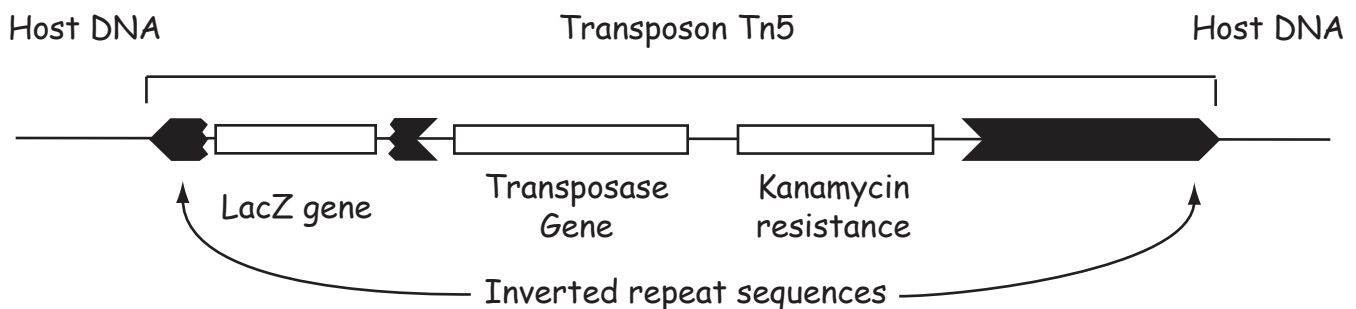


Lecture 13

Transposable elements

Transposons are usually from 10^3 to 10^4 base pairs in length, depending on the transposon type. The key property of transposons is that a copy of the entire transposon sequence can at a low frequency become inserted at a new chromosomal site. The mechanism by which transposons insert into new sites differs from one kind of transposon to another, but the details are not important to understand how transposons can be used. It is worth contrasting the recombination events that occur during transposition to the homologous recombination events that we have considered in meiosis. In homologous recombination, crossovers occur between like sequences. While this type of recombination can generate new combinations of alleles the arrangement of genes is left undisturbed. In contrast, transposition involves recombination between unrelated sequences, namely the ends of the transposon and a site in the target sequence. Transposition therefore results in a new arrangement of genes along the chromosome.

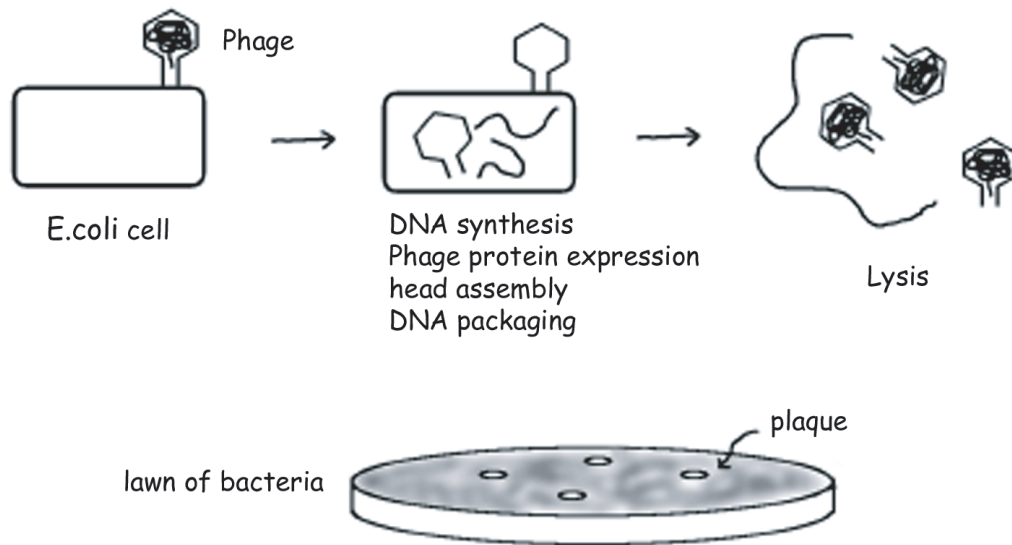
The generic structure of a transposon looks like this:



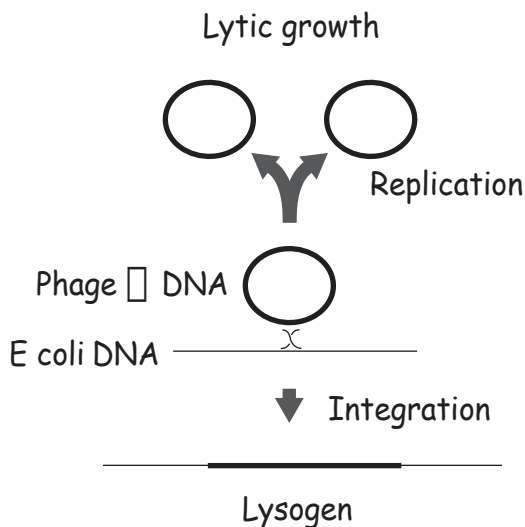
Transposon Element	Function
Transposase	An enzyme that cuts the target DNA more or less at random and splices the transposon ends to the target sequences. Other steps in transposition are performed by host enzymes.
Inverted Repeats	These sequences direct transposase to act at the ends of the transposon. Note that because the sequences are inverted, the two ends have identical sequence.
Selectable Marker(s)	Transposons are thought to have evolved by providing a selective advantage to the host cell. Many transposons carry genes that confer antibiotic resistance or some other benefit to the host.

The study of transposition mechanism and the biology of transposons is an interesting subject in genetics but for our current purposes we are going to concentrate on how transposons can be used for bacterial genetic analysis. For this purpose we will focus on the transposon Tn5 which can function in *E. coli* as well as a wide variety of other bacterial species. The selectable marker in Tn5 is a gene that confers resistance to the antibiotic kanamycin. Thus bacteria without Tn5 are sensitive to kanamycin (Kan^S), whereas bacteria that have Tn5 inserted into the chromosome are resistant to kanamycin (Kan^r).

One of the best ways to introduce a transposon into *E. coli* is by using a virus that infects bacteria known as a phage. For this purpose we will use a well studied type of phage known as λ .



Once the DNA from phage λ enters a cell it circularizes and then can undergo two possible fates.



In the lytic growth mode phage DNA is replicated in preparation for packaging into new phage particles. Among the phage genes required for replication is the P gene.

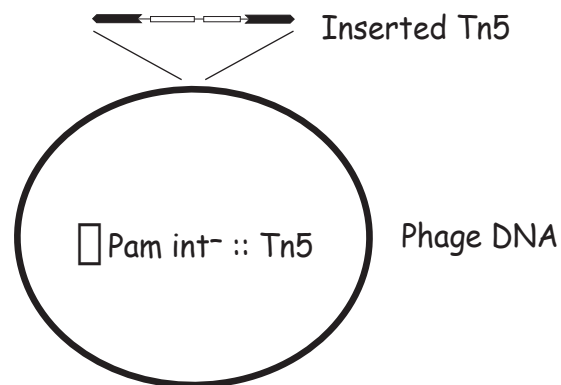
In the lysogenic mode phage DNA is integrated into the chromosome. This quiescent state the phage is replicated along with the chromosome in a state known as a lysogen. The phage gene required for integration is the **Int** gene.

To introduce random insertions of Tn5 into the *E. coli* chromosome we will start with Tn5 carried on a special λ phage vector: λ Pam $int^-::Tn5$.

Pam designates an amber (nonsense) mutation in the λ phage P gene. When λ Pam phage infect *E. coli* with an amber suppressor (Su^+) the phage multiply normally, which allows propagation of these mutant phage. But when λ Pam phage infect a nonsuppressing host (Su^-) the phage DNA cannot replicate.

int^- is a mutation in the λ integrase gene. Phage with this mutation can not integrate into the host chromosome to make a stable prophage.

$::Tn5$ designates that the λ phage carries an inserted copy of Tn5.



When λ Pam $int^-::Tn5$ infects a wild type ($Su^- Kan^S$) *E. coli* host, the phage DNA cannot replicate (Pam) nor can it integrate (int^-) thus the only way for the *E. coli* to become Kan^R is for Tn5 to transpose from the λ DNA to some location on the *E. coli* chromosome. This type of transposition is an inherently rare process and will occur in about one out of 10^5 phage-infected *E. coli* cells.

This is how a transposon mutagenesis can be done:

- 1) Infect 2×10^9 wild-type *E. coli* cells with λ Pam $int^-::Tn5$ so that each cell receives at least one phage chromosome.
- 2) Select for Kan^R by plating on medium that contains kanamycin. There should be a total of about 2×10^4 Kan^R colonies. Each of these should have Tn5 inserted into a different site on the *E. coli* chromosome.

The genes of *E. coli* are densely spaced along the chromosome and about half of the Tn5 insertions will lie in one gene or another. There are 4,200 genes in *E. coli* so our collection of 2×10^4 random Tn5 insertions will likely contain at least one insertion in each gene. (Note that insertions in genes that are essential for *E. coli* growth such as the genes for RNA polymerase or ribosomal subunits will not be recovered because these insertion mutants will not form colonies on the kanamycin plates).

Let's say that we are interested in the *E. coli* genes that are involved in synthesis of histidine. To find insertion mutants that can not synthesize histidine (His^-) we could screen amongst our collection of 2×10^4 random Tn5 insertions to find those that are His^- . The easiest way to do this would be to plate out the collection of insertions at a density of 200 colonies per plate (100 plates total). Each of these master plates would then be replica plated (first by transfer to a sterile piece of velvet) to a plate that contains histidine and also to a plate that lacks histidine. His^- insertion mutants would be identified as colonies that can not grow on the plates that lack histidine. Note that the same collection of random Tn5 insertions can be screened multiple times to find interesting mutations with different phenotypes.

3) Identify His^- Tn5 insertion mutants by replica plating to find colonies that specifically can not grow on plates that don't contain histidine.

Once we have a set of His^- insertion mutations (in the present example, one might expect to find 10-20 different His^- mutants), the affected gene(s) can be identified by the simple fact that they will be "tagged" by the inserted Tn5 sequences. The easiest way to identify the site of insertion is by performing a special PCR amplification of the DNA fragment that corresponds to the novel junction between Tn5 and the bacterial chromosomal sequences. Ordinarily PCR reactions are carried out using two DNA primers, each of which corresponding to an end of the sequence to be amplified. When we want to amplify a junction fragment we can use as one of the primers a sequence that lies near the end of Tn5 but we won't yet know the relevant chromosomal sequence to allow the other primer to be designed. There are several tricks that can be used to circumvent this problem, which are too complicated to describe here. Suffice it to say that there are ways that the junction fragment can be amplified by PCR using only sequences defined by the Tn5 portion of the junction fragment.

4) Use the known sequence of the end of Tn5 to PCR amplify a fragment that spans the junction between the end of Tn5 and the *E. coli* chromosomal site that was the target for insertion. DNA sequencing of the amplified junction fragments will give the identity of the target sequences. Since we know the DNA sequence of the entire *E. coli* chromosome, the gene that was the target for Tn5 insertion can be identified unambiguously.

5) The DNA sequence of the junction fragments will identify all of the genes that have been inactivated to give the His^- phenotype.

The procedure just outlined can be used to isolate and characterize a wide variety of useful mutations. A major limitation of this method is that as stated earlier, transposon mutations usually completely disrupt the target gene and therefore lead to a complete inactivation of the gene product. Often we will want to work with point mutations (such as temperature sensitive mutations or nonsense mutations). In the next lecture we will see how transposons can also be used to facilitate analysis and manipulation of point mutations.