**Bacterial virulence gene detection**

**1. Wide range mechanisms of bacterial pathogenesis**

 Bacteria exploit a number of common molecular mechanisms to achieve a range of different objectives during infection. This lecture describes how bacterial molecular genetics has been used to uncover these key mechanisms of microbial pathogenicity.

From genome sequence information, it is clear that virulence genes often occur in clusters and that these regions are absent from closely related non-pathogenic bacteria. Furthermore, based on the%GCcontent of these regions compared to the rest of the genome, it became apparent that these large sections (e.g. 30–50 kb) of DNA had been acquired from other organisms. Together, these observations gave rise to the concept of pathogenicity islands which suggests that bacteria can acquire, en masse, complete systems that expand their ability to exploit different host environments. An example is Vibrio cholerae, where the toxin genes are located on a pathogenicity island which was subsequently shown to be an integrated bacteriophage.

**2. Detection of virulence genes**

Many bacterial pathogens have separate free-living and pathogenic life cycles, so they encounter very different environments and require very different functions for survival. As a consequence, pathogens must be able to recognize signals in the host that convey the need to express virulence genes. The expression of certain virulence gene functions in Shigella, for example, is triggered at body temperature (37 8C) but not at environmental temperatures (< 30 \_C). Therefore, if the genes which are only expressed during infection can be identified, then the key virulence traits can also be identified. **Reporter genes have been especially useful for this purpose**.

**In Vivo Expression Technology (IVET)**

One of the most widely used methods of using reporters to identify virulence genes is known as In Vivo Expression Technology (IVET). This is a technique that selects bacterial promoters which are only expressed in the host and which thus drive the expression of virulence traits. In most examples of IVET, random fragments of DNA from the bacterial host are inserted adjacent to a promoterless reporter gene whose product confers a phenotype that can be positively selected for in the host. For example, a purA mutant of Salmonella typhimurium is unable to survive in an animal model because purine biosynthesis is essential for Salmonella in this environment. This defect can be complemented by a plasmid carrying a functional purA gene, but the purA gene on the plasmid will only be expressed if a promoter is inserted which is functional in vivo (Figure 1). (The construct shown also contains a lacZ gene whose expression is also controlled by the inserted promoter so that constitutively expressed promoters can be identified and excluded).

When animals are infected with a pool of clones each carrying a different DNA fragment, only those clones containing an active promoter are able to survive and to be recovered from the animals, thus providing direct selection for those promoters which are active during infection. Identification of these promoter fragments then leads to identification of the genes that are normally regulated by them, which are likely to include genes which are essential for the virulence of Salmonella.



**Signature tagged mutagenesis**

 As shown previously, the classical approach to identifying the genes responsible for a given phenotype would be to generate mutants that are defective in that phenotype. For virulence genes, the phenotype of the mutant would be the inability to survive in an in vivo model for the disease. Isolation of the mutants cells will prove to be difficult as they are the very cells which do not survive. To circumvent this obstacle, a procedure called signature tagged mutagenesis (STM) has been invented. This is essentially a negative selection technique derived from transposon mutagenesis in which each individual mutant is labelled with a unique DNA signature. By comparing the mutants present in the initial inoculum with those that are recovered after infection of the model, it is possible to identify those that did not survive and thus identify the genes that are required for virulence.



An overview of STM is given in (Figure 2). A library of mutants is made for the bacterium of interest and each mutated gene tagged with a section of DNA containing a unique central region and two flanking arms which share their sequence with all of the other tags. The key to STM is that each individual mutant can be distinguished from every other mutant based on the possession of its own unique tag. The mutants are then stored individually in ordered arrays and DNA from each one spotted onto a membrane in an grid-like manner. The mutants are then pooled and inoculated into a relevant animal model and the bacteria that are able to survive and establish infection are recovered. PCR is then used to amplify all the tags present in the recovered bacteria and the mixed product is used to probe the gridded membrane. Mutants that fail to survive in the infected animal (those which are defective in virulence) can be identified since their tags will not be present in the output pool. They can be recovered from the original stored arrays for further study. This system has been widely used to identify novel virulence factors that are involved in colonization, immune system evasion and attachment to human cells in a number of bacterial pathogens.

**Specific mutagenesis**

conventional mutational techniques have been considered – producing mutants with an altered phenotype and then identifying the genes affected and determining their functions. These techniques can be complemented by the use of recombinant DNA technology. In contrast to conventional genetics,the recombinant DNA approach starts with an hypothesis that a specific gene is involved. This gene is then modified or deleted and the resultant phenotype characterized. So, while conventional genetics starts with the phenotype and works towards identifying the nature and function of the genes involved, the molecular approach starts by altering the gene and works towards an analysis of the phenotype.

**1. Gene replacement**

A key technique for determining the function of a specific gene is to inactivate it by a process known variously as gene replacement, allelic replacement or gene knock-out. **Essentially, this uses homologous recombination to remove all or part of a specific gene or to replace it with an altered or inactivated gene**.

An example of how this can be done is shown in (Figure 3). In this case, a plasmid has been constructed in which the central part of the cloned gene has been removed and replaced by an antibiotic resistance gene (aph, aminoglycoside phosphotransferase which confers resistance to kanamycin). The kanamycin resistance gene is flanked by regions of DNA that are the same as those in the host strain and it is within those regions that homologous recombination will occur. Note that two recombination events (a double crossover) are needed for gene replacement.

Recombination at a single site (a single crossover) will merely integrate the plasmid into the chromosome. Additional techniques are needed to ensure thata double crossover is achieved.

The plasmid that is used is one that cannot replicate in the chosen host cell. Thus after transformation, selection for kanamycin resistance will isolate cells in which the aph gene has been inserted into the chromosome by homologous recombination with the target gene, thereby inactivating that gene. Tests can be carried out to ascertain whether the expected phenotype is produced. For example, if it has been assumed that the target gene is necessary for survival and growth within macrophages, then testing the mutant for a deficiency in this respect will confirm or deny the assumption.



if the target gene is part of an operon, the gene knockout may affect the expression of other genes within the operon (in other words, the mutation may be polar). This possibility can be partly eliminated by the use of complementation.A plasmid carrying the wild-type gene (in this case using a plasmid that can replicate in this host) can be introduced into the mutant and if the original gene knock-out was responsible for the observed effect (e.g. loss of ability to grow within macrophages), then the introduced plasmid will restore the original, wildtype phenotype. Successful complementation therefore supports the contention that the product of this gene is necessary for survival in macrophages.

**2. Antisense RNA**

One problem with gene knock-outs is that if the gene is essential for the growth of the cell in the laboratory, then the complete loss of that gene would be lethal and therefore no recombinants would be obtained. It may therefore be useful to partially reduce the expression of the gene or the functionality of its product –thus leaving enough activity to cope with the comfortable conditions of normal laboratory growth but not enough to deal with the stress conditions that may be imposed on it subsequently. One alternative strategy is to use antisense RNA.

For this purpose, part of the gene would be cloned in an expression vector, so that it is transcribed from a promoter on the vector, but the insert would be deliberately put in the wrong orientation. The insert will therefore be transcribed in the opposite direction from normal – or in other words, the opposite strand will

be transcribed. The RNA produced (the antisense RNA) will be complementary to the normal mRNA and will pair with it to produce a double-stranded RNA molecule. This may interfere with translation of the mRNA and thus reduce the level of the protein that is made. Using a stronger or a weaker promoter (or even

better, using a promoter that can be switched on and off) will lead to different amounts of the antisense RNA being made and hence will alter the extent of the reduction in the amount of the protein product formed.

