

MICROBIAL GENETICS

Genetic exchange: DNA transfer mechanisms in bacteria

Abhijit A. Sardesai and J. Gowrishankar
Laboratory of Bacterial Genetics
Centre for DNA Fingerprinting and Diagnostics
ECIL Road, Nacharam
Hyderabad 500 076

17 May 2006 (Revised April 2007)

CONTENTS

[Introduction](#)

[A bird's-eye view of a bacterial cell](#)

[Conjugation \(DNA transfer\) system of the F factor of *E. coli*](#)

[Historical perspective on bacterial conjugation](#)

[Mechanism of F-mediated conjugation](#)

[Hfr state](#)

[Formation of F-prime \(F'\)](#)

[Transduction of DNA](#)

[Transformation](#)

[Integrative and conjugative elements \(ICEs\)](#)

[Implications of DNA transfer mechanisms](#)

Keywords

Conjugation, F (fertility) factor, F'(F prime), Hfr (high frequency of recombination), transduction, bacteriophage, competence, transformation, horizontal gene transfer, integrative conjugative elements

1.0 Introduction

This chapter focuses on the mechanisms by which the genetic material (mainly DNA) is transferred (naturally) from one bacterium to the other or can be transferred (experimentally) into a particular bacterium. Though the process of DNA transfer is usually one way, it is directed towards the pathway of genetic exchange, ensuring genetic variability in the subsequent generation. Experimenters have thus adopted a convention to describe the two entities (bacteria) participating in the process of DNA transfer with one designated as a donor and the other as the recipient. By now the reader should be able to start gauging the message of this chapter going by the opening few sentences which is “that at least in the microbial (bacterial) world the genetic material unique to a particular bacterium does not appear to be confined within, rather there exist mechanisms mediating DNA transfer between bacteria”. This process is a naturally occurring one and the mechanisms governing DNA transfer reside in the donor bacterium. Genetic exchange provides the basis for the hypothetical phenomenon known as “**horizontal gene transfer**” (HGT). HGT refers to the phenomenon or process, which results in parts of the genetic material of two organisms being more (or less) similar to one another than their ancestral phylogenetic relationships would indicate. Sometimes it is also referred to as “lateral” gene transfer. The phenomenon of HGT will be dwelt upon in the later sections of this chapter but for the present it will suffice to say that HGT is thought to be one factor, which has contributed to the widespread phenomenon of antibiotic resistance displayed by bacteria, a major impediment in treatment of bacterial infections.

2.0 A bird's-eye view of a bacterial cell

At this point, it may be relevant to survey elements of bacterial cell structure and the prototype bacterial cell discussed here is that of the bacterium *Escherichia coli* (*E. coli*). *E. coli* under the microscope appear as rod shaped cells with flagella (locomotory appendages) emanating from the cell periphery (Figure 1). The cytoplasm of *E. coli* is bounded by two membranes, the inner (cytoplasmic) membrane and the outer membrane. A complex of sugars and lipids, lipopolysaccharide (LPS) coats the surface of the outer membrane. The space between the two membranes is known as the periplasm. The cytoplasm contains the genetic material, a single continuous piece of double-stranded DNA, which is circular and also known as the chromosomal DNA. Frequently though one finds the presence of additional double-stranded DNA which is circular and such DNA is termed as a plasmid. A plasmid is a genetic element that is capable of existing as an autonomously replicating entity, meaning that it replicates alongside the bacterial chromosome but remains spatially separated. Such a configuration of co-residence of a plasmid with the chromosomal DNA is not unique to *E. coli* and many bacterial species are known to possess such “**extrachromosomal**” elements. A special kind of plasmid is called as an episome. Episomes have the capacity to replicate “**autonomously**” wherein the episome replicates alongside the chromosome but the two remain as distinct entities, separated within the cell. In addition, episomes can integrate into the bacterial chromosome and the integrated episomal DNA is replicated along with the bacterial chromosome. Lambda prophage and F factor are examples of episomes. It may be said that despite the above-mentioned distinguishing feature between episomes and plasmids the two terms are used synonymously. Another general point about bacteria is that bacteria are haploid organisms meaning that each gene on the bacterial chromosome exists in single copy, at least in resting or nondividing cells.

Knowledge on the mechanisms of DNA transfer has come from studies on bacteria like *E. coli* and *Bacillus subtilis* wherein investigators have had at their disposal several experimental tools to dissect a given biological process. Discussed below are some well-studied examples of DNA transfer.

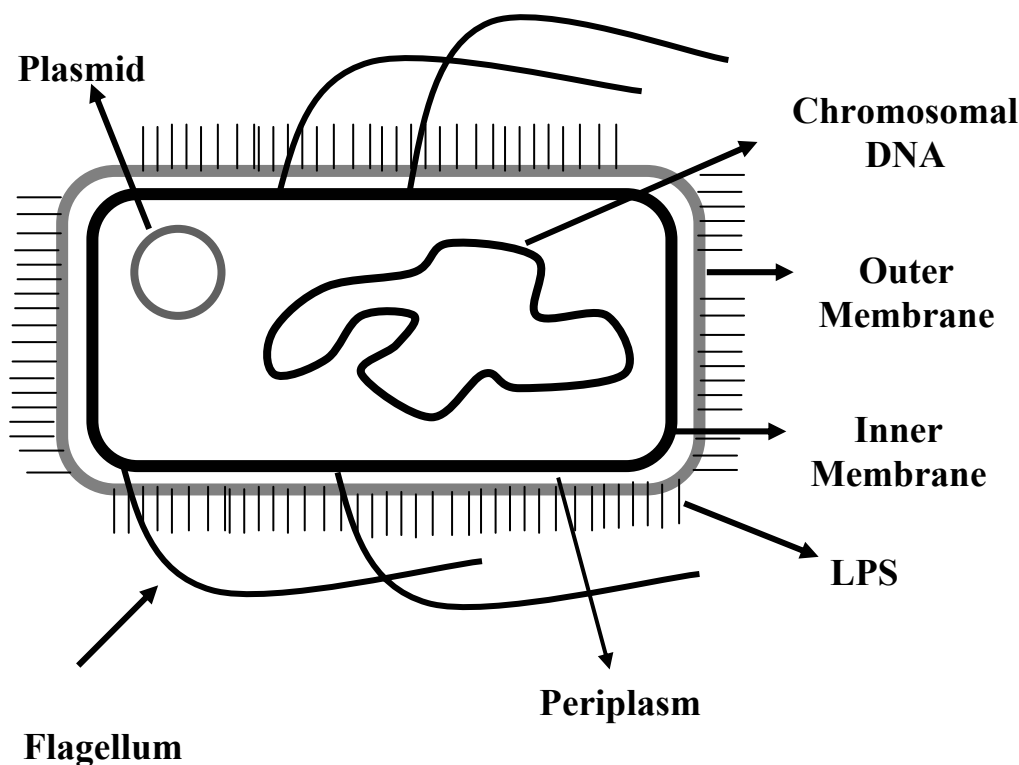


Figure 1: Elements of bacterial cell structure (A schematic representation of a “cut-away” section of a typical *E. coli* cell is shown. The various cytological features are indicated)

3.0 The conjugation (DNA transfer) system of the F factor of *E. coli*

Some strains of *E. coli* contain a plasmid called as the F (fertility) factor, also referred to as F-plasmid. F is a circular double-stranded DNA molecule approximately 10^5 basepairs (100 kilobases, kb) in size. F factor can exist in the cell either as free autonomously replicating entity or can reside in an integrated state within the bacterial chromosome. Thus F is an example of an episome. The F factor can transfer itself from one cell to a cell lacking F in a process termed as conjugation (see references 10 and 15 for information on conjugation). The F factor comprises the following features (Figure 2):

1. In cells not engaged in conjugation (that is during vegetative growth), the F factor replicates autonomously and F replication is initiated from a specific sequence on the F called *oriV* (vegetative origin). The *oriV* mediated replication is bi-directional. There is a second origin

of unidirectional replication *oriS*, but for the most part *oriV* serves as the major origin of replication.

2. The *oriT* (transfer origin) region is the point from which replication involved in the transfer process of F DNA is initiated.

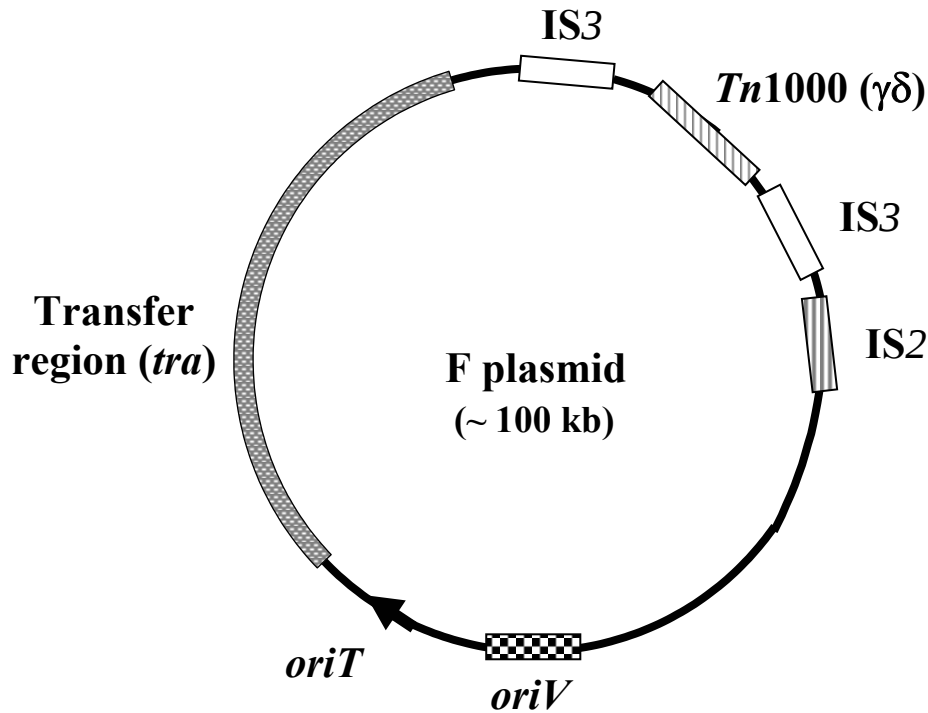


Figure 2: Structure of the circular F plasmid: Map of the circular double-stranded DNA molecule of the F factor, showing the important regions.

3. A region designated, as *tra* comprises genes that encode components required for the process of DNA transfer.
4. The F DNA bears a set of unique DNA sequences known as transposable elements. These include single copies each of the transposon *Tn1000* (also known as $\gamma\delta$) and the insertion sequence *IS2* and two copies of the insertion sequence *IS3*. Transposable elements may simply be defined as DNA sequences capable of moving from one location in the DNA to another. Furthermore, transposable elements are categorized into two families namely i) those which only contain genes required for the process of transposition and ii) those bearing additional genes other than the ones involved in transposition. Members of the former class are designated as insertion sequences (IS) while those from the latter class are termed transposons, prefixed with the letters Tn. It is worth mentioning here that in biology of the F factor the role the transposable elements play, is not always related to their ability to transpose. In fact, they impart on the F its ability to integrate into the bacterial chromosomal DNA.

Within the cell, the F factor can exist in two physiological states. In the one state the F factor exists as an autonomously replicating entity meaning that it replicates, using *oriV*, independently of the bacterial chromosome and the two remain spatially separated. There is a variant of the F

autonomous state, in which the F factor exists as an independently replicating entity but the F carries some part of the bacterial chromosomal DNA on it. Such an F factor is known as an F-prime (F'). In the second state, the F episome integrates into the bacterial chromosome and is replicated along with the chromosome. This state of F is also known as the Hfr state. Hfr is an acronym for high frequency of recombination. Hfr formation through Tn1000 is by transposition and that by the IS2/IS3 elements is by homologous recombination. Strains bearing either an F (F') or an Hfr are referred to as F⁺ (male or donor) cells, while those lacking either are designated as F⁻ (female or recipient) cells. The F factor can transfer itself into the recipient cell, and while in the integrated state the covalently attached bacterial chromosome is also mobilized as a result of F transfer. The process of F DNA transfer (or of transfer mediated by F-like elements) is known as conjugation.

3.1 A historical perspective on bacterial conjugation

In higher eukaryotic organisms bearing multiple linear chromosomes, the entire genetic material of a cell exists in a state known as the diploid state, which means that any given gene is present in two copies, on a pair of identical (homologous) chromosomes. The transfer of genetic material from one generation to another occurs via the process of “**sexual reproduction**”. In this process, specialized cells of an individual organism undergo the process of meiosis, leading to the precise halving of the genetic complement. Cells bearing such a halved genome are known as gametes and they are in the haploid state. The fusion of two gametes, one each derived from the male and female individuals, leads to the formation of a new individual with the restoration of the diploid state. This process is termed “**sexual**” because the genetic material from two individuals undergoes mixing following fusion of the two gametes.

At this point a quick primer on bacterial cell division is introduced. During the growth of a bacterium, its chromosomal DNA replicates with replication initiating at a fixed sequence of DNA known as the chromosomal origin of replication (*oriC*) and proceeds in a bi-directional fashion (clockwise and counter clockwise) around the circular chromosome, and terminates at a sequence known as the replication terminus. The terminus is located almost directly opposite the *oriC* region on the chromosome, so that the distance traversed by the two (clockwise and counterclockwise) replication arms is nearly equal. After semi-conservative replication is completed, the two newly formed chromosomes, each bearing one DNA strand derived from the parental chromosomal DNA used as a template for replication and a complementary newly synthesized strand, separate from each other towards the ends of the cells known as the cell poles. The cellular machinery then lays down a cell envelope partition known as the cell septum between the two separated molecules of DNA, leading to the formation two new daughter cells. It is therefore said that bacteria propagate themselves by the mechanism of “**asexual reproduction**” and no net transfer/mixing of genetic material takes place (Figure 3).

Early studies were initiated on bacterial conjugation because the investigators wanted to find out whether in organisms with haploid genomes and an asexual mode of propagation, sexual (conjugal) transfer of genetic material could occur. Critical to the discovery of conjugation were the availability of two techniques to the investigators. One was the formulation of a defined medium in which bacteria could be cultivated. This medium is known as minimal or synthetic medium, of known composition. The second technique available at that time was a method to isolate *E. coli* strains bearing multiple nutritional requirements. A standard wild type *E. coli*

strain can grow vigorously on minimal medium consisting only of glucose, ammonium sulfate, buffered potassium phosphate, and appropriate amounts of magnesium and calcium. A nutritional mutant of this strain cannot grow on minimal medium unless the deficiency caused by the mutation is overcome by providing a particular chemical as a nutritional supplement. For example a *trp*⁻ *E. coli* strain cannot grow on minimal medium unless the amino acid tryptophan is added to this medium. This strain, because of a mutation in a gene coding for an enzyme involved in tryptophan biosynthesis cannot make tryptophan and thus needs tryptophan for growth. Such a mutant is said to be a tryptophan **auxotroph** whereas the wild-type strain is said to be a **prototroph**. Similarly a *met*⁻ strain is a methionine auxotroph and therefore needs methionine as a nutritional supplement in minimal medium. The availability of strains bearing multiple nutritional deficiencies was an important resource.

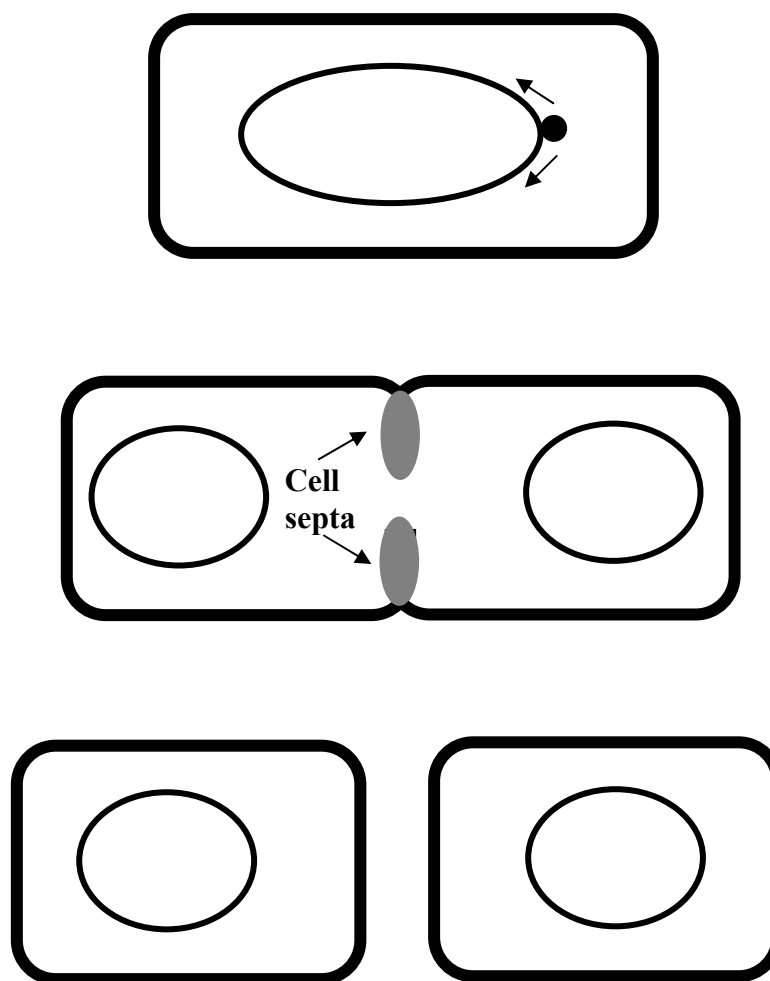


Figure 3: Bacterial cell division: A simplified view of the *E. coli* cell division process is shown. Bidirectional chromosome replication initiating at the *oriC* sequences (black circle) is indicated with a pair of outwardly directed arrows. The growing cell envelopes or septa are depicted as gray ellipses.

In the mid 1940s Joshua Lederberg posed the following question: “Can sexual transfer of DNA be demonstrated in bacteria”? Lederberg worked on the premise that if two strains of *E. coli* were mixed and cultured, and if some sort of DNA transfer mechanism existed, then the progeny (zygote) derived from such a cross would contain genes from both the organisms. The problem was to identify such zygotes. For this Lederberg and Edward Tatum employed two multi-auxotrophic strains of *E. coli*. Strain A had a nutritional requirement for the amino acids threonine, leucine and the vitamin thiamin, and strain B had a nutritional requirement for the vitamin biotin, and the amino acid methionine. In genetic parlance, the genotype of strain A is $thr^- leu^- thi^-$ and that of strain B is $bio^- met^-$. Both strains cannot grow on minimal medium unless their nutritional requirements are met by exogenous supplementation with the required amino acids or growth factors. It may be noted that strain A can also be referred as $bio^+ met^+$ while strain B is $thr^+ leu^+ thi^+$. If DNA transfer (conjugation) were to occur from strain A to strain B or vice versa then, the nutritional deficiencies of one strain (because of mutated genes) would be overcome by the presence of the normal copy of the respective genes obtained from the donor strain. Such a strain (recombinant) would be the only strain capable of growing on minimal medium after cultures of strain A and B were mixed and plated. Minimal medium plays the role of a “**selective medium**”, where only the prototrophic recombinants can grow. Such recombinants were indeed obtained. There was an assumption made in this experiment that the transferred DNA from the donor would recombine into the genome of the recipient strain. It may be said that, at the time this experiment was done, although the phenomenon of transfer of genetic material was detected, Lederberg and Tatum did not seem to have an explanation on the mechanism by which the transfer could have taken place. William Hayes, as described below provided the answer to this. Importantly, Lederberg and Tatum recognized that a physical contact between the two strains was necessary for this process of conjugation.

It was William Hayes who solved the mystery of conjugation through a series of experiments and demonstrated that: (i) in the process of conjugation, one strain served as a donor of genetic material while the other served as a recipient, (ii) the process was unidirectional, (iii) survival of the recipient cell was necessary for the formation of the zygote, (iv) the donor (male) cell contains a factor (F or fertility factor) which is transferred to the recipient (female) cell, and (v) in the Lederberg and Tatum experiment, the population of donor cells most probably bore some cells in which an F factor had integrated into the genome of the donor cell and F-mediated transfer had resulted in transfer of the donor’s genetic material into the recipient cytoplasm (described in detail in section 3.3).

3.2 Mechanism of F-mediated conjugation

The process of F DNA transfer requires the gene products of the F *tra* (transfer) region and it is estimated that around 40 gene products participate in the initiation and completion of transfer of F DNA from the donor to the recipient cell. Mutations in any of the *tra* region genes lead to an inability of the F to transfer itself, hence they are designated with the prefix *tra* e.g *traA*, *traD*, *traS* and so on. Steps occurring in the F transfer are described below (for reviews on F-mediated conjugation see references 6, 8 and 18).

A) Formation and stabilization of pairs of donor and recipient cells

In the initial phase of conjugation, F^+ and F^- cells are often found to be associated in close proximity to each other and form what is known as the “mating pair aggregate”. Over a period of time, such associations strengthen so that an individual aggregate becomes resistant to shear (turbulent) forces. The *traA* gene product on the F plasmid encodes the protein pilin, which forms the subunit of a homopolymeric (comprised of identical components) assembly of proteins, generating the F pilus. The F pilus is the entity responsible for the formation and stabilization of the mating pair. Initially the tip of the F pilus makes a contact with the LPS layer on the outer membrane of the F^- cell. Following the contact the F pilus traverses the outer membrane of the recipient cell followed by its depolymerization. The step of pilin depolymerization in effect serves to bring the pair of conjugating cells in close proximity.

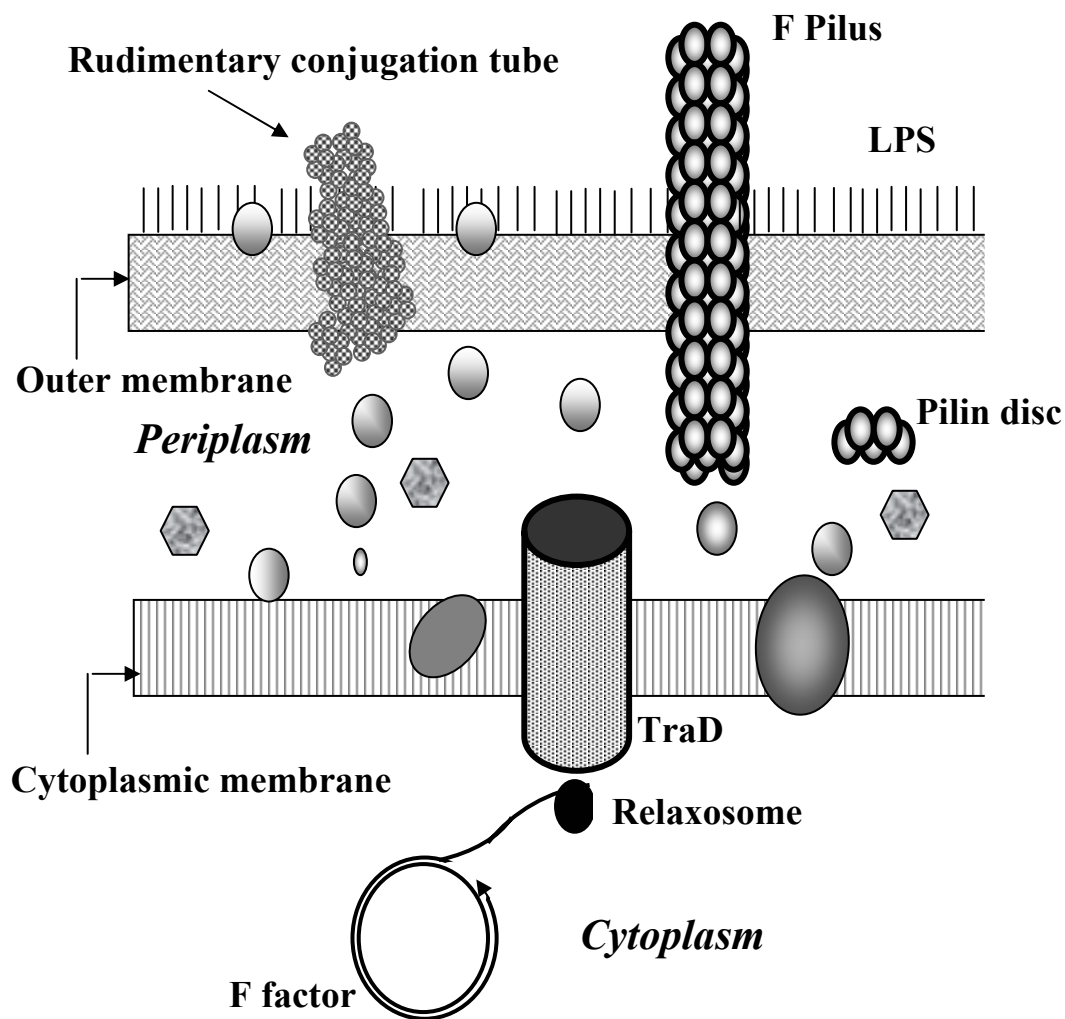


Figure 4: F DNA transfer apparatus: A schematic representation of components of the F transfer machinery and their spatial locations in the donor cell. A limited but relevant set of components are shown. Unlabelled objects in the periplasm represent products of various *tra* genes involved in pilus assembly, conjugation tube formation and surface exclusion. (Figure adapted from reference 6)

F pilus is a multiprotein aggregate whose individual subunits are made up of one protein namely pilin. Five molecules of pilin associate with each other to form a “disc” like structure and multiple “pilin discs” stack upon each other to form a cylindrical stalk like structure known as the pilus (Figure 4). The stacking of pilin discs is not exact, rather one disc appears rotated at a particular angle with respect to its counterpart located below, giving a helical appearance to the pilus. Pilin initially is synthesized in the cytoplasm of the F^+ cell as a precursor polypeptide. Functions of multiple gene products of the F *tra* region are called upon for the maturation and transport of the pilin subunit from the donor cell cytoplasm to its periplasm. Once inside the periplasm, additional *tra* region-encoded gene products participate in the assembly of the pilus.

B) The F DNA transfer apparatus and initiation of DNA transfer

As the tip of the pilus contacts the outer membrane of the recipient, a “mating signal” is generated leading to the initiation of the DNA transfer process. The nature of such a mating signal is, however, not known. The F DNA transfer is initiated by the occurrence of a DNA strand break (a nick) at a specific sequence within the *oriT* region of the F episome. The F-encoded TraI protein nicks one strand of DNA. TraI has the additional property of being a helicase, an enzyme capable of unwinding the broken strand of DNA and this serves to separate the broken *oriT* strand from its paired counterpart on the double-stranded F DNA. TraM, TraY and TraZ proteins act as accessory factors and help speed up this process and form along with TraI a protein complex known as the “**relaxosome**”. As the TraI generates the nick, it becomes covalently bound to one end of the broken strand (the 5' phosphate bearing DNA end) and begins to unwind that strand. The other end bears a 3' hydroxyl group and serves as a primer for DNA replication of the F factor in the donor. In effect, the F episome undergoes rolling circle replication with concomitant displacement of the TraI-bound *oriT* strand that is thought to be mobilized into the recipient cell (Figure 5). As the associations between the donor and the recipient cells strengthen, an intercellular junction known as the “conjugation tube” is apparently formed. Evidence for this junction appears from electron microscopic studies on pairs of interacting cells as an electron dense (that is, appearing as a dark object) region. At this point, it must be stated that details on the later stages of the process of conjugation are still somewhat sketchy, and the scheme presented herein may be termed as a “working model”. How is the displaced F DNA strand in the donor cell brought to the conjugation tube? The answer to this question, in part, comes from the observation that the F encodes an inner membrane protein, TraD, which has been shown to possess non-specific DNA binding ability. It is therefore suspected that TraD mediates transfer of the displaced strand through the conjugation tube into the recipient cell. TraD is also known to physically associate with the TraM protein (a component of the relaxosome complex) thereby serving, in addition, to bring the displaced F strand towards the surface of the donor cell. TraD, is therefore, known as the coupling protein. Examples of “TraD-like” proteins are found also in DNA transfer systems elaborated by the R100 episome and the Ti plasmid of *Agrobacterium tumefaciens*.

The end of the displaced F *oriT* strand is thought to remain located at the intercellular junction, and upon further unwinding the F DNA behind it is transferred into the recipient cytoplasm wherein it undergoes a discontinuous mode of (lagging strand) replication. In the donor cell, the F is replicated in a continuous fashion, from the 3' hydroxyl end of the *oriT* nick. Transfer

process apparently reaches the end stage when the *oriT* region in both the donor and recipient is replicated followed by recircularization of the donor and recipient F factors.

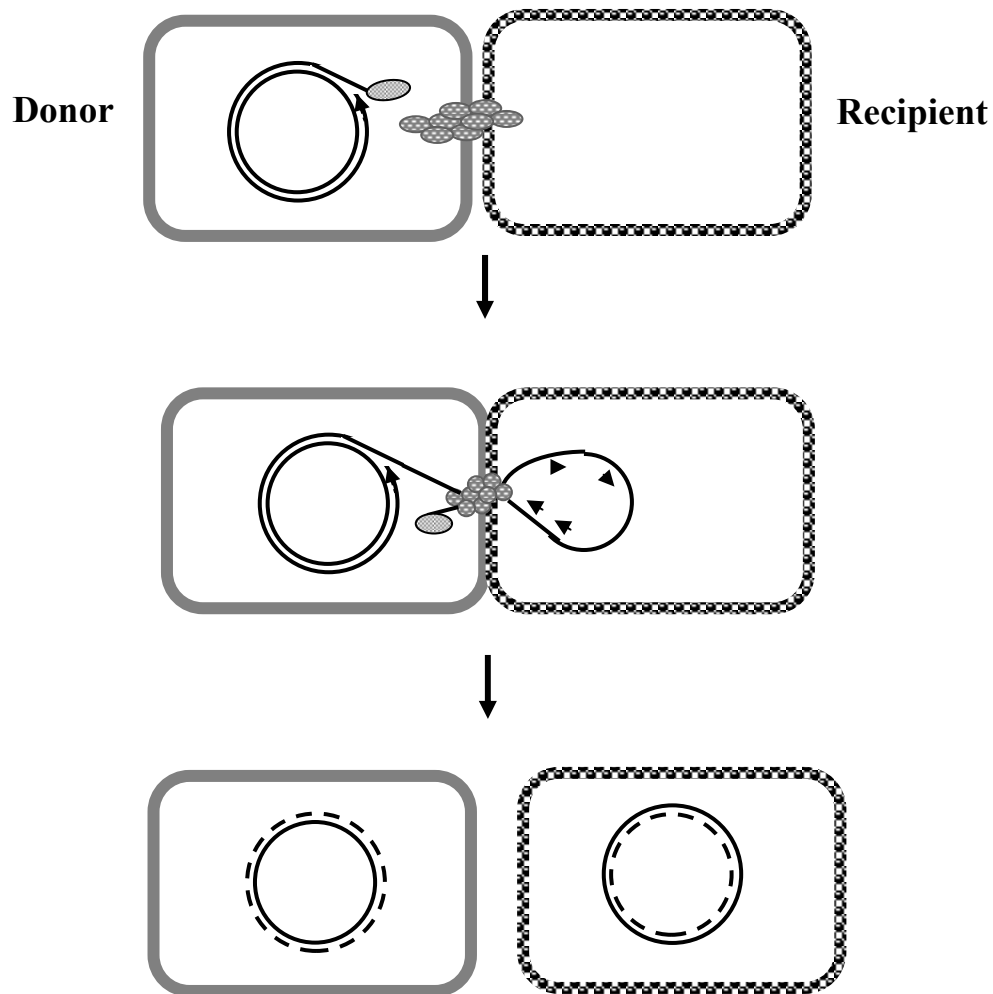


Figure 5: DNA transfer in F mediated conjugation: A model depicting DNA transfer during conjugation. The 5'phosphate of the nick generated at *oriT* is bound by the components of the relaxosome (gray ellipse), and recruited to the cell membranes from where the displaced single-stranded DNA is thought to be transferred into the recipient cytoplasm. In the recipient the displaced DNA is replicated in a discontinuous fashion. In the donor the 3'OH end of the nicked *oriT* primes continuous F replication. Resolution of the two DNA molecules occurs after one complete round of replication in the donor and the recipient takes place. (Figure adapted from reference 6)

C) The phenomenon of surface exclusion

The F-bearing cell is known to be incapable of transferring its F factor to another cell bearing F, and the underlying mechanism is known as surface exclusion. This property is attributed to the presence of two F borne genes *traS* and *traT*, whose products interfere in establishment of the mating process. TraS, an inner membrane protein is thought to interfere in the generation of the

mating signal. TraT, on the other hand, is a lipoprotein, which decorates the outer membrane of the F-bearing cell and prevents the formation of the mating pair aggregate. For reasons not clearly understood, surface exclusion can be overcome by particular environmental conditions, for example, by growth of the F-bearing “recipient” cultures with vigorous agitation to late stationary phase at a temperature of 25°C before mixing with the F-bearing “donor” cells that have been grown with slow agitation to mid-exponential phase at 37°C.

3.3 *Hfr* state

It may be recalled (section 3.0) that the F factor can also exist as well, in an integrated state in the donor cell genome. How does the F factor integrate? Credit for suggesting a mechanism for F integration is given to Allan Campbell. Campbell proposed that if two circular genomes shared a common sequence, then using such a site of homology a simple break-and-rejoining reaction (that is, a single-crossover event) would lead to the fusion of two genomes taking a final form of a larger circular entity known as a cointegrate (Figure 6). Such a process would require an enzyme capable of mediating **homologous recombination**. Another outcome of this proposal is that the sequence used for recombination gets duplicated on either side of the integrated DNA and is aligned in a fashion termed as a “direct repeat”. Campbell’s proposal forms the basis for the integration of the F episome and also for the site-specific integration of bacteriophage lambda (see section 4.2), into the bacterial chromosomal DNA. As per the “Campbell model”, F integration requires a host cell enzyme RecA and bacteriophage lambda requires the phage-derived enzyme Int, for promoting homologous recombination.

As described earlier (section 3.0 and Figure 1), the F factor contains a single copy of IS2 and two copies of IS3, which collectively are known as insertion sequences (ISs). Similar ISs are found scattered on the circular *E. coli* genome. The F factor integrates using these ISs as sites of homology, and the RecA enzyme facilitates the integration process. It should, therefore, be apparent that since the ISs are scattered in the genome of *E. coli*, there are several places where the F factor can integrate. Usually, sequences, which act as sites of homology, are “polar” in nature, which means that the individual nucleotides of this sequence are arranged in a definite order. For recombination to occur, the two participating sequences (one from each DNA molecule) must align properly before recombination takes place. In figure 7 (Panel A), the participating sequences (say, the IS2 from F and that from the bacterial chromosome) initially must align in the fashion depicted. Following recombination, the F becomes a part of the bacterial DNA, resulting in formation of an Hfr strain. In panel B (Figure 7), the IS2 sequence at a second location is oriented on the bacterial chromosome in an order opposite to that found in panel A. The F will now integrate in the opposite orientation. These points are relevant because the orientation of the *oriT* region in panel A will be opposite to that seen in panel B. Thus, during Hfr-mediated chromosome mobilization (discussed below), if the F exists in the chromosome in the configuration seen in panel A, markers (genes) located to the left, namely A, B and C will be transferred early with C being the first to enter. In the other configuration (panel B) genes to the right namely S, T and U will be transferred early with S being the earliest to enter the recipient cell. In a population of cells bearing an F factor, few cells will be in the Hfr state (with F integrations at different sites in the chromosome), and thus this population on the whole can transfer chromosomal genes at a low frequency (originally detected by Lederberg and Tatum). However if the Hfr-bearing cell can be isolated in a pure form, then this strain will transfer

chromosomal markers, from a fixed site and in a fixed order at a high frequency. The acronym Hfr thus stands for **high frequency of recombination**.

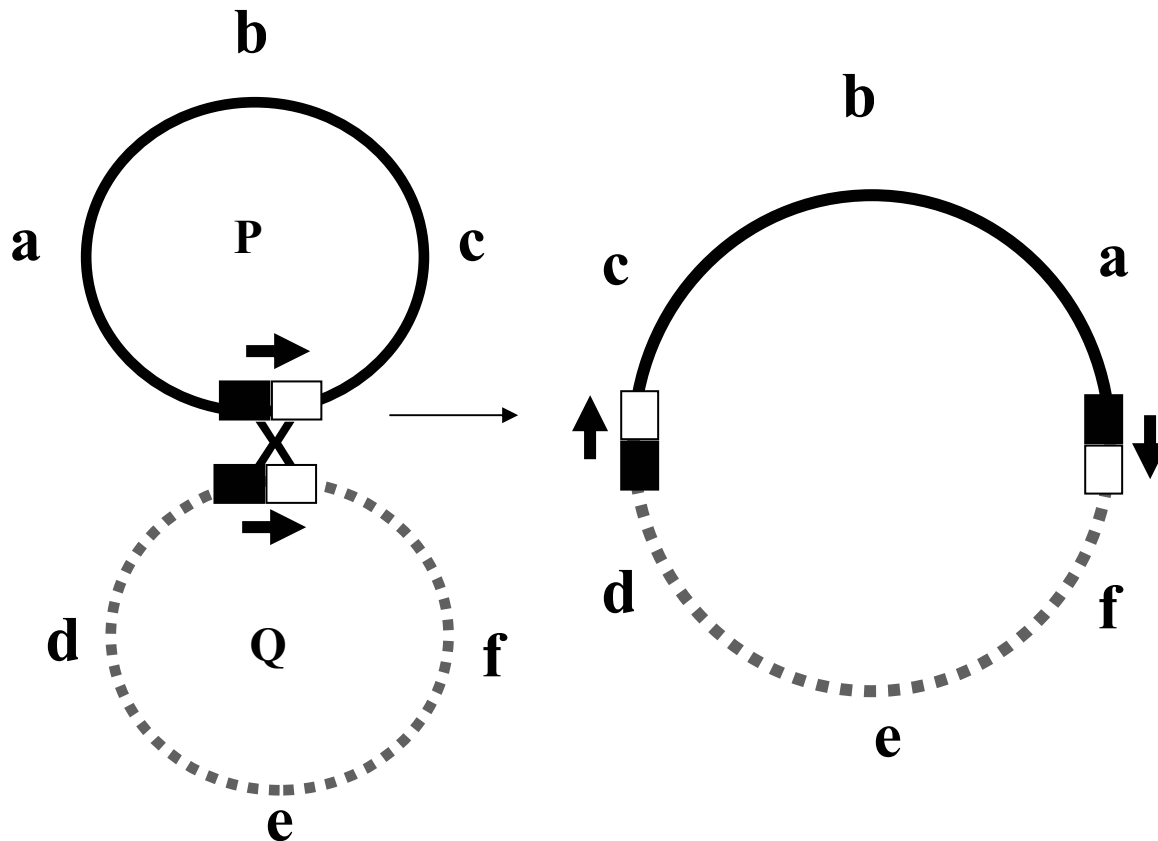


Figure 6: The Campbell scheme of DNA integration: Shown are two double-stranded circular genomes P and Q recombining (solid cross) with each other using a shared sequence (black and white box). The polarity of the shared sequence is depicted by a bold arrow. Small case letters denote arbitrary genetic markers on the two genomes. The larger circle on the right is the cointegrate, product of the PQ recombination. Note, the duplication and the “direct repeat” arrangement of the shared sequence in the cointegrate.

A) Chromosome mobilization by an Hfr

An F factor in the Hfr state can mobilize itself in the same way that an autonomously replicating F does. However, since it exists as a part of the bacterial chromosome, bacterial DNA located adjacent to the integrated F is also mobilized. A cell in an Hfr state is thus F^+ that is, a donor. The trademark nicking of the *oriT* region in F occurs and the displaced strand enters the cytoplasm of the recipient (F^-) strain. All the steps in conjugation remain the same as described earlier. The displaced strand is capable of recombining into the genome of the recipient strain, within a homologous region, a process requiring the activity of the RecA protein (Figure 8). It is worth reiterating that as an Hfr mobilizes itself, bacterial genes located adjacent to it will also be

mobilized. Those located close to the nicked *oriT* strand will be transferred early. Geneticists refer to these as “**early markers**”. Genes located in front of the *oriT* nick will be the last to enter the recipient cell and are referred to as “**late markers**”. This forms the basis for mapping bacterial genes using Hfrs .

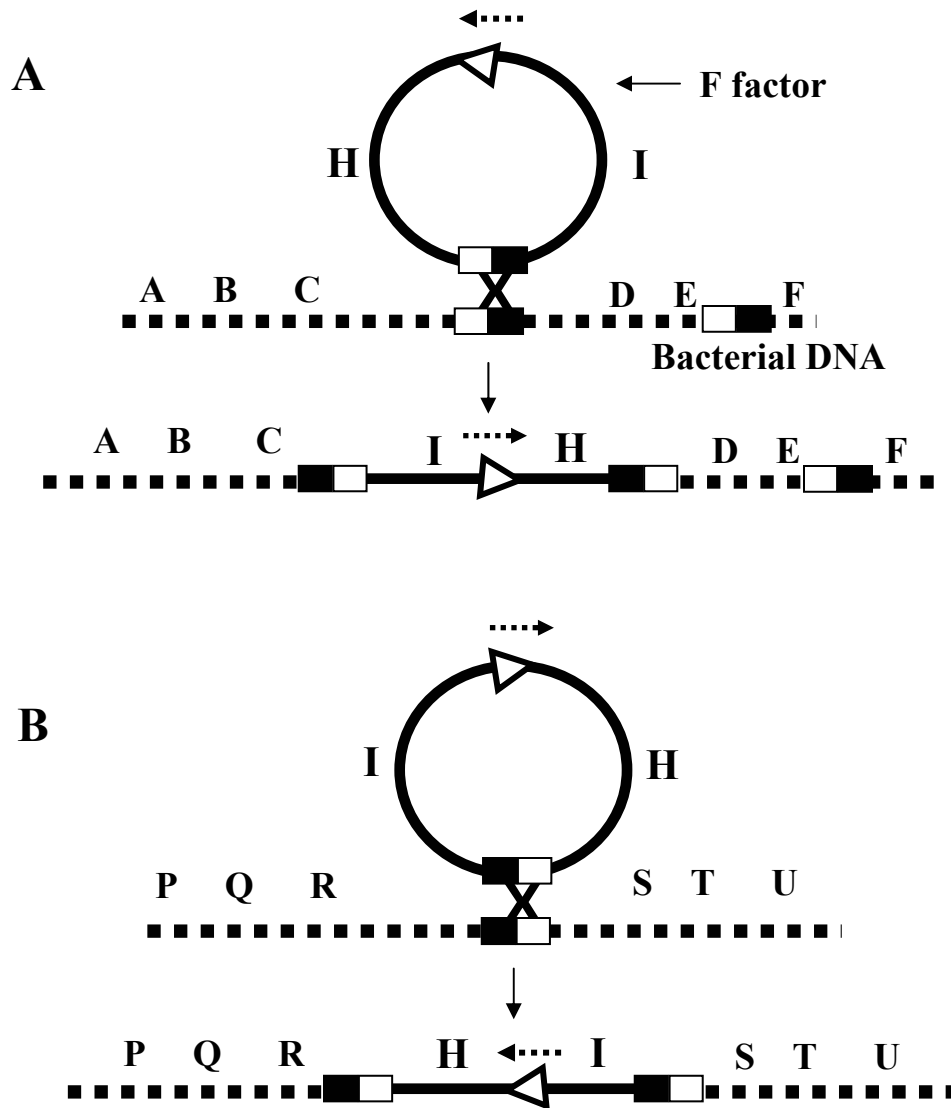


Figure 7: F integration on the chromosome (Hfr formation): The scheme for F factor integration, at two distinct sites and in two different orientations (panels A and B) on the bacterial chromosome (interrupted lines) is shown. Here the IS2 element on F (black and white box) is being employed to integrate (solid cross) within a similar sequence on the chromosome. In panel A, a second IS2 element located between chromosomal markers E and F is shown, whose relevance will be apparent in figure 11. H and I are arbitrary markers on the F factor while other roman letters denote contiguous genetic markers at two positions on the bacterial chromosome. The *oriT* region on F is shown with a white triangle and a dashed arrow indicates the direction of F DNA transfer.

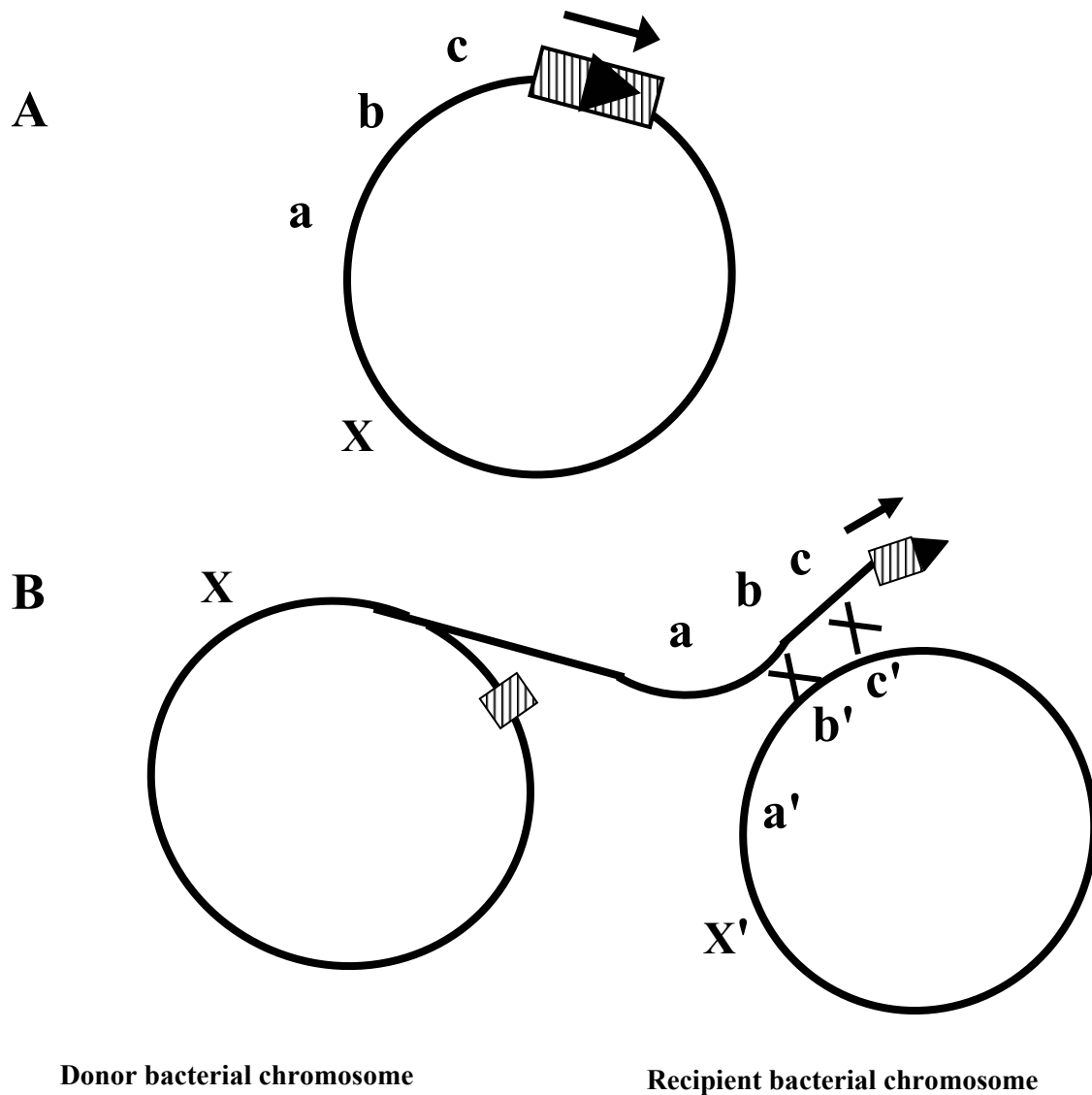


Figure 8: Hfr mediated chromosome mobilization: An Hfr donor strain contains an F factor (hatched box) integrated at a location in the bacterial DNA (panel A). The direction of F DNA transfer is indicated with an arrow. DNA processing reactions at the *oriT* region (black triangle) of the integrated F lead to the transfer of single-stranded DNA into the recipient cell. The donor markers (genes) a, b and c will enter the recipient early with c being the first to enter. The donor marker X is the late marker to be transferred. The transferred DNA bearing the donor genes a, b and c can recombine with the homologous region in the recipient replacing the recipient genes a', b' and c' and producing different types of recombinants

B) Mapping genes on the bacterial chromosome using Hfrs

In the formative years of development of bacterial genetics in the 1950s and 60s, it was suspected that the *E. coli* genome was circular. Where did the genes encoding proteins involved in different metabolic processes lie on this circular chromosome? Hfrs came in as handy tools to work out positions of different genes. Allied to this was the availability of multi-auxotrophic strains (described in section 3.1) of *E. coli*. A standard approach for gene mapping using Hfrs known as the **interrupted mating experiment** is described here (see also references 10 and 15). In a typical Hfr-based mapping experiment, the donor strain A is an Hfr that transfers chromosomal DNA in “clockwise” direction, while the recipient strain B is a multi-auxotrophic strain and is F^- . For example, strain B cannot biosynthesize threonine, proline, and tryptophan nor can it utilize lactose as the sole source of carbon for growth, whereas the donor strain is fully prototrophic. The recipient strain is also resistant to streptomycin whereas the donor strain is sensitive, a property which allows for the selection of “exconjugants” (that is, recipient cells which may bear transferred donor DNA). In genetic parlance, the donor is termed Str^s that is streptomycin sensitive (it is also thr^+ , pro^+ , lac^+ and trp^+) whereas the recipient is thr^- , pro^- , lac^- , trp^- and Str^r (streptomycin resistant). The donor and recipient cells are mixed, and the process of conjugation is initiated. At intervals of say, 5 minutes, aliquots of the conjugation mixture are removed and subjected to vigorous agitation, a process that disturbs the mating pair formation and so stops the DNA transfer.

The following assumptions / expectations arise:

1. DNA is transferred from the donor to the recipient at a fixed (arbitrary) rate, which remains constant over a period of time.
2. Disruption of the mating process early will lead to transfer of only the genes mobilized early (early markers) and no transfer at all of a gene mobilized late will occur.
3. Disruption of the mating process in the later stages will lead to transfer of more of the early markers followed by a gradual decline in the transfer of subsequent markers. It must be noted, however, that even without the step of vigorous agitation, there is spontaneous disruption of mating-pair that occurs during conjugation.

After disrupting the mating pairs at various time intervals following initiation of conjugation, the mixture of donor and recipient cells is plated on various selective media containing streptomycin. Addition of streptomycin ensures growth of only the recipient cells and the following results are obtained:

- 1) In the first time interval of 5 minutes, the number of thr^+ (threonine prototrophs, obtained by selection on minimal medium containing streptomycin and all amino acids except threonine) exconjugants is low which increases as the conjugation is allowed to proceed for longer time. Strains capable of proline, tryptophan biosynthesis and lactose utilization namely pro^+ , lac^+ and trp^+ exconjugants are not yet found amongst the thr^+ prototrophs (also known as thr^+ exconjugants).

Conclusion: The gene(s) for threonine biosynthesis is transferred as an early marker by the Hfr strain A and is positioned at 1 to 5 min (minutes) from the chromosomal site of F integration in the donor strain. The gene (s) for proline, tryptophan biosynthesis and lactose utilization may be located distal to those for threonine biosynthesis.

2) *lac*⁺ exconjugants (obtained by selection on minimal medium containing streptomycin and all amino acids, but with lactose instead of glucose as carbon source) begin to appear 10 to 15 minutes after initiation of conjugation.

3) *pro*⁺ exconjugants (obtained by selection on minimal medium containing streptomycin and all amino acids except proline) also begin to appear at around the same time and many *lac*⁺ exconjugants are *pro*⁺.

Conclusion: The gene(s) for lactose utilization is located at around the 10 min position and clockwise with respect to *oriT*. Since majority of the *lac*⁺ exconjugants have also become *pro*⁺, the gene(s) for proline biosynthesis must be located very close to the lactose utilization gene(s). Whether the *pro* gene(s) are positioned clockwise or counterclockwise with respect to the *lac* gene(s), cannot be deduced from this experiment.

4) 25 to 30 minutes following the initiation of conjugation, *trp*⁺ exconjugants (obtained by selection on minimal medium containing streptomycin and all amino acids except tryptophan) start to appear.

Conclusion: The gene(s) for tryptophan biosynthesis is transferred after the gene(s) for proline biosynthesis and is located at the 30 min position from *oriT*, clockwise with respect to the threonine, proline biosynthesis genes and the lactose utilization gene(s).

The results of the above mentioned experiment can be graphically represented as in Figure 9. It is noticeable that the slopes of the curves depicting entry of various markers differ, because when conjugation occurs, there is a spontaneous disruption of the mating-pair formation. The further a marker is located from the origin of transfer on an Hfr, the greater will be the effect of spontaneous mating-pair disruptions during its transfer. Consequently the apparent rates of recovery of recombinants for markers located distal to the origin of transfer on an Hfr are less than those for markers located closer.

This permits the construction of a temporal genetic map, depicted in figure 10. It is known that within 100 minutes the entire bacterial chromosome can be transferred, although at a very low efficiency (given the occurrence of spontaneous mating-pair disruption). Hence, if a hypothetical gene “X” is located clockwise with respect to the *pro*⁺ marker and lies between the *pro*⁺ and the *trp*⁺ markers in the donor strain Hfr described in figure 10, then transfer of gene X will occur after the transfer of the *pro*⁺ marker and before transfer of the *trp*⁺ marker. Strange as it might sound, as new genes were being discovered they were assigned map positions based upon the time it took for their transfer! Using Hfrs located at different positions on the genome, and based on the knowledge of the markers transferred early by each Hfr it became possible to establish that the *E. coli* genome is circular. Over a period of time, the map has been progressively refined and today the positions of all *E. coli* genes (there are around 4000 of them) are known. Despite this seemingly encyclopedic compendium, if one were to ask a dedicated *E. coli* researcher “Where are the genes for arabinose operon located”? pat will come the reply “between 1 and 2 minutes”!! We end this section on Hfrs by posing a question, which is as follows:

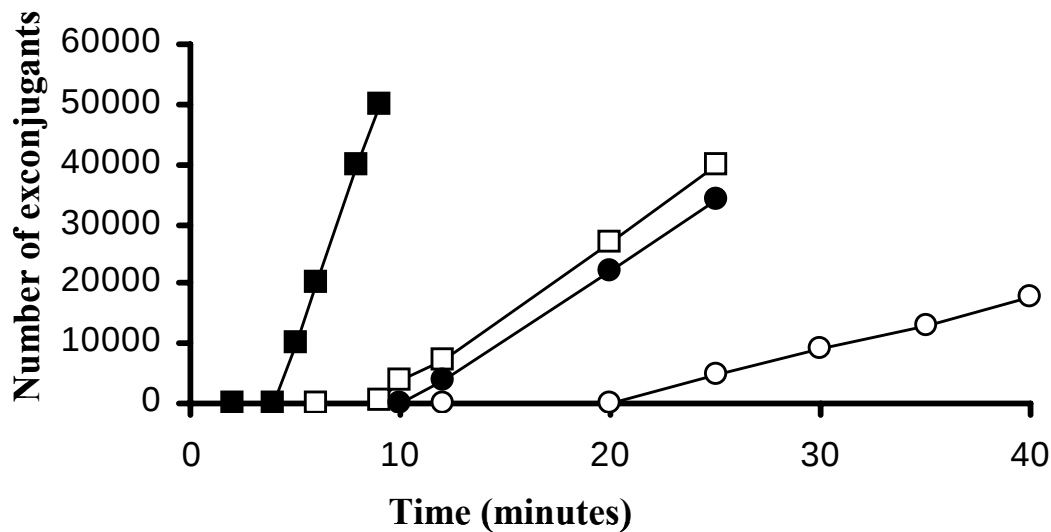


Figure 9: Graphical representation of the time of entry of various genetic markers employed in the interrupted mating experiment: In the experiment described in section 3.3 B an Hfr donor (strain A) mobilizes the *thr*⁺, *pro*⁺, *lac*⁺ and the *trp*⁺ markers into a recipient (strain B), which is *thr*⁻, *pro*⁻, *lac*⁻, and *trp*⁻ under conditions of interrupted mating. At intervals of 5 minutes aliquots of the conjugation mixture are withdrawn and subjected to vigorous agitation to terminate the mating process and plated on various selective media supplemented with streptomycin. The numbers of exconjugants so obtained in 0.1 ml of the conjugation mixture, are plotted as a function of time. Symbols: *thr*⁺ exconjugants (■); *lac*⁺ exconjugants (□); *pro*⁺ exconjugants (●) and *trp*⁺ exconjugants (○). Note: each marker appears in the exconjugants at a characteristic time.

It is known that when an F⁺ cell is mated to an F⁻ cell, the exconjugants are always F⁺. What is the F status of the exconjugants derived from a mating between an Hfr and an F⁻ cell? The answer is that almost all exconjugants will be F⁻. Notice (Figure 7), that in the integrated state of F (an Hfr), the *oriT* lies almost in the middle of the F factor. Since the transfer process is unidirectional, the region behind *oriT* will be transferred early in comparison to the region in front which will be the last to be transferred (comprising the *tra* region). If the latter happens this will lead to a complete transfer of an entire chromosome. Usually in a cross between an Hfr and an F⁻ cell, the probability that an entire chromosome will be transferred is finite but extremely low. Hence an Hfr, F⁻ mating will usually yield exconjugants which are F⁻ (see reference 11 for experiments with Hfrs).

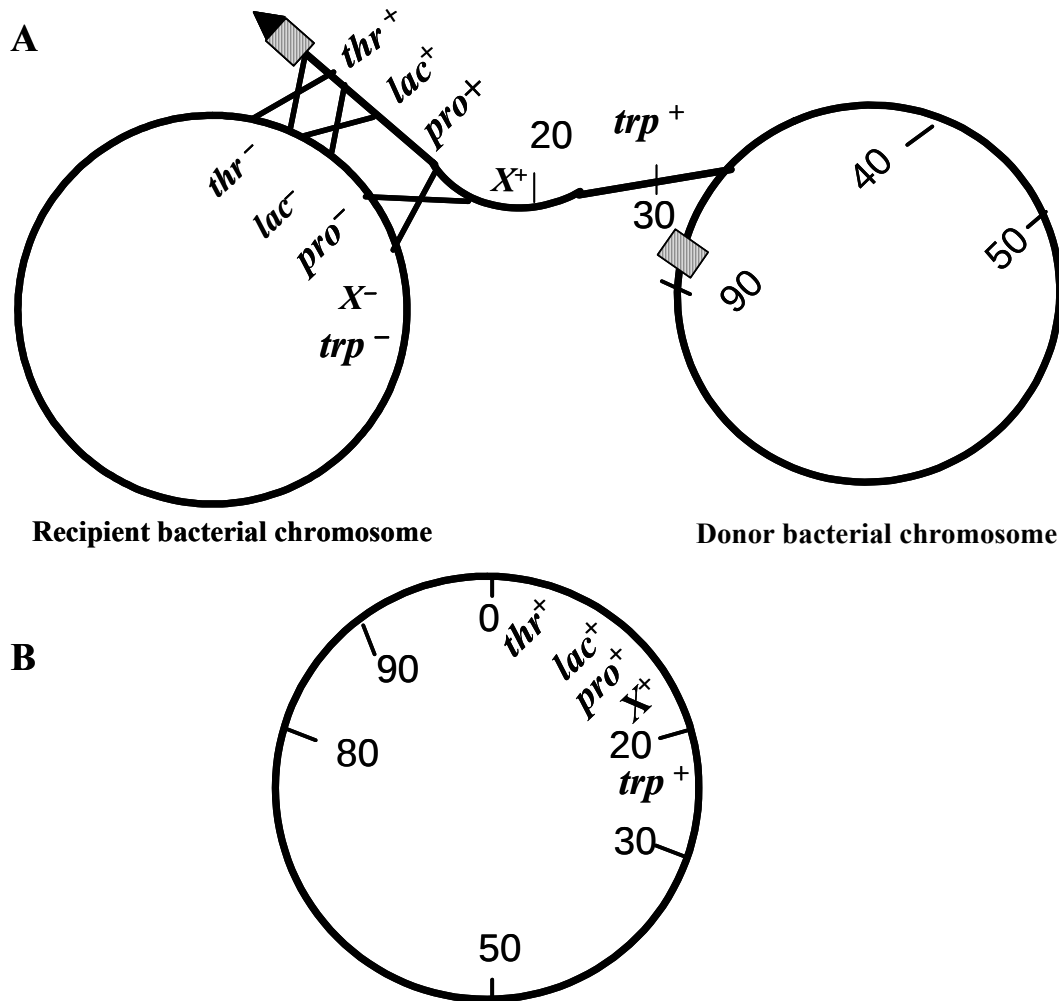


Figure 10: Construction of a gene map with the interrupted mating experiment: In the experiment described in section 3.3 B an Hfr donor, strain A (panel A) mobilizes the *thr*⁺, *pro*⁺, *lac*⁺, *X*⁺ and the *trp*⁺ markers into a recipient, strain B, which is *thr*⁻, *pro*⁻, *lac*⁻, *X*⁻ and *trp*⁻ under conditions of interrupted mating. The *thr*⁺ marker is the earliest to be transferred (within the first 5 minutes) while the transfer of the *lac*⁺ and *pro*⁺ genes takes place during 10 minutes. Transfer of gene X (*X*⁺) and *trp*⁺ will take approximately 20 and about 24 minutes respectively. This permits the construction of a preliminary gene map shown in panel B wherein genes are positioned based upon the time it takes for them to be transferred into a recipient cell.

3.4 Formation of F-prime (F')

In most cases, an F factor in the Hfr state can convert back to the autonomously replicating F⁺ state by the reversal of the mode by which it integrates into the bacterial chromosome. It may be recalled that the Campbell mode of integration leads to the duplication of the sequence used by the F factor to get into the bacterial chromosome such that this sequence flanks the integrated F

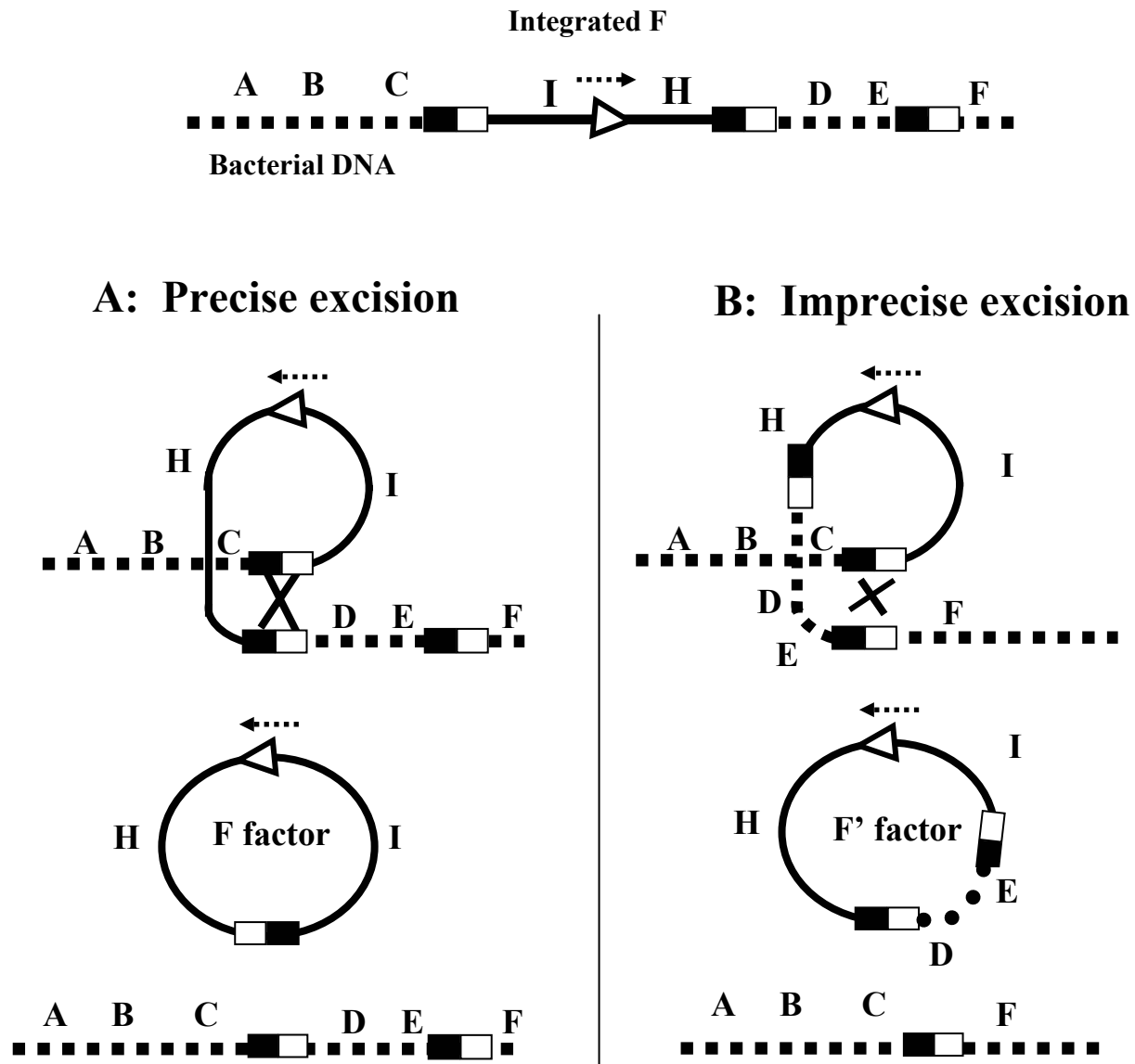


Figure 11: Formation of an F' factor: Precise (panel A) and imprecise excisions (panel B) of an integrated F factor result in the formation of F and F' factors respectively. Precise excision results from recombination (solid cross) between directly repeated copies of the insertion sequence (black and white box). F' formation occurs via illegitimate recombination (solid cross) between an insertion sequence on F and a similar insertion sequence in the adjoining bacterial chromosomal DNA, located between markers E and F. Here the imprecisely excised F now carries the bacterial genes D and E which are lost on the bacterial chromosome.

and is arranged in an order termed as direct repeats. It is apparent (Figure 11 A) that a correct realignment of the two repeats will promote reversal of the integrated state of F to that of an autonomously replicating one, a process that again requires the action of the host cell enzyme RecA. When this happens the Hfr is said to have undergone a precise excision. Occasionally the F can excise in an imprecise fashion, a product of illegitimate recombination between sequences

on F and those on the adjacent bacterial DNA. This is a rare event, with the outcome being that the excised F factor now carries a part of the adjacent bacterial chromosomal DNA. Such an F factor is designated as an F' (F-prime). Note that the excision event in the bacterial chromosome can occur on either side of an integrated F. Often, such, imprecise excision occurs by recombination between one IS element on the F and another similar IS element located in the adjacent bacterial DNA (which was not involved in the original F integration event- Figure 11 B). An F' is capable of transferring itself in a manner identical to that displayed by an F factor except that now the limited set of bacterial genes covalently associated with the F sequence are also transferred at high frequency. In colloquial terms, in an F', adjoining bacterial genes “piggy back” on the F and are transferred from cell to cell. An F'-donor cell can also “mobilize” at a lower frequency (still considerably higher than that of F alone), chromosomal genes in the same order and orientation as the ancestral Hfr from which the F' was derived - a process requiring functional RecA protein.

3.5 Can the process of conjugation be termed “sexual”?

When conjugation was discovered in bacteria scientists displayed a ready eagerness to term this process as “sexual transfer of DNA”. This was due to the fact that now haploid organisms like bacteria were seen to possess an innate capacity to transfer genetic material. Moreover the two types of bacterial cells involved in the process can be compared to different mating types (F/ Hfr/ F' and F⁻). However, in recent times, this notion has undergone a reappraisal. Strictly speaking (and literally so), in conjugation, there is actually a lateral (sideways) movement of DNA from a donor to a recipient. This has led to conjugation being bracketed as a category in a set of process collectively known as “horizontal/lateral gene transfer” (H/LGT). It is fairly obvious that since the F' and the Hfr states of the F can transfer genetic material of the host cell, in nature such transfer events (genetic exchange) may contribute to the appearance of organisms endowed with donor specific or recombinant genetic traits. Evidence for intra- as well as inter-species conjugal DNA transfers exists, that explains the generation of genetic variability in a bacterial population (reviewed in reference 16).

3.6 Why does a bacterial cell possess an F factor?

The F factor is a member of a large family of naturally- occurring episomes found in enterobacteria (bacteria residing in the human gut), with a mechanism of DNA transfer, which is “F-like”. Many of these are capable of causing disease and the F-like episomes in them have been known to possess genes conferring resistance to many antibiotics, as well as genes that are involved in production of toxins, haemolysins and surface antigens collectively known as virulence factors (see reference 18). Clearly these genes provide a bacterium with a “selective growth advantage” and an added mechanism, which aids in the establishment of disease against a constant backdrop of the human body's innate defense response. The F factor of the harmless laboratory strain of *E. coli* (*E. coli* K-12) does not seem to provide any obvious selective growth advantage, but *E. coli* K-12 strains bearing F are known to be resistant to infection with a virus, bacteriophage T7, a property contributed by genes on F. Another feature of F (and other F-like episomes), is that F bears a mechanism which ensures its continuous propagation. This means that as F bearing cells divide the occasional daughter cells, which do not receive a copy of the F episome, die (more dramatically stated, “commit suicide”), a phenomenon also known as plasmid addiction. This behavior is attributed to a pair of gene products encoded by F, which

mediate the killing of the cell lacking the F. Since a mechanism like this exists in the F episome of *E. coli* K-12, it may in itself be a sufficient explanation for the question raised in the title of this subsection: that is, the reason why the bacterial cell has an F factor is not because the latter confers any selective growth advantage but simply because it functions as “selfish DNA” which kills any daughter cell that fails to inherit it. Alternatively, of course, (and as stated above) the advantage conferred by F in evolution may be that of imparting sexuality and opportunities for lateral gene transfer to bacterial cells.

4.0 Transduction of DNA

Transduction is another mechanism by which DNA is transferred from one bacterium to another. As seen in the case of conjugation, the terms donor and recipient retain their usual meanings. Transduction is mediated by (bacteriophages) phages that are viruses, which infect bacteria. Within the bacterial cytoplasm, phages undergo replication, become mature viral particles via encapsulation of viral DNA by viral coat proteins, lyse (break open) the host cell and become free to infect another cell. It is during this vegetative growth that a phage can pick up bacterial genes from one cell (donor) and transfer them to another cell (recipient). Such a phenomenon is known as transduction. Transduction comes in two forms: **generalized transduction**, in which any region of the host cell genome can be transferred, and **specialized transduction**, in which a **limited** but a **specific part** of the host cell genome is transferred.

4.1 Generalized transduction

The process of generalized transduction is typified by phage P1, which infects *E. coli* and phage P22, which infects *Salmonella typhimurium* (*Salmonella enterica* serovar Typhimurium). The P1 virus particle (encapsulating the viral DNA) initially adsorbs to the outer membrane of *E. coli*. As mentioned above, the outer membrane of *E. coli* is decorated with LPS and P1, uses the LPS as receptor to adsorb to the cell surface. Inside the virus particle, viral DNA is present as a linear double-stranded molecule of an approximate size of 100 kb. In the post-adsorption phase, the viral DNA is injected into the cell. Once the viral DNA enters the cell, it circularizes. The unique feature of the genome of phage P1 and many other phages is that the ends of the P1 linear DNA are usually identical, a feature known as “terminal redundancy”. This feature arises, as we shall discuss later because of the way in which P1 DNA gets packaged. Essentially it means that the two terminally redundant ends of P1 are believed to recombine, converting the linear P1 DNA into a circular entity within the host cell.

P1 is a “temperate phage” and can exist in two physiological states. In one state, it remains as a passive extrachromosomal element, more accurately described as the P1 plasmid. In this state, the cell is said to be a P1 “lysogen”. In the second state P1 DNA undergoes the rolling-circle mode of replication and enters the “lytic state” (Figure 12). In the lytic phase, initially P1 DNA is replicated and then proteins comprising the P1 coat are made. The viral DNA is encapsulated in the P1 coat in a fairly elaborate process of packaging. All along the progression of the P1 life cycle, host cell growth ceases mostly due to the effects of P1- encoded proteins. After the packaging events, the host cell is broken apart and a new population of P1 phage particles is released into the cell exterior.

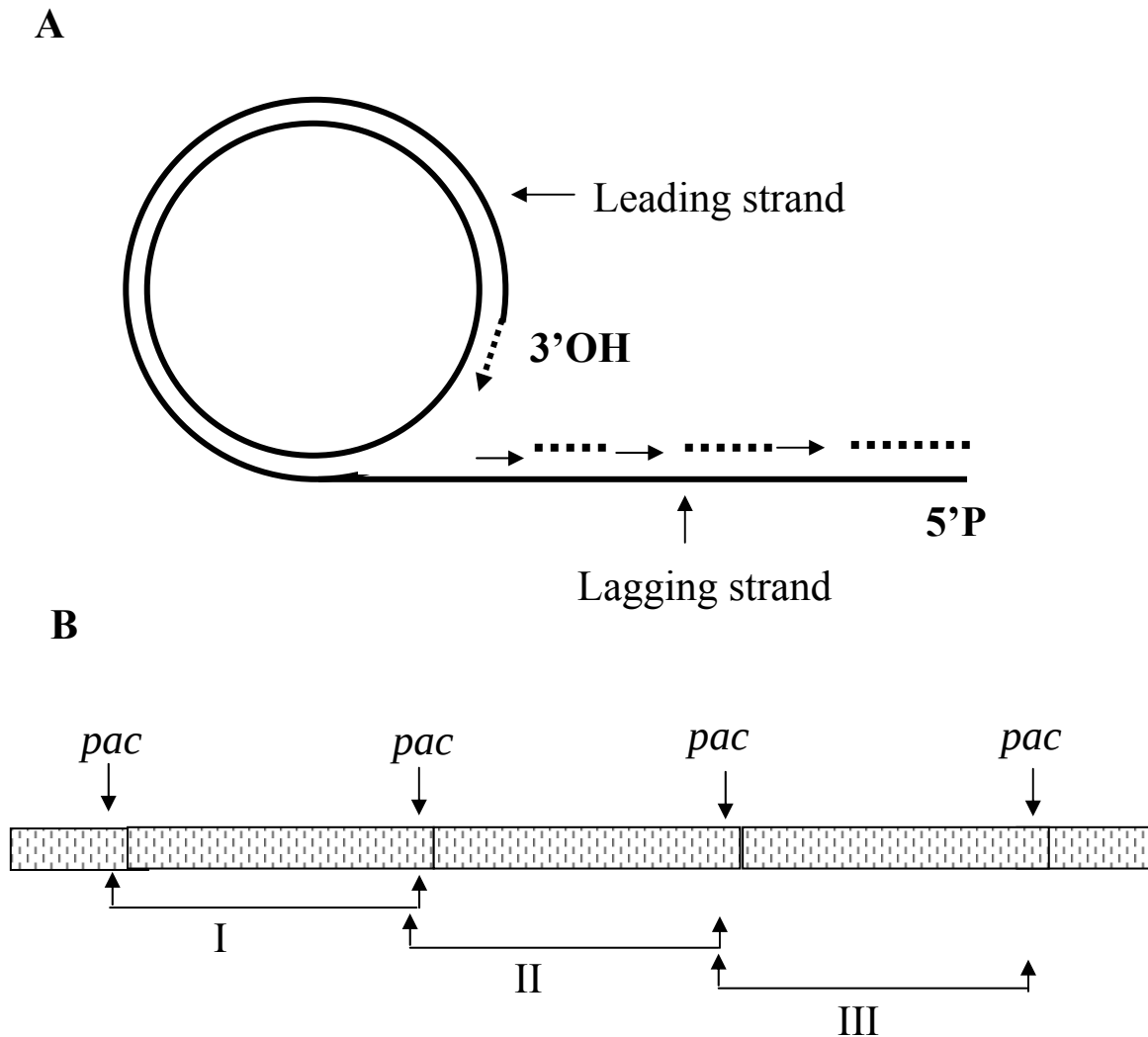


Figure 12: P1 DNA replication and packaging: The rolling circle mode of P1 phage DNA replication (panel A) leads to the formation of a linear P1 DNA concatemer (panel B) in which P1 DNA molecules are joined in a head to tail fashion. The leading and the lagging strand DNA syntheses in rolling circle replication are indicated. In panel B three individual units (I, II and III) approximately 100 kb in size, of the concatemer are shown. The left and the right regions of the concatemer represent, respectively portions of the preceding and succeeding units of the DNA concatemer. Each individual genome unit for headful packaging are indicated.

The process of packaging of P1 DNA is relevant to the phenomenon of generalized transduction. During the rolling circle replication phase of the P1 lytic cycle, newly replicated P1 DNA, from its circular template, is “concatemeric” in nature meaning that new copies of P1 DNA are arranged tandemly in a covalently joined “head to tail” fashion (Figure 12). The packaging machinery of P1 recognizes a particular sequence on this DNA concatemer known as the “*pac*” site. One copy of *pac* site is present in the P1 DNA. Packaging begins in the region around a *pac* site and follows a “**headful mechanism**” in which around 100 kb of the concatemeric substrate

is packaged and the process is repeated for packaging of successive phage heads. This sets the limit to the amount of P1 DNA which can be packaged in the interior of the phage head and hence the term “**headful packaging**”. Since the unit P1 genome size is somewhat smaller than the packaged size, the net result is that a linear double stranded P1 DNA molecule with terminally redundant ends is internalized within the P1 coat.

At a much lower frequency (about 1:1000), the packaging machinery of P1 may turn its attention onto the bacterial chromosomal DNA and begins to package the host DNA in the headful packaging way. It is thought that this probably is an outcome of nonspecific recognition of bacterial DNA as a substrate by the P1 packaging machinery. It is by these transducing particles in which random fragments of the host chromosomal DNA are packaged that generalized transduction is achieved.

When P1 is added to a host bacterial culture, the culture initially grows but with phage multiplication it gradually begins to lyse. The clear supernatant derived from this culture is known as a P1 lysate. A vast majority of the phage particles in the P1 lysate contain P1 DNA (mature phages) but 0.01% of such particles contain *E. coli* DNA. In such phage particles around 100 kb of contiguous segments of *E. coli* DNA are packaged, one discrete segment per phage particle. Usually a P1 lysate contains 10^9 phage particles per ml of which around 10^5 particles contain segments of host cell DNA. As the bacterial DNA packaging is random or non-specific, it may result in the entire host cell genome being represented in such “transducing” phage particles. Phage particles bearing *E. coli* DNA instead of P1 DNA are actually “crippled”, incapable of further multiplication. Since their phage coats are identical to that of the mature P1 particles they are capable of performing the initial stages of P1 infection, namely adsorption and injection of host cell DNA. However their lytic cycle is terminated because of a total absence of P1 phage DNA. In effect, they serve as carriers of bacterial DNA, which is derived from the host cell (donor) on which the P1 phage was propagated. The DNA released by such transducing particles can in another cell (recipient) recombine at the homologous region resulting in generalized transduction. For this to happen the recipient must bear the enzyme RecA, which acts as a catalyst in promoting homologous recombination.

Generalized transduction is a valuable tool in delineating linkage relationships between genes, a process referred to as transductional mapping. An example of transductional mapping is described (Figure 13). Consider two strains A and B. Strain A bears an insertion of an antibiotic resistance element conferring resistance to the antibiotic tetracycline, in the *lacZ* gene coding for the enzyme β -galactosidase. As a result the *lacZ* gene is disrupted (rendered non-functional). The disruption does not affect the ability of strain A to grow on minimal medium containing glucose as the sole carbon source. Strain B, on the other hand, has an intact *lacZ* gene but is unable to biosynthesize proline because it bears a mutation in the gene *proC*, which codes for an enzyme involved in proline biosynthesis. Thus, strain B is incapable of growing in minimal medium without proline supplementation. The genotype of strain A is thus *lacZ*⁻ (*lacZ*::Tet) and *proC*⁺, and that of strain B is *lacZ*⁺ *proC*⁻. Both the strains are able to grow on standard nutrient medium as all their nutritional requirements are satisfied. If one wants to find out how close are the *lacZ* and the *proC* genes or, in other words the linkage between *lacZ* and the *proC* genes, transduction can be employed. In a typical experiment one propagates P1 phage on strain A (donor) and uses the phage lysate to transduce strain B (recipient) at a multiplicity of infection (MOI, that is phage : cell ratio) of around 0.1. A vast majority of phages in the P1 lysate are “normal” phages, which

can mediate the killing of recipient cells. To prevent the phage particles released, from re-infecting and killing those cells which have received transducing phage particles, special measures are taken such as allowing phage adsorption for a limited period of time and removing metal ions (required as cofactors for infection) immediately after phage adsorption. Alternatively MOI is kept low so that a cell can either be infected by a transducing phage or a normal (virulent) phage but not both. The transduced cells are allowed to grow on standard nutrient medium containing tetracycline. By employing tetracycline (Tet)-containing medium, only those cells of strain B will grow which received from phage P1 packaged bacterial DNA of strain A from the region containing and surrounding the *lacZ* disruption, followed by recombination of the disrupted *lacZ* allele (*lacZ::Tet*) and the adjoining DNA, into the chromosome of the recipient strain. The *lacZ::Tet* marker acts as a **selectable marker**. In such tetracycline resistant transductants the original *lacZ*⁺ gene plus some amount of adjoining DNA of the recipient will now have been replaced with the corresponding region (*lacZ::Tet*) from the donor strain A. Note that other bacterial cells of strain B will receive different regions of strain A's chromosomal DNA, because the P1 lysate contains phage particles bearing different bits of host (donor) strain DNA. These however will be sensitive to tetracycline and hence will not form colonies on Tet-containing media.

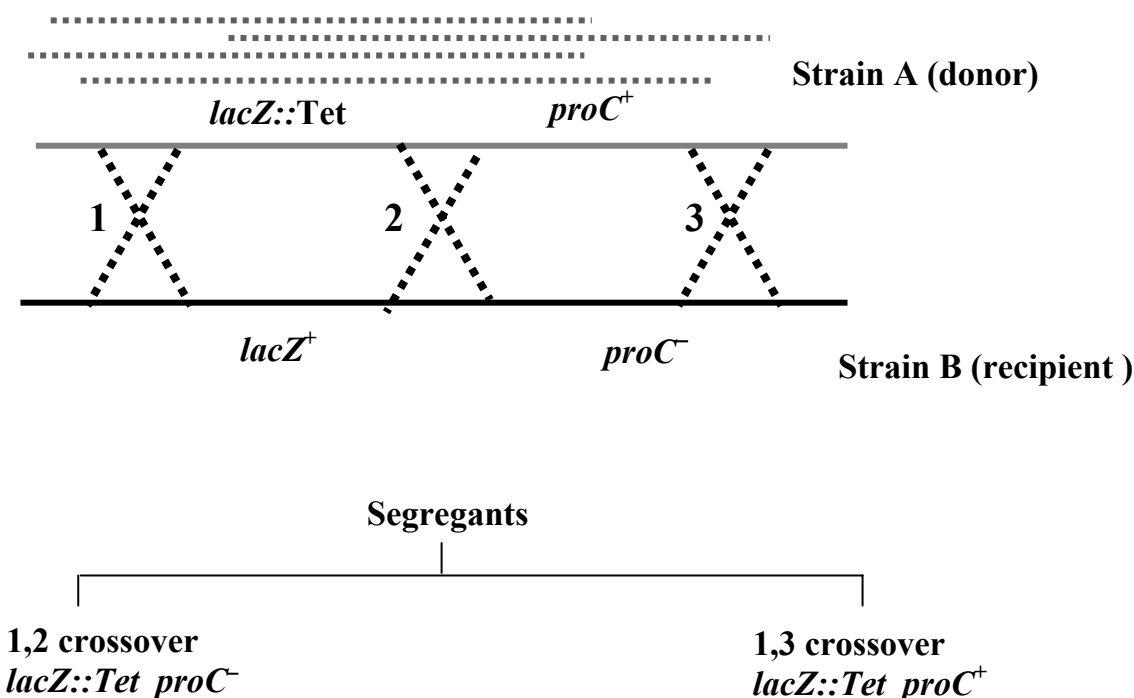


Figure 13: Linkage analysis by P1 mediated generalized transduction: A P1 lysate prepared on strain A (donor- *lacZ::Tet proC*⁺) is used to infect strain B (recipient- *lacZ*⁺ *proC*⁻), with a primary selection for tetracycline resistant (Tetr) colonies. Depending upon the distance between the *lac* and the *pro* genes, the Tetr transductants (segregants) will be of two genotypes namely *lacZ::Tet , proC*⁺ (products of 1, 3 crossover) or *lacZ::Tet , proC*⁻ (products of 1, 2 crossover). If the *lac* and the *pro* genes are very close to each other then the proportion of *lacZ::Tet proC*⁺ transductants will be high amongst the tetracycline resistant colonies. During the preparation of

the P1 lysate of strain A, overlapping segments of the *lac* (*lacZ::Tet*) region are packaged, indicated as interrupted horizontal lines.

Individual Tet resistant colonies (Tet^r), of strain B are now picked up and tested for their ability to grow on minimal medium without any proline supplementation. This is a **screening** procedure. One may ask the question, “What percentage of the *lacZ::Tet* bearing colonies have also inherited the *proC*⁺ gene of the donor strain A”? It is apparent that the closer the *proC*⁺ gene to the *lacZ::Tet* in the donor strain, the greater is the probability that the *proC*⁺ gene will be co-inherited (co-transduced) along with the *lacZ::Tet* allele. In such a case, the majority of Tet^r transductants will also have become *proC*⁺, that is regained the ability to biosynthesize proline. If the *proC*⁺ gene is extremely close to the *lacZ::Tet* allele in the donor, then transductants (segregants) of only one genotype will be recovered namely *lacZ::Tet* and *proC*⁺. However if the *proC*⁺ marker is located some distance away from the *lacZ::Tet* marker then, transductants with the genotype of *lacZ::Tet* and *proC*⁻ will also be recovered (Figure 13).

The co-transduction frequency of the *lac* (*lacZ::Tet*) and the *proC*⁺ alleles is obtained by dividing the number of Tet^r transductants that have also become *proC*⁺, by the total number of Tet^r transductants. By multiplying the co-transduction frequency by 100, one obtains the percent co-transduction between two markers. For example, if, in the above mentioned experiment the total number of Tet^r transductants obtained is 180, of which 27 transductants are of the genotype *lacZ::Tet*, *proC*⁺ (the rest being *lacZ::Tet*, *proC*⁻), then the co-transduction frequency of the *lac* and the *proC* markers is 0.15. In other words the *lac* and the *proC* markers are linked but not very closely and are 15% co-transducible. Based on the co-transduction frequency between the two markers one can estimate the genetic distance between two markers, which in turn is based on the Wu formula. It is known that a generalized transducing P1 particle can accommodate approximately 100 kb equivalent bacterial chromosomal DNA, corresponding to a genetic distance (length, L) of 2 minutes. The Wu formula for estimating the distance (d) between two markers, given the length of the packaged chromosomal DNA (L = 2) and the co-transduction frequency (F) between the said markers is as follows:

$$F = (1 - d/L)^3 \text{ or } d = L [1 - (F)^{1/3}]$$

Using the value of the co-transduction frequency, of the *lac* and the *proC* markers, obtained in the above-mentioned experiment, which is 0.15 and applying the Wu formula, one arrives at an estimate of the genetic distance between the two markers, which is 0.94 minutes. As mentioned above, P1 generalized transducing phage contain 2 minutes equivalent of chromosomal DNA. Thus the obtained value of 0.94 of the genetic distance between the *lac* and the *proC* markers corresponds to a physical distance of 47 kb. It is apparent, that if two markers are very closely linked such that is their co-transduction frequency (F) is 1, then the genetic distance (d) obtained from the Wu formula attains a value of 0, an indicator of very close physical association between the two markers. Conversely, if two markers are unlinked, that is F equals 0, then d attains a value of 2, meaning that the genetic distance between the said markers is greater than 2 minutes and thus the two markers cannot be co-transduced. Thus in all estimations of genetic distances between a given pair of markers the value of d will range between the two extremes namely 0 and 2.

It is worth noting that since in generalized transduction the incoming linear DNA recombines into the circular chromosome of the recipient strain, an even number of exchanges (depicted as hatched crosses in figure 13) will result in the inheritance of the transducing fragment with the concomitant replacement of a homologous region of the recipient genome. Odd number of exchanges can and will occur as well but in such cases no transductants will be recovered as these will result in the breakage of the circular bacterial chromosome which is lethal. Finally, another point to note is that when P1 is packaging the *lacZ::Tet* bearing region (or any other region) of the donor strain A, a series of overlapping packaging reactions occur, resulting in discrete but overlapping fragments of the said region being packaged within individual phage coats (depicted as interrupted horizontal lines in figure 13). A more detailed description on deriving gene linkage relationships via generalized transduction can be obtained from reference 11.

4.2 Specialized transduction

The process of specialized transduction is also mediated by a bacteriophage but as the name suggests, a limited but a specific region of the bacterial chromosomal is transduced. Bacteriophage lambda is the prototype mediator of this process. Lambda, like P1 is a temperate phage with a genome size of approximately 50 kb. Infection of *E. coli* by lambda is achieved by adsorption of the phage to an outer membrane host protein, LamB. This is followed by the release of lambda DNA into the host cell cytoplasm. Similar to phage P1 lambda can exist in two mutually exclusive physiological states namely, the lytic and the lysogenic states. More literature is available on the elaborate (yet elegant) phage encoded regulatory circuits, which govern the establishment, maintenance and interconversions between the two states (reviewed in references 9, 10, 13 and 15). Phage lambda in the lysogenic state exists in an integrated state at a specific site or locus on the circular bacterial chromosome, called *attB* or the bacterial (phage) attachment site. When phage lambda injects its DNA into the cell, its linear DNA circularizes. The circularization is achieved by annealing of the two ends of the phage linear DNA which are single-stranded protrusions and are complementary to each other to reconstitute the *cos* (acronym for cohesive) site (Figure 14). The circular lambda genome serves as a template for its replication, which occurs to begin with in a bi-directional θ (theta) mode but converts later to the rolling circle mode. Concatemers of lambda DNA are formed and the lambda packaging machinery packages precisely DNA from one *cos* site to the next. In the circular topological state, lambda can integrate following the Campbell scheme of homologous recombination into the *attB* site on the chromosome. Lambda bears a DNA sequence, *attP*, which is homologous to a region in *attB*, and using this sequence it recombines into the *attB* site on the *E. coli* chromosome. Lambda codes for an enzyme Int, which promotes recombination between the *attP* and the *attB* sequences utilizing, in addition, the services of a host protein IHF (Figure 14). Lambda therefore displays the property of “site-specific recombination” and *attB* is by far the most preferred place in the bacterial chromosome where lambda integrates. In the integrated state, all the functions required for manifestation of the lytic state are repressed, the phage resides in the genome as a neutral entity referred as prophage, and a bacterial strain with an integrated lambda is said to be a “lambda lysogen”. The *attB* site on the bacterial chromosome is located between genes encoding the products required for galactose catabolism (*gal*) and those involved in promoting biosynthesis of the growth factor biotin (*bio*).

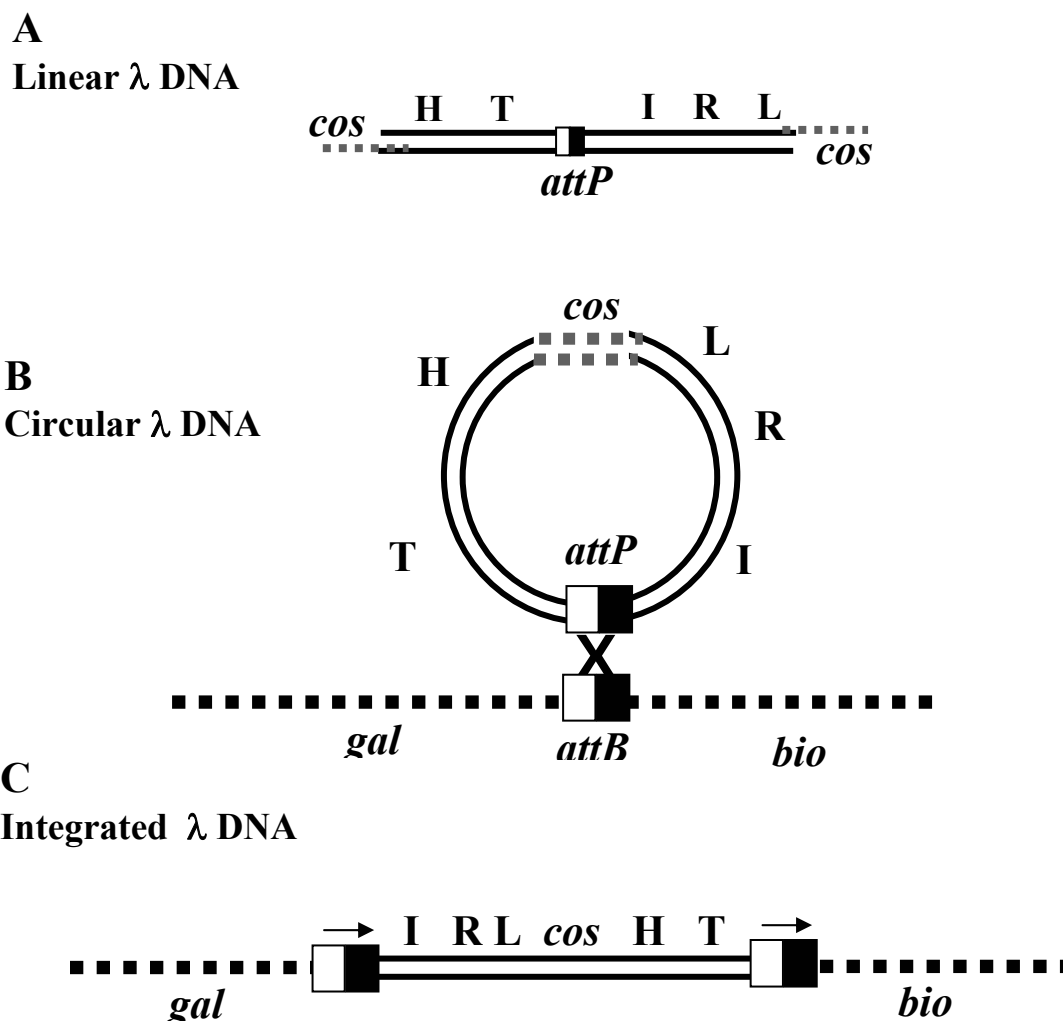


Figure 14: Site specific recombination of bacteriophage Lambda: Linear lambda DNA (panel A), inside the cell circularizes via *cos* site homology (panel B). Circular lambda DNA integrates using its *attP* sequence (black and white box), in a site specific manner, within the *attB* site located in the bacterial DNA (interrupted line), between the *gal* and the *bio* genes, to form a lambda lysogen (panel C). I, R, L and H represent lambda genes. Notice the circular permutation of phage genes in the integrated state.

Depending upon the physiological state of the cell, a lambda lysogen can excise from its location within the *attB* sequence on the chromosome, getting converted into a circular form with the derepression of lytic state functions. Similar to an Hfr to F conversion, this process is a simple reversal of the Campbell mode of integration and requires the assistance of two phage- encoded proteins namely, Int and Xis. More often than not, the excision is precise but at a very low frequency, imprecise excision events may also occur. These are products of illegitimate recombination between sequences on lambda and those on the adjacent bacterial DNA, and as a result imprecisely excised lambda phage particles carry either the *gal* genes (λgal) or the *bio* genes (λbio) (Figure 15).

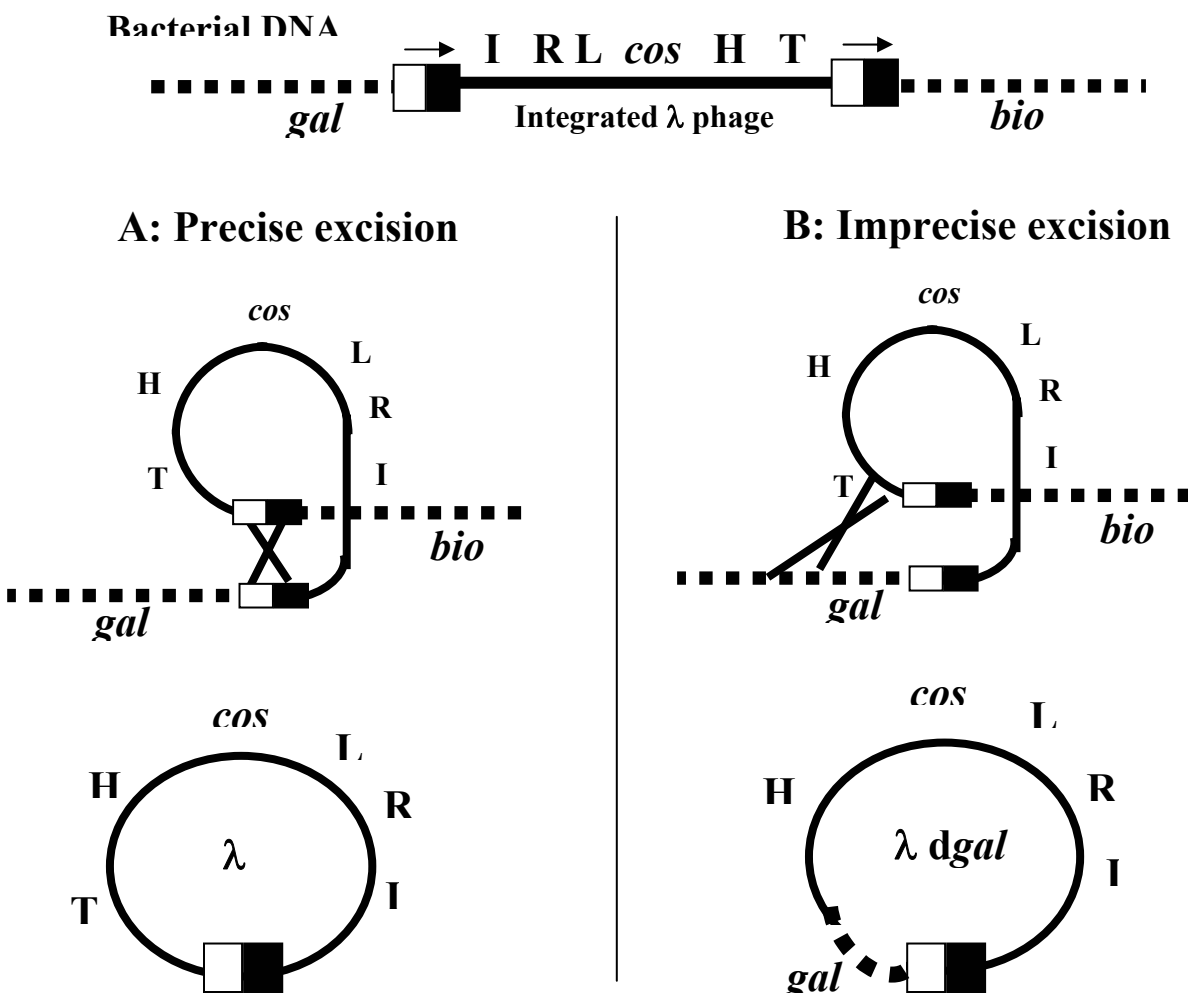


Figure 15: Formation of specialized transducing lambda phage: Precise (panel A) and imprecise (panel B) excisions of an integrated lambda phage result in the formation of mature and specialized transducing phage respectively. Precise excisions are an outcome of recombination (solid cross) between directly repeated *att* sites (black and white boxes). Illegitimate recombination (solid cross) between phage and adjoining bacterial DNA sequences (interrupted line) leads to formation of specialized transducing phages. Here the product of imprecise excision carries the genes of the galactose operon of *E. coli*. Note: that in order to compensate for the gain of *gal* genes λ will lose a portion of its genome from the opposite end, indicated by the absence of the region marked T on λ *dgal*.

When excision by specific chemicals or environmental conditions is induced in a culture of a lambda lysogen, the growth rate of the culture begins to decline until the culture lyses and a population of lambda particles are released to the cell exterior. The clear supernatant contains a vast majority of phage lambda particles bearing precisely excised lambda DNA. However, a low proportion of lambda particles may contain imprecisely excised lambda DNA bearing some part of the adjoining bacterial DNA, either the *gal* or the *bio* genes. These phages are referred to as

specialized transducing phages because when they lysogenize recipient strains they mediate the transfer of either the *gal* or the *bio* genes.

A slightly more detailed consideration on the process of imprecise excision is presented here. If one takes a look at the process by which lambda integrates into the bacterial genome, then the following observations can be made:

1. Inside the cell, linear lambda DNA circularizes as its two termini reanneal, regenerating a site known as the *cos* site. This site is analogous to the P1 packaging site, *pac* and constitutes the nucleotide sequence from where packaging of lambda DNA is initiated (Figure 14).
2. Notice, the circular disposition of the hypothetical lambda genes H, I, L, R, and T in figure 14, panel B.
3. Integration of lambda occurs utilizing the *attP* site.
4. In the integrated state, notice now that the order of lambda genes is a circular permutation of the order seen in the circular form of lambda (Figure 14, panel C).

It is known that on the lambda genome regions can be identified which are essential or non-essential for lytic growth. In a lambda lysogen, because of the phenomenon of circular permutation of gene order, these regions get repositioned with reference to the bacterial DNA. Examination of the prophage (integrated lambda) map shows that when imprecise excision takes place, viral DNA is always lost from the end opposite the newly acquired bacterial DNA, a feature imposed by the packaging limit of approximately 50 kb for lambda. If illegitimate recombination occurs, for example, between sequences on the lambda and those located towards the *gal* side and the lambda *cos* site is retained then this product will retain all properties to enable it to be packaged within the lambda coat proteins to produce viral particles. Any excision event not retaining the *cos* site will not produce specialized transducing phage, simply because the excised DNA will not be packaged. Notice that specialized transducing phage retaining the *cos* site will also retain an intact copy of the *attP/B* sequence, thereby enabling this phage to be proficient in lysogenizing a suitable recipient strain. However such λ *gal* phages will lack tail functions, whose absence renders these phages incapable of entering the lytic phase. Such phages are proficient for lysogeny but deficient for lysis, hence designated as λ *dgal*, the prefix “d” standing for “defective”. This is depicted in figure 15 wherein the λ *dgal* phage is shown missing the lambda region marked T. Similar excision events can generate λ *bio*, which will lack lambda genes involved in recombination and may also be those involved in establishment of the lysogenic state (lambda regions I and R depicted in figures 14 and 15). Lambda *bio* phages are proficient for establishment of the lytic stage but are unable to enter into the lysogenic state. Because they retain all the functions required for lytic growth λ *bio* phages are not designated as “defective”. Packaging of lambda DNA, like that of P1 DNA follows the headful mechanism but with the size limit of approximately 50 kb. Thus there is a limit to the amount of flanking DNA, which can be packaged. Such imprecise excision events, therefore lead to formation of lambda phages that are missing some essential phage encoded functions, a feature imposed by the packaging limit. Specialized transducing phage so generated can act as vehicles for transfer of bacterial genes (*gal* or *bio*) located adjacent to an integrated lambda in a lysogen. Their multiplication, however, will be possible only if they are infected along with a normal (helper) lambda phage into a suitable recipient. The helper lambda phage provides the necessary functions that an imprecisely excised specialized transducing phage lacks.

In a phage lysate obtained by inducing a culture of a lambda lysogen (*gal*⁺ donor strain), the presence of specialized transducing phage can be inferred as follows. Typically, a strain (recipient strain) defective for galactose catabolism (*gal*⁻) is infected with such a lysate and a selection for ability to grow on minimal plates supplemented with galactose as the only carbon source is imposed. Bacteria capable of growing on this medium are those that have been lysogenized with λ *gal* specialized transducing phages bearing intact galactose catabolism genes from the donor strain. It can then be shown that the ability to utilize galactose as a carbon source, displayed by the colonies so recovered is due to the presence of the λ *dgal* phage.

5.0 Transformation

Genetic exchange between bacteria can also be mediated by the process of transformation. The phenomenon of transformation was first discovered by Griffith in pneumococci in the 1920s, and was later shown in mid-1940 (through the work of Avery, Macleod and McCarty) to be mediated by DNA. Using two strains of *Pneumococcus* (*Streptococcus pneumoniae*), that differed in their ability to cause disease (virulence) in mice, these workers established that DNA from a virulent strain, could be introduced into an avirulent strain imparting upon the latter the property to cause disease in mice. The avirulent strain was said to be “transformed” from a benign state to a virulent state. This experiment was also the forerunner in demonstrating that the genetic material of a cell is DNA.

There are two routes to achieve transformation in bacteria: (i) artificial (chemical induced) and (ii) natural. In the experiments of Avery and coworkers, the recipient bacterium had an innate ability to become “competent” to undergo transformation, depending upon its physiological state. It was noticed that pneumococcal cultures reached a growth stage where they became proficient in uptake of exogenously supplied DNA. Such bacteria are said to be naturally competent and the transformation process is governed by the action of specific gene products present in the recipient cell. The biology of natural transformation is discussed below.

5.1 Natural transformation

The innate ability of bacteria to develop a state of “competence” such that they become capable of internalizing DNA is seen for many bacterial species. Bacteria such as *Bacillus subtilis*, *Haemophilus influenzae*, *Streptococcus pneumoniae*, and *Neisseria gonorrhoeae* display the ability of natural transformation. Probably the best studied example is that of *Bacillus subtilis* (*B. subtilis*), which is discussed below.

As a culture of *B. subtilis* grows and reaches a high cell density (stationary phase), two proteins namely ComX and CSF (competence and sporulation stimulating factor) are secreted by cells into the culture medium. These are small proteins (peptides) and are also designated as peptide pheromones. As a result, a subpopulation of *B. subtilis* cells respond to the effects of these pheromones and enter a physiological state of competence wherein they become receptive for uptake of exogenous DNA. The said pheromones activate an intracellular signaling system, comprising a set of enzymes leading to the activation of a protein ComK. ComK in turn stimulates the expression of genes coding for components of the DNA uptake machinery. It may

be mentioned that the activation of ComK is an elaborate process, but for the sake of brevity it suffices here to say that the ComK activation pathway acts by freeing ComK from proteins which form a complex with it and prevent it from exerting its stimulatory effects on competence gene expression (reviewed in reference 17). The cells are then rendered ready to uptake exogenous DNA.

The uptake machinery comprises of the bacterial membrane proteins ComEA, ComEC, ComFA and a set of proteins required for the biogenesis and assembly of “pilus like” proteins. The process of DNA uptake is initiated with the binding of double-stranded DNA to the outer surface of the cell, the peptidoglycan layer. A characteristic feature of this step is that the bound DNA is resistant to separative processes such as centrifugation. The obvious implication of this is that somewhat robust interactions occur between DNA and the outer surface of the cell. In the state of competence, *B. subtilis* elaborates the expression, synthesis and assembly of “pilus like” structures, which are associated with the cell wall. Strains defective in pilus formation are blocked in the early step of transformation, namely DNA association. It appears that pili via, an unknown mechanism, deliver the DNA to the cytoplasmic membrane where the components of the DNA transfer machinery reside.

The bound DNA then undergoes cleavage into shorter fragments, promoted by the cytoplasmic membrane-embedded protein NucA. A large part of the NucA protein appears to lie external to the cytoplasmic membrane and this appears consistent with its role as a DNA degrading enzyme. The cytoplasmic membrane lies internal to the peptidoglycan layer, and clearly some mechanism appears to exist which allows the initial entry of double-stranded DNA towards the cytoplasmic membrane. The clipped DNA is thought to interact with the cytoplasmic membrane protein ComEA, which has a known non-specific DNA binding activity. It is believed that the ComEA may transfer in some way its bound DNA to a large membrane protein ComEC. ComEC is thought to be capable of forming a pore through which DNA may pass. *B. subtilis* strains bearing mutated versions of either the *comEA* or the *comEC* genes display the common phenotype of inability to be transformed. In both *comEA* or *comEC* mutants, DNA binding occurs but bound DNA can be degraded by DNA digesting enzymes (nucleases) whereas in the wild type strain, bound DNA is resistant to the action of nucleases. These observations suggest that the ComEA and ComEC proteins participate in the later stages of DNA uptake, namely DNA entry into the cytoplasm. ComFA is another cytoplasmic protein known to participate in DNA uptake. It appears to be tethered to the membrane but a large fraction of it is located at the cytoplasmic face of the membrane. The ComFA amino acid sequence bears similarities to sequences found in proteins, which can translocate along DNA, namely DNA helicases. It is speculated that ComFA might be the protein involved in pulling the single stranded DNA as it enters the recipient cytoplasm (Figure 16).

At this point, things begin to get a little unclear and there is some uncertainty on the exact mechanism by which DNA is internalized into the cytoplasm. It is known that during DNA uptake, single-stranded DNA is found internalized and short oligonucleotides (bits of DNA) are found released in the medium. It appears somewhat implicit that one strand of DNA (the non-transforming strand) is degraded. How is this achieved? It is not quite clear how this happens but it is suspected that the ComEA, ComEC, and ComFA proteins at the cytoplasmic membrane may form a complex, which recruits an unknown protein(s) capable of DNA processing. The internalized DNA is thought to be bound initially by SSB (single-strand DNA-binding protein)

and then coated by the RecA enzyme, promoting its recombination at a homologous region on the bacterial genomic DNA (see references 2, 4 and 5 for reviews).

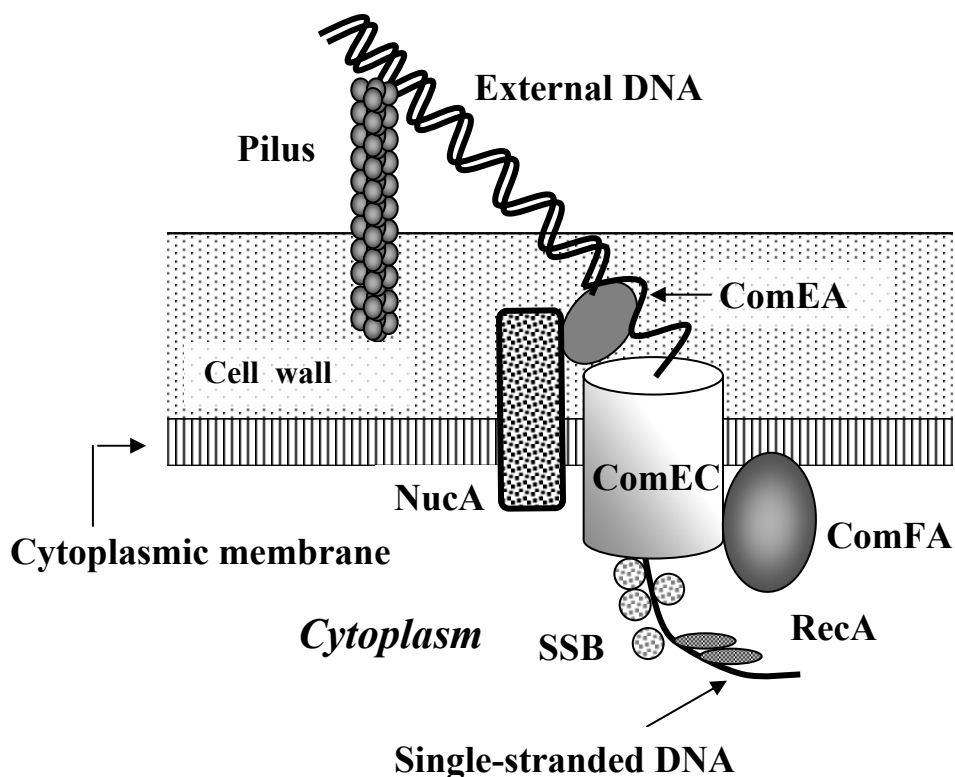


Figure 16: DNA transfer apparatus of *B. subtilis*: Figure depicting the proposed scheme for DNA uptake by *B. subtilis*. External DNA is brought to the cytoplasmic membrane wherein it is subjected to the action of components of the uptake machinery. NucA appears to be the nuclease responsible for the initial degradation of the incoming DNA. Conversion of the DNA from the double-stranded to the single-stranded (ssDNA) form and the degradation of the non transforming strand of DNA, occurs via an unknown mechanism and the ssDNA is thought to enter the cytoplasm through a channel formed by membrane located ComEC. ssDNA within the cytoplasm is presumed to be coated with SSB and RecA, prior to its integration in the recipient genome. (Figure adapted from reference 2)

Why does the phenomenon of natural transformation exist? It is speculated that this phenomenon, may allow bacteria a means to acquire genes from the environment so as to improve their “fitness”, but this is not further discussed here (reviewed in reference 5). It may suffice to say that fitness can be defined as an ability to sustain and propagate in a given environmental condition.

5.2 Chemical transformation

Bacteria such as *E. coli* and *Salmonella typhimurium*, though not naturally competent, can be rendered competent by treatments with chemicals, notably those bearing divalent cations. Such bacteria then become receptive for the uptake of plasmid DNA. Chemical transformation

constitutes one of the basic experiments in molecular biology and is an integral component of recombinant DNA technology. Competent *E. coli* can also take up linear DNA, but linear DNA is rapidly degraded inside the cell by the action of a DNA-digesting enzyme, the RecBCD nuclease. A linear piece of DNA bearing a gene of interest, however, can be covalently joined with a linearized plasmid, through enzymatic manipulations. The joining reactions are accomplished using an enzyme known as DNA ligase. The net result is that the gene of interest becomes a part of the plasmid and reconstitutes the circular form of the plasmid. The recombinant DNA molecules are mixed with *E. coli* cells that have been conditioned (made competent) to accommodate them, a step that is carried out at low (0 to 40°C) temperature. After a suitable time interval, the cells are exposed to a high temperature (42°C) pulse, and then allowed to recover by growing them in a growth medium. The cells are then plated on a medium, which permits growth of only the transformed cells (a selective medium). The heat pulsing step increases the efficiency of transformation. Since plasmids possess the property of autonomous replication, the gene of interest thus gets established as part of the plasmid (reviewed in references 14 and 15).

It may be said that while knowledge exists for making competent *E. coli* cells to take up DNA, the exact mechanisms by which plasmid DNA traverses the twin membranes of the envelope of *E. coli* are at best speculative. It is believed that the two membranes might not be completely separated from each other rather they might be fused in some regions, forming what are known as zones of adhesion. Furthermore, there does not seem to be any sort of transport entity mediating passive uptake of DNA in *E. coli*. This is an inferred notion because the size range of plasmids, which can be successfully transformed, is very large. Plasmids ranging from a few kb to 100 kb can be transformed though the transformation efficiency decreases with the increase in size of the plasmid. It, therefore, appears improbable that an uptake protein with such a diverse size range for its substrate can exist. But the method works!!

Despite the existing uncertainty on the mechanisms of chemical transformation, factors required for effecting efficient transformation have been recognized. Divalent cations, notably calcium ions and to some extent magnesium and manganese ions, are critical for efficient transformation. Transformation occurs with high efficiency if the recipient strain culture contains cells, which are still actively growing, prior to their exposure to the conditioning regimen. The conditioning of bacteria at low temperatures is also vital to the transformation process (for a detailed review on bacterial transformation see reference 7).

An alternative method of transformation, known as electroporation, is also available. If a mixture of competent *E. coli* cells and plasmid DNA are subjected to a brief pulse of high voltage, then plasmid DNA can enter and establish itself in the bacterial cytoplasm. The mechanism of such “electro-transformation” of DNA is again not well understood. A plausible explanation is that a transient exposure to a high-strength electric field seems to increase the porosity of the bacterial cell envelope facilitating thereby internalization of plasmid DNA. This is a very high- efficiency method that is used to transform *E. coli* and other gram-negative and gram-positive bacteria.

6.0 Integrative and conjugative elements (ICEs)

ICEs, formerly known as conjugative transposons, constitute a recently discovered class of DNA elements mediating genetic exchange. ICEs are somewhat chimerical in nature because they are endowed with properties, which are “phage lambda-like”, “transposons-like” and “F episome-like”. The element Tn916 from *Enterococcus faecalis* and CtnDOT from *Bacteroides thetaiotaomicron* represent the early examples of ICEs, but subsequent studies have shown that the occurrence of ICEs is very widespread. ICEs can exist in an integrated state on the bacterial chromosome. Integration is promoted by Int and Xis proteins, which share similarity with their counterparts from bacteriophage lambda. Tn916 shows a considerable amount of flexibility in its choice of the chromosomal target site for integration. Consequently the Tn916 element is found to be integrated at different sites on the chromosome (hence the designation Tn for transposable element). In scientific jargon, Tn916 is said to display “promiscuous” integration. The mechanism by which Tn916 integrates is somewhat beyond the scope of this chapter and is therefore not discussed (reviewed in reference 3). However, many ICEs use the Campbell mode of integration to achieve site-specific integration on the bacterial chromosome, and they share with the bacterial chromosome, common sequences which will facilitate integration. The lambda nomenclature is usually employed and the ICE and the chromosomal sequences participating in ICE integration are designated attP and attB, respectively (Figure 17).

ICEs integrate within a nonessential region on the chromosome. For example the SXT- ICE from *Vibrio cholerae* integrates at a sequence located at the 5' end of the prfC gene, accompanied by the expected duplication of the target site. Excision of the ICE usually occurs, following the lambda mechanism, leaving behind intact copies of attB on the chromosome and attP on the ICE. Once freed from its genomic location, the circular ICE is thought to mobilize itself in a process similar to F conjugation; in this regard, ICE displays “F-like” properties. ICEs are therefore capable of “self transmission”. Embedded within the ICE are genes specifying the synthesis of components of the mating machinery, such as pili and DNA transfer proteins. It is thought that generally, the prelude to DNA transfer is the excised circular form of the ICE. The SXT-ICE from *V. cholerae* can also act as an Hfr and can transfer chromosomal markers if it is made to integrate within the *E. coli* genome, suggesting that SXT element has the potential to transfer chromosomal markers of its natural host *V. cholerae* as well. The initiation of DNA transfer occurs in much the same way as the F DNA transfer is initiated (Figure 17). In some cases, oriT like sequences within the ICE have been recognized, and it is therefore thought that the transfer of an ICE may follow a mechanism similar to F factor transfer in conjugation.

Perhaps the most striking feature of ICEs is that they endow the host cell with new gene functions (reviewed in reference 1). Many ICEs bear genes conferring resistance to different antibiotics. Take the case of the SXT element, which contains determinants conferring resistance to sulfamethoxazole, streptomycin, chloramphenicol and trimethoprim, clearly presenting a major impediment in treating *V. cholerae* infection. Some ICE genes elaborate virulence factors thereby aiding a pathogen in its establishment in a host cell, while others render the bacterium capable of tolerating extreme environments. As an example of the latter instance, the clc ICE of *Pseudomonas* sp strain B13 carries a cluster of genes namely the clcRABDE genes, imparting a property of chlorocatechol degradation, upon the bacterium. The clc element chooses a sequence within the gene coding for the glycine tRNA, a gene whose disruption by the clc element does not lead to deleterious effects on cell growth, for integration.

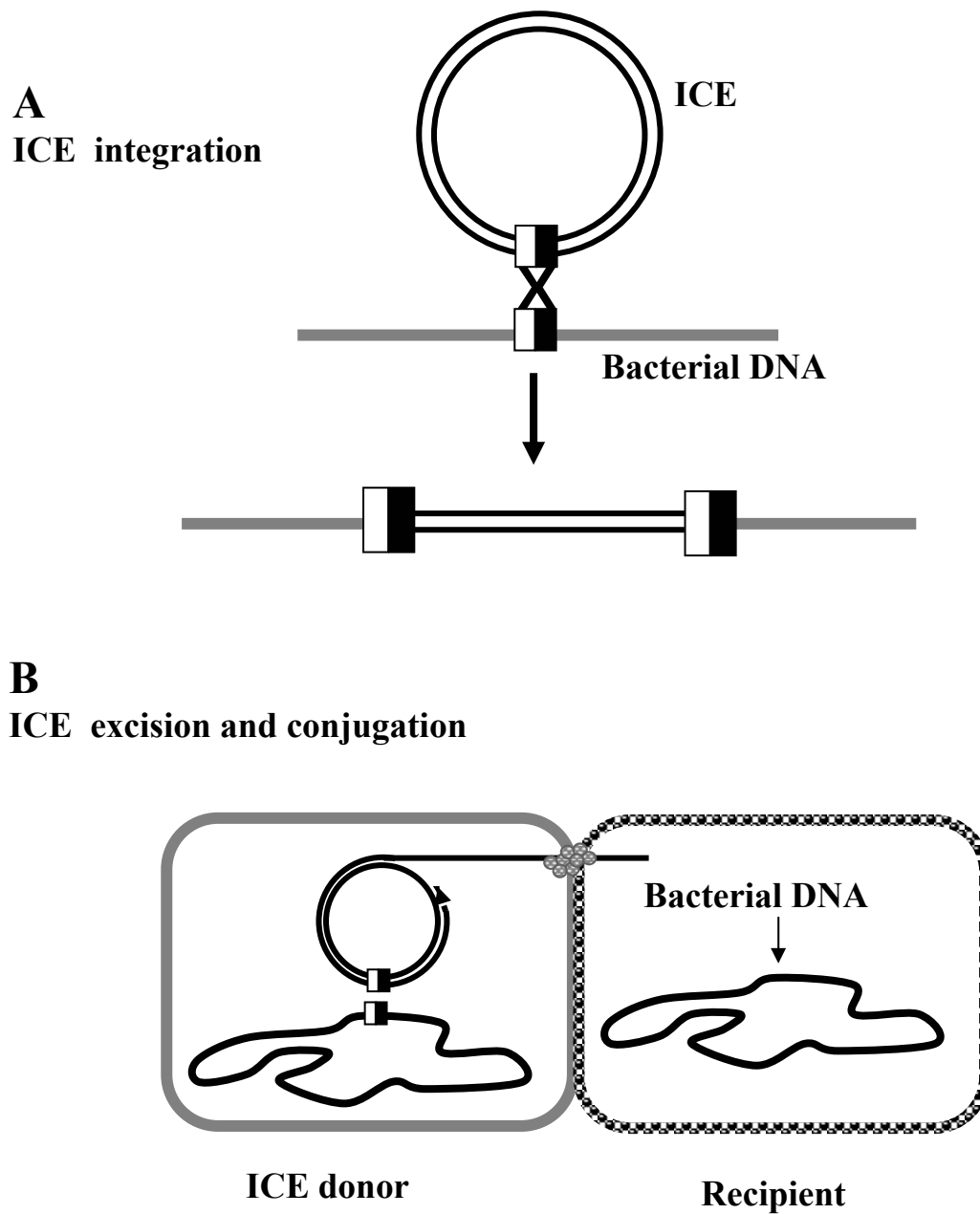


Figure 17: Physiological states of an Integrative conjugative element (ICE): Integration (panel A) of an ICE into the bacterial DNA occurs usually via the “lambda type” site specific recombination. ICE excision (panel B) is a reversal of the integration process whereas ICE transfer from one cell to another is thought to occur in a manner analogous to that seen for the F factor.

7.0 Implications of DNA transfer mechanisms

So far, mechanisms have been reviewed which can enable transport of DNA across bacterial membranes. These may be termed as prototype mechanisms of DNA transfer. There exists some amount of diversity in each category of DNA transfer mechanism but the underlying principle remains the same. We turn our attention from the confines of a laboratory wherein individual mechanisms are usually studied in relative isolation, to the wild (nature), wherein different kinds of bacteria coexist. It is now recognized that DNA transfer mechanisms represent the driving force in horizontal gene transfer (HGT) and the accompanying genetic exchange is thought to enable an organism to optimize its fitness in a given environment. By being able to acquire and retain transferred genes, an organism is presented with a means to utilize benefits (usually in the form of an added property) of the said gene and so be able to propagate successfully in the said environment. The benefit that a bacterium derives from HGT or more precisely the phenotypes it displays could be diverse. For example, the ability to degrade toxic compounds, to withstand antibiotics in the environment, or to establish an infection through the actions of virulence genes, constitutes some of these phenotypes. It may be noted that not all horizontal DNA acquisitions will be beneficial. Some will be “silent”, meaning that the transferred gene does not confer any obvious benefit to the cell. In such a case, it is thought that the benefit (if any) is likely to be manifest under certain or in as yet unknown environments. Other DNA transfer events may impart a harmful effect to the recipient cell, leading to the decline of the population bearing the deleterious gene, because, the deleterious gene lowers the “fitness” of the transformed cells relative to the non-transformed counterparts, which therefore propagate faster and predominate within the population.

A recent review by Thomas and Nielsen (2005) deals with the barriers to the establishment of a successful HGT and a précis of the said review (reference 16) is described below. Two main barriers that need to be overcome have been identified. The first is the barrier imposed by cell membranes of the recipient organism. During the process of F conjugation, the formation of the mating pair aggregate is a critical step and is mediated by the contact established between the F pilus from the donor and the recipient cell membrane. For this to happen the LPS layer of the outer membrane of the recipient cell is first contacted. F has a very limited host range because of this specificity. Thus, recipient cells possessing an altered LPS are not contacted and therefore F-mediated conjugation, even within the same species of *E. coli*, is more of an exception than the rule. On the other hand, the RP4 plasmid, which has a mechanism of DNA transfer similar to that of F, is able to establish itself in different bacteria. One reason for this is that an RP4-bearing cell can elaborate the required mating contacts with cell walls of both gram-positive or gram-negative type bacteria. If the mobilized episome/plasmid finds suitable conditions for its replication it will get established in the recipient cell.

The second barrier to successful HGT is termed as “host restriction”. It is known that bacteria contain mechanisms, which enable them to recognize their DNA (“self”) from “foreign” DNA. Foreign DNA upon entering the cell is subjected to degradation whereas the native genetic material is untouched. This is due to the activity of the restriction modification (RM) systems. RM systems are usually comprised of two enzymatic activities namely restriction and modification, which exert their action on short but specific DNA sequences known as RM sequences. The modification activity modifies the RM sequences in the endogenous bacterial DNA, so rendering it immune from the DNA degradative activity of the restriction system. If a

particular RM sequence on foreign DNA escapes the endogenous host cell modification, then the host restriction system cleaves the said sequence and degrades the foreign DNA (see references 10 and 12 for information on RM systems). The probability that a DNA transferred will have a sequence that is an RM substrate increases with increasing length of the DNA. Conjugal transfer systems can partly overcome the host restriction barriers in two ways. First, by maintaining a smaller size, the probability that an episome carries a specific RM sequence would be low. Size reduction can take place if, spontaneously, random deletions occur on an episome during its propagation in a cell. Sometimes, a small deletion on the episome could lead to the removal of a DNA sequence, which is an RM substrate. Episomes bearing such deletions are able to escape the effects of RM systems more frequently than their normal-sized counterparts and so come to predominate in nature.

The second way, in which the effects of RM systems could be tolerated, is related to the form of the transferred DNA upon entry in the recipient cytoplasm. Because the transferred DNA, in conjugal transfer, enters the recipient in a single-stranded form, it is relatively resistant to the action of RM systems, which prefer double-stranded DNA as substrates. Natural transformation mechanisms also render the incoming DNA partly immune to the effects of the RM systems because DNA enters (in most cases) the recipient in a single-stranded form. This might be a probable reason as to why natural transformation and conjugal DNA transfers seem to occur between unrelated bacteria. This is not to say that RM systems are rendered completely useless; they may act at a later stage following DNA transfer, say, during replication of a newly transferred episome in the recipient when double-stranded DNA becomes available. Once the transferred DNA escapes the RM systems it has the potential to establish itself within the recipient.

DNA transfer mechanisms, such as HGT events that involve natural transduction by bacteriophages (in which DNA is transferred in the double-stranded form) are most susceptible to the restriction barrier imposed by the host. Natural transduction by bacteriophages is also limited to occur only between closely related bacteria, since any given recipient must possess the cognate cell surface receptor for phage entry.

How is HGT detected? For episome-based transfers, researchers look for characteristic features associated with the transfer elements such as DNA sequences capable of encoding Tra or Pili proteins. If phage-mediated or ICE transfers are suspected, then DNA sequences resembling those for the *int* genes or the characteristic target site duplication are sought. Another feature that is employed especially if HGT of a particular gene is suspected in an organism possessing the ability to take up DNA, is the analysis of its GC content. It is known that the content of the two bases of DNA guanine (G) and cytosine (C) is characteristic for a given organism. For example, the GC content of the bacterium *Mycobacterium tuberculosis* is a little over 60 percent. This means that 60 percent of the entire genome of *Mycobacterium tuberculosis* contains GC base pairs. Similarly, the GC content of the malarial parasite *Plasmodium falciparum* is around 20 percent. In this way, GC contents of many bacterial species are known. Typically, the suspicion of the involvement of HGT increases if the DNA sequence of a given gene or region bears a GC content strikingly different from the rest of the genome. Researchers then try to match the GC content of the said gene with known GC contents of other microbial genomes. For example, a gene X from *Haemophilus influenzae* could show a GC content similar to that seen for *E. coli* genes but very different from its own. If it turns out that the nucleotide sequence of gene X is

very much identical to that of a homologous gene from *E. coli*, then it is considered as fairly good evidence for HGT of gene X from *E. coli* to *H. influenzae*.

In conclusion, it may be said that mechanisms exist to promote genetic exchange between bacteria, which in turn, contribute to the evolution of bacterial genomes so as to enable them to adapt to a variety of environmental conditions. An understanding of the mechanisms by which genetic exchange is mediated, in model organisms such as *E. coli* and *B. subtilis* has led to a greater recognition and knowledge of exchange mechanisms operating in nature.

Suggested Readings

1. Burrus, V., and M. K. Waldor. 2004. Shaping bacterial genomes with integrative and conjugative elements. *Res. Microbiol.* 155: 376-386.
2. Chen, I., P.J. Christie, and D. Dubnau. 2005. The ins and outs of DNA transfer in bacteria. *Science* 310: 1456-1460.
3. Clewell, D. B., and C. Gawron-Burke. 1986. Conjugative transposons and the dissemination of antibiotic resistance in *Streptococci*. *Ann. Rev. Microbiol.* 40: 635-659.
4. Dubnau, D. 1999. DNA uptake in bacteria. *Annu. Rev. Microbiol.* 53: 217-244.
5. Dubnau D, and R. Provvedi. 2000. Internalizing DNA. *Res. Microbiol.* 151: 475-480.
6. Firth N., K. Ippen-Ihler, and R.A. Skurray. 1996. Structure and function of the F factor and mechanism of conjugation. In F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. American Society for Microbiology, Washington, D.C.
7. Hanahan, D., and F. R. Bloom. 1996. Mechanisms of DNA transformation. In F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. American Society for Microbiology, Washington, D.C.
8. Ippen-Ihler, K. A., and E. G. Minkley, Jr. 1986. The conjugation system of F, the Fertility factor of *Escherichia Coli*. *Ann. Rev. Gen.* 20: 593-624.
9. Lewin, B. 1997. *Genes VI*. Oxford University Press.
10. Maloy, S. R, J. E Cronan, Jr and D. Freifelder. 1994. *Microbial Genetics*, 2nd ed. Jones and Bartlett Publishers, U.S.A.
11. Miller, J. H. 1992. *A short course in bacterial genetics: a laboratory manual and handbook for Escherichia coli and related bacteria*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
12. Murray, N. E. 2000. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). *Microbiol. Mol. Biol. Rev.* 64: 412-434.
13. Ptashne, M. 1992. *A genetic switch*. Cell Press and Blackwell Publications, Cambridge, U.S.A.
14. Sambrook, J., and D. W. Russell. 2001. *Molecular cloning: a laboratory manual*, 3rd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. Srivastava, S., and P. S. Srivastava. 2003. *Understanding bacteria*. Kluwer Academic Publishers, The Netherlands.
16. Thomas, C. M., and K. M. Nielsen. 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Rev. Microbiol.* 3: 711-721.
17. Tortosa, P., and D. Dubnau. 1999. Competence for transformation: a matter of taste. *Curr. Opin. Microbiol.* 2(6): 588-592.
18. Willets, N., and R. Skurray. 1980. The conjugation system of F-like plasmids. *Ann. Rev. Genet.* 14: 41-76.