

## Extra chromosomal Agents

### Transposable elements

Mobile genetic elements are segments of DNA able to move from site to site in the genome, or between genomes in the same cell. Found in both prokaryotes and eukaryotes, they are a diverse group differing in structure, mechanism of mobilization, distribution, freedom of movement .

Those elements which cannot replicate outside the host genome, but which control their own mobility within it, are termed transposable elements . The transposable elements are the largest elements and demonstrate great diversity in structure.

#### 1. Insertion sequences

Insertion sequences (ISs) constitute an important component of most bacterial genomes. Insertion of a DNA fragment into a gene will usually result in the inactivation of that gene, and it is by the loss of that function that such events were initially recognized. A number of genetic elements, including some phages and plasmids , can be inserted into the bacterial chromosome.

The simplest of these genetic elements are known as **Insertion Sequences (IS)**. Bacterial transposable elements were discovered as a consequence of their ability to cause unstable but strongly polar mutations in *E. coli*. About 100 different IS elements have been identified, the majority in enteric bacteria and their plasmids.

#### Structure of insertion sequences

There are many IS elements known. They differ in size and other details, but the overall structure of most such elements is similar. IS is 768 bases long but many other IS elements are longer (usually 1300–1500 bases). The central region of an IS element codes for a protein (known as a **transposase**) which is necessary for the movement of the element from one site to another. At the ends of the insertion sequence are almost perfect inverted repeat (IR) sequences, which in IS1 consist of 23 nucleotides. A minority of elements, such as IS900 from *Mycobacterium paratuberculosis* do not have inverted repeat ends. Since DNA sequences are often presented as just one of the two strands, an inverted repeat of the sequence CAT will appear as ATG.

In addition to the inverted repeats, inspection of a DNA region containing an insertion sequence usually shows a further short sequence that is duplicated – but this sequence is repeated in the same orientation and is therefore referred to as a direct repeat (DR). This is not part of the IS, but arises from duplication of

the DNA at the insertion site and therefore different copies of IS1 will have different target sequence repeats depending on the point of insertion.

Insertion sequences have been identified in most bacterial genera, although the presence and the number of copies of any one element often varies from strain to strain.

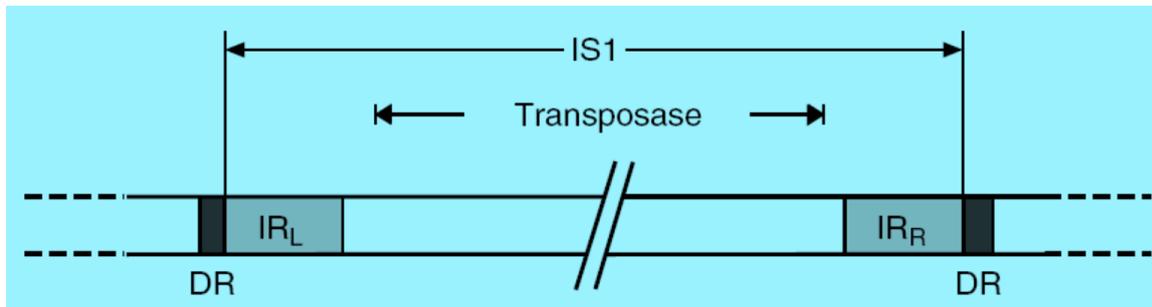


Figure 19 : Structure of the insertion sequence IS1. DR, direct repeat (duplicated target sequence); IR, inverted repeats [Dale and Park,2004].

## 2. Transposons

Transposons were originally defined as bacterial transposable elements which carried not only those genes required for transposition, but also nonessential genes, e.g. for antibiotic resistance.

When resistance plasmids were first discovered, there was much speculation as to how a single element could have evolved to carry a number of different antibiotic resistance genes and in particular how apparently related plasmids could have different combinations of such genes (or, conversely, how otherwise dissimilar plasmids could carry related resistance genes). It was assumed that a basic plasmid, having the ability to replicate independently but not carrying any other information, had somehow picked up a resistance gene from the chromosome of a resistant host strain. Transfer of this plasmid to an otherwise sensitive strain then produces a selective advantage for that strain, and therefore indirectly a selective advantage for this 'new' plasmid. As the plasmid moves from one organism to another it has the opportunity to acquire additional resistance genes, thus giving rise to a family of plasmids containing different combinations of resistance genes.

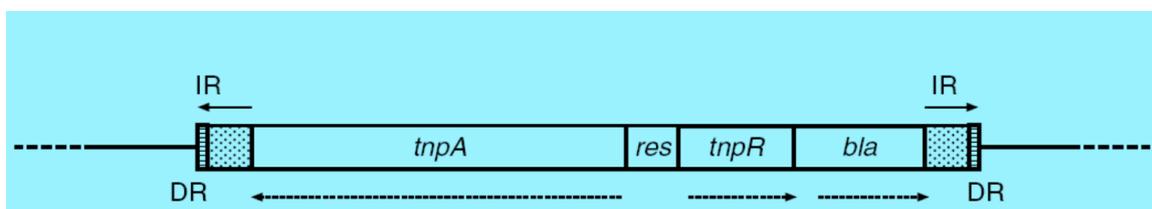
When a transposable element becomes inserted into a target DNA, a short sequence in the target at the site of integration is duplicated during the insertion process. The duplication arises because single-stranded DNA breaks are generated by the transposase. The transposon is then attached to the single-

stranded ends that have been generated, and repair of the single-strand portions results in the duplication.

If the duplicated DNA is sufficient to have duplicated an entire gene or group of genes, the organism will contain multiple copies of these particular genes. Such gene-duplication events are thought to fuel microbial evolution. This is because mutations occurring in one copy of the gene(s) do not affect the other copy; the function of the faulty protein is still "covered" by the product of the unmutated duplicate gene. However, beneficial mutations in one copy can lead to production of a protein that has superior properties to the normal protein and thereby increase fitness. In this way, evolution can "experiment" with one copy of the gene while the identical copy provides the necessary backup function. Genomic analyses have revealed numerous examples of protein-encoding genes that were clearly derived from gene duplication.

### Structure of Transposons

The structure of a simple transposon, Tn3 consists of about 5000 base pairs and has a short (38 bp) inverted repeat sequence at each end. It is therefore analogous to an insertion sequence, the distinction being that a transposon carries an identifiable genetic marker – in this case the ampicillin resistance gene (*bla*, *b*-lactamase). Tn3 codes for two other proteins as well: a transposase (*TnpA*), and *TnpR*, a bifunctional protein that acts as a repressor and is also responsible for one stage of transposition known as resolution (this is explained more fully later on). As with the insertion sequences, there is a short direct repeat at either end of the transposon (five base pairs in the case of Tn3).



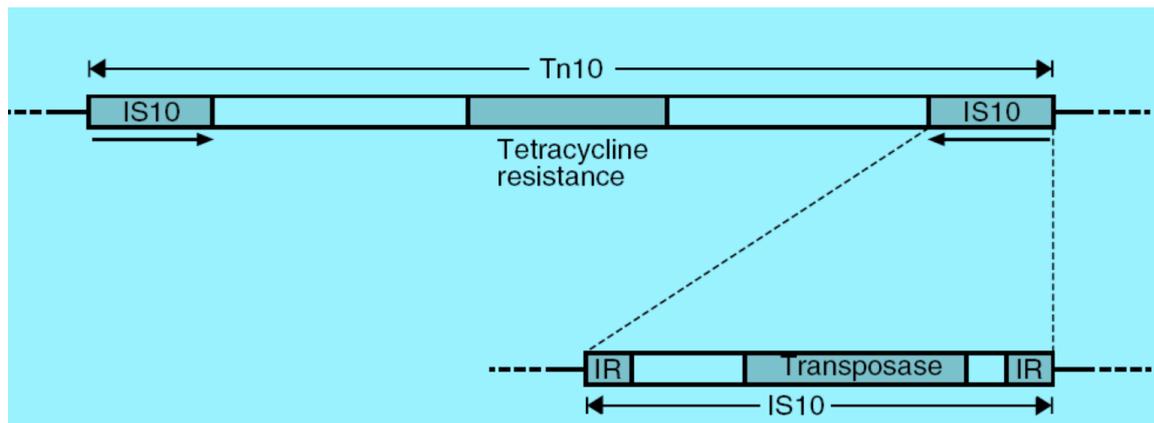
**Figure 20:** Structure of the transposon Tn3. DR, five-base pair direct repeat (target duplication); IR, 38-base pair inverted repeats; *res*, resolution site; *tnpA*, transposase; *tnpR*, resolvase; *bla*, *b*-lactamase (ampicillin resistance) [Dale and Park,2004].

Some transposable elements have a more complex structure than Tn3. These composite transposons consist of two copies of an insertion sequence on either side of a set of resistance genes. For example the tetracycline resistance transposon Tn10, which is about 9300 bp in length, consists of a central region

carrying the resistance determinants flanked by two copies of the IS10 insertion sequence in opposite orientations. IS10 itself is about 1300 bp long with 23-bp inverted repeat ends and contains a transposase gene.

Composite transposons may have their flanking IS regions in inverted orientation or as direct repeats. For example Tn10 and Tn5 both have inverted repeats of an IS (IS10 and IS50 respectively) at their ends, while Tn9 has direct repeats of IS1. The transposition behaviour of such composite elements can be quite complex; the insertion sequences themselves may transpose independently or transposition of the entire region may occur. Furthermore, recombination between the IS elements can occur, leading to deletion or inversion of the region separating them .

Even more complex arrangements can occur. For example, Tn4 appears to be related to Tn21 but contains a complete copy of Tn3 within it. The ampicillin resistance gene of Tn4 can thus be transposed as part of the complete Tn4 transposon or by transposition of the Tn3 element.



**Figure 21:** Structure of a composite transposon Tn10[Dale and Park,2004].

The composite transposons such as Tn10 are also known as class I transposons, while transposons such as Tn3, flanked by inverted repeats rather than IS elements, are referred to as class II transposons.

### 3. Integrons

Integrons are independent, mobile elements that encode genes for protein functions, and encode additional DNA to guarantee the integron's expression and integration into the bacterial genome.

The integron consists of a regulatory region, an open reading frame encoding a site-specific recombinase and a target site for the recombinase, *attI*, which can

integrate a variety of antibiotic resistance gene cassettes carrying a 59-base element, which is also recognized by the recombinase.

Integrations effectively generate widespread antibiotic resistance by donating antibiotic resistance genes to any strain of bacteria. Ironically, it is widely believed that integrations evolved only recently in response to antibiotic selection pressure. In other words, the use of antibiotics advanced the widespread occurrence of antibiotic resistance.

## **Mechanisms of transposition**

### **Replicative transposition**

In considering mechanisms of transposition, it is not necessary to distinguish between insertion sequences and transposons as the same mechanisms apply to both types of transposable elements. Transposons such as Tn3 transpose by a replicative mechanism: a copy of the element is inserted at a different site (on the chromosome or on a plasmid) while the original copy is retained.

The outcome of this stage is the formation of a larger plasmid known as a cointegrate, which consists of the complete sequence of both plasmids fused together, but now with two copies of the transposon, in the same orientation.

With some naturally-occurring plasmids, cointegrate molecules such as this can be readily isolated; in other cases, the intermediate is rapidly resolved into two separate plasmids, each of which contains a copy of the transposon, with the target sequence on the recipient plasmid being duplicated on either side of the inserted transposon.

### **Non-replicative (conservative) transposition**

Not all transposons show exactly the same behaviour. In particular, some transposons and some insertion sequences, do not replicate when they transpose, exhibiting a mode of transposition known as conservative (or non-replicative) transposition. This occurs with the insertion sequence IS10 and the related transposon Tn10.