

REVIEW ARTICLE

Spot the difference: applications of subtractive hybridisation to the study of bacterial pathogens

CRAIG WINSTANLEY

Department of Medical Microbiology and Genito-Urinary Medicine, University of Liverpool, Duncan Building, Daulby Street, Liverpool L69 3GA, UK

Comparison of DNA from virulent strains of bacterial pathogens with DNA from less virulent or avirulent close relatives allows the identification of those genomic regions that are present only in virulent strains. Such regions are often associated with pathogenicity islands (PIs) and their characterisation can lead to a greater understanding of the pathogenesis of infectious diseases. There is now a large database of bacterial genomic sequences that provides useful reference information with which to compare the genomes of strains that exhibit variations in virulence or host preferences. Subtractive hybridisation (SH) and its sister method, suppression subtractive hybridisation (SSH), are techniques designed to identify those regions present in one genome but absent from another. The application of these techniques has led to the identification of PIs, mobile genetic elements and variations in virulence gene expression in a range of bacterial pathogens.

Introduction

One of the biggest questions facing those who study bacterial pathogenicity is – what makes some strains more pathogenic than others? The greater virulence of some bacterial strains when compared with close relatives is often associated with pathogenicity islands (PIs), regions of the genome that are absent from less virulent strains and contain additional genetic information that contributes to pathogenicity. PIs have been identified in both gram-positive and gram-negative bacterial pathogens [1–9] and many share common characteristics, including atypical G + C mol% content and the presence in flanking regions of repeated or insertion sequences (IS) or DNA with homology to phage integrase genes or plasmid origins of replication [10].

In recent times a wealth of nucleotide sequence data from the genomes of bacterial pathogens has become available. Comparative analysis of bacterial genomes provides valuable information on the physiology and evolution of bacteria and allows for detailed comparisons between related bacteria on a scale hitherto beyond reach [11]. The availability of whole genome

sequences has led to the development of micro-array technology for gene expression profiling (transcriptomics) [12] and proteomics [13]. Information regarding the many complete and ongoing genome sequence projects can be obtained from <http://www.tigr.org/tdb/mdb/mdbcomplete.html>. Not only do these databases help in the identification of potential PIs, they also serve as reference genomes for the identification of regions that are either absent from or additional to the genomes of related strains. Although in some cases the choice of strain has included highly virulent or epidemic strains, many of the bacterial genomes sequenced are from strains that were chosen on the basis of common laboratory use. There are some examples of comparative studies in which genome sequences of close relatives have been targeted. However, it is unreasonable to expect genome sequencing to provide all the answers unless every strain of interest is targeted. As two close relatives would have the majority of their genome sequence in common, this would be a wasteful and expensive approach. For that reason, several methods aimed at identifying differences between the genomes of close relatives have been developed. Whilst micro-array technology and proteomics are both valuable techniques for such studies, they are expensive and require considerable expertise, and commercially available microbial micro-arrays are few in number. Relatively simple and cheap alternatives are, therefore, attractive. This review discusses the application of subtractive hybridisation (SH) to the

Received 2 Aug. 2001; revised version received 19 Dec. 2001; accepted 17 Jan. 2002.

Corresponding author: Dr C. Winstanley (e.mail: C.Winstanley@liverpool.ac.uk).

study of genomic variations and the implications for the identification of virulence factors.

Subtractive hybridisation

A method called genomic subtraction, for the isolation of DNA absent in deletion mutants, was first reported in 1990 and applied to the study of yeast [14]. The method incorporates the main features of SH, namely the hybridisation of genomic sequences from two closely related sources, followed by the isolation of those sequences unique to one of the sources. This initial study involved the removal from wild-type DNA of sequences present in both wild-type and deletion mutant yeast genomes. DNA corresponding to the deleted region was enriched by allowing a mixture of

denatured wild-type and biotinylated mutant DNA to re-associate. After re-association, avidin-coated beads were used to remove the biotinylated sequences. Following the ligation of short oligonucleotide sequences (adaptors), the recovered DNA was amplified by using adaptor sequence as a primer in the PCR. This amplified DNA was available for sequencing or for use as a probe to screen strains and identify sequences in genomic libraries.

Although the methodology has been amended, and has appeared under more than one name, the principles of SH remain the same. A typical scheme for SH is presented in Fig. 1. Essentially, genomic DNA extracted from a 'driver' strain is hybridised with DNA extracted from a 'tester' strain with a view to isolating those sequences that are present in the tester strain but

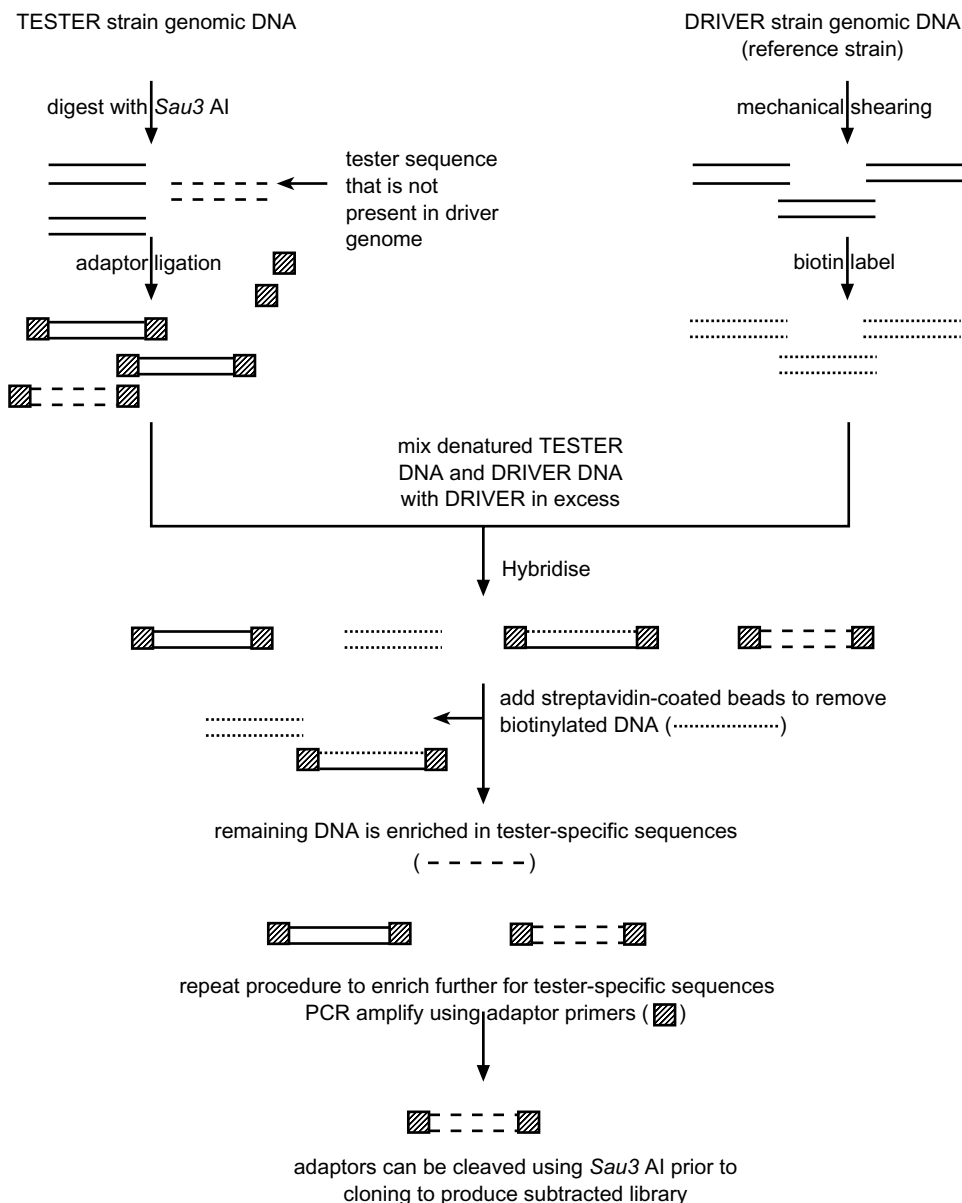


Fig. 1. Generalised scheme of subtractive hybridisation.

absent from the driver strain. To facilitate the process, the driver strain DNA should be present in excess, and it may be necessary to repeat the subtraction. In addition, the DNA is generally cleaved either by ultrasonication or by restriction enzymes before hybridisation. It is especially important to ensure that tester DNA sequences are short. In the typical scheme presented (Fig. 1), tester DNA is digested with frequently cutting restriction endonucleases (such as *Sau3AI*) and driver DNA is subjected to mechanical shearing. In the example used, *Sau3AI* can be used to cleave the adaptors and facilitate cloning into a vector cleaved with *Bam*HI, which produces complementary ends.

It is important that a method should be available for the separation of non-hybridising tester DNA from other sequences in the mixture. This can be achieved by several approaches. One modification on the original method involves the ligation of adaptors to the tester DNA and the labelling of driver DNA with biotin, both before hybridisation [15]. After hybridisation, hybrids that contain driver DNA can be removed by capture of biotin with beads coated with streptavidin. A biotin-based approach for selective enrichment of specific DNA sequences was reported as early as 1986 [16]. As in the initial study, adaptor sequences can be targeted for PCR amplification of those remaining sequences unique to the tester strain. PCR amplicons can be cloned directly with one of the many vectors available, or digested with the relevant restriction endonuclease (e.g., *Sau3AI*) before cloning. This enables the production of a subtracted library of sequences present in the tester DNA but absent from the driver DNA. One variation on the basic strategy, used in a study to identify *Listeria monocytogenes*-specific sequences, involved poly(A)-tailing of the tester DNA with terminal transferase and asymmetrical amplification after subtraction [17]. Another variation on the SH technique is representational difference analysis (RDA), which has been applied to the study of variations amongst strains of *Neisseria meningitidis* and *N. gonorrhoeae* [18–20]. This technique also involves the ligation of adaptors to restriction-digested tester DNA before hybridisation but employs two rounds of PCR amplification and hybridisation to enrich for tester-specific sequences.

A modification of SH, involving PCR amplification of tester and driver sequences before hybridisation, was reported in 1992 and used to identify strain-specific *Rhizobium* DNA sequences [21]. The modification also included the ligation of different adaptors to *Sau3AI*-digested tester and driver sequences to facilitate the PCR amplification steps. An interesting addition to the strategy was the inclusion of dUTP instead of dTTP in the driver DNA PCR mixture. This enabled the removal of any contaminating driver DNA after hybridisation by digestion with uracil DNA glycosylase [21].

Some of the earlier reports of SH applied to bacteria concerned its use with the plant pathogen *Ralstonia (Pseudomonas) solanacearum*, the causal agent of bacterial wilt. Cook and Sequeira [22] reported the development of a probe as a useful diagnostic tool for the detection of race 3 strains of the organism. Seal *et al.* [23] also used SH to isolate a *R. solanacearum*-specific DNA probe. This review will focus on the application of SH to the study of human and animal pathogens.

Suppression subtractive hybridisation (SSH)

SSH was initially reported in 1996 [24] and was first applied to bacteria in a study of *Helicobacter pylori* in 1998 [25]. As a kit for SSH is now available commercially (Clontech PCR-Select Bacterial Genome Subtraction Kit, BD Clontech UK, Hampshire), this represents the most accessible version of the technique available for inexperienced users. The kit has been applied to several different bacterial pathogens including *Aeromonas hydrophila* [26], *Burkholderia pseudomallei* [27], *Escherichia coli/Salmonella* [28, 29] and *Klebsiella pneumoniae* [30]. Fig. 2 gives an outline of the method for SSH. Briefly, tester strain DNA is digested with a restriction endonuclease (such as *Rsa*I) and separated into two portions, each of which is subjected to a ligation reaction to attach a different adaptor sequence to the 5' ends. The two portions are then separately hybridised to driver strain DNA (in excess). Any sequences that hybridise with the driver DNA should be 'mopped up', leaving only tester-specific single-stranded sequences. When the two portions are eventually mixed and hybridised together, allowing homologous single-stranded DNAs to hybridise, only those sequences unique to the tester strain will have different adaptors present on each strand. These sequences are detected by PCR amplification with primers designed to bind to adaptor sequences. The key to the success of this strategy is that sequences containing the same adaptor sequence at both ends are unable to amplify because of a suppression effect whereby a secondary structure forms to prevent primer annealing (Clontech PCR-Select™ Bacterial Genome Subtraction Kit User Manual: PT3170-1). Consequently, only those sequences now carrying both adaptors (one at each end) will amplify. These PCR products are cloned into a suitable vector to produce a subtracted DNA library. Although the procedure is not entirely effective, >50% of clones should be tester-specific. This can be tested by labelling clones and probing the tester and driver strains.

Applications

Table 1 gives an indication of the range of bacteria targeted by SH and the potential applications. SH has been particularly useful for the identification of

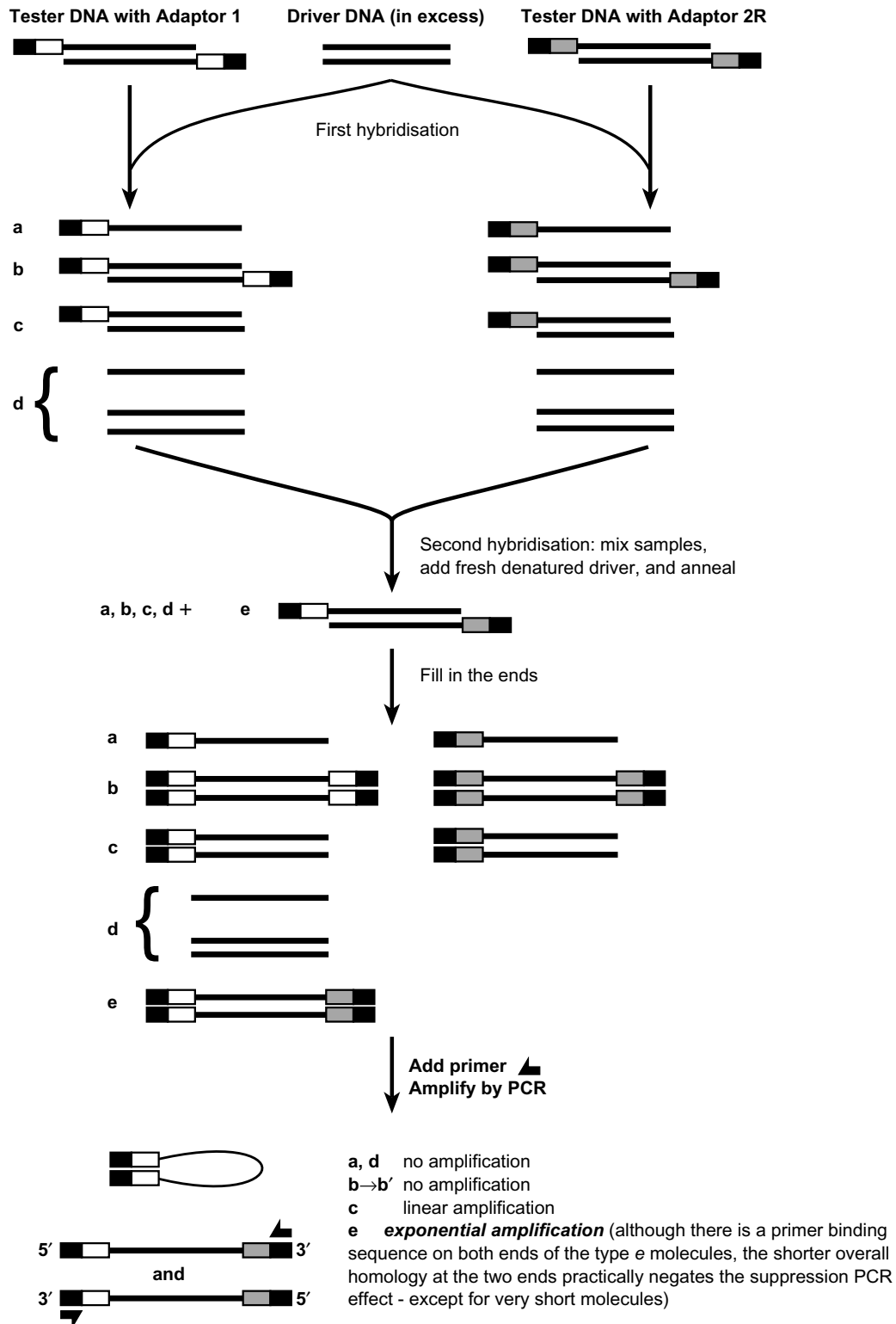


Fig. 2. Schematic diagram of PCR-Select suppression subtractive hybridisation. The figure is taken from the Clontech PCR-Select™ Bacterial Genome Subtraction kit User Manual (reproduced with the permission of BD Biosciences Clontech). 'Fill in the ends' refers to filling in the missing strands of adaptor sequences by pre-incubation at 72°C in the presence of *Taq* polymerase and nucleotides.

transposable elements or markers specific to individual strains or clones. SH has also played an important role in the identification of genomic islands implicated in virulence and in studies relating to variations in gene expression.

Identification of strain- or clone-specific markers

There have been several examples in which SH has been used to find sequences unique to a particular strain or clone to identify a useful target for diagnostic purposes. Once strain differences have been identified,

Table 1. Examples of SH applied to the study of pathogenic bacteria

Pathogen	Application/outcomes	References
<i>Aeromonas hydrophila</i>	Identification of genomic islands implicated in virulence	26
<i>Burkholderia pseudomallei</i>	Identification of genomic island implicated in virulence; identification of transposable elements	27, 39, 50
<i>Chlamydia psittaci</i>	Strain-specific diagnostic probe	31
<i>Escherichia coli</i>	Identification of variation between <i>E. coli</i> and <i>S. Typhimurium</i> ; identification of genomic islands implicated in virulence	28, 29, 41
<i>Helicobacter pylori</i>	Identification of clone-specific sequences; identification of transposable element	25, 40
<i>Klebsiella pneumoniae</i>	Identification of genomic island implicated in virulence; identification of transposable element	30
<i>Listeria monocytogenes</i>	Probes for diagnostic purposes	17
<i>Mycobacterium</i> spp.	Probes for diagnostic purposes; variations in virulence gene expression	32, 33, 45, 46
<i>Neisseria meningitidis</i> and <i>Neisseria gonorrhoeae</i>	Identification of genomic islands implicated in virulence and varying between different pathogenic species	18–20
<i>Porphyromonas gingivalis</i>	Identification of insertion sequence	38
<i>Pseudomonas aeruginosa</i>	Variations in gene expression	47
<i>Salmonella</i> spp.	Identification of genomic islands implicated in virulence; variations in gene expression	42, 43
<i>Staphylococcus aureus</i>	Identification of clone-specific markers	35
<i>Streptococcus</i> (group B)	Identification of clone-specific markers	36
<i>Treponema pallidum</i>	Identification of genomic island implicated in virulence	44

the candidate diagnostic sequence is sought in a large number of strains by hybridisation or PCR amplification to assess its specificity.

SH has been used to identify genomic regions as diagnostic targets in various animal and human pathogens [17, 31–37]. Strains of *Chlamydia psittaci* that can invade and colonise the placenta of sheep cause enzootic abortion in ewes (EAE). SH was used to identify unique fragments in the genome of an abortifacient strain (S26/3) of *C. psittaci* [31]. One such sequence, encoding a putative helicase repeated throughout the EAE genome, was proposed as a strain-specific diagnostic probe for the detection of EAE. By subtracting *Mycobacterium avium* genomic DNA from *M. bovis* genomic DNA, a novel *M. tuberculosis* complex-specific sequence was identified. Mahairas *et al.* [33] used SH to analyse genetic differences between *M. bovis* BCG (live attenuated vaccine strain) and virulent *M. bovis*, with a view to better understanding the nature of the BCG attenuation. The authors concluded that the loss of virulence by BCG is due to the loss of the genomic region RD1, which was implicated in the repression of several proteins. By re-introducing RD1 into BCG, proteins similar to those of virulent strains were obtained [33]. The differences identified by SH also offer the possibility of the development of diagnostic tests to distinguish BCG vaccination from tuberculosis infection.

El Adhami *et al.* [34] used SH to identify a 40-bp *TaqI* DNA fragment (WASA) that was present in 84% of clinical isolates of a clonal line of *Staphylococcus aureus* from hospitals in major eastern Australian cities compared with 10% of community isolates. In this example of SH, a representative strain of the clonal line was hybridised with randomly fragmented pooled DNA from 10 community isolates. The WASA

sequence was used as a probe to identify two open reading frames (ORFs) that may be related to virulence. In a subsequent report, El Adhami *et al.* [35] identified further DNA sequences associated with the clinical, clonal populations of *Staph. aureus*.

Group B streptococci (GBS) can be subdivided into three major phylogenetic lineages (III-1, III-2 and III-3), with most neonatal invasive disease being caused by III-3 strains, implying that III-3 strains are more virulent than III-2 or III-1 strains. Bohnsack *et al.* [36] used SH to identify nine short III-3-specific sequences as a first step to identifying genes that might account for differences in virulence.

Identification of mobile genetic elements

One feature that often allows closely related strains of the same bacterium to be distinguished is variation in mobile genetic elements. SH is a technique ideally suited to identifying such variation [30, 38–40]. For example, Sawada *et al.* [38] used SH to identify an insertion sequence (IS1598) that was specific for virulent strains of *Porphyromonas gingivalis*, a potent pathogen in periodontal disease. The IS1598 was present within a gene encoding an outer-membrane protein. The authors suggested that interruption of synthesis of this protein might result in changes in the bacterial outer membrane.

Members of the genus *Burkholderia* are known to be rich in insertion sequences. *B. pseudomallei* causes the disease melioidosis but is closely related to *B. thailandensis*, which appears to be less virulent or avirulent. SH of *B. pseudomallei* genomic DNA with *B. thailandensis* genomic DNA led to the identification of three clones with sequence similarity to transposable elements, only one of which, IS407, had already been

identified in *B. pseudomallei*. All three sequences were subsequently shown to be present in multiple copies in a number of *B. pseudomallei* strains [39].

Identification of PIs/genomic islands

Perhaps the most interesting application of SH is the identification of genes or genomic islands present in virulent strains but absent from related avirulent strains. The basic strategy employed is similar regardless of the pathogen being studied. Initial rounds of SH are used to identify short sequences present in virulent strains but absent from avirulent strains. Often this sequence is labelled and used as a probe against a panel of representative strains to check the specificity. Once it has been established that the sequence is specific to virulent strains the sequence is extended by cloning of larger fragments that include flanking sequences, screening of cosmid or other gene libraries, or by PCR-based techniques such as the Genome-Walker kit (Clontech). These extended sequence data should enable the identification of ORFs that can be used in BLAST searches for homologous nucleotide or predicted protein sequences in WWW sites such as <http://www.ncbi.nlm.nih.gov/BLAST/> or <http://www.ebi.ac.uk/> or individual genome sequence project DNA specific to the organism in question. Further analysis often involves specific mutation of individual genes to study the effects of such mutations on colonisation or virulence in animal models, or on phenotypic properties such as adhesion to cells, invasion or cytotoxicity.

In *Escherichia coli*, unique genomic regions associated with the virulence of an avian pathogenic strain were identified by SH between the strain and *E. coli* K-12 [41]. DNA identified by this method was used to screen for cosmid clones carrying unique DNA from a library of the avian pathogenic strain. To assess the role each of 12 unique regions might play in virulence, mutants were constructed and assessed for inability to colonise the air sac and cause septicaemia in 2-week-old white Leghorn chickens. Two avirulent mutants were identified, both of which could be complemented to partial virulence by cosmid clones carrying the relevant wild-type sequence. Janke *et al.* [29] used SSH to identify genomic regions present in uropathogenic *E. coli* (strain 536) but absent from *E. coli* K-12. Five of 22 DNA fragments that were specific to strain 536 had homology to known virulence determinants, whilst four fragments matched genes for lipopolysaccharide (LPS) or capsule biosynthesis and a siderophore receptor. Screening of other wild-type strains indicated that only two fragments were highly specific for uropathogenic *E. coli*.

While *Salmonella enterica* serovar Typhimurium causes systemic and fatal infection in inbred mice, *S. Typhi* is avirulent for mammals other than man. Emmerth *et al.* [42] employed SH to look for genes present in *S. Typhimurium* but absent from *S. Typhi* and identified a

novel fimbrial operon, *stf*, present in *S. Typhimurium* and several other serotypes, but absent from *S. Typhi*, *S. Arizonae* and *S. Bongori*. This fimbrial operon represents a genomic insertion. Amongst the other sequences identified by SH were F plasmid-related sequences from the *S. Typhimurium* virulence plasmid, putative sugar transporter genes and a number of DNA fragments with homology to lambdoid prophages. Interestingly, the majority of subtracted sequences mapped to three distinct locations in the *S. Typhimurium* chromosome, around centisomes 5, 27 and 57. Morrow *et al.* [43] reported similar findings, indicating the reproducibility of SH as a technique.

SSH has been used to identify virulence genes and genetic differences among strains of *A. hydrophila* [26]. The two strains chosen for SSH differed in their virulence in fish. The study identified 69 genomic regions specific to the virulent strain, of which 23 showed homology to known proteins. The majority of these regions were found in all virulent strains tested but absent from most avirulent strains. Five of the genomic regions contained genes for known virulence factors. Genes for O antigen synthesis and type II restriction/modification system were also identified.

SSH was also used to compare virulent *B. pseudomallei* with the closely related but avirulent *B. thailandensis* [27]. This led to the identification of a cluster of genes encoding capsular polysaccharide production in *B. pseudomallei*. When the putative glycosyltransferase gene identified by SSH was insertionally inactivated, values for the LD50 of *B. pseudomallei* in animal models resembled those obtained for *B. thailandensis*, suggesting that this gene and capsular production generally is important in virulence. Interestingly, the subtractive clones identified by Reckseidler *et al.* [27] all had G + C mol% values that were far lower than is typical for *B. pseudomallei*, suggesting that the method is particularly effective at identifying sequences probably acquired from other organisms by gene transfer. Atypical G + C mol% content is a feature of PIs, and is something that has also been observed in my laboratory, where we have been using SSH to study differences between strains of *Pseudomonas aeruginosa*, another GC-rich pathogen.

With SH and differential immunological screening of a *Treponema pallidum* genomic library, Centurion-Lara *et al.* [44] identified a cluster of genes encoding targets for opsonic antibody and protective immunity. One gene, *tprK*, encodes a major sheath protein homologue. The authors demonstrated that antibodies directed to purified recombinant regions of TprK can opsonise *T. pallidum* for phagocytosis. Furthermore, recombinant TprK elicited a protective immune response against *T. pallidum* infection in rabbits, suggesting that the *tprk* gene plays an important role in pathogenesis and immunity during syphilis infection [44].

RDA has been used to identify eight genetic islands that distinguish *N. meningitidis* from *N. gonorrhoeae* [19]. One of the islands identified by this approach was required for bacteraemia in a rat model of infection and encodes a putative siderophore receptor. Other previously described factors required for colonisation and survival at the port of entry were also detected with this approach [20].

Variations in gene expression

SH allows a comparison of gene expression between strains, or of the same strain grown under different conditions. The only difference in this approach is that complementary DNA (cDNA) must first be obtained following isolation of total RNA.

Dasgupta *et al.* [45] used SH to identify genes that are expressed at higher levels in a virulent strain of *M. tuberculosis* than in an avirulent strain. This led to the identification of a two-component system, *devR-devS*, in which *devR* encodes a response regulator and *devS* encodes a histidine protein kinase. The *devR-devS* locus was conserved in *M. tuberculosis* and *M. bovis* BCG but was absent from other mycobacteria. Further expression studies confirmed that levels of transcription varied between virulent and avirulent strains, suggesting an important role for the *devR-devS* system in the regulation of virulence. A similar approach was used to study the induction of gene expression in *M. avium* following phagocytosis by human macrophages [46]. cDNA prepared from *M. avium* grown in human-derived macrophages was subtracted from cDNA prepared from cells of the same strain grown in broth culture. Thus, genes expressed in both environments were eliminated, leaving only those specifically expressed during growth within macrophages.

This approach has been applied to the study of genes transcribed by *S. Typhimurium* following macrophage phagocytosis [43]. This led to the identification of a transcriptional regulator of the LysR family. The regulator was absent from *S. Typhi* and *S. Paratyphi*, both human-adapted serotypes, and was transcribed by *S. Typhimurium* only when inside macrophages. However, mutation studies indicated that the gene was required neither for virulence nor for survival within macrophages. It may be that the absence of particular genes accounts for the inability of some serotypes to cause disease in hosts other than man and may have contributed to the evolution of *S. Typhi* into a highly successful human pathogen.

In cystic fibrosis patients, most aminoglycoside-resistant isolates of *P. aeruginosa* exhibit resistance due to aminoglycoside impermeability. Westbrook-Wadman *et al.* [47] used SH to study differences in gene expression between strain PAO1 and isogenic, spontaneous aminoglycoside-resistant mutants of the impermeability phenotype. The authors discovered that

amrAB genes, encoding an efflux system, were expressed at higher levels in the mutants. Deletion of *amrB* from the mutant strains resulted in the restoration of aminoglycoside sensitivity to a level comparable with strain PAO1.

Discussion and conclusions

It is clear from the literature that SH, and particularly SSH, is becoming increasingly popular as a tool for the identification of novel genomic regions. The availability of kits makes the procedures simpler and more reproducible and interesting results have been achieved. The validity of this approach is strengthened by the fact that SH has proved effective in identifying previously characterised regions implicated in virulence and there is no doubt that this technique is able to identify PIs. In G + C mol % -rich bacteria, SH tends mainly or exclusively to identify regions with atypical, lower G + C mol % values. In many ways this is an advantage, because it selects for regions that may be implicated in horizontal gene transfer, and these are often virulence-related PIs. However, there is also the possibility that important variations might be missed. For example, by means of a completely unrelated approach, we have identified a large type III secretion gene cluster in *B. pseudomallei* that is absent from *B. thailandensis* [48, 49]. None of the SH studies comparing these organisms identified this region [27, 50], which has only a marginally reduced G + C mol % content compared with the average for the whole genome. Yet type III secretion systems are important in virulence in a number of bacterial pathogens and may prove to be so in *B. pseudomallei*. There are often problems with background hybridisation with GC-rich DNA in conventional methods such as dot-blot hybridisation. Similar non-specific hybridisation between GC-rich regions may make it more difficult to identify some regions of sequence variation. It should be possible to overcome such difficulties by varying the conditions used during the experimental procedures. In particular, higher hybridisation temperatures are required for GC-rich bacteria.

It might be the case that SH does not lead to the identification of all regions of variation between two bacterial genomes and other techniques should also be considered. For example, differential hybridisation of arrayed libraries of cloned fragments has been used to identify a genomic island (named PAGI-1) present in the majority of pathogenic isolates of *P. aeruginosa* [51]. In this procedure, DNA from strain PAO1 was labelled and used as the probe against an arrayed library representing genomic DNA from a strain that caused a urinary tract infection. Clones that did not hybridise with the PAO1 probe were considered to be unique to the pathogenic strain and were studied further to determine their prevalence among a larger number of isolates. However, these approaches are not mutually

exclusive. Indeed, in my laboratory, SSH was also able to pick out the PAGI-1 island in a cystic fibrosis epidemic strain, following comparison with strain PAO1.

With the bacterial genome database growing all the time, SH is an attractive and relatively rapid procedure for spotting differences between a strain of interest and a reference strain that has been sequenced. It may not provide all the answers, but it does make for a good starting point in the identification and characterisation of those genetic factors that account for differences in virulence or host preferences between otherwise closely related strains.

References

- McDaniel TK, Kaper JB. A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on *E. coli* K-12. *Mol Microbiol* 1997; **23**: 399–407.
- Blanc-Potard A-B, Groisman EA. The *Salmonella selC* locus contains a pathogenicity island mediating intramacrophage survival. *EMBO J* 1997; **16**: 5376–5385.
- Carniel E. The *Yersinia* high-pathogenicity island: an iron-uptake island. *Microbes Infect* 2001; **3**: 561–569.
- Wood MW, Jones MA, Watson PR, Hedges S, Wallis TS, Galyov EE. Identification of a pathogenicity island required for *Salmonella* enteropathogenicity. *Mol Microbiol* 1998; **29**: 883–891.
- Atherton JC. *H. pylori* virulence factors. *Br Med Bull* 1998; **54**: 105–120.
- Karaolis DKR, Johnson JA, Bailey CC, Boedeker EC, Kaper JB, Reeves PR. A *Vibrio cholerae* pathogenicity island associated with epidemic and pandemic strains. *Proc Natl Acad Sci USA* 1998; **95**: 3134–3139.
- Kao J-S, Stucker DM, Warren JW, Mobley HLT. Pathogenicity island sequences of pyelonephritogenic *Escherichia coli* CFT073 are associated with virulent uropathogenic strains. *Infect Immun* 1997; **65**: 2812–2820.
- Fitzgerald JR, Monday SR, Foster TJ *et al.* Characterization of a putative pathogenicity island from bovine *Staphylococcus aureus* encoding multiple superantigens. *J Bacteriol* 2001; **183**: 63–70.
- Brown JS, Gilliland SM, Holden DW. A *Streptococcus pneumoniae* pathogenicity island encoding an ABC transporter involved in iron uptake and virulence. *Mol Microbiol* 2001; **40**: 572–585.
- Hacker J, Blum-Oehler G, Mühldorfer I, Tschäpe H. Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol Microbiol* 1997; **23**: 1089–1097.
- Fraser CM, Eisen J, Fleischmann RD, Ketchum KA, Peterson S. Comparative genomics and understanding of microbial biology. *Emerg Infect Dis* 2000; **6**: 505–512.
- Lucchini S, Thompson A, Hinton JCD. Microarrays for microbiologists. *Microbiology* 2000; **147**: 1403–1414.
- Washburn MP, Yates JR. Analysis of the microbial proteome. *Curr Opin Microbiol* 2000; **3**: 292–297.
- Straus D, Ausubel FM. Genomic subtraction for cloning DNA corresponding to deletion mutations. *Proc Natl Acad Sci USA* 1990; **87**: 1889–1893.
- Schmidt KD, Schmidt-Rose T, Romling U, Tummeler B. Differential genome analysis of bacteria by genomic subtractive hybridization and pulsed field gel electrophoresis. *Electrophoresis* 1998; **19**: 509–514.
- Welcher AA, Torres AR, Ward DC. Selective enrichment of specific DNA, cDNA and RNA sequences using biotinylated probes, avidin and copper-chelate agarose. *Nucleic Acids Res* 1986; **14**: 10027–10044.
- Wu FM, Muriana PM. Genomic subtraction in combination with PCR for enrichment of *Listeria monocytogenes*-specific sequences. *Int J Food Microbiol* 1995; **27**: 161–174.
- Tinsley CR, Nassif X. Analysis of the genetic differences between *Neisseria meningitidis* and *Neisseria gonorrhoeae*: two closely related bacteria expressing two different pathogenicities. *Proc Natl Acad Sci USA* 1996; **93**: 11109–11114.
- Klee SR, Nassif X, Kusecek B *et al.* Molecular and biological analysis of eight genetic islands that distinguish *Neisseria meningitidis* from the closely related pathogen *Neisseria gonorrhoeae*. *Infect Immun* 2000; **68**: 2082–2095.
- Perrin A, Nassif X, Tinsley C. Identification of regions of the chromosome of *Neisseria meningitidis* and *Neisseria gonorrhoeae* which are specific to the pathogenic *Neisseria* species. *Infect Immun* 1999; **67**: 6119–6129.
- Bjourson AJ, Stone CE, Cooper JE. Combined subtraction hybridization and polymerase chain reaction amplification procedure for isolation of strain-specific *Rhizobium* DNA sequences. *Appl Environ Microbiol* 1992; **58**: 2296–2301.
- Cook D, Sequeira L. The use of subtractive hybridization to obtain a DNA probe specific for *Pseudomonas solanacearum* race 3. *Mol Gen Genet* 1991; **227**: 401–410.
- Seal SE, Jackson LA, Daniels MJ. Isolation of a *Pseudomonas solanacearum*-specific DNA probe by subtraction hybridization and construction of species-specific oligonucleotide primers for sensitive detection by the polymerase chain reaction. *Appl Environ Microbiol* 1992; **58**: 3751–3758.
- Diatchenko L, Lau YF-C, Campbell AP *et al.* Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci USA* 1996; **93**: 6025–6030.
- Akopyants NS, Fradkov A, Diatchenko L *et al.* PCR-based subtractive hybridization and differences in gene content among strains of *Helicobacter pylori*. *Proc Natl Acad Sci USA* 1998; **95**: 13108–13113.
- Zhang YL, Ong CT, Leung KY. Molecular analysis of genetic differences between virulent and avirulent strains of *Aeromonas hydrophila* isolated from diseased fish. *Microbiology* 2000; **146**: 999–1009.
- Reckseidler SL, DeShazer D, Sokol PA, Woods DE. Detection of bacterial virulence genes by subtractive hybridization: identification of capsular polysaccharide of *Burkholderia pseudomallei* as a major virulence determinant. *Infect Immun* 2001; **69**: 34–44.
- Bogush ML, Velikodvorskaya TV, Lebedev YB *et al.* Identification and localization of differences between *Escherichia coli* and *Salmonella typhimurium* genomes by suppressive subtractive hybridization. *Mol Gen Genet* 1999; **262**: 721–729.
- Janke B, Dobrindt U, Hacker J, Blum-Oehler G. A subtractive hybridisation analysis of genomic differences between the uropathogenic *E. coli* strain 536 and the *E. coli* K-12 strain MG1655. *FEMS Microbiol Lett* 2001; **199**: 61–66.
- Lai Y-C, Yang S-L, Peng H-L, Chang H-Y. Identification of genes present specifically in a virulent strain of *Klebsiella pneumoniae*. *Infect Immun* 2000; **68**: 7149–7151.
- Creelan JL, Bjourson AJ, Meehan BM, McCullough SJ. Characterisation of strain-specific sequences from an abortifacient strain of ovine *Chlamydia psittaci* using subtraction hybridisation. *FEMS Microbiol Lett* 1999; **171**: 17–25.
- Hughes MS, Beck L-A, Skuce RA, Neill SD. Development of mycobacterial species-specific DNA probes by subtraction hybridization. *FEMS Microbiol Lett* 1997; **156**: 31–36.
- Mahairas GG, Sabo PJ, Hickey MJ, Singh DC, Stover CK. Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. *J Bacteriol* 1996; **178**: 1274–1282.
- El Adhami WA, Stewart PR, Matthaeci KI. The isolation and cloning of chromosomal DNA specific for a clonal population of *Staphylococcus aureus* by subtractive hybridisation. *J Med Microbiol* 1997; **46**: 987–997.
- El Adhami W. Expression of a clone specific DNA sequence from *Staphylococcus aureus* in *Escherichia coli*. *J Biotechnol* 1999; **73**: 181–184.
- Bohnsack JF, Takahashi S, Detrick SR *et al.* Phylogenetic classification of serotype III group B streptococci on the basis of *hylB* gene analysis and DNA sequences specific to restriction digest pattern type III-3. *J Infect Dis* 2001; **183**: 1694–1697.
- Townsend KM, Frost AJ, Lee CW, Papadimitriou JM, Dawkins HJS. Development of PCR assays for species- and type-specific identification of *Pasteurella multocida* isolates. *J Clin Microbiol* 1998; **36**: 1096–1100.

38. Sawada K, Kokeguchi S, Hongyo H *et al.* Identification by subtractive hybridization of a novel insertion sequence specific for virulent strains of *Porphyromonas gingivalis*. *Infect Immun* 1999; **67**: 5621–5625.
39. Brown NF, Lew AE, Beacham IR. Identification of new transposable genetic elements in *Burkholderia pseudomallei* using subtractive hybridisation. *FEMS Microbiol Lett* 2000; **183**: 73–79.
40. Kersulyte D, Mukhopadhyay AK, Shirai M, Nakazawa T, Berg DE. Functional organization and insertion specificity of IS607, a chimeric element of *Helicobacter pylori*. *J Bacteriol* 2000; **182**: 5300–5308.
41. Brown PK, Curtiss R. Unique chromosomal regions associated with virulence of an avian pathogenic *Escherichia coli* strain. *Proc Natl Acad Sci USA* 1996; **93**: 11149–11154.
42. Emmerth M, Goebel W, Miller SI, Hueck CJ. Genomic subtraction identifies *Salmonella typhimurium* prophages, F-related plasmid sequences, and a novel fimbrial operon, *stf*, which are absent in *Salmonella typhi*. *J Bacteriol* 1999; **181**: 5652–5661.
43. Morrow BJ, Graham JE, Curtiss R. Genomic subtractive hybridization and selective capture of transcribed sequences identify a novel *Salmonella typhimurium* fimbrial operon and putative transcriptional regulator that are absent from the *Salmonella typhi* genome. *Infect Immun* 1999; **67**: 5106–5116.
44. Centurion-Lara A, Castro C, Barrett L *et al.* *Treponema pallidum* major sheath protein homologue Tpr K is a target of opsonic antibody and the protective immune response. *J Exp Med* 1999; **189**: 647–656.
45. Dasgupta N, Kapur V, Singh KK *et al.* Characterization of a two-component system, *devR-devS*, of *Mycobacterium tuberculosis*. *Tuber Lung Dis* 2000; **80**: 141–159.
46. Plum G, Clark-Curtiss JE. Induction of *Mycobacterium avium* gene expression following phagocytosis by human macrophages. *Infect Immun* 1994; **62**: 476–483.
47. Westbrook-Wadman S, Sherman DR, Hickey MJ *et al.* Characterization of a *Pseudomonas aeruginosa* efflux pump contributing to aminoglycoside impermeability. *Antimicrob Agents Chemother* 1999; **43**: 2975–2983.
48. Winstanley C, Hart CA. Presence of type III secretion genes in *Burkholderia pseudomallei* correlates with Ara⁻ phenotypes. *J Clin Microbiol* 2000; **38**: 883–885.
49. Winstanley C, Hales BA, Hart CA. Evidence for the presence in *Burkholderia pseudomallei* of a type III secretion system-associated gene cluster. *J Med Microbiol* 1999; **48**: 649–656.
50. Brown NF, Beacham IR. Cloning and analysis of genomic differences unique to *Burkholderia pseudomallei* by comparison with *B. thailandensis*. *J Med Microbiol* 2000; **49**: 993–1001.
51. Liang X, Pham X-QT, Olson MV, Lory S. Identification of a genomic island present in the majority of pathogenic isolates of *Pseudomonas aeruginosa*. *J Bacteriol* 2001; **183**: 843–853.