Molecular genetic approaches for the study of virulence in both pathogenic bacteria and fungi

Michael Hensel† and David W. Holden

Understanding pathogenesis

An understanding of the mechanisms of bacterial and fungal pathogenesis is dependent on the identification and characterization of the microbial gene products that influence the progression of infection. The complexity of host–pathogen interactions, particularly in the context of changing cellular and tissue environments encountered by invading pathogens, has made this task very difficult to approach by physiological or biochemical experimentation. A much more productive approach has been to use genetics and molecular biology. With the development over the past two decades of a variety of molecular genetic screening and selection methods, it is now possible to identify and determine the functions of the major virulence determinants of many pathogens. This is clearly an important milestone because our understanding of pathogenicity has often lagged behind knowledge of host responses to infection, particularly mammalian immunity. A detailed understanding of the sequence of events in pathogenesis can also be expected to result in a more informed approach to the development of new antibacterial and antifungal drugs, the need for which has been brought into sharp focus with the emergence of multiple-drug-resistant bacteria and the rising numbers of immunosuppressed patients who are often at risk of opportunistic infections which are difficult to treat (Cohen, 1996).

Here we review the principles behind popular approaches to the identification and isolation of bacterial and fungal virulence genes, discuss their merits and limitations, and assess some potentially powerful techniques which have been developed recently (for a summary, see Table 1). We have incorporated examples involving both bacterial and fungal pathogens of plants and animals because the principles of the approaches can often be applied to both prokaryotes and eukaryotes. The words virulence and pathogenicity have different connotations in the various sub-disciplines of microbial pathogenesis. For the purpose of this review, virulence is regarded as a quantitative trait, and pathogenicity as a qualitative trait: a pathogen has the ability to cause disease; the degree to which this occurs is referred to as virulence.

Gene expression

A popular approach to the identification of virulence genes is based on the tenet that genes which are important for promoting growth and host colonization by pathogens must be expressed at some stage during infection. Many of these genes are only activated in the appropriate environment, often in response to signals from the host, and therefore are only expressed when required. The differential expression of such genes has provided the basis for a number of screening techniques.

cDNA approaches

The ability to synthesize cDNA from RNA populations isolated from infected tissues permits differential screening to identify genes that are specifically expressed during infection. The strength of this approach is illustrated by work on the interaction of the pathogenic fungus Magnaporthe grisea with its host the rice plant, in which a cDNA library was made from heavily infected rice leaves (Talbot et al., 1993). Replica filters of this library were probed with labelled cDNAs derived from three sources: the pathogen grown in culture, uninfected rice plants and infected leaves. This led to the identification of many fungal clones that hybridized preferentially to the probe from infected plants. Knowledge of the amount of biomass contributed by the fungus in an infected leaf was used to estimate the increase in levels of mRNAs corresponding to these clones, and one clone that showed strong induction was found to contain a gene (MPGI)
encoding a hydrophobin-like product. Subsequent deletion of the gene and virulence tests with a null mutant confirmed the importance of MPG1 in the development of fungal infection structures and for fungal pathogenesis (Talbot et al., 1993). The ability to achieve a high fungal load in the infected host tissue undoubtedly contributed to the success of this approach: in infected rice plants, the fungal load reached 10% of the total biomass. Such a high percentage may not be achievable in other host-pathogen interactions, particularly in mammalian tissues, which may limit the broader application of this technique.

A subtractive cDNA hybridization approach has been used to identify mRNAs induced following host cell phagocytosis of Mycobacterium avium. Bacterial mRNAs recovered after infection of human macrophages with M. avium were converted to cDNAs, and biotinylated cDNAs prepared from M. avium grown in broth culture were used to remove non-macrophage-specific cDNAs from the infected cells. Three rounds of subtraction led to the identification of one gene that was highly induced in macrophages (Plum & Clark-Curtiss, 1994).

Differential display of eukaryotic messenger RNAs (Liang & Pardee, 1992) offers another means of isolating genes that are expressed during fungal infection. This approach would appear to be most directly applicable to fungi in which developmental transitions, that may be important for pathogenesis, can be induced in culture, thereby avoiding the presence of host RNA. The method, in which many PCR-amplified segments of cDNAs are displayed on sequencing gels, has been used to isolate genes that are developmentally regulated in the plant-pathogenic fungus Ustilago maydis (Bohmann et al., 1994). Two genes were specifically induced following transition from yeast-like to hyphal growth (which is essential for pathogenicity), and one gene was isolated that appeared to be specifically transcribed during growth in planta. However, the virulence of strains carrying mutations in both genes was apparently unaltered. The transition from yeast-like to hyphal growth of the human pathogen Candida albicans is also thought to be an important aspect of its virulence. Using differential display, Cormack & Falkow (1995) isolated six genes expressed during this phenotypic switch, including those encoding a pH-regulated cell-surface glycoprotein, a novel hyphal-induced superoxide dismutase, and a C. albicans homologue of mammalian leucocitene A4 hydrolase (LTA4 hydrolase); deletion of the LTA4 hydrolase gene did not reduce its virulence (B. Cormack, personal communication).

These examples illustrate both the advantages and disadvantages of expression-based approaches. On one hand they are often feasible in organisms that have poorly developed molecular genetics, or in which standard genetic analysis is impracticable, and they have the capacity to identify genes that are specifically induced during infection. However, although expressed in the host, genes identified in this way may not necessarily be absolutely required for, or have any significant role in, virulence.

**Transcriptional and translational fusions**

Several approaches based on gene fusions have been developed for identification of host-induced genes of bacterial pathogens of plants and animals. In 1987 a
promoter–probe system was described for the plant-pathogenic bacterium *Xanthomonas campestris*, based on the introduction of random bacterial chromosomal DNA fragments upstream of a promoterless chloramphenicol acetyl transferase gene. Introduction of these gene fusions into *X. campestris* and subsequent inoculation onto chloramphenicol-treated plant seedlings resulted in the growth of only those bacteria expressing chloramphenicol resistance and permitted the isolation of 14 putative plant-inducible promoter fragments (Osbourn et al., 1987). More recently, a technique has been developed in which the promoter fusions are made to a promoterless housekeeping gene whose expression is essential for bacterial survival in host tissue (Mahan et al., 1993). The system, termed in vivo expression technology (IVET) was developed in *Salmonella typhimurium* which causes a systemic typhoid-like illness in mice. IVET has the advantage of incorporating the lacZ gene downstream of the housekeeping gene (parA) to enable the relative transcriptional activity of the promoter fusions to be monitored easily. Chromosomal fusions were introduced into a parA mutant strain of *S. typhimurium* and pools of integrated random-DNA–parA–lacZ fusion strains tested simultaneously for ability to survive and multiply in the mouse. Bacteria recovered from the spleens of infected animals had a greater proportion of LacZ cells compared with the inoculum, and allowed the selection of strains that were LacZ in culture but virulent in mice. The screen resulted in the isolation of three genes that were subsequently shown to be important virulence determinants by mutational and lethal dose 50 (LD50) analysis. A potential drawback of this version of IVET is that the promoter of interest would have to be expressed to some degree during many of the stages of infection in order to be recovered from the spleen. Therefore, virulence genes whose promoters are tightly regulated might be missed by this approach. Another version of IVET was described recently (Mahan et al., 1995) which resembles the system developed by Osbourn et al. (1987) in that the chloramphenicol acetyltransferase gene is used to select for host-inducible promoters in mice that have been treated with the antibiotic. This is likely to be more widely applicable than the original version of IVET in that it does not require the use of an attenuated bacterial mutant and does not require a cloned complementing gene. Yet another version of IVET is based on transcriptional fusions to a gene encoding β6 resolvase which mediates recombination-based excision of an antibiotic resistance marker (Camilli et al., 1994). The attraction of this particular version is that there is in principle no necessity for constitutive expression of the promoter in the host, allowing the detection of genes that are only transiently expressed during infection.

Yet another way of isolating genes that are selectively expressed in the infected host is based on the use of antibodies obtained from the serum of mice either actively infected with *Borrelia burgdorferi* (the cause of Lyme disease), or from mice challenged with heat-killed spirochaetes, to screen an expression library of *B. burgdorferi* (Suk et al., 1995). Three clones were identified that reacted specifically with sera from infected animals. As pointed out by the authors, the method will only work if the induced gene product is sufficiently immunogenic. However, the attraction of this approach is that it will be applicable to a wider variety of pathogens, as it does not require sophisticated molecular genetics.

### Gene transfer

Many virulence genes have been identified by the transfer of cloned DNA from a pathogen either to a closely related species which is non-pathogenic, or to a non-pathogenic strain of the same species or to an avirulent mutant of the same strain, and then selecting for the acquisition of virulence in the recipients.

One of the earliest examples of this approach was the restoration of virulence of rough mutants of *Streptococcus pneumoniae* by co-infection with heat-killed wild-type cells (Griffith, 1928). More recently, the virulence of *Yersinia pseudotuberculosis* was investigated using a bank of genomic DNA from *Y. pseudotuberculosis* maintained in *Escherichia coli* (Isberg & Falkow, 1985). In this case, the presence of a single *Y. pseudotuberculosis* gene conferred on *E. coli* the ability to adhere to and invade cultured epithelial cells. In similar approaches, a non-invasive strain of *S. typhimurium* was rendered invasive for cultured epithelial cells through use of a DNA library from an invasive strain (Galan & Curtiss, 1989), leading to the identification of an invasion gene cluster in *S. typhimurium*, while Elsinghorst et al. (1989) reported the identification of invasion genes of *S. typhi* by transfer of a genomic DNA region of 33 kb (containing a cluster of invasion genes) to *E. coli* HB101, which enabled this strain to invade human epithelial cells. An obvious prerequisite of this approach is that the biosynthetic machinery of the host organism should allow the expression of the transferred genes. In this context the experiments of Arruda et al. (1993) are remarkable. The ability to invade epithelial cells and to survive within macrophages was transferred from *Mycobacterium tuberculosis* to a non-invasive *E. coli* strain by transformation with a cloned DNA fragment. In view of the long evolutionary distance between these two species (Woese, 1987), this experiment suggests that in some cases *E. coli* can appropriately express genes from distantly related bacterial species.

A similar approach has been used to identify virulence determinants of the fungus *Candida albicans* (Barki et al., 1993). A gene conferring adhesion and aggregation was identified after transferring a *C. albicans* genomic library to *Saccharomyces cerevisiae* and the selection of transformants able to adhere to polystyrol surfaces.

An advantage of the gene transfer approach is that it does not require assumptions to be made about the involvement of specific genes in functions such as invasion or adherence. Furthermore, the experimental requirements mainly consist of the ability to generate a genomic library of the donor strain, and a suitable screening method. On the other hand, a limitation is imposed by the size of DNA
that can be transferred between donor and host strain. As the maximal size is usually the capacity of a cosmid, this will be in the region of 40 kb. Where a virulence trait requires the action of several genes, complementation will not be possible if the genetic loci concerned are scattered throughout the chromosome of the donor. Virulence is often governed by tightly regulated gene expression and coordinated interaction of virulence gene products in response to biochemical signals from the host organism (Mekalanos, 1992). Clearly, the larger the evolutionary distance between recipient and pathogenic donor in a gene transfer experiment, the less likely the recipient will be capable of expressing virulence products in an appropriate context.

Although most screening procedures involving gene transfer approaches have utilized cell culture assays, it is also possible to exploit living hosts. Rakin et al. (1994) reported the cloning of the *Yersinia enteroocolitica* pesticin/yersiniabactin receptor, which is important for survival in mice, using an in vivo enrichment approach. After transfer by conjugation of a genomic library to a non-virulent strain, mice were infected with pools of transconjugants and those able to survive and multiply within the host were isolated. Pascopella et al. (1994) and Collins et al. (1995) used shuttle vectors to transfer banks of genomic DNA of *M. tuberculosis* and *Mycobacterium bovis*, respectively, into attenuated strains of the species. Subsequently, an in vivo enrichment approach allowed the identification of genes which restored the pathogenicity of the attenuated strain in animal models of tuberculosis.

**Genome comparison**

An alternative approach to identifying pathogenicity genes is to compare directly by hybridization analysis the genomes of pathogens and non-pathogenic relatives. In a number of bacterial pathogens, some of the virulence genes are clustered together on 'pathogenicity islands'. By screening an *S. typhimurium* DNA library by hybridization to *E. coli*, several clones containing sequences that were apparently unique to salmonellae were identified. Although mutational analysis of one such region failed to establish a role for it in virulence (Groisman et al., 1993), other regions probably correspond to two pathogenicity islands, which have been shown by mutational analysis to be important for virulence (Mills et al., 1995; Shea et al., 1996). The utility of this approach is illustrated by the recent identification of a virulence-related locus in the causative agent of ovine footrot, *Dichelobacter nodosus*, for which transformation and conjugation methods are not yet available (Haring et al., 1995).

**Mutation-based approaches**

Mutational analysis of one form or another has probably been the most popular and successful basis for identifying microbial virulence determinants. Since a comprehensive review is beyond the scope of this article, we give only a few examples in this section to illustrate the principles involved.

**Directed mutagenesis**

Mutation-based approaches can be divided into those that are directed and those that are random. In a directed approach, a certain trait of a pathogen is postulated to be important for virulence, based on knowledge of the pathogenesis of either the microbe itself, or a related species. To investigate the role of a putative virulence determinant, a gene necessary for its expression is cloned, and a mutant strain is constructed by gene disruption or gene replacement. The virulence of the mutant strain is then compared with that of the parental strain. This approach has been useful in establishing the importance of certain suspected virulence determinants. For example, Bowyer et al. (1995) used gene disruption to inactivate a gene of *Gaumannomyces graminis* encoding an enzyme capable of detoxifying saponin, an antifungal compound produced by plants. The resulting mutant was no longer able to infect oats, a saponin-producing host of *G. graminis*. Targeted gene disruption has also been used to confirm that capsule production by the human pathogen *Cryptococcus neoformans* is required for virulence (Chang & Kwon-Chung, 1994). On the other hand, this approach can also refute long-held hypotheses about the roles of some microbial products in virulence. Work in our laboratory on virulence determinants of *Aspergillus fumigatus* (which causes pulmonary infection in immunocompromised patients) led us to investigate the role in pathogenesis of an extracellular protease and a cytotoxin, known to be produced in lung tissue and thought to be involved in host-tissue invasion and degradation. Virulence studies using strains carrying relevant gene knockouts indicated that neither the secreted protease nor the cytotoxin are important in *A. fumigatus* pathogenesis (Smith et al., 1994). These examples illustrate two important points. First, if the case for a putative virulence determinant is based on circumstantial evidence alone, then no matter how plausible the candidate may seem, its role in virulence is still uncertain. The second point is that just because certain genes are transcriptionally active during infection does not mean that they are actually required for infection to occur. Directed approaches are performed by reverse genetics. This usually involves the isolation and characterization of the protein thought to be important for pathogenicity, the cloning of the corresponding gene, its inactivation and tests comparing the virulence of the mutant strain with the wild-type strain from which it was derived. If diploid pathogens (such as *C. albicans*) are under investigation, a further step is required to inactivate the second copy of the gene. Although techniques such as protein microsequencing and PCR have greatly aided reverse genetics, the pathway from protein to gene can be very labour-intensive and frustrating if the gene turns out not to be important for virulence.

**Random mutagenesis**

The great advantage of random mutagenesis as an approach is that no previous assumptions have to be made about the roles of specific genes during pathogenesis. Following random mutagenesis, mutants can be analysed...
either individually or in pools for attenuated or increased virulence. The mutant gene(s) of interesting strains are then cloned and characterized. Clearly, random mutagenesis is only appropriate for the analysis of pathogens that are haploid during the infective stage. The major problem associated with random mutagenesis by UV or chemicals is the difficult and labor-intensive identification of mutagenized genes. Another complication is the possibility of inducing multiple mutations, although this can often be resolved by standard genetic crosses prior to gene cloning. The simplest way to identify the mutated virulence gene is to restore virulence to the mutant strain by complementation with an intact copy of the gene. However, having to screen a bank of mutants for the complementation of virulence can introduce additional problems which are discussed below. Nevertheless, several virulence genes (particularly from plant-pathogenic fungi) have been identified by this conventional genetic method (see Perpetua et al., 1994 for example).

**Transposon mutagenesis**

Much of our current understanding of the mechanisms of bacterial pathogenesis is based on the use of transposons for random insertional mutagenesis. Not only is it possible to generate large numbers of single-site mutations in a relatively short time period by transposon mutagenesis, but as each mutant is tagged with the transposon, subsequent mapping and subcloning of the mutagenized gene is greatly facilitated.

Transposons of Gram-negative bacteria have been extensively modified in recent years to facilitate their use in mutagenesis. This is because unmodified transposons are often bulky, have a shortage of convenient restriction sites and often carry, together with genes for transposition, unwanted genes for antibiotic resistance. Furthermore, unmodified transposons can be unstable, and often exhibit preferred sites of transposition or ‘hotspots’, thus reducing the randomness of the mutagenesis. The mini-transposons developed by the groups of Kleckner and Timmis have eliminated most of these undesirable characteristics (Kleckner et al., 1991; de Lorenzo & Timmis, 1994), for example by providing the transposase functions in trans (thereby promoting stable single-copy integration of a transposon into the target genome), reducing their size (facilitating the handling of the transposon and subcloning of transposon-tagged genes), providing polylinkers and using insertion sequences optimized for random integration.

One drawback of insertional mutagenesis is that insertion of a vector may terminate the transcription of genes downstream from its insertion point or influence the stability of a polycistronic mRNA (‘polar effects’). Therefore, additional experiments including the construction of non-polar mutations are often required to provide proof that the interrupted ORF is itself involved in the virulence function.

Generally, insertional mutagenesis will result in a loss-of-function phenotype. Another variation of the transposon mutagenesis approach (Lee et al., 1992) addresses the regulation of virulence. In this example, the tightly regulated process of invasion of epithelial cells by *S. typhimurium* was investigated using a transposon carrying a constitutive promoter, which generates promoter fusions to chromosomal genes. Mutants were then screened for gain-of-function mutations which display increased ability to enter epithelial cells. Subsequent analysis of one particular mutant led to the identification of the hyperinvasion locus (bii) (Lee et al., 1992).

An important group of virulence factors are those which are secreted in order to interact with cells or other components of the host organism. A further modification of the transposon mutagenesis approach has allowed genes encoding secreted or surface-associated proteins to be identified. In this version (TrpHo.A), the transposon carries the coding region of a gene encoding alkaline phosphatase, but lacks the promoter and signal sequence. As the alkaline phosphatase is only active if secreted, mutagenesis with this transposon can lead to fusions which result in phosphatase-secreting cells. These can be detected easily by a plate assay, providing a useful screen for genes whose products are surface-associated or extracellular (Maniol & Beckwith, 1985). This subset of mutants can then be tested for virulence defects (Miller et al., 1989; Donnenberg et al., 1990).

Another modification of the transposon mutagenesis approach allows mutagenesis of bacteria for which there are no useful naturally occurring transposons. In shuttle mutagenesis, a bank of genomic DNA of the bacterium of interest is maintained in plasmids in *E. coli*. This gene bank serves as the target for transposon mutagenesis performed in *E. coli*. The mutagenized plasmids are then pooled and transferred back to the pathogen, for example by electroporation of *Mycobacterium* spp. (Kalpana et al., 1991), or by utilizing a naturally occurring DNA uptake mechanism, such as that of *Neisseria gonorrhoeae* (Kahrs et al., 1994). The presence of a selectable marker in the transposon allows for selection of mutants following integration of the mutagenised DNA into the recipient genome. Shuttle mutagenesis has been applied to *N. gonorrhoeae* (Kahrs et al., 1994), to identify genes encoding secreted proteins thought to be involved in pathogenesis. However, the success of shuttle mutagenesis is dependent on two factors: the efficiency of homologous recombination in the target organism, and the ability to maintain genomic DNA from this organism in *E. coli*. In applying this method to *Mycobacterium* spp., Kalpana et al. (1991) found that allelic exchange was observed in the faster-growing species *M. smegmatis*, while integration of transformed DNA occurred predominantly by illegitimate recombination in the slow-growing *M. tuberculosis*. The incorporation of a second marker for the selection against illegitimate (ectopic) integration (Sander et al., 1995) may circumvent this problem. Although shuttle mutagenesis has also been described for *Saccharomyces cerevisiae* (Seifert et al., 1986; Hoekstra et al., 1991), mutagenesis of pathogenic fungi by this approach has not been reported so far. The significant frequency of ectopic integration events which follows transform-
Insertional mutagenesis of eukaryotic pathogens has been hampered by two main problems. First, insertional mutagenesis will only be useful in organisms that are haploid during the infective stage of the life-cycle. Second, mobile genetic elements that have the necessary characteristics for insertional mutagenesis are generally not available. However, recently a tool has become available which appears to circumvent the second limitation. Non-homologous integration of transforming DNA does not normally occur in *S. cerevisiae*, but when DNA fragments were used to effect transformation in the presence of the restriction enzyme BamHI, single-copy integrations occurred at different genomic GATC sites (Schiestl & Petes, 1991). This process, referred to as restriction enzyme-mediated integration (REMI), now forms the basis of insertional mutagenesis systems in *Dictyostelium* (Kuspa & Loomis, 1992) and several filamentous fungi (Fig. 1). For example, Lu et al. (1994) applied REMI using a hygromycin resistance gene to obtain two toxin-deficient mutants among 1300 transformants of *Cochliobolus heterostrophus*. Both mutants arose as a result of single-copy integration events at different but closely linked sites, and in each case the hygromycin resistance co-segregated with the toxin-deficient phenotype. REMI has also been used successfully to tag pathogenicity genes of *Ustilago maydis* (Bölker et al., 1995), and preliminary work suggests it may also be useful in *Aspergillus fumigatus* (M. Hensel & D. W. Holden, unpublished data). Although many parameters of mutagenesis by REMI still have to be defined, e.g. the randomness of the integration, or the extent of ‘silent’ genomic rearrangements that might occur during the process of mutagenesis, REMI certainly offers great promise for generating insertional mutant banks of eukaryotic pathogens. It is also conceivable that REMI might provide a means of generating mutant banks of bacteria that have poorly developed transposon- or shuttle-mutagenesis systems, such as the mycobacteria.

**Screening procedures for mutant banks**

The generation of a mutant bank of a given pathogen by one of the methods outlined above is often a relatively straightforward procedure. The screening approach used to identify mutants affecting virulence is however more difficult, principally because the insertional mutagenesis will usually result in the loss of a particular virulence trait, and therefore a ‘negative’ phenotype.

The simplest way to assay for virulence is to subject each individual mutant to a virulence test in an appropriate model of infection. This brute force approach has been applied to some plant pathogens, but because it requires a large number of host organisms it is not practicable for pathogens of animals. Moreover, the very large numbers of mutants necessary to saturate even a bacterial genome makes this approach unrealistic as a means of comprehensively screening an entire genome. Alternatively one can preselect for mutants with certain desirable characteristics, such as *TnphoA* insertions in genes for secreted proteins (Miller et al., 1989). Another option is to screen for attenuated mutants using a cell culture model representing one aspect of the infection process. For example, a bank of mutants can be screened for the loss of ability to adhere to, to invade, or to grow within infected cultured epithelial cells, or to survive within macrophages after phagocytosis. It has also been possible to regenerate polarized epithelial layers and to assay mutants for their ability to transcytose such layers (Finlay et al., 1988). A comprehensive description of such assays can be found in *Methods in Enzymology* vol. 236 (1994). These techniques have been used widely for a range of pathogens and have resulted in the identification of a number of genes involved in the interaction of the pathogen and the host cell. For example, the invasin of *Y. pseudotuberculosis* (Isberg & Falkow, 1985), the invasion genes of *S. typhimurium* (Galan & Curtiss, 1989) and *Shigella flexneri* (Maurelli et al., 1988), and the internalin of *Listeria monocytogenes* (Gaillard et al., 1991) were identified and cloned using cell culture assays. Obviously, this approach is only possible when a cell culture assay is available for the pathogen and requires some knowledge of the sequence of events in the infection process itself, i.e.
which cells or tissues are infected by the pathogen. More importantly, cell cultures cannot be expected to provide all the host-derived signals which are received by the pathogen during the particular stage of infection under study, and established cell lines can be physiologically very different to their natural counterparts in the host. Many pathogens go through several distinct stages of infection and interact with various types of host cells during these stages. Only a subset of virulence genes is likely to be required at a certain stage of infection, and it is therefore unlikely that all genes important for virulence can be identified by cell culture assays. Despite these limitations, cell culture assays have proved to be invaluable tools for the analysis of interactions between...
pathogens and host cells, and have the advantages of a defined functional defect to correspond to the mutation, and avoiding the use of large numbers of animals.

**Signature-tagged mutagenesis (STM)**

One of the major obstacles to efficient screening for, or selection of virulence genes has been the inability to identify a mutant with attenuated virulence within a large population of different mutants. Recently, our laboratory has developed a technique, termed signature-tagged mutagenesis (STM) which combines the power of insertional mutagenesis with the ability to follow simultaneously the fate of a large number of different mutants within a single animal (Hensel *et al.*, 1995). In this process, every mutation that is generated carries a unique tag which allows mutants to be differentiated from each other. The tags are short segments of DNA, containing a 40 bp variable central region flanked by invariant 'arms' which facilitate the amplification and labelling of the central portion by PCR. Mutants are assembled into microtitre dishes, from which replica DNA colony blots are prepared. The mutants are then combined to form the 'input pool' prior to inoculation into an animal. After infection is established, bacteria are isolated from the animal and pooled to form the 'recovered pool'. The tags in the recovered pool and the tags in the input pool are separately amplified, labelled, and then used to probe DNA colony blots of the inoculum. Mutants with attenuated virulence will not be recovered from the animals and will be therefore identified as colonies that hybridize when probed with tags from the inoculum pool but not when probed with tags from the recovered pool (Fig. 2). An important feature of this technique is the use of an animal model for the selection of attenuated mutants, thus reproducing 'natural' conditions during an infection and avoiding the limitations of cell culture assays. Using STM, we have been able to identify a new pathogenicity island in the *S. typhimurium* chromosome (Shea *et al.*, 1996).

Although STM has proved to be a very useful means of isolating virulence genes from *S. typhimurium*, it remains to be seen whether it can be applied successfully to other bacterial and fungal pathogens. As an approach it suffers from the same limitations as conventional transposon mutagenesis, such as the potential for polar effects on genes downstream from the transposon insertion point, and the tendency of some transposons to integrate in a non-random fashion. STM also requires that the inoculum should infect as a mixed rather than clonal population, and that representatives of each mutant in the inoculum pool have the opportunity to multiply within the host if they are capable of doing so. In addition, some mutants that have a virulence defect might be capable of proliferation in the host if other mutants in the pool can provide the missing virulence function. It seems clear that the size and complexity of the inoculum mixture will have to be determined empirically for each different pathogen-host combination. In addition to the analysis of bacterial pathogens, STM in combination with REMI may also be applicable to fungal pathogens.

**Analysing virulence determinants of human pathogens in plants**

A novel approach to the identification of bacterial virulence genes comes from work on *Pseudomonas aeruginosa*, some strains of which are both plant pathogens and opportunistic pathogens of humans. Rahme *et al.* (1995) discovered that mutations in three different genes encoding a phospholipase, a protein synthesis inhibitor (both of which are virulence determinants in animals) and a regulator of secreted proteins (a virulence determinant in plants) resulted in attenuated virulence in both *Arabidopsis* and a model of skin-burn infection in mice. This finding suggests that for some opportunistic pathogens, screening mutants on plants may be a useful surrogate for an animal model of infection, and that the molecular basis of virulence, at least for opportunistic pathogens, may be closer than is suggested by the evolutionary distance between their hosts. Another example in support of this is provided by the type III secretion systems of bacteria, which are required for export of harpins (hypersensitive response-inducing proteins) from some plant pathogens, and a variety of important virulence determinants from several human pathogens, including *Shigella* spp., *Salmonella* spp. and *Yersinia* spp. (van Gijsegem *et al.*, 1993).

**Perspective**

The foregoing discussion gives some idea of the variety of screening approaches that can be used to identify the important virulence determinants of microbial pathogens. The versatility of these techniques means that virulence genes can now be isolated from the vast majority of pathogenic microbes. Approaches that use a living host organism to select for the genes of interest, such as IVET and STM, have inherent power because of their capacity to identify genes that act at different stages of infection. Both of these methods were developed using *S. typhimurium*. However there is no reason why they could not be adapted to other pathogens, although for some Gram-positive bacteria, mycobacteria and fungal pathogens, improvements in the basic molecular genetics of the pathogen will be required. The development and improvement of animal models of human infections may also expand the range of pathogens which can be analysed. The availability of efficient screening techniques combined with the prospect of complete nucleotide sequences for genomes of some of the most important bacterial pathogens (Moxon, 1995) suggests that the identity of most of the virulence genes of these pathogens will be known in the near future. This in turn will provide the basis for novel and more rational approaches to drug development, and through functional analysis of these genes, a more comprehensive understanding of the process of pathogenicity.

**Acknowledgements**

M. H. was supported by a grant from the European Community Human Capital and Mobility Programme. Work in D. H.'s laboratory is also supported by grants from the Medical Research Council and The Wellcome Trust. We are grateful to...
C. Tang, J. Holden, N. Gow and members of our laboratory for their helpful comments on the manuscript.

References


