microarray analysis. *Proceedings of the National Academy of Sciences USA* **99**, 8956–8961.

Price CW (2000) Protective function and regulation of the general stress response in *Bacillus subtilis* and related Gram-positive bacteria. In *Bacterial Stress Responses*, pp. 179–197 [G Storz and R Hengge-Aaronis, editors]. Washington, DC: ASM Press.

Pupo GM, Lan R & Reeves PR (2000) Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics. *Proceedings of the National Academy of Sciences USA* **97**, 10567–10572.

Qi SY, Moir A & O'Connor CD (1996) Proteome of *Salmonella typhimurium* SL1344: identification of novel abundant cell envelope proteins and assignment to a two-dimensional reference map. *Journal of Bacteriology* **178**, 5032–5038.

Rabsch W, Tschape H & Baumler AJ (2001) Non-typhoidal salmonellosis: emerging problems. *Microbes and Infection* **3**, 237–247.

Rhodius V, Van Dyk TK, Gross C & LaRossa RA (2002) Impact of genomic technologies on studies of bacterial gene expression. *Annual Review of Microbiology* **56**, 599–624.

Richmond CS, Glasner JD, Mau R, Jin H & Blattner FR (1999) Genome-wide expression profiling in *Escherichia coli* K-12. *Nucleic Acids Research* **27**, 3821–3835.

Riehle MM, Bennett AF & Long AD (2001) Genetic architecture of thermal adaptation in *Escherichia coli*. *Proceedings of the National Academy of Sciences USA* **98**, 525–530.

Rocha EPC, Matic I & Taddei F (2002) Over-representation of repeats in stress response genes: a strategy to increase versatility under stressful conditions? *Nucleic Acids Research* **30**, 1886–1894.

Rosenberger CM, Pollard AJ & Finlay BB (2001) Gene array technology to determine host responses to *Salmonella*. *Microbes and Infection* **3**, 1353–1360.

Rosenberger CM, Scott MG, Gold MR, Hancock RE & Finlay BB (2000) *Salmonella typhimurium* infection and lipopolysaccharide stimulation induce similar changes in macrophage gene expression. *Journal of Immunology* **164**, 5894–5904.

Roth FP, Hughes JD, Estep PW & Church GM (1998) Finding DNA regulator motifs within unaligned noncoding sequences clustered by whole-genome mRNA quantitation. *Nature Biotechnology* **16**, 939–945.

Rouquette C, Ripio M-T, Pellegrini E, Bolla JM, Tascon RI, Vazquez-Boland JA & Berche P (1996) Identification of a ClpC ATPase required for stress tolerance and in vivo survival of *Listeria monocytogenes*. *Molecular Microbiology* **21**, 977–987.

Sails AD, Wareing DR, Bolton FJ, Fox AJ & Curry A (1998) Characterisation of 16 *Campylobacter jejuni* and *C. coli* typing bacteriophages. *Journal of Medical Microbiology* **47**, 123–128.

Schena M, Shalon D, Davis RW & Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470.

Schlech WF III (2001) Foodborne listeriosis. *Clinical Infection and Disease* **32**, 1518–1519.

Schuchhardt J, Beule D, Malik A, Wolski E, Eickhoff H, Lehrach H & Herzel H (2000) Normalization strategies for cDNA microarrays. *Nucleic Acids Research* **28**, E47.

Shalon D, Smith SJ & Brown PO (1996) A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Research* **6**, 639–645.

Shevchenko A, Loboda A, Ens W & Standing KG (2000) MALDI quadrupole time-of-flight mass spectrometry: a powerful tool for proteomic research. *Analytical Chemistry* **72**, 2132–2141.

Singh-Gasson S, Green RD, Yue Y, Nelson C, Blattner F, Sussman MR & Cerrina F (1999) Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nature Biotechnology* **17**, 974–978.

Smelt JPPM (1998) Recent advances in the microbiology of high pressure processing. *Trends in Food Science and Technology* **9**, 152–158.

Soini H & Musser JM (2001) Molecular diagnosis of mycobacteria. *Clinical Chemistry* **47**, 809–814.

Sokurenko EV, Hasty DL & Dykhuizen DE (1999) Pathoadaptive mutations: gene loss and variation in bacterial pathogens. *Trends in Microbiology* **7**, 191–195.

Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB, Brown PO, Botstein D & Futcher B (1998) Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by mircoarray hybridization. *Molecular Biology of the Cell* **9**, 3273–3297.

Sperandio V, Torres AG, Giron JA & Kaper JB (2001) Quorum sensing is a global regulatory mechanism in enterohemorrhagic

Escherichia coli 0157:H7. *Journal of Bacteriology* **183**, 5187–5197.

Steemers FJ, Ferguson JA & Walt DR (2000) Screening unlabeled DNA targets with randomly ordered fiber-optic gene arrays. *Nature Biotechnology* **18**, 91–94.

Strockbine NA, Marques LR, Newland JW, Smith HW, Holmes RK & O'Brien AD (1986) Two toxin-converting phages from *Escherichia coli* O157:H7 strain 933 encode antigenically distinct toxins with similar biologic activities. *Infection and Immunity* **53**, 135–140.

Talaat AM, Hunter P & Johnston SA (2000) Genome-directed primers for selective labeling of bacterial transcripts for DNA microarray analysis. *Nature Biotechnology* **18**, 679–682.

Taton TA, Mirkin CA & Letsinger RL (2000) Scanometric DNA array detection with nanoparticle probes. *Science* **289**, 1757–1760.

Tavazoie S, Hughes JD, Campbell MJ, Cho RJ & Church GM (1999) Systematic determination of genetic network architecture. *Nature Genetics* **22**, 281–285.

Tirado C & Schmidt K (2001) WHO surveillance programme for control of foodborne infections and intoxications: preliminary results and trends across greater Europe. World Health Organization. *Journal of Infection* **43**, 80–84.

van Belkum A, Scherer S, van Alphen L & Verbrugh H (1998) Short-sequence DNA repeats in prokaryotic genomes. *Microbiology and Molecular Biology Reviews* **62**, 275–293.

Vazquez-Boland JA, Kuhn M, Berche P, Chakraborty T, Dominguez-Bernal G, Goebel W, Gonzalez-Zorn B, Wehland J & Kreft J (2001) *Listeria* pathogenesis and molecular virulence determinants. *Clinical Microbiological Reviews* **14**, 584–640.

Voelker U, Maul B & Hecker M (1999) Expression of the σB-dependent general stress regulon confers multiple stress resistance in *Bacillus subtilis*. *Journal of Bacteriology* **181**, 3942–3948.

Volokhov D, Rasooly A, Chumakov K & Chizhikov V (2002) Identification of listeria species by microarray-based assay. *Journal of Clinical Microbiology*. **40**, 4720–4728.

Walker SJ, Archer P & Banks JG (1990) Growth of *Listeria monocytogenes* at refrigeration temperatures. *Journal of Applied Bacteriology* **68**, 157–162.

Washburn MP & Yates JR III (2000) Analysis of the microbial proteome. *Current Opinion in Microbiology* **3**, 292–297.

Wassenaar TM & Blaser MJ (1999) Pathophysiology of *Campylobacter jejuni* infections of humans. *Microbes and Infection* **1**, 1023–1033.

Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler MI, Krichevsky MI & Truper HG (1987) Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *International Journal of Systematic Bacteriology* **37**, 463–464.

Wei Y, Lee JM, Richmond C, Blattner FR, Rafalski JA & LaRossa RA (2001) High-density microarray-mediated gene expression profiling of *Escherichia coli*. *Journal of Bacteriology* **183**, 545–556.

Westin L, Xu X, Miller C, Wang L, Edman CF & Nerenberg M (2000) Anchored multiplex amplification on a microelectronic chip array. *Nature Biotechnology* **18**, 199–204.

Woese CR (1987) Bacterial evolution. *Microbiological Reviews* **51**, 221–271.

Zheng M, Wang X, Templeton LJ, Smulski DR, LaRossa RA & Storz G (2001) DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide. *Journal of Bacteriology* **183**, 4562–4570.

Zhu G, Spellman PT, Volpe T, Brown PO, Botstein D, Davis TN & Futcher B (2000) Two yeast forkhead genes regulate the cell cycle and pseudohyphal growth. *Nature* **406**, 90–94.

Zhu H, Cong JP, Mamtora G, Gingeras T & Shenk T (1998) Cellular gene expression altered by human cytomegalovirus: global monitoring with nucleotide arrays. *Proceedings of the National Academy of Sciences USA* **95**, 14470–14475.

https://doi.org/10.1079/NRR200358 Published online by Cambridge University Press