Review on

Agrobacterium-mediated gene transfer in higher plants

Subhas Chandra Roy
Department of Botany, University of North Bengal

1.1 Introduction:

The genetic engineering of plants actually started with the discovery of a pathogenic bacterium, *Agrobacterium tumefaciens*, causing a disease known as crown gall (tumor). The bacterium transfers a set of genes from its plasmid into infected plant cells, where the transferred genes are expressed and produced gall at the junction of root and stem (*i.e.*, crown region of the plant). This discovery has revolutionized the plant transformation method through which any gene can be introduced into plant genome to change their genetic make-up for commercial benefits, spaning from disease resistance to the production of pharmaceutically important proteins. Various methods of gene transfer have been developed to transform the higher plant. These are broadly grouped into the following categories:

I. Vector-mediated gene transfer methods:

- i) *Agrobacterium*-mediated gene transfer
- ii) Virus-vector mediated gene transfer
  - a) CaMV based vector
  - b) Gemini virus based vector

II. Vectorless or Direct gene transfer methods:

- i) Physical based technique
  - a) Microprojectile Bombardment (Biolistic)
  - b) Silicon carbide fibers
  - c) Electroporation and Electrophoresis
  - d) Microinjection
  - e) Sonication
  - f) Macroinjection
- ii) Chemical based technique
  - a) Liposome fusion and lipofection
  - b) PEG-mediated DNA delivery

1.2 Vector-mediated gene transfer methods:

Vector-mediated gene transfer method is achieved either by *Agrobacterium*-mediated gene transfer or by use of plant viruses as transformation vectors. Main focus has been given to the *Agrobacterium*-mediated gene transfer method in the present context.

Natural Genetic Engineer of Plant.

Agrobacterium is valuable in many ways. It is the only prokaryote known that is capable of transfering DNA to eukaryotic cells (Bundock *et al*., 1995). In 1983, the era of plant transformation was initiated when *Agrobacterium*-mediated gene delivery system was used to develop transgenic tobacco (*Nicotiana tabacum*), a dicotyledonous plant (Fraley *et al*, 1983). This system has an obvious limitation that its inability to transform monocotyledonous plants. It

has been observed in case of corn (*Zea mays* L) that the secondary metabolite *dimboa* has an inhibitory effect on *Agrobacterium* growth and consequently not allowing to infect corn, a monocots (Sahi *et al*, 1990). More recently, so many monocotyledonous cereal crops such as rice, wheat have been transformed using *Agrobacterium*-mediated gene delivery system. Now, it is possible to transform a wide range of plants for desire agronomic importance. Among the dicots, soybean, cotton have been transformed with genes coding for herbicide resistance, insect resistance, viral protection, improved seed storage proteins and increased starch.
Agrobacterium tumefaciens is a soil-borne, Gram negative, sporulating, phytopathogenic bacterium. It is rod shaped, motile and belongs to the bacterial family Rhizobiaceae, present in the rhizosphere around the roots. There are three species of *Agrobacterium*-

Agrobacterium tumefaciens - Causes crown gall disease in dicotyledonous plants and gymnosperms. Strains are as such - Ach5, A6, B6, C58, T37, Bo542 and B023955.

Agrobacterium rhizogenes- Causes hairy root disease; strains are as such- NCPPB1855, ATCC23834 and A4.

Agrobacterium radiobacter- It is an avirulent strain, can be used as biocontrol agent to control crown gall disease.

Agrobacterium tumefaciens is a causative agent of *"Crown-gall"* disease, infecting a wide range of broad leaved dicotyledonous plants that was identified by Erwin F. Smith and Townsend of the U.S. Department of Agriculture (USDA) in the year 1907. It causes considerable damage to perennial crops notably in grapes, peach, walnuts, apples, roses, ornamental plants, other flowers and fruit trees. Crown gall was the first tumor in any organism, for which the causative agent was identified before the discovery of 'ROUS' sarcoma virus by several years. The ability to cause *'crown-gall'* (tumor growth) depends on the transfer ability of tumor inducing genes from bacterium into the plant genome. It has been demonstrated that the infective process is a natural forms of genetic engineering to develop disease on host plant by transferring bacterial genes into the plant genome. This natural process of gene transfer of *Agrobacterium* has been used widely as a major tool to manipulate the plant genome by which a new gene can be put into plants for expressing thereon to give rise a new variety of the plant with desired traits (Horsch *et al*, 1984). In the 1980s, Jeff Schell and Mare Van Montagu and their groups at the State University of Ghent, had shown that a large size of plasmid present in all virulent but not in avirulent strains of *A. tumefaciens*, which has clearly shown that the plasmid is responsible for tumor formation. This plasmid can be transferred to avirulent strain through conjugation. Virulence is lost when the bacteria are cured of plasmid. Both the virulent strains of *A. tumefaciens* and *A. rhizogenes* carry the characteristics large megaplasmids. The plasmid found in *A. tumefaciens* is named as Ti-plasmid (***T****umor* ***i****nducing* plasmid), which is responsible for the development of *'crown-gall'* tumors on appropriate host plants (Nester *et al*, 1984). *A. rhizogenes* strain carrying Ri-plasmid (***R****oot* ***i****nducing*-plasmid) is the causative agent of hairy root disease. Ri-plasmid can also be used as a gene transfer vectors to develop healthy genetically modified plants. The Ti-plasmid is today the only successful gene vector for higher plants.

1.3 Structural features of Ti-plasmid:

The Ti-plasmid of *Agrobacterium* is a autoreplicating, large double stranded circular DNA molecule of 200kb in size, which is about 3-8% of the *Agrobacterium* chromosome. Different strains of *Agrobacterium* have different types of Ti-plasmid (octopine, nopaline and agropine

types) but containing some common structural features (Figure 1.), which are as follows–

- It has own origin of replication *i.e.* autoreplicating genetic element
- A genetic element responsible for conjugative transfer (*tra*-region)
- Opine synthesizing gene (nopaline synthase, octopine synthase)
- Opine catabolizing gene (octopine and nopaline catabolizing genes)
- Contains T-DNA (Transfer-DNA) region, one or more than one to transfer into the plant genome which are flanked by left and right border
- A virulence region (*vir*-genes)

Among the above genetic components, two are more important, which are required for gene transfer in plants, the virulence (*vir*) region, and T-DNA region. Beside these a set of gene products encoded by bacterial chromosomal genes (chvA, chvB, chvE, pscA, attR genes) are also required for T-DNA transfer into plant cells.

1.3.1 Main structural features of T-DNA:

Nopaline T-DNA: The T-DNA of nopaline type Ti-plasmid (pTiC58) is a single continuous segment of DNA of about 22 kb long. The T-DNA has a specific boundary known as border, at both the ends, referred to as left border (LB) and right border (RB), which are 25bp imperfect direct repeats (Zambryski *et al,* 1982). The border sequences of nopaline T-DNA is as follows- left border 5′ TGGCAGGATATATTGTGGTGTAAAC 3′, right border 5′ TGACAGGATATATGGCGGGTTAAAC3′ (Figure 2). The border sequences act as a *cis*-element signal for the transfer apparatus. Any foreign DNA is placed between the borders are transferred into the plant cell. Through the investigation it has been demonstrated that this portion of the Ti-plasmid is integrated into the nuclear genome of the plant cell, and subsequently termed as 'transferred DNA' (T-DNA), which is responsible for tumour induction in plant.

Octopine T-DNA: The T-DNA of octopine type Ti-plasmid is divided into two different segments and designated as transferred left T-DNA segment (T_L-transfer left) and transferred right T-DNA (T_R-transfer right). The T_L is of 14 kb in size containing eight ORFs and T_R segment is of 7-kb in size. Each of the T-DNA is flanked by separate 25bp border sequences as found in the nopaline T-DNA. The T_R and T_L can independently be transferred and integrated into the plant nuclear genome due to the presence of these boundary sequences. Left border of T_L T-DNA is as follows 5′ CGGCAGGATATATTCAATTGTAAAT 3′ and right border like this 5′ TGGCAGGATATATACCGTTGTAATT 3′. The T_L DNA carries the tumor inducing function (oncogenes) but T_R-DNA contains opine synthesizing genes (Figure 2). T-DNA of octopine Ti-plasmid contains an overdrive or enhancer sequence outside of the T-DNA but close to right border required for optimal T-DNA transfer. (*e.g.* octopine plasmid; pTiAch5). The T_L T-DNA is similar to that of T-DNA of nopaline Ti-plasmid. It is a uniqueness of the T-DNA genes that although carried by the bacterium but only express in plant cells.

1.3.2 Eukaryotic features of T-DNA:

T-DNA genes are as like as the eukaryotic genes. Some molecular evidence support this features. The alpha-amanitin inhibits the transcription of T-DNA genes, suggesting that T-DNA genes are transcribed by RNA polymerase II, which is available in the eukaryotic nucleus. Eukaryotic structure of the T-DNA genes later have been confirmed by S1 mapping to know the mRNA initiation and termination sites. T-DNA genes show structurally common features with the eukaryotic genes. As like as the eukaryotic genes, a TATA box is present at –30 promoter

region in the initiation site, which is a eukaryotic characteristics, whereas in prokaryotic gene this consensus sequences are located in the -10 and -35 regions. There is also another conserved sequence present in the T-DNA genes at –80 region, which is known as CAAT box, an important feature of eukaryotic promoters, which is involved in correct initiation of mRNA transcription. Besides these TATA and CAAT boxes, there are one specific region present on the *nos* promoter at -100 to -170 spanning some direct repeat as well as some inverted repeat sequences. Like a eukaryotic gene, the 3' end of the mRNA of *nos* gene contains an untranslated region containing one or more consensus polyadenylation signal sequence AATAAA, which is also a feature of eukaryotic genes. The T-DNA genes do not contain introns just like a many other eukaryotic genes.

1.3.3 Tumor inducing oncogenes:

T-DNA contains two categories of genes, the oncogenic genes and the opine synthesizing genes. Oncogenic gene consisting of auxin and cytokinin synthesizing genes along with the tumor size determining gene. The oncogenic region of transferred DNA includes the two auxin synthesizing genes namely ***tms1***; ***tms2*** and one cytokinin synthesizing gene, ***tmr*** and one tumor size determining gene, ***tml***.The *tms1* (auxA, also known as iaaM) encodes tryptophan mono-oxygenase and *tms2* (auxB or iaaH) encodes indole-3-acetamide-hydrolase, responsible for the biosynthesis of plant hormone auxin (IAA) from the amino acid tryptophan. The enzyme tryptophan mono-oxygenase converts tryptophan to indole-3-acetamide, which is again converted to indole-3- acetic acid (auxin) by indole-3-acetamide hydrolase encoded by *tms2* gene. Apart from the genes for auxin synthesis, T-DNA also contains gene necessary for cytokinin synthesis. The *tmr* (also known as *cyt* or *ipt*) encodes an enzyme isopentenyl transferase that catalyzes the most essential step in cytokinin (isopentenyl adenosine-5'-monophosphate) production from isopentenyl pyrophosphate. Both the phytohormones auxin and cytokinin induces the tumor formation *i.e.* crown gall in plants and hence designated as 'oncogenes'.

1.4 The mechanism of T-DNA transfer and crown gall tumourigenesis:

The process of T-DNA transfer and integration into the plant genome is illustrated diagrammatically in figure 3 and which has been described in the following stages:

1.4.1 Contact between bacterial and plant cells:

Attachment to plant cells: Wounded plant cells release some kinds of chemical compounds to heal the wound region and also release some other types of compounds to resist pathogenic attack as a defense strategy, these are the derivatives of phenolic compounds such as acetosyringone, vanillin, ethylferulate, sinapinic acid and some monosaccharides such as D-mannose, D-galactose *etc.* They act as a chemo-attractants to attract the *Agrobacterium* cells near the wounded region (*i.e.*, rhizosphere). *Agrobacteria* are attracted to the wound sites of the plant by sensoring these signal molecules, such as phenolics, sugars that are released by the wounded plant cells. Attracted bacteria multiply in the wound sap and attach to the walls of plant cells. Monosaccharides also acts as effective chemoattractants and are responsible for the long distance attraction of virulent and non-virulent strains of *Agrobacterium* to the rhizosphere of wounded plants.The phenolic compounds such as acetosyringone (AS) acts as specific chemoattractants for virulent *Agrobacterium* strain only. The product of virA and virG genes are also involved in this chemotaxis mechanism, these two proteins regulate the signal transduction

mechanism to make a close contact between plant cells and bacterial cells. It has been observed by genetic analysis that two constitutively expressing neighboring genes *chv*A (1.5 kb) and *chv*B (5 kb) are responsible for bacterial attachment to the plant cells, which are present on bacterial chromosome. Genes *chv*A and *chv*B are involved in the synthesis of cyclic beta-1,2 glycans and excreted out from the bacterial cells, which is required for attachment of the bacterium to the plant cells. The gene *chv*E, is involved in the synthesis of glucose/galactose transporter for the enhancement of the sugar concentration, which subsequently induces the *vir* genes. The *psc*A (*exo*C) gene is responsible for the synthesis of polysaccharide but only induces in presence of succinoglycan which is synthesized by the gene *att*R, and ultimately produces polysaccharide that facilitates the initial attachment between bacteria and plant cells. Subsequently the *Cel-gene* produces cellulose fibril, which forms a mesh like structure at the attachment site by which bacteria can clump together and firmly anchor to the plant cells. Mutation in these two chromosomal virulence genes (*chv*A and *chv*B) can only make defective attachment. So, constitutively expressed products of several chromosomal genes (*chv*A, *chv*B, *chv*E, *psc*A and *att*R) in *Agrobacterium sp* are responsible for their attachment to the plant cells and the specific attachment of *Agrobacterium* to plant cells is a prerequisite for subsequent transfer of T-DNA of *Ti/Ri* plasmid. It is observed that monocots are not generally susceptible to infection by *Agrobacterium tumefaciens* because of failure of attachment to the plant cells.

1.4.2 Induction of bacterial virulence genes:

The T-DNA itself has no ability to transfer it from bacterial cell to plant cell but a set of genes are responsible for this transfer. These genes are clustered in a region out-side of the T-DNA boundary and referred to as virulence region (*vir*-region) containing virulence genes (*vir*-genes). The *vir*-region is of 30-40 kb in size in both the *Ti/Ri*-plasmids. Vir-region is consisting of at least six essential operons (*vir*A, *vir*B, *vir*C, *vir*D, *vir*E, *vir*G) and three non-essential operons (*vir*F, *vir*H, *vir*J). They are co-regulated and thus form a regulan. The *vir*A and *vir*G genes are expressed constitutively and they are monocistronic, all other *vir*-genes are inducible and polycistronics. The *vir*A, *vir*G have only one gene, *vir*E, *vir*C, *vir*H have two genes and *vir*D and *vir*B have four and eleven genes respectively. The *vir*-genes mostly responsible for plant cell recognization, bacterial attachment, excision and transfer of T-DNA and possibly the integration of T-DNA into the targeted plant genome (see Table 1). The *vir*A gene is expressed constitutively under all conditions. The *vir*G gene is expressed at low levels in vegetative cells but is rapidly induced to higher expression levels by exudates from wounded plant cells. Exudates may contain small phenolic compounds such as acetosyringone (AS), hydorxy-acetosyringone (OH-AS), vanillin, ethylferulate, sinapinic acid, lignin and some kinds of monosaccharide (D-mannose, D-galactose *etc.*). Lignin is a cell wall component of the vascular plant required to repair the wounded cells which are synthesized from sinapinic acid. Normal leaf tissue produces all the components in a lesser quantity but in the wounded tissues they are produced at least ten-fold higher amounts than the intact tissues. The natural physiological role of these compounds (AS, OH-AS, sinapinic acid, flavonoids *etc*) may be associated with wound healing processes and may also involve in the defense mechanism as a phytoalexin to inhibit the growth of the invading pathogens during wound. Some of the flavonoids inhibit the growth of some of the pathogens and act as a phytoalexins (Darvill and Albershein, 1984). The *vir*-genes are exclusively plant inducible and can be induced by the small phenolic compounds like acetosyringone (most effective inducer, this is a monocyclic phenolic compound), released by wounded plant tissue,

and induces the *vir*-genes. Virulence inducing signals are first recognized by *vir*A protein. The *vir*A is a transmembrane dimeric sensor protein kinase that detects signal molecules, mainly small phenolic compounds, such as acetosyringone, released from wounded plant cells. The signal molecules may be other phenolic compounds, or certain class of monosaccharides which act synergistically with phenolic compounds. After receiving signal, *vir*A protein autophosphorylates and subsequently phosphorylate the virG protein by transferring its phosphate group to a conserved aspartate amino residue of *vir*G proteins and activate them. Phosphorylated active *vir*G acts as a transcriptional factor (a DNA binding protein) and up regulates the expression of all the virulence genes *vir*B, *vir*C, *vir*D, *vir*E, *vir*F, *vir*H, *vir*J including *vir*A and *vir*G gene. This transcriptional factor binds to the operator region of the *vir*-genes and induces their expression. The C-terminal part is responsible for the DNA binding activity while the N-terminal part acts as a phosphorylating domain. Certain monosaccharide particularly, glucose, galactose, xylose enhances *vir*-genes induction. The induction by this system is only possible if the periplasmic sugar (glucose/galactose) transporter protein *chv*E is expressed to import the sugar inside the bacterial cells and which in turn interacts with virA to induce the other vir-genes. Temperature and pH condition may play important role in vir-gene induction. At temperature greater than 32°C, the vir-genes are not expressed because of a conformational change occurred in the folding of virA protein which inactivate its functional properties. The expression of these vir-genes are tightly regulated and only expressed when the *Agrobacterium* infect a plant. Wound sap induces the *vir*-genes and most effective inducers are monocyclic phenolic compounds such as acetosyringone. The *vir*A and *vir*G form a two component regulatory system activating the transcription of the other vir-genes. Virulence inducing signals are first recognized by *vir*A protein, which then activates *vir*G by phosphorylation. Both the *vir-genes*, *vir*A and *vir*G are expressed constitutively but the virG gene can be induced by AS to higher levels of expression. The reason is that un-induced *vir*G gene transcribes a shorter mRNA in compare to induced mRNA transcript because the constitutively transcribed mRNA is without Shine-Dalgarno sequences, which is essential for efficient translation. The feature can be compared with the CI gene regulation by pRE and pRM in the bacteriophage lambda (lambda Phage), and considered as a reminiscent of the same. Inducible *vir-genes* do not have –35 consensus sequences in the promoter but they carry a hexanucleotide sequences common to the entire region at –35 and act as *cis*-acting regulatory sequences for the *vir*-gene induction. It has been reported by Ashby *et al* (1987) that *Agrobacteria* are positively attracted and showing chemotactic response to AS at an optimum concentration of 10^{-7} M. The most monocots are not producing AS at a concentration sufficient to induce *vir*-gene expressions otherwise they would have been also infected with the *Agrobacterium sp*. The host range of *Agrobacterium* may be artificially extended to the monocotyledonous plant also by adding AS to the *Agrobacterium*-monocot cell *co-cultivation* inoculation method and it has been achieved.

1.4.3 Production of ssT-DNA transfer complex and its transportation to nucleus:
As soon as the *vir*-genes are expressed and produced the respective *vir*-proteins they start to generate a single stranded transfer DNA (ssT-DNA) molecule to transfer it to the plant cell (Stachel and Zambryski, 1986). This ssT-DNA molecule represents the copy of the bottom strand of the T-DNA. The proteins virD1 and virD2 encoded by the *vir*D locus starts the T-DNA excision by recognizing the T-DNA border sequences and nicking (endonuclease activity) the bottom strand at each border. The virD2 protein has the specificity to cut within the conserved

25 bp T-DNA borders, the origin of transfer (oriT) in association with virD1 protein(Wang *et al*,1984; 1990). VirD2 protein recognizes these border repeats (LB and RB) and makes single stranded nicks between the third and fourth bases in the bottom strand of the T-DNA from the 3' end, with the help of virD1, virC, and also virD3 proteins. The VirD1 protein has topoisomerase activity and assisting the endonuclease cleavage in this T-DNA excision mechanism, that helps the virD2 to cleave the supercoiled double stranded T-DNA. Then the both ended nicked bottom strand is displaced by synthesizing a new DNA strand in the 5' to 3' direction using upper T-DNA strand as a template through rolling circle mechanism of DNA replication. DNA repair mechanism is also involved in this ssT-DNA formation. The strand displacement is controlled by the helicase activity, which is a inherent property of virD2 protein. The newly synthesized DNA strand is remained as a bottom strand with the Ti-plasmid. The displaced single-stranded T-DNA (ssT-DNA, The ssT-DNA molecule is also designated as T-strands *i.e.* transfer strand of T-DNA, which is an intermediate molecule in the transfer mechanism of the T-DNA.) is remained covalently attached with virD2 protein at the 5'-end only and the T-strand is coated with virE2 protein, which protect the T-strand from being attack of exo-endonucleases. The single–stranded T-DNA (ssT-DNA) associated with virD2 and virE2 form a transfer complex (T-complex) and is exported to the plant cell. During this transfer virF proteins also transferred to the plant cells. The virD2 acts as a capping protein to determine the polarity and also guides the T-strand mobilization towards the plant nucleus, during plant cell transformation (Wang *et al*, 1990). This ssT-DNA-virD2-complex is coated by the 69kDa virE2 proteins (600 molecules per 20kb T-DNA) to form the transfer complex (T-complex). The virE2 is the single stranded DNA-binding proteins, which protects the ssT-DNA from being nuclease attack and reduces the complex diameter of the ssT-DNA-virD2-complex at least 2nm, maintaining the structure somewhat in a linear fashion and make it easier to pass through membrane channels. The virE1 protein acts as a chaperon of virE2 proteins to give proper structural conformation to the virE2 protein. VirB loci (11 cistrons) produce various products which are involved in the formation of transfer membrane channel in association with other proteins like virD4. The virB and virD4 proteins form a membrane channel (*i.e.* T-pilus, or conjugal pore) which is the type IV secretion system between the bacterium and the plant cell (Zupan *et al*., 2000). The virE2, virD2 and virF are transported out from the bacterial cells along with other proteins involved in this mechanism (Zupan *et al*, 1995;1996). The ssT-DNA protein complex travels through the T-pilus, and reached into the cytoplasm of the infected plant cells. Some of the proteins of the membrane channel acts as a ATPase activity, probably the B11 protein, which provides the essential energy required for the passage of the T-strand by its autophosphorylating property. Although virD4 is a transmembrane protein but predominantly present at the cytoplasmic side of the cell membrane. Some of the virB proteins are assembled as a membrane spanning channel covering both the membranes, but other virB proteins may be extra cellular. The virB1 protein is always present extracellularly on surface and virB2 is with the extracellular functions. VirB4 and virB11 are with the ATpase activity, they provide necessary energy for active T-DNA transfer. VirB3 has the likely functional activity with that of virB4, whereas virB7 provides the correct conformation to the T-DNA complex during transfer. VirB9 is unstable, but stable when makes heterodimer with virB7 through disulfide bridges. The virB7-virB9 heterodimer is assumed to stabilize other *vir* proteins during the formation of functional transmembrane channels. VirB1 is a transglycosidase and may be involved in the formation of transmembrane channel. Both the associated proteins (virD2 and virE2) of the ssT-DNA

complex interact with the different plant proteins and facilitate the entry of T-DNA-complex into the cell cytoplasm.

The role of T-DNA borders in T-DNA transfer: The T-DNA of *Ti/Ri* plasmids is flanked by imperfect direct repeats of 25 bp and designated as the left border (LB) and right border (RB). T-DNA transfer process is prevented if the RB is deleted but not the left border (LB). It is also observed that the 25 bp right border alone can promote the T-DNA transfer in one direction only (5' to 3' direction of the T-strand), without any size restriction (upto 50 kb of DNA). These findings allow the researchers to modify *Agrobacterium* Ti-plasmid to be used as transfer vehicle to transform the plant cells. The right border nick provides the initiation of DNA replication whereas the left border nick gives the signal for termination of the replication and also transfer the same. If the left border is deleted, termination process will occur but at a distal site. In principle, as both the LB and RB contain the similar 25 bp repeat sequences and can function in similar way and can initiate transfer of T-DNA to the left of its, as a result there will be no T-DNA transfer (with foreign gene) into the plant genome, consequently there was no tumourigenesis. Instead of the left border, the right border is preferred site for initiation of T-DNA transfer as because of the presence of ***enhancer sequence*** (sometimes called as ***overdrive sequence***) adjacent to the RB of the Ti-plasmid outside of the T-DNA and enhances the transfer efficiency. It is consisting of 24bp long DNA sequences and which is like this 5' TAAGTCGCTGTGTATGTTTGTTTG 3' and is located next to the right border outside of the T-DNA region. The *'overdrive sequence'* is recognized by the virC1 protein and enhances its ability and polarity of the T-DNA transfer in a 5' to 3' direction, initiating from the nicked end in the right border but not transferring itself into the plant cell. Some genetic experiments have proved that these 25 bp repeat sequences, particularly those on the right border of the T-DNA, are absolutely required for T-DNA transfer and that functions in *cis*-condition and also directs the polarity of the T-DNA transfer. It is also evident that any DNA sequence can be transferred to plant cells if it is flanked by the 25 bp left and right borders, and which will be transported to the plant cell in a correct orientation (Rubin, 1986). The transferred DNA into the plant genome is identified by the presence of these 25 bp repeat sequences (partially or fully).

1.4.4 Integration of T-DNA into the plant nuclear genome and their expression:

T-DNA Integration: After entering into the plant cytoplasm, the ssT-DNA-virD2-virE2 complex (transfer complex) is correctly localized into the nucleus because virD2 contains a nuclear localization signal (NLS) in its amino acid compositions. The nuclear localization signal (NLS) present on the virD2 is recognized by cytoplasmic proteins importins which facilitates the import of ssT-DNA transfer complex to the nucleus through the nuclear pore complex (NPC). NLS also present on the virE2 proteins but two in numbers and which is also recognized by plant protein importins with the help of other type of cytoplasmic protein named as VIP1 and VIP2. If VIP1 is associated with the T-DNA-virD2-virE2 complex then NLS present on the virE2 is recognized by importins and facilitates its passage through NPC. The NLS present on virE2 protein mainly helps to transport the large DNA molecules. VirE2 proteins in the T-DNA-virD2-complex provides the correct conformation to the T-DNA-virD2-complex and also protects it from nuclear degradation. Then the T-DNA-virD2 complex interact with the another virE2 interacting proteins, VIP2, which direct this complex towards the chromatin and plays vital role in integration to the nuclear genome. During the integration of the T-strand, the virD2 protein is released from the 5' -end and provide energy by cleaving the phosphodiester bond between

Tyr29 residue and the first nucleotide of T-strand, and allow the 5' end of the T-strand to integrate into the plant DNA through illegitimate recombination (Chilton *et al*, 1977; Gheysen *et al*, 1989). In this recombination, pairing of few bases between the T-strand and plant DNA is necessary, and termed as micro-homologies (Tinland *et al*., 1995). Although the mechanism of T-DNA integration *via illegitimate recombination* is not fully known but possibly it is required some kinds of plant proteins, recombinases and DNA polymerases in association with other plant proteins such as cyclophilin, importin, histoneH2A *etc*. The phenomenon of T-DNA integration into the nuclear genome not only involving the bacterial encoded proteins and enzymes but also they involve proteins and enzymes coming from the plant sources. Because *Agrobacterium* infects only the wounded cells and wounded cells must undergo cell division for wound healing purposes. The replication of the DNA is necessary to proceed towards the cell division and the necessary enzymes for DNA replication and repair are available during this cell division and some of the proteins and enzymes of the cell divisions can also be utilized in the T-DNA transfer and integration into the plant genome. T-DNA of octopine type Ti-plasmid contains two accessory *vir* operons, virF and virH. virF encodes a protein of 23 kDa and involved in the integration of the T-DNA to the nuclear genome and control the cell cycle by prolonging the S-phase. On the other side, *vir*H has two genes encoding virH1 and virH2 proteins, which are involved in the detoxification of certain toxic compounds produced by the plant cells that might otherwise adversely affect the growth of *Agrobacterium*. That means these two genes recognize the host range specificity for bacterial infection (Kanemoto *et al*., 1989). Tumor formation efficiency is increased by bacteria having the virH genes as compared to those lacking the same. *Vir*F can also discriminate the host and non-host plant, by interacting with the T-DNA transfer complex and determine the host range. The T-DNA derived from *Ti/ Ri* plasmid is covalently integrated into the nuclear genome either as a single copy or several copies (on an average 2-3 copies) randomly through out the genome. The random integration process may insert multiple copies of the T-DNA including direct or inverted tandemerization or truncation of T-DNA or make aberrant integration. After its integration, T-DNA adopts eukaryotic characteristics of chromatin organization in order to make nucleosome structure and insensitive to DNase I. Recent investigations have suggested that integration is not occurred randomly instead integration occurs preferentially at the region of chromatin, which is transcriptionally active. It is being observed through the transgenic production that the highly expressed open chromatin region is being the targeted region for preferential recombination. Successful transformation leads to the transcription of the transferred oncogenic genes. The tms1, tms2, and tmr genes are expressed in a coordinated manner to produce the respective enzymes for the biosynthesis of phytohormones-auxin and cytokinin at a high level, which triggers the high mitotic activity and leads to the formation of neoplastic growth (*i.e.* crown-gall). Auxin synthesis is autoregulated by a product of T-DNA gene, indole-3-lactate from tryptophan which is a antagonist of auxin and control the over production of the auxin and thus prevent toxic effects of the auxin overdose.

1.5 Basic features of *Ti*-plasmid required to be used as a plant transformation vectors: Natural *Ti-plasmid* as such cannot be used as a plant transformation vectors because of the following reasons– i) Large size (200 kb) Ti-plasmid can not be used as a suitable vector for direct DNA cloning and transfer; high yield circular plasmid cannot be prepared, more over it is without multiple cloning sites (MCS), where the foreign gene(s) are to be inserted to transfer to

plasmid is referred to as binary vector. The necessary genetic elements for replication, conjugal transfer are derived from plasmid, pRK252. The virulence genes containing helper Ti-plasmid supply the vir genes in *trans* arrangement and termed as plasmid for virulence (pVIR). This helper plasmid has no T-DNA region and is not capable to transfer any foreign DNA into the plant nuclear genome. In contrast to that, the other plasmid is considered as a binary vector, which is smaller in size, at least 20 times less than the helper Ti-plasmid because they don't have to carry any *vir* genes, which is supplied in *trans* conditions by the helper Ti-plasmid (pVIR). There is no need of any kind of genetic manipulation to the helper Ti-plasmid so it can be stably maintained in a suitable *Agrobacterium* strain. The binary plasmid vector (pBV) contains T-DNA region along with the left and right border sequences, which is needed to transfer the transgene into the plant cells. The bacterial selectable marker genes are located on the binary vector outside of the T-DNA region for selection in *E. coli* and *Agrobacterium sp.* A plant selectable marker gene is located on T-DNA region to select the transformed plant cells. This binary vector plasmid (pBV) is normally maintained in *E coli* host cell for *in vitro* manipulation (to clone the foreign gene *i.e.* transgene into this plasmid) (Klee *et al*, 2000).
The pBIN19 was the first binary vector used to transform the plant cells in the early 1980s, (Figure 4) which was one of the most widely used general purpose plant vectors (Bevan., 1984). The pBIN19 is 11.77kb in size, and an antibiotic resistant gene kanamycin (*npt*-II) is located close to the right border (RB). The multiple cloning site derived from pUC19, is housed within the LacZα region of β-galactosidase on the T-DNA, which facilitates the blue/white screening on IPTG/X-gal plates to distinguish recombinant from non-recombinant colonies. The MCS contains seven restriction enzyme cutting sites for the insertion of passenger DNA (foreign/transgene). The pBIN19 contains a prokaryotic kanamycin resistance gene (APH-I) for the selection of bacterial transformation outside of the T-DNA. The T-DNA contains a plant transformation selectable marker, kanamycin resistance gene (*npt*-II) derived from Tn5, which is placed under the control of the *nos* promoter (nopaline synthase) and *nos* 3' polyadenylation signal close to the right border. The origin of replication is derived from a wide host range plasmid pTiT37. The helper Ti-plasmid is derived from a octopine type plasmid pTiAch5. The helper Ti-plasmid can also be obtained from nopaline (T37-SE), octopine (pAL4404), and succinamopine strains (pEHA101).

Advantages of Binary vectors:

i) Agrobacterium-mediated gene transfer to plants mainly are obtained using the binary vectors than co-integrative vectors, due to small size, they are convenient to manipulate *in vitro.* ii) In case of binary vector system, only the foreign DNA along with plant selectable markers are introduced into the plant genome but in co-integrative vector, unwanted bacterial plasmid DNA also introduced into the plant genome.

1.6 *Agrobacterium*-mediated plant Transformation Techniques:

The purpose of most plant transformation experiments in plant biotechnology is to produce whole, transgenic plants. *Agrobacterium*-mediated technique is the most widely used for the transformation of plants and regeneration of transgenic plants. The explants used in transformation experiments must therefore be capable of producing whole plants by regeneration and must be competent for transformation. So, tissue culture techniques have been playing important roles in most of the transformation methods available to date (only exception *in planta*-transformation). The tissue culture based conventional transformation techniques are

referred to as *in vitro* transformation. Generally, *Agrobacterium*-mediated *in vitro* plant transformation techniques have been adopted to inoculate the plant explants to transfer the desired gene into the plant cells. A variety of inoculation methods are available to transform plant cells for transferring foreign gene(s) into the higher plants through *Agrobacterium*-mediated gene delivery methods, these are as follows-

i) Co-cultivation method
ii) Leaf- disk transformation method
iii) *Agroinfection* method and
iv) *In planta* transformation

Advantages of *Agrobacterium*-mediated gene transfer method:

i) Agrobacterium is capable of infecting any plant explants (plant cells, tissues or organs). ii) It is as like as natural gene transfer in higher plants, iii) Large DNA segment (up to 50kb) can be efficiently transferred into the plant cells, iv)Without substantial rearrangements of the genomic DNA., v) Integration of the T-DNA is not so random but it is precise., vi) The stability of the transferred gene is reasonably good., vii) Transformed plant explants can be regenerated into whole plant effectively.

Disadvantages:

i) Some important cereal crops cannot be infected with *Agrobacterium* because they are belonging to monocotyledons. But recently, this limitation has been overcomed by using the ***super virulent*** strain of *Agrobacterium*, which can infect wide host range of plant species.

ii) More easily regenerable plant explants, the embryogenic cells are difficult to transform with the *Agrobacterium* because they are laid in deep layers of the tissue system and are not easily available for *Agrobacterium* infection or simply is not ideal target for T-DNA transfer.

Table 1. The virulence genes (*vir*-genes) of *Agrobacterium sp* and their role in T-DNA transfer.

Vir genes	Functional role	Nature of the gene
Vir A (2.5 kb)	Transmembrane phenolic chemosensor kinase detect the small phenolic compounds released by wounded plant cells, after sensing, the vir-A protein autophosphorylate itself and consequently phosphorylates virG-protein.	Monocistronic, expressed constitutively
VirG (1.2 kb)	VirG is a transcription factor, activates all other vir-genes	Monocistronic, inducible (also low level constitutive)
VirJ	T-DNA export	-
VirD (4.5kb) VirD1 Vir D2	 Regulate the virD2 activity and essential for T-DNA processing Recognize the 25bp border sequences with the help of virD1 and make a single stranded nick in the bottom strand of each border and remains tightly associated with the 5'end of the ssT-DNA to protect them from nucleases. Nuclear localization signal is present on the virD2 protein for targeting the T-DNA into the nucleus. Interact with plant protein such as importin and cyclophilins during the ssT-DNA transfer.	Polycistronic (4-cistron), inducible
VirF(23 kDa)	Helps in T-DNA integration into the plant genome, by regulating the cell cycle, prolonged S-phase	Monocistronic, inducible.
VirE VirE1 (2kb) VirE2	 VirE1, acts as a chaperone of virE2 protein, required to maintain its structure and function. Single stranded DNA binding protein, protect ssT-DNA from nuclease degradation, it also carries NLS, for nuclear targeting of the ssT-DNA and involve in passage through nuclear pore complex (NPC). Interacts with plant protein VIP1 and VIP2, to facilitate the T-DNA to move near the active chromatin region for integration.	Polycistronic (2-cistrons), inducible
VirC1 (2kb)	Recognize the overdrive sequence of the Ti-plasmid and enhances the efficiency of ssT-DNA transfer.	Polycistronic (2-cistrons), inducible.
VirB1-B11(9.5 kb)	Make a transfer complex apparatus at the junction of bacterial and plant cell surface for transfer of T-DNA.	Polycistronic (11-cistrons), inducible.
VirH(virH1 & H2)	Detoxification of certain plant compounds	Polycistronic, (2-cistrons), inducible

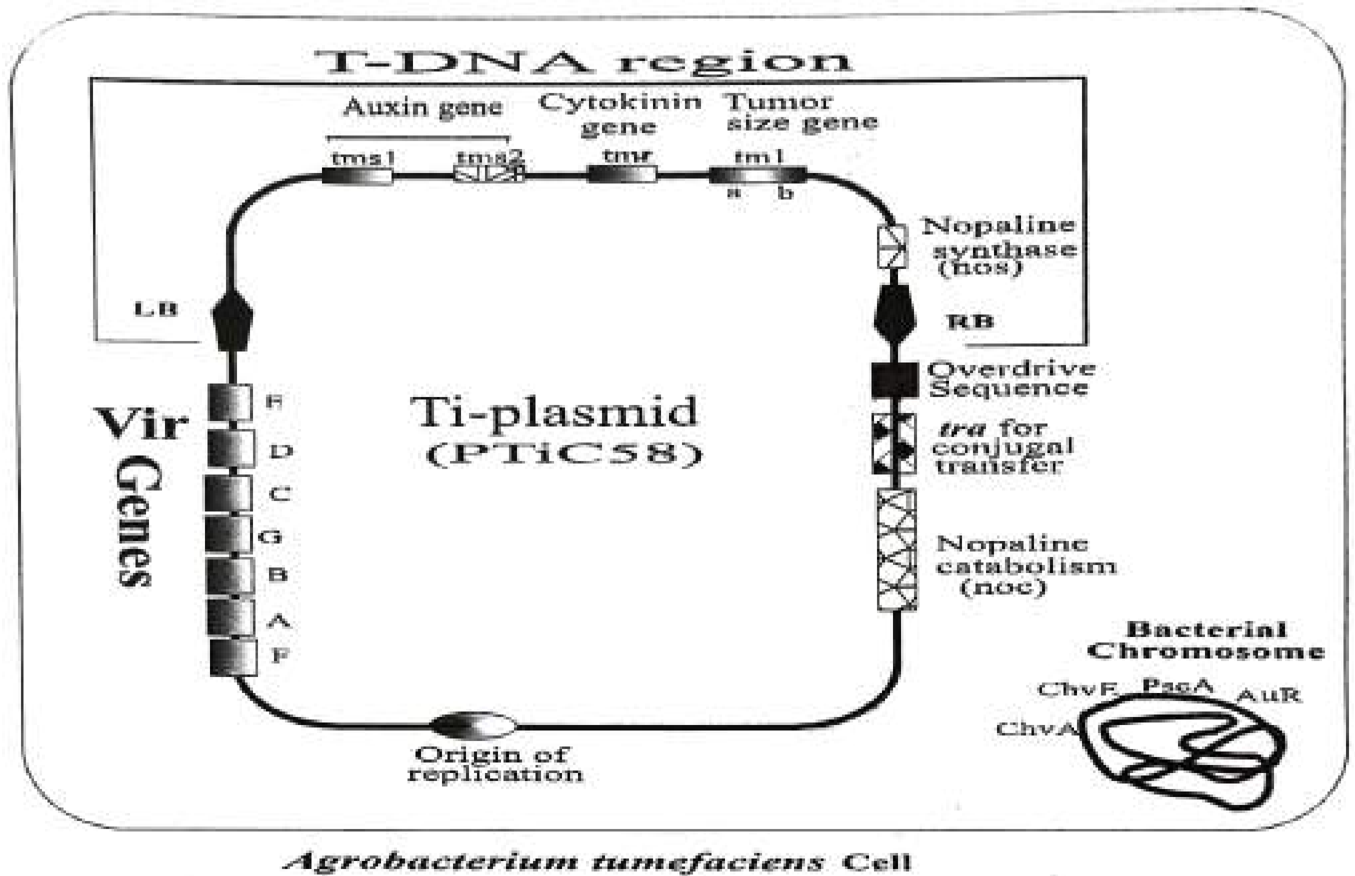

Figure 1. Nopaline type Ti-plasmid showing all the structural components necessary for gene transfer in higher plants.

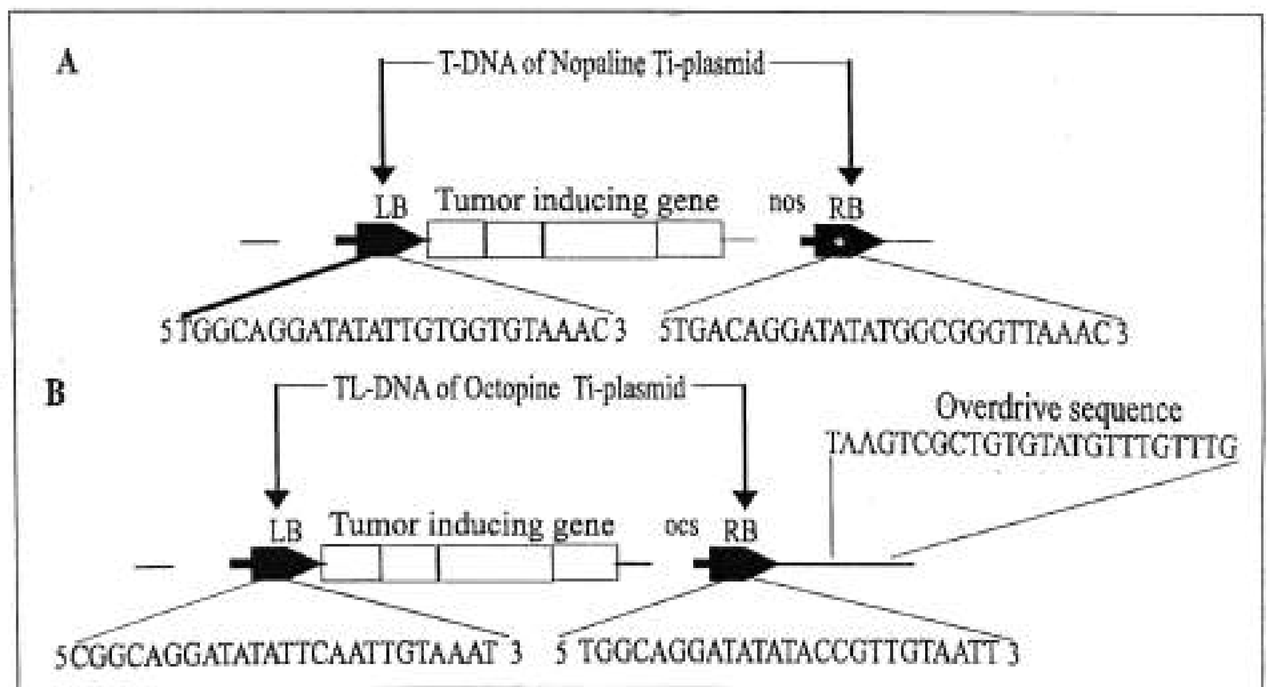

Figure 2. Structure of the T-DNA of nopaline Ti-plasmid (panel A) and TL-T-DNA of octopine Ti-plasmid (panel B) has been depicted diagrammatically.

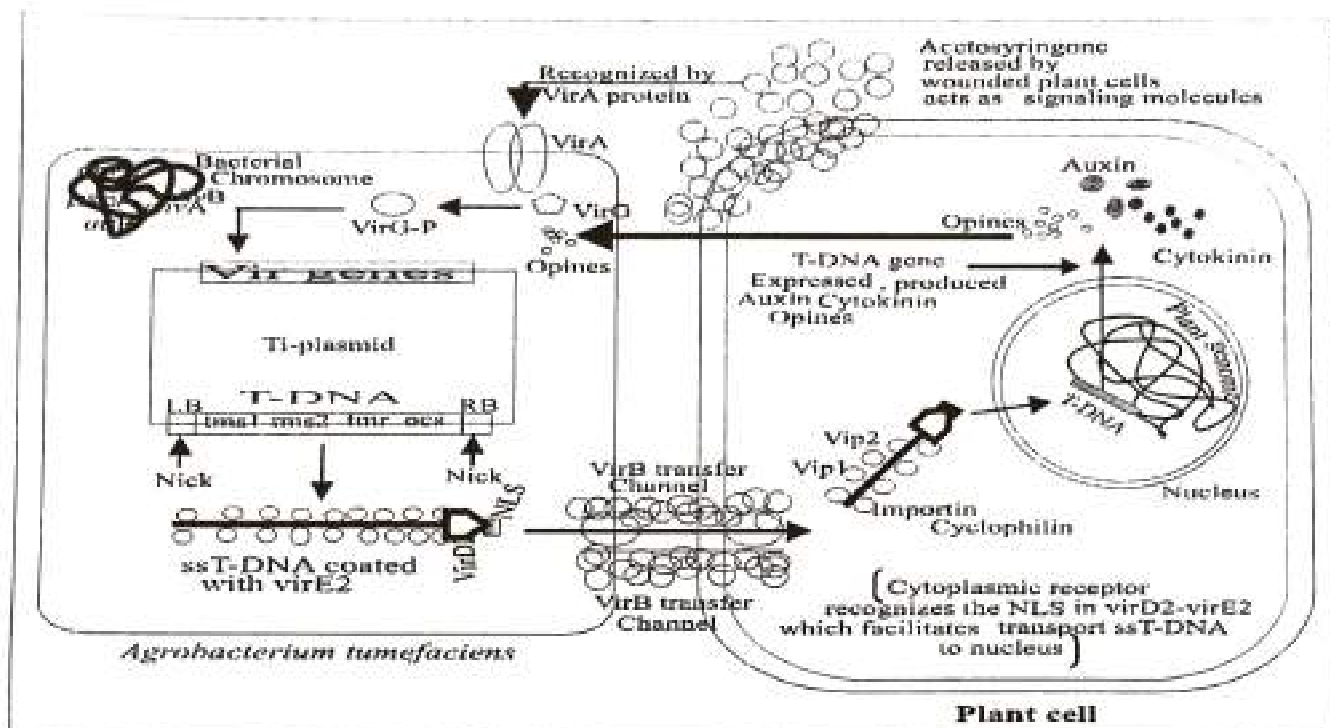

Figure 3. A generalized diagramatic representation of T-DNA transfer and integration into the plant nuclear genome.

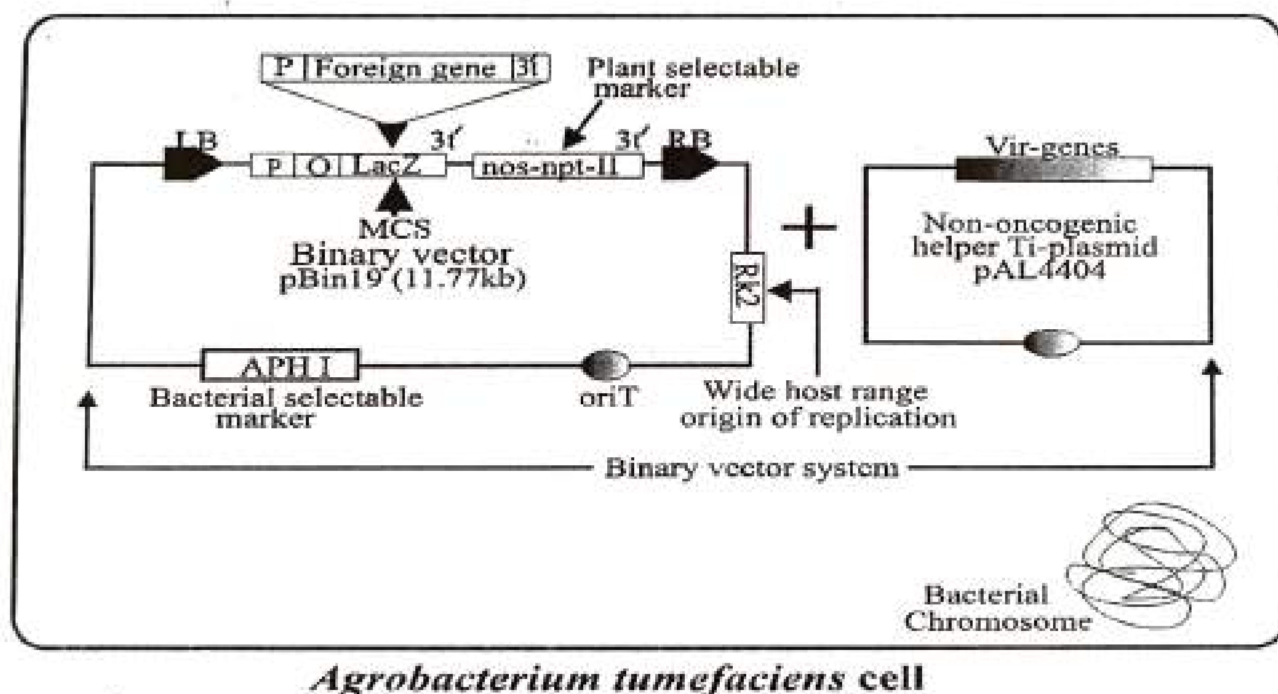

Figure 4. Diagrammatic representation of a typical binary vector system to transfer foreign gene into plant cell. Any binary vector system contains a pair of plasmid vector, one acts as a binary vector as such p Bin 19 and second one provide the virulence function in *trans*-arrangement and referred to as non-oncogenic helper Ti-plasmid pAL4404. Both the plasmid vectors reside on a single *Agrobacterium* cell.

References:

1. Ashby AM, Watson MD, Shaw CH 1987. A Ti-plasmid determined function is responsible for chemotaxis of *Agrobacterium tumefaciens* towards the plant wound product Acetosyringone. *FEMS Microbial Lett.* 41: 189-192
2. Bevan M 1984. *Agrobacterium* vector for plant transformation. *Nucl.Acid Res.*12: 8711-8721
3. Bundock PA, den Dulk-Ras A, Beijerbergen A, PJJ Hooykaas 1995. Trnas-kingdom T-DNA transfers from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae. EMBO journal.* 14:3208-3214.
4. Chilton MO, Drummond MH, Merio DJ, Gordon MP 1977. Stable incorporation of plasmid DNA into higher plant cells. The molecular basis of crown gall tumorigenesis. *Cell.* 11: 263.
5. Darvill, AG, Albersheim PA 1984. Phytoalexins and their elicitors-a defense against microbial infection in plants. *Ann.Rev.plant physiol.* 35: 243-275.
6. Fraley RT, Rogers SG, Horsch RB 1986. Genetic transformation in higher plants. *Crit. Rev.Plant Sci.* 4:1.
7. Gheysen G, Villarroel R, Van Montagu 1985. Illegitimate recombination in plants; a model for T-DNA integration. *Genes Development.* 5: 287-297.
8. Klee H, Hellers R, Mullineaux P 2000. A guide to *Agrobacterium* binary Ti vector. *Trends in Plant Science.* 5: 446-451.
9. Wang K, Herrera-Estrella A, Van Montagu M 1990. A bacterial peptide acting as a plant nuclear targeting signal: the amino-terminal portion of *Agrobacterium* VirD2 protein directs the beta-galactosidase fusion protein into tobacco nuclei. *PNAS.* 87:9534-9537.
10. Hoekema A, Hirsch PR, Hooykaas PJ, Schilperoorts RA 1983. A binary plant vectors strategy based on separation of Vir and T-regions of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature.* 303:179.
11. Horsch RB, Fry JE, Hoffman, NL, Wlroth M, Fraley RT 1984. A simple and general method for transferring gene into plants. *Science.* 227: 1229-31.
12. KanemotoRH,PowellAT,AkiyoshiDE,RegierDA, Kerstetter RA, Nester EW, Haws MC, Gordon MP 1989. Nucleotide sequence and analysis of the plant-inducible locus pinF from *Agrobacterium tumefaciens. Journal of Bacteriology.* 171: 2506-2512.
13. Nester EW, Gordon MP, Amasino, RM, Yanofsky MF 1984. Crown gall: a molecular and physiological analysis. *Annual Review of Plant Physiology.* 35: 387-413.
14. Rubin RA 1986. Genetic studies on the role of octopine T-DNA border regions in crown gall tumor formation. *Molec.gen.genet.* 202:312-320.
15. Sahi SV, Chilton MD, Chilton WS 1990. Corn metabolites affect growth and virulence of *A. tumefaciens*. *PNAS*,USA. 87: 3879-3883.
16. Smith EF, Towsend CO 1907. A plant tumor of bacterial origin. *Science.* 25: 671-673.
17. Stachel SE, Zambryski PG 1986. Generation of single stranded T-DNA molecule during the initial stage of T-DNA transfer from *Agrobacterium tumefaciens* to plant cell. *Nature.* 322: 706.
18. Tinland B, Schoumacher F, Gloeckler V, Bravo AM, Angel M, Hohn B 1995. The *Agrobacterium tumefaciens* virulence D2 protein is responsible for precise integration of T-DNA into the plant genome. *EMBO journal.* 14: 3585-3595.
19. Wang K, Herrera-Estrella L, Van Montagu M, Zambryski P 1984. Right 25 bp terminus

sequence of the nopaline T-DNA is essential for and determines direction of DNA transfer from *Agrobacterium* to the plant genome. *Cell.* 38: 455.

20. Zambryski P, Depicker A, Kruger K, Goodman HM 1982. Tumur induction by *Agrobacterium tumefaciens*: Analysis of the boundaries of T-DNA. *J.molec. appl.Genet.* 1: 361-370.
21. Zupan J, Muth TR, Draper O, Zambryski P 2000. The transfer of DNA from *A. tumefaciens* into plants: A feast of fundamental insight. *Plant Journal.* 23: 11-28.
22. Zupan JR, Zambryski PC 1995. Transfer of T-DNA from *Agrobacterium* to the plant cell. *Plant Physiology.* 107: 1041-1047.
23. Zupan JR, Citovsky V, Zambrysky PC 1996. *Agrobacterium* VirE2 protein mediates nuclear uptake of single-stranded DNA in plant cells. *PNAS.* 93: 2392-2397.