

## Gene cloning

1

### *Gene cloning (DNA cloning)*

The basic events in gene cloning:

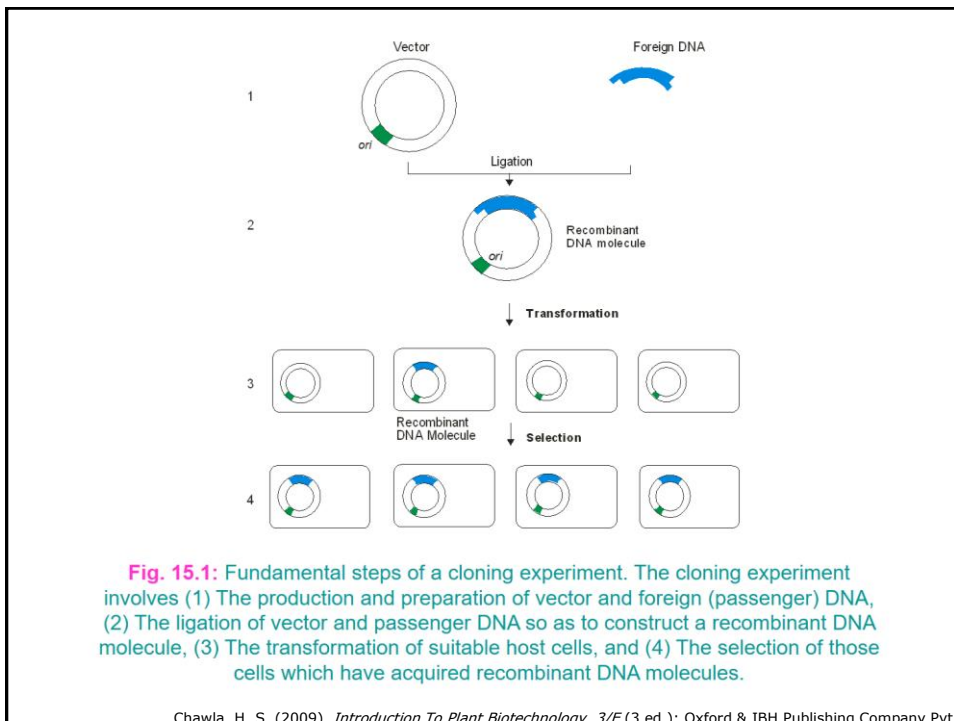
- 1) Isolation of the gene of interest.
- 2) Incorporation a fragment of DNA into a small replicating (usually circular) DNA molecule called a vector (*E. coli* plasmid, virus, cosmid,...) – The vector with an incorporated gene is called a **recombinant vector or molecule**.
- 3) The recombinant vector is introduced into a host cell by transformation.
- 4) Cells that have acquired recombinant DNA molecule are selected.
- 5) Recombinant DNA molecule is multiplied within the host cell to produce a number of identical copies of the cloned gene.

*Gene cloning can be used for genome sequencing – e. g. A gene library, a random collection of cloned fragments in a suitable vector that ideally includes all the genetic information of that species*

Essential components:

- Enzymes for cutting and joining the DNA fragments into vector molecules.
- Cloning vehicles or vectors.
- DNA fragments, i.e. gene libraries.
- Selection method of a clone of transformed cells that has acquired the recombinant chimeric DNA molecule.

2



3

## Plasmids

A plasmid vector used for cloning is specifically developed by adding certain features:

- Reduction in size of vector to a minimum to expand the capacity of vector to clone large fragments (less efficient transformation of plasmids larger than 15 kb, easier manipulation with small molecules) – therefore they should have 3-4 kb.
- It should contain an origin of replication
- Introduction of selectable markers
- Introduction of synthetic cloning sites termed polylinker , restriction site bank, or polycloning sites that are recognized by restriction enzymes
- They should have the ability to be present in cell with several copies.

For the expression of cloned DNA, the vector DNA should contain suitable control elements, such as promoters, terminators, and ribosome binding sites.

4

### Biology of *E. coli* K-12

- Stanley Cohen and Herb Boyer chose *Escherichia coli* strain K-12 for their molecular biology experiments in the early 1970s because it was easy to grow and amenable to metabolic studies.
- *E. coli* K-12 has proved to be an innocuous biological host for the propagation of recombinant DNA molecules
- The attenuated *E. coli* K-12 strain does not thrive outside of the laboratory environment and it is unable to compete against the more genetically robust *E. coli* serotypes normally found in the human intestine
- *E. coli* cells undergo cell division every 20 minutes
- *E. coli* K-12 is a safe, nonpathogenic bacterium, but specific laboratory practices are always warranted and specific biosafety guidelines must be followed for recombinant DNA research.
- *E. coli* K-12 strains used for gene cloning contain mutations in the *recA* gene. *recA*-enhances the biosafety of *E. coli* K-12 strains because *recA* - strains are sensitive to ultraviolet light.

5

### Plasmids

- Plasmids are double-stranded, closed circular DNA molecules, which exist in the cell as extrachromosomal units.
- They are self-replicating, and found in a variety of bacterial species, where they behave as accessory genetic units.
- There are three general classes of plasmids: (i) virulence plasmids, which encode toxin genes, (ii) drug resistance plasmids, which confer resistance to antibiotics, and (iii) plasmids, which encode genes required for bacterial conjugation (F plasmid, *tra* genes)).
- Plasmids range in size from 1 to 200 kb
- The best-studied plasmid replication system is that of *ColE1*, which is a small *E. coli* plasmid encoding antibacterial proteins called colicins.

6

## Vectors for *E. coli*

### Plasmid pBR322

- One of the first artificial cloning vectors to be constructed (1977)
- It is a 4363 bp double stranded cloning vector.
- This plasmid vector has been put together from fragments originating from three different naturally occurring plasmids.
- It contains two antibiotic resistance genes, one is the ampicillin resistance (*amp<sup>R</sup>*) gene coding for a  $\beta$ -lactamase (which modifies ampicillin into a form that is nontoxic to *E. coli*) and the other is the tetracycline-resistance (*tet<sup>R</sup>*) gene (a set of genes coding for enzymes that detoxify tetracycline) (**SELECTABLE MARKER GENES**).
- The plasmid contains 20 unique recognition sites for restriction enzymes. Six of these sites, i.e. *EcoRV*, *Bam*HI, *Sph*I, *Sal*I, *Xma*III and *Nru*I, are located within the gene coding for tetracycline resistance; two sites for *Hind* III and *Cla*I lie within the promoter of the tetracycline resistance gene; and three sites for *Pst*I, *Pvu*I and *Sca*I lie within the  $\beta$ -lactamase gene.

7

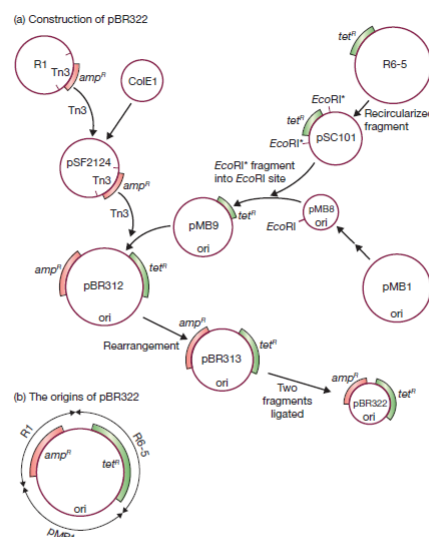


Figure 6.2

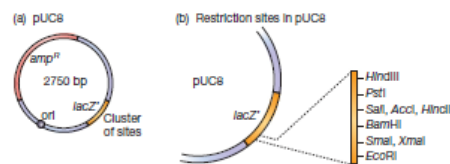
The pedigree of pBR322. (a) The manipulations involved in construction of pBR322. The *amp<sup>R</sup>* gene was obtained from Tn3, a type of **transposable element** carried by the R1 plasmid. The *tet<sup>R</sup>* gene was excised from pSC101 by treatment with *Eco*R1 in a low-salt solution, which decreases the specificity of the enzyme so that, as well as cutting at its standard GAATTC recognition sequence, it also cuts at related sequences such as TAATTC. This is called **star activity**, and the related sequences are referred to as *Eco*R1\* sites. (b) A summary of the origins of pBR322.

Brown, T. A. (2010). *Gene Cloning and DNA Analysis: An Introduction*. Wiley.

8

### pUC8 – a *Lac* selection plasmid

- pUC (University of California, Norrander in sod. 1983)
- 2.7 kb and possess the *ColE1* ori of replicaton
- It includes a new gene *lacZ*, which was derived from the *lac* operon of *E. coli* that codes for beta-galactosidase.
- *amp<sup>R</sup>* has been changed so that it no longer contains the unique restriction sites. All of these cloning sites are now clustered into a short segment of the *lacZ'* gene.
- 500-700 copies
- identification of recombinant cells can be achieved
- by a single-step process, by plating onto agar medium containing ampicillin plus X-gal



**Figure 6.3**

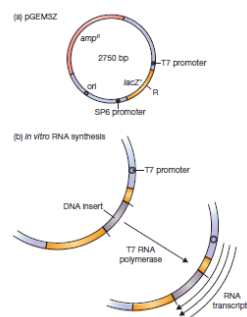
The pUC plasmids. (a) The structure of pUC8. (b) The restriction site cluster in the *lacZ'* gene of pUC8.

Brown, T. A. (2010). *Gene Cloning and DNA Analysis: An Introduction*: Wiley.

9

### pGEM3Z – in vitro transcription of cloned DNA

- pGEM3Z is very similar to a pUC vector. The distinction is that pGEM3Z has two additional short pieces of DNA, each of which acts as the recognition site for attachment of an RNA polymerase enzyme (not recognized by the *E. coli* RNA polymerase, but by RNA polymerase coded by T7 bacteriophage and SP6 phage – they are synthesized during the infection of **E. coli** and are responsible for transcribing the phage genes).
- These two promoter sequences lie on either side of the cluster of restriction sites used for the introduction of new DNA into the pGEM3Z molecule (MCS, multiple cloning site).



**Figure 6.4**

pGEM3Z. (a) Map of the vector. 'R' indicates a polylinker containing restriction sites for EcoRI, SacI, KpnI, Aval, SmaI, BamHI, XbaI, SalI, AccI, HincII, PstI, SphI, and HindIII. (b) In vitro RNA synthesis from the T7 promoter.

Brown, T. A. (2010). *Gene Cloning and DNA Analysis: An Introduction*: Wiley.

10

Screening for pBR322 recombinants by insertional inactivation

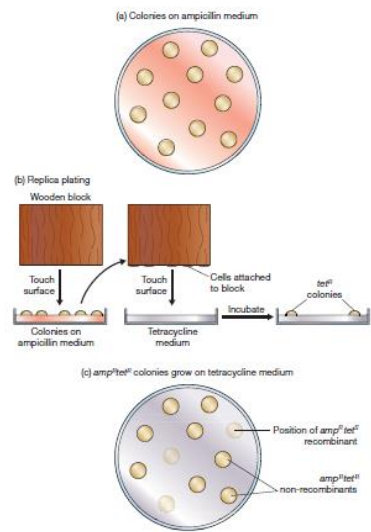


Figure 5.8  
Screening for pBR322 recombinants by insertional inactivation of the tetracycline resistance gene. (a) Cells are plated onto ampicillin agar: all the transformants produce colonies. (b) The colonies are replica plated onto tetracycline medium. (c) The colonies that grow on tetracycline medium are amp<sup>r</sup>/ter<sup>r</sup> and therefore non-recombinants. Recombinants (amp<sup>r</sup>/ter<sup>s</sup>) do not grow, but their position on the ampicillin plate is now known.

Brown, T. A. (2010). *Gene Cloning and DNA Analysis: An Introduction*: Wiley.

11

Screening for pUC recombinants by insertional inactivation

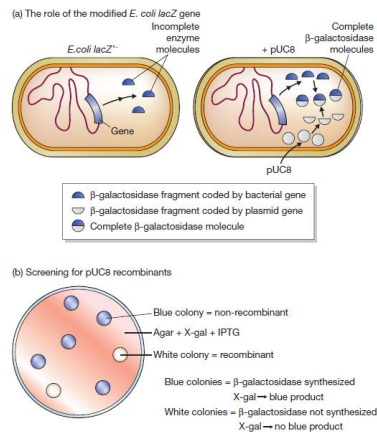
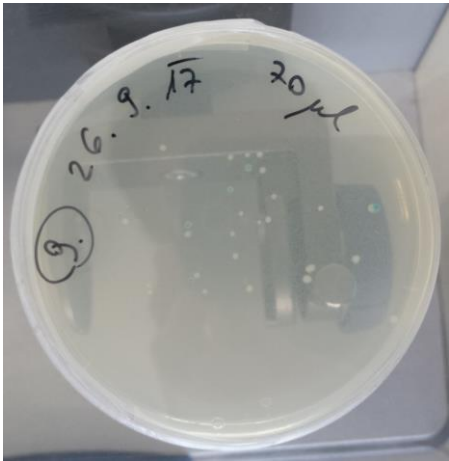


Figure 5.10  
The rationale behind insertional inactivation of the *lacZ* gene carried by pUC8. (a) The bacterial and plasmid genes complement each other to produce a functional β-galactosidase molecule. (b) Recombinants are screened by plating onto agar containing X-gal and IPTG.

Brown, T. A. (2010). *Gene Cloning and DNA Analysis: An Introduction*: Wiley.

12

Agar plates 24 hours after transformation.



13

BACTERIOPHAGES OR PHAGES

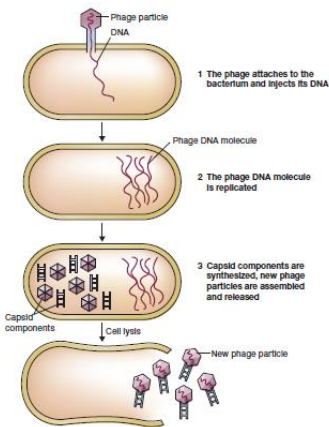


Figure 2.6  
The general pattern of infection of a bacterial cell by a bacteriophage.

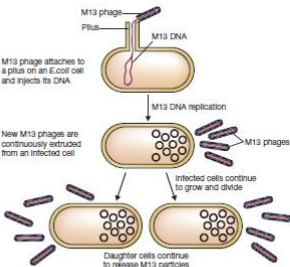


Figure 2.8  
The infection cycle of bacteriophage M13.

14

### Lambda bacteriophage as a vector

- Typical example of head-and-tail phage
- Genome – 48, 5 kb packed in the polyhedral head structure
- *Helper virus* (supply only the missing genes for the coat protein)
- At either end of the molecule is a short 12 nt stretch of single stranded DNA (cohesive or sticky ends); they are complementary and they can form circular DNA.
- An insert 37 do 52 kb

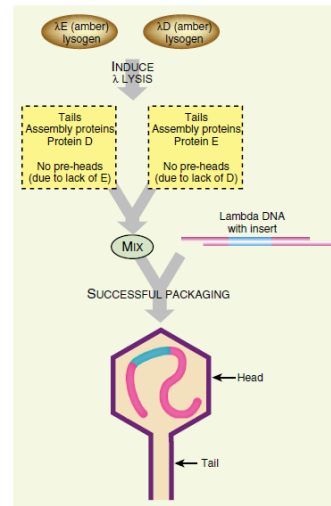


FIGURE 3.17 In Vitro Packaging

FIGURE 3.17 In Vitro Packaging A lambda cloning vector containing cloned DNA must be packaged in a phage head before it can infect *E. coli*. First, one culture of *E. coli* cells is infected with a mutant lambda that lacks the gene for one of the head proteins called E. A different culture of *E. coli* is infected with a different mutant, which lacks the phage head protein D. The two cultures are induced to lyse, which releases the tails, assembly proteins, and head proteins, but no complete heads because of the missing proteins. When these are mixed with a lambda replacement vector, the three spontaneously form complete viral particles containing DNA. These are then used to infect *E. coli*.

Clark, D. P., & Pazdernik, N. J. (2015). *Biotechnology*. Elsevier Science.

15

### Artificial chromosomes hold the largest pieces of DNA.

These include:

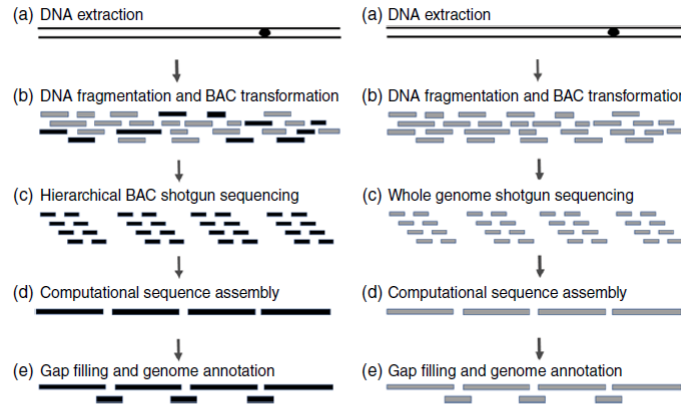
- Yeast artificial chromosomes (YACs) – up to 2000 kb,
- bacterial artificial chromosomes (BACs), and
- P1 bacteriophage artificial chromosomes (PACs).

They are used to contain lengths of DNA from 150 kb to 2000 kb. YACs hold the largest amount of DNA, up to about 2000 kb. YACs have yeast centromeres and yeast telomeres for maintenance in yeast. BACs can be circularized and grown in bacteria; therefore, they have a bacterial origin of replication and antibiotic resistance genes.

16



## Genome sequencing with BACs

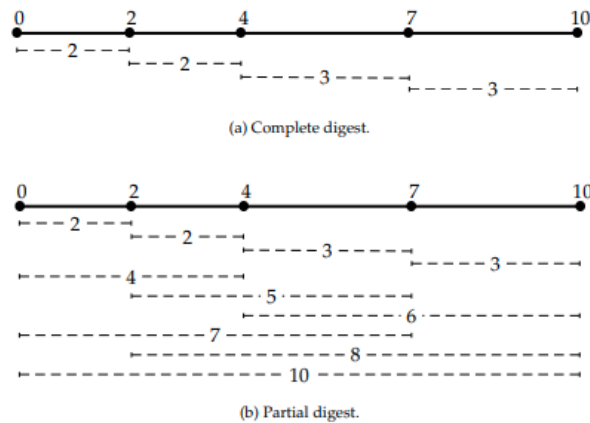


**Figure 7.9.** Overview of a genome sequencing project showing a hierarchical bacterial artificial chromosome (BAC) method (left) or whole genome shotgun sequencing (right), which contain many of the same steps: (a) Sequencing an entire genome involves the genomic DNA extraction, (b) genomic DNA fragmentation by random shearing and vector construction, (c) shotgun sequencing, (d) computational assembly of the sequences, and (e) gap filling and genome annotation (and verification). The hierarchical BAC shotgun sequencing (left) differs from whole genome shotgun sequencing (right) by selecting a minimal set of overlapping BACs from physical or genetic maps for shotgun sequencing. Putting genomic DNA into BACs is an organizational tool that can be bypassed using the whole genome shotgun approach.

Stewart, C. N. (2016). *Plant Biotechnology and Genetics: Principles, Techniques, and Applications*: Wiley.

17

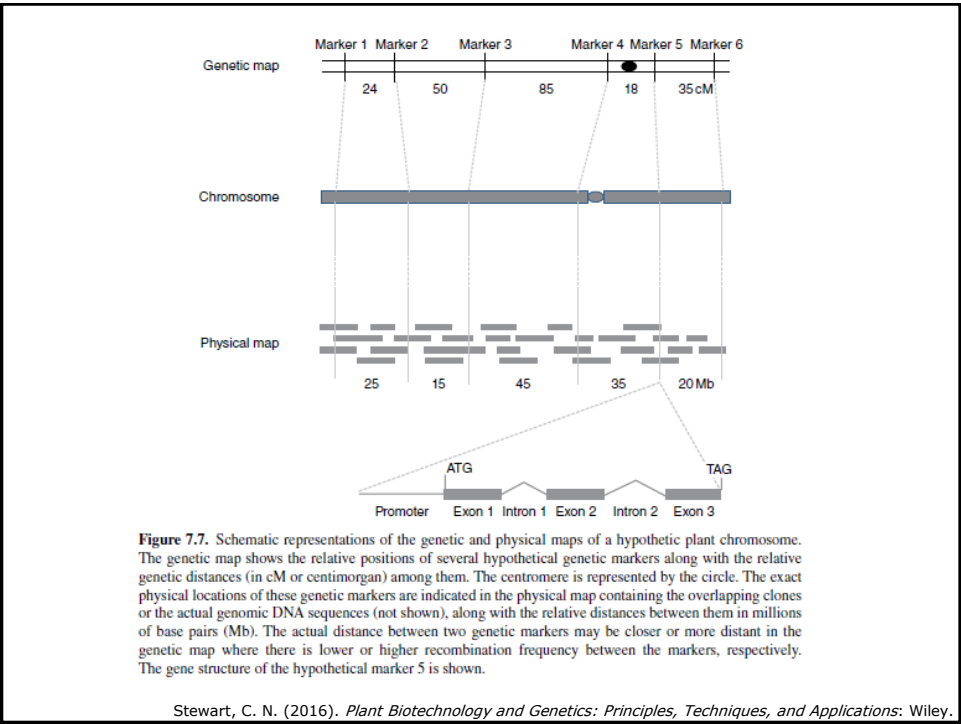
## Physical mapping



**Figure 4.1** Different methods of digesting a DNA molecule. A complete digest produces only fragments between consecutive restriction sites, while a partial digest yields fragments between any two restriction sites. Each of the dots represents a restriction site.

<http://www.cs.ukzn.ac.za/~hughm/bio/docs/IntroToBioinfAlgorithms.pdf>

18



19

DNA Vectors for plant transformation and gene transfer in plants

20

## Gene transfer methods

Various gene transfer techniques are grouped in two main categories:

- 1) Vector mediated gene transfer
- 2) Vectorless or direct DNA transfer

### Vector mediated gene transfer

- *Agrobacterium* (*A. tumefaciens*, *A. rhizogenes*),
- Viruses mediated gene transfer (Caulimoviruses, Gemini Viruses, RNA viruses) and
- transposable elements

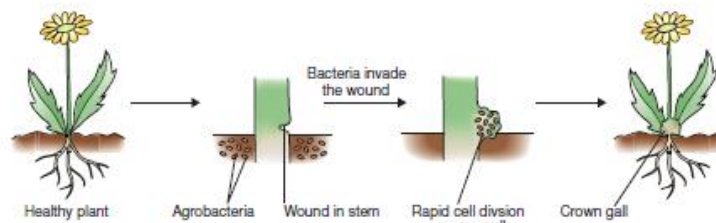
### Direct DNA transfer (naked DNA – as plasmid or linear construct)

- Physical gene transfer methods (electroporation, particle bombardment/microprojectile/biolistics, macroinjection, microinjection,...)
- Chemical gene transfer methods (PEG,...)

21

### *Agrobacterium tumefaciens*

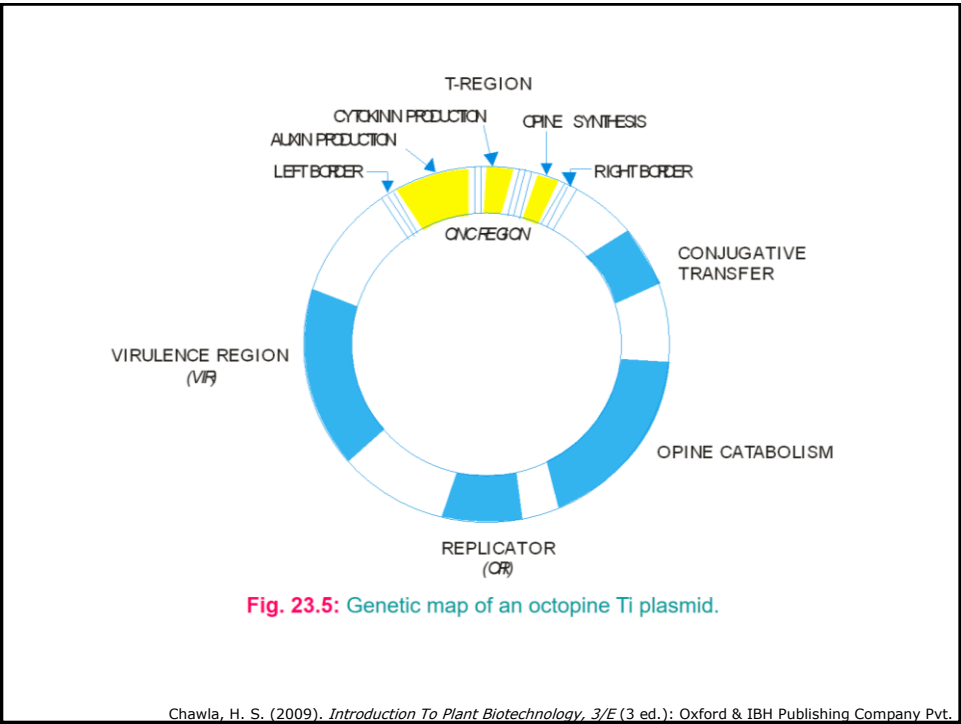
- Ti plasmid
- Up to 200 kb!
- T-DNA – T-DNA is the transferred DNA or transforming DNA of Ti plasmid (15 - 30 kb)
- The T-DNA is flanked by 25-bp direct repeats, known as the left border (LB) and right border (RB) sequences.



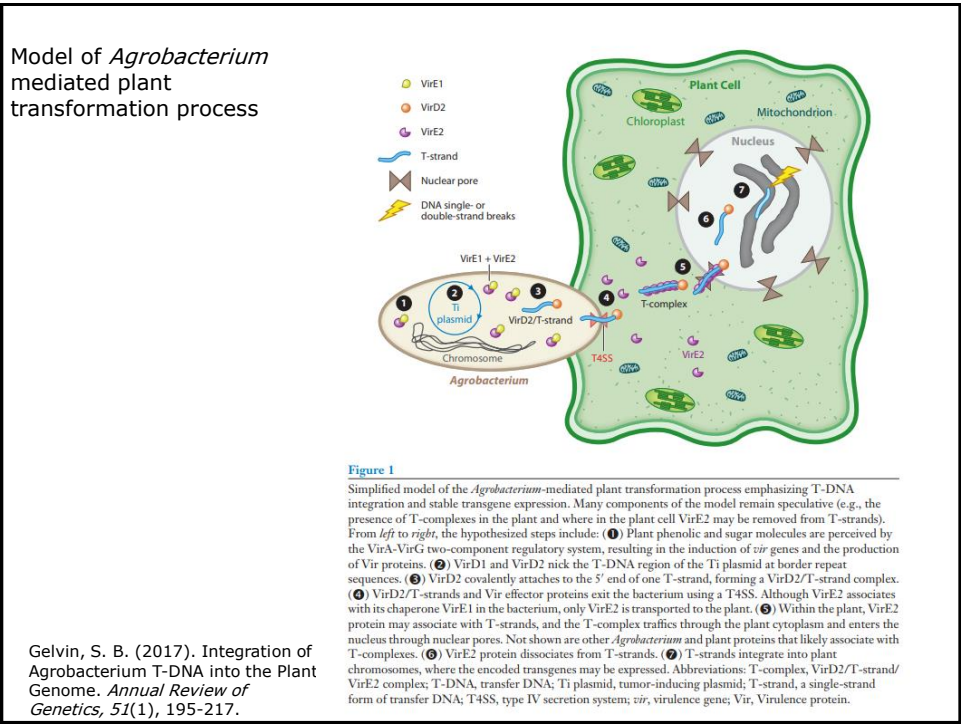
**Figure 7.9**  
Crown gall disease.

Brown, T. A. (2010). *Gene Cloning and DNA Analysis: An Introduction*. Wiley.

22



23

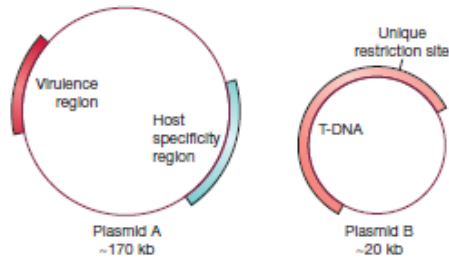


24

## BINARY VECTOR SYSTEM – two plasmids are required

The binary vector system consists of two autonomously replicating plasmids within *A. tumefaciens*:

- A shuttle (more commonly referred to as a binary) vector that contains gene of interest between the T-DNA borders and
- A helper Ti plasmid that provides the vir gene products

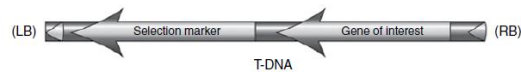


**Figure 7.11**

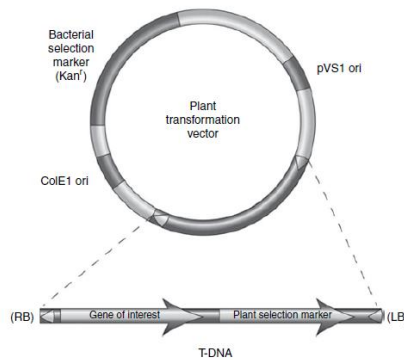
The binary vector strategy. Plasmids A and B complement each other when present together in the same *A. tumefaciens* cell. The T-DNA carried by plasmid B is transferred to the plant chromosomal DNA by proteins coded by genes carried by plasmid A.

Brown, T. A. (2010). *Gene Cloning and DNA Analysis: An Introduction*: Wiley.

25



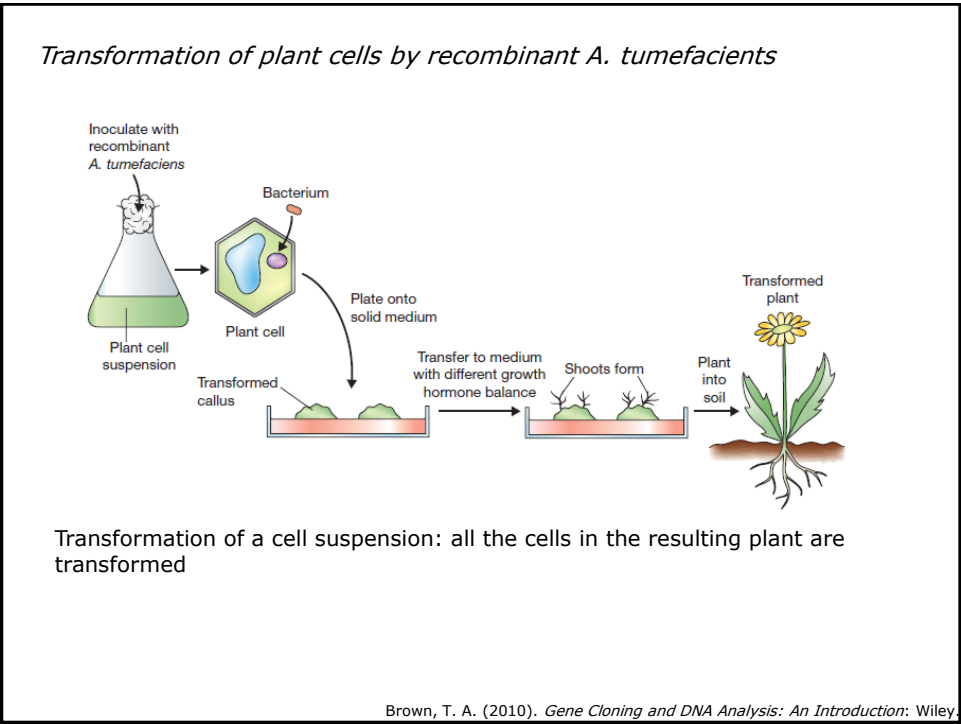
**Figure 8.7.** T-DNA used to genetically engineer plants frequently contains a selectable marker gene under the transcriptional control of a constitutively and ubiquitously active promoter to ensure gene expression in all tissues at all stages of development, together with the gene of interest (GOI) providing a novel phenotype for the plant.



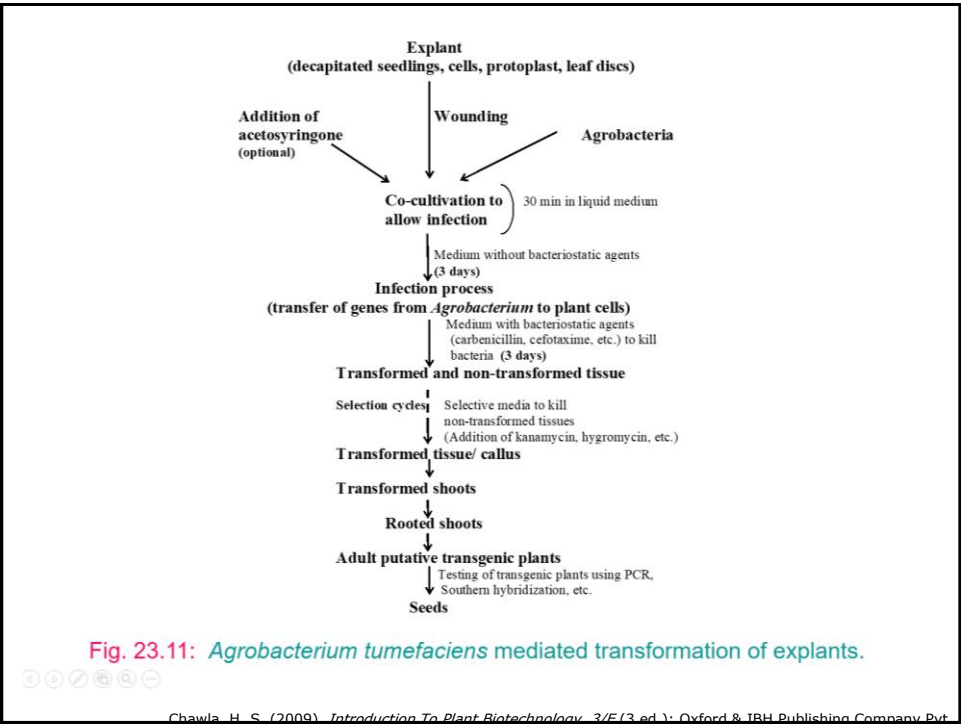
**Figure 8.8.** A generic plant binary vector with two origins of replication, the pVS1 ori for propagation in *Agrobacterium* and the ColE1 ori for propagation in *E. coli*. The backbone of the vector contains an antibiotic resistance gene for bacterial selection (kanamycin resistance), and the T-DNA contains a plant selectable marker and the gene of interest (GOI).

Stewart, C. N. (2016). *Plant Biotechnology and Genetics: Principles, Techniques, and Applications*: Wiley.

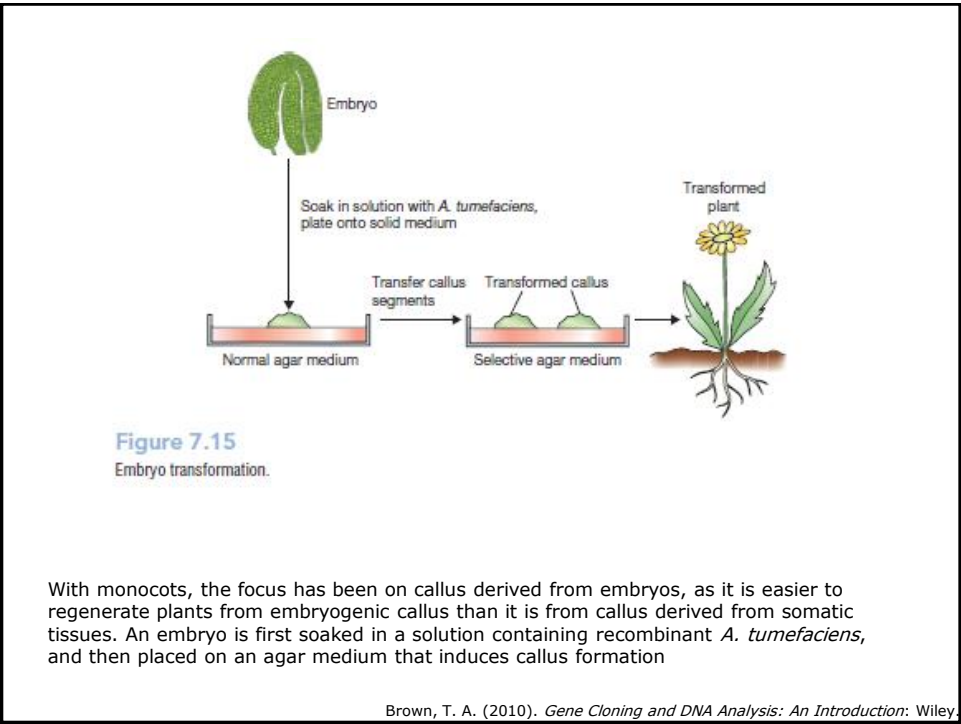
26



27



28

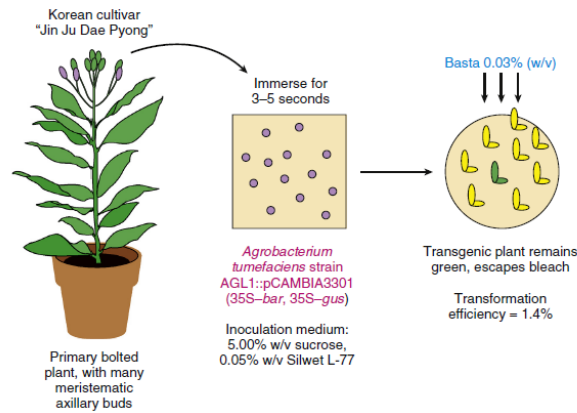


29



30

## Floral dip method

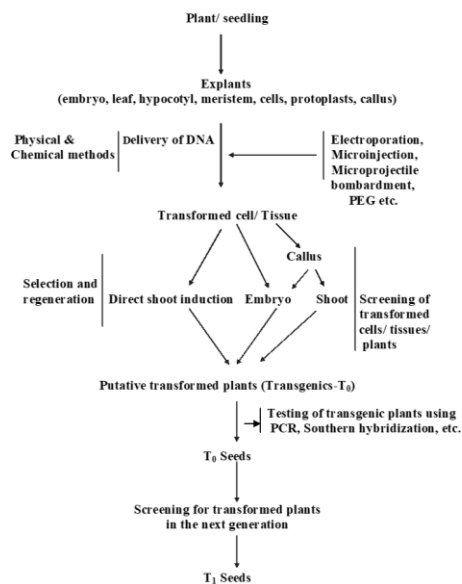


**FIGURE 15.9 Floral Dip Method of Plant Transformation**  
Flower buds exposed to *Agrobacterium* containing modified T-DNA can result in the production of transgenic seeds. Adapted from Curtis IS (2003). The noble radish: past, present and future. *Trends Plant Sci* 8, 305-307.

The floral dip methods results in generation of independently transformed seeds, as a result of *Agrobacterium*-mediated transformation of the female gametophyte or the egg.

Clark, D. P., & Pazdernik, N. J. (2015). *Biotechnology*. Elsevier Science.

31



**Fig. 23.14: A general scheme for the production of transgenic plants using direct gene delivery methods.**

Chawla, H. S. (2009). *Introduction To Plant Biotechnology*. 3/E (3 ed.): Oxford & IBH Publishing Company Pvt

32



### Particle bombardment

Particle bombardment refers to a method where heavy metal particles ( $\sim 1 \mu\text{m}$  gold or tungsten) are coated with DNA, accelerated toward the target tissue, and penetrate the cell wall to rest either adjacent to or directly in the nucleus. The DNA on the particles somehow finds its way to the native DNA of the target cell, where it becomes integrated into the chromosome to become a permanent addition to the genome.



Figure 11.7. Two different particle bombardment devices: The commercially available PDS1000 He (BioRad) (left) and the non-commercial Particle Inflow Gun (right).

Stewart, C. N. (2016). *Plant Biotechnology and Genetics: Principles, Techniques, and Applications*: Wiley.

33

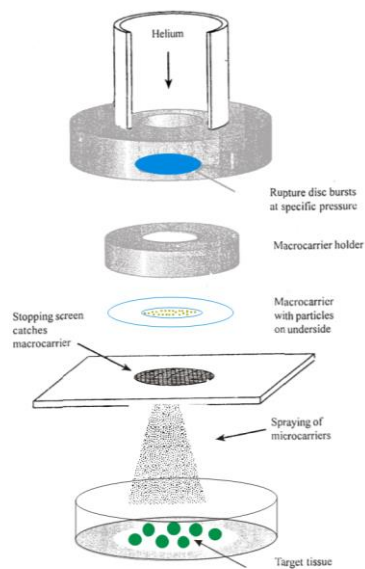


Fig 23.15: Schematic representation of the particle bombardment process.

3 2 1 0

Chawla, H. S. (2009). *Introduction To Plant Biotechnology*. 3/E (3 ed.): Oxford & IBH Publishing Company Pvt

34

### Virus mediated gene transfer

The potential of plant viruses as cloning vectors has been explored for several years, but without great success. One problem here is that the vast majority of plant viruses have genomes not of DNA but of RNA. RNA viruses are less useful as potential cloning vectors because manipulations with RNA are more difficult to carry out. Only two classes of DNA virus are known to infect higher plants – the caulimoviruses and geminiviruses – and neither is ideally suited to gene cloning.

#### Caulimovirus vectors

- the capacity for carrying inserted DNA is still very limited.
- extremely narrow host range of caulimoviruses. This restricts cloning experiments to just a few plants, mainly brassicas such as turnips, cabbages and cauliflowers.

#### Geminivirus vectors

- during the infection cycle the genomes of some geminiviruses undergo rearrangements and deletions, which would scramble up any additional DNA that has been inserted
- **virus-induced gene silencing (VIGS)**, a technique used to investigate the functions of individual plant genes.

35

### Modern cloning strategies

36



37

Modern cloning strategies – they are not based on restriction enzymes and ligases:

- 1) Strategies that rely on site-specific DNA recombination techniques (e.g. Gateway Cloning (ThermoFisher))
- 2) Strategies that rely on DNA end-linking assembly techniques (e.g. Gibson Assembly Cloning (New England Biolabs))

38

GATEWAY® CLONING VECTORS

A newer cloning system uses the lambda phage integration and excision sites for cloning genes from one vector to another.

The integration reaction occurs when integrase makes staggered cuts in the center of the phage attP site and in the center of the bacterial site attB (after integration sequences are different than the original).

This reaction can be reversed, but since the two sites are different after integration, another enzyme called Xis, or excisionase, removes the inserted DNA and relegates the broken DNA.

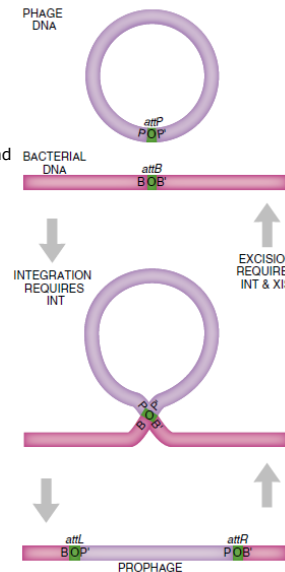


FIGURE 3.27 Integration of Lambda DNA

Phage DNA has an attachment sequence called attP. Bacterial DNA has an attachment sequence called attB. Bacterial DNA and λ-phage DNA align at the “O” region of the attachment sequences. During integration, int protein induces two double-stranded breaks that are resolved, resulting in the integration of the phage DNA into the bacterial DNA. The process is reversible, and requires int protein and xis protein to excise the phage DNA from the bacterial DNA. Notice that the integrated phage DNA “O” site is flanked with one side from the phage and one side from the bacteria. These are called the attL and attR sites.

Clark, D. P., & Pazdernik, N. J. (2015). *Biotechnology*. Elsevier Science.

Site-specific recombination involves breakage and reunion

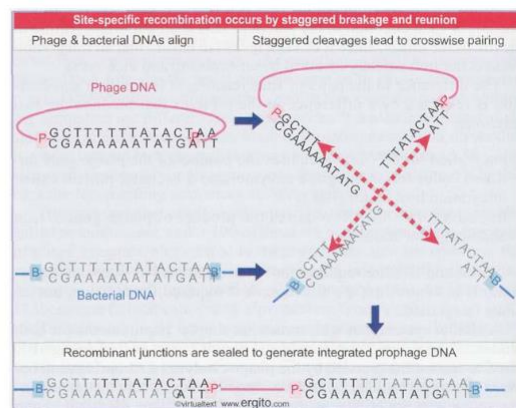


Figure 15.27 Staggered cleavages in the common core sequence of attP and attB allow crosswise reunion to generate reciprocal recombinant junctions.

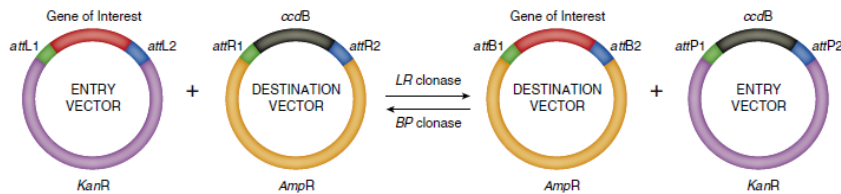
Levin. 2003. Genes VIII

The first step of the system is to get the gene of interest between two attL sites. This can be done by cloning the gene into a multicloning site found in the entry clone. The entry clone has a gene called ccdB in the middle of the multicloning site, which encodes a toxin that kills the host when expressed unless the bacteria has a corresponding gene for an antitoxin.

BP clonase (integrase) – recombination of attB and attP sites and attL in attR sites are created.  
LR clonase (excisionase) – reverse reaction

Modified att sites ensure that the orientation of the DNA fragment is maintained during the excision and integration process.

Only bacteria containing the desired recombinant construct that lacks the ccdB gene and contains the appropriate antibiotic resistance marker gene can survive.



**FIGURE 3.29 Gateway® BP and LR Reactions**

Moving a gene of interest from the entry clone to the destination vector is done in the LR reaction. The excisionase and integrase enzymes work to remove the gene of interest in the entry clone by cutting at the attL and attR sites of the entry clone and destination vector. The gene of interest and ccdB swap positions, therefore changing the att site to become attB and attP. The BP reaction works in reverse, moving the gene of interest back into the entry clone.

Clark, D. P., & Pazdernik, N. J. (2015). *Biotechnology*. Elsevier Science.

41

**An Entry clone** contains your gene of interest flanked by attL sequences, which are then used to recombine with attR sequences to create your desired expression clone.



**FIGURE 3.28 Gateway® Entry Clone**

The entry clone for the Gateway® system has an origin of replication for growing in bacteria, an antibiotic resistance gene for selecting bacteria with the vector, a multicloning site containing the gene ccdB in between two attL sites (attL1 and attL2). The gene of interest replaces the ccdB gene during standard cloning using unique restriction enzyme sites. The ccdB gene produces a toxin that kills its host bacteria unless the bacteria has a corresponding gene for an antitoxin.

#### Building an Entry clone

The three possible methods that lead to the Entry clone are depicted in Figure 1. Using TOPO vectors and PCR amplification/restriction-enzyme vectors are the most common ways to construct your own Entry clone.

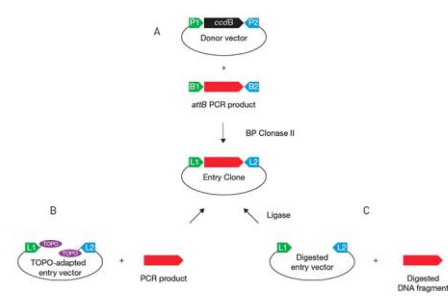


Figure 1. Strategies to build the Entry clone. The three possible methods that lead to the Entry clone are depicted: (A) BP cloning, (B) TOPO cloning, and (C) restriction enzyme and ligation cloning. Red arrows represent the fragment of interest. Adapted from Kallan F (2007) Expert Opin Drug Discov 2(4):571-588.

<https://www.thermofisher.com/si/en/home/life-science/cloning/gateway-cloning/entry-clones.html>

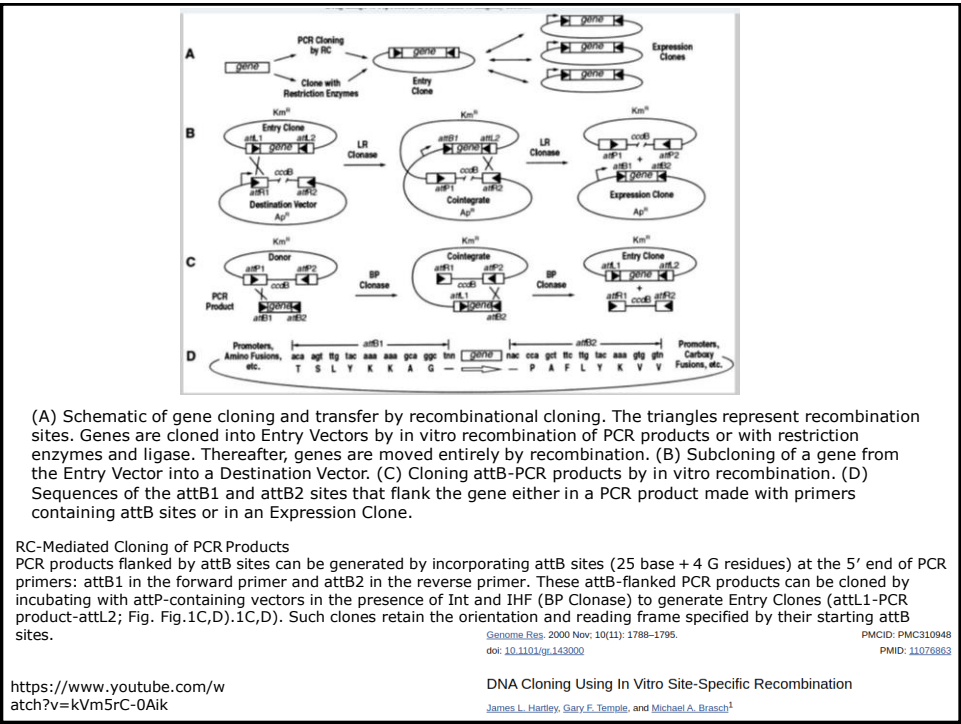
#### TOPO CLONING

<https://www.thermofisher.com/si/en/home/life-science/cloning/topo/topo-resources/the-technology-behind-topo-cloning.html>

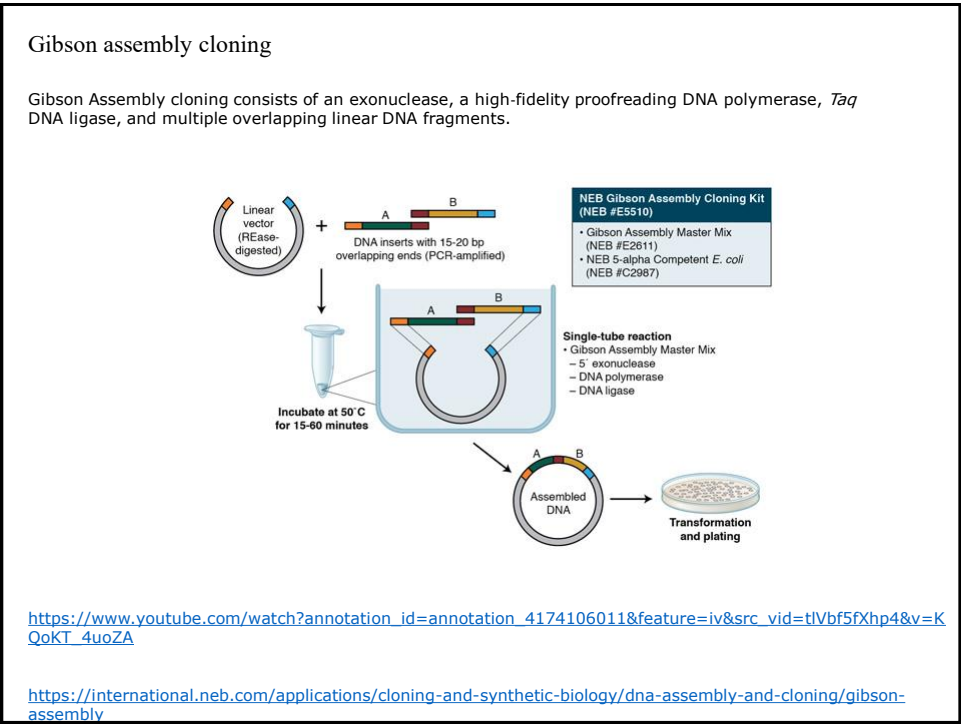
Additional explanation (addgene):

<https://blog.addgene.org/plasmids-101-gateway-cloning>

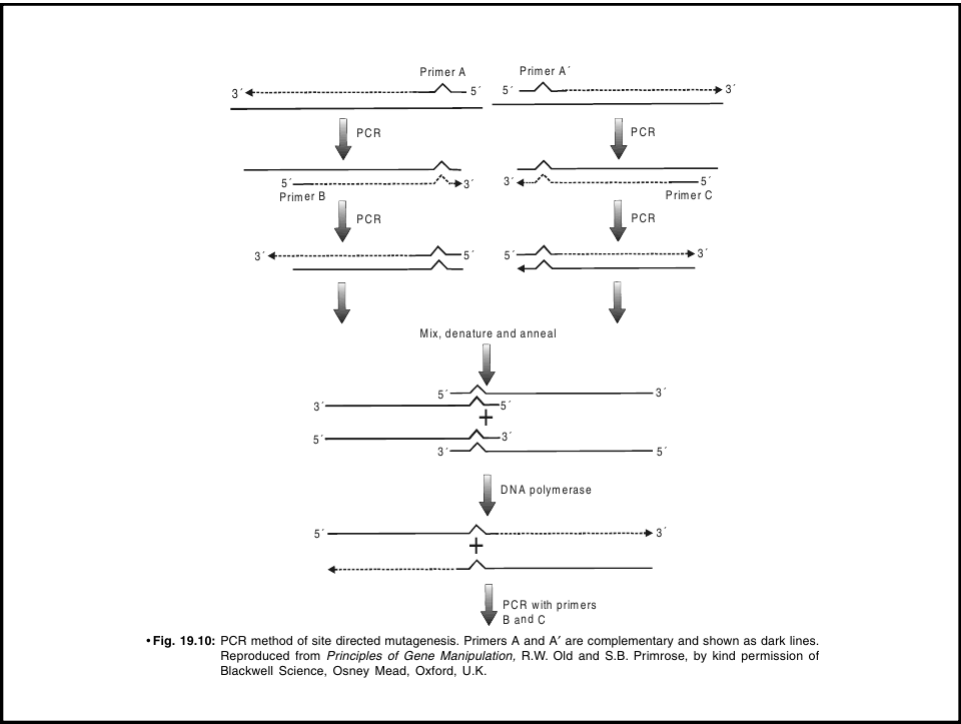
42



43



44



45

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7958140/>

[Genome Med.](#), 2021; 13: 43. PMCID: PMC7958140  
 Published online 2021 Mar 15. doi: [10.1186/s13073-021-00859-1](https://doi.org/10.1186/s13073-021-00859-1) PMID: [33722288](https://pubmed.ncbi.nlm.nih.gov/33722288/)

**SARS-CoV-2 vaccine ChAdOx1 nCoV-19 infection of human cell lines reveals low levels of viral backbone gene transcription alongside very high levels of SARS-CoV-2 S glycoprotein gene transcription**

[Abdulaziz Almuqrin](#),<sup>1,2</sup> [Andrew D. Davidson](#),<sup>1</sup> [Maia Kavanagh Williamson](#),<sup>1</sup> [Phillip A. Lewis](#),<sup>1</sup> [Kate J. Heesom](#),<sup>3</sup> [Susan Morris](#),<sup>4</sup> [Sarah C. Gilbert](#),<sup>4</sup> and [David A. Matthews](#)<sup>int1</sup>

► [Author information](#) ► [Article notes](#) ► [Copyright and License information](#) [Disclaimer](#)

*Hence the vector construct is replication defective. Expression of the SARS-CoV-2 spike glycoprotein in the vector construct was posited under control of the Tet-responsive CMV [cytomegalovirus] promoter (Loew et al. 2010), one of the strong eukaryotic promoters.*

46

## Case studies on genetically modified organisms (GMOs): Potential risk scenarios and associated health indicators.

de Santis B<sup>1</sup>, Stockhofe N<sup>2</sup>, Wal JM<sup>3</sup>, Weesendorp E<sup>2</sup>, Lallès JP<sup>4</sup>, van Dijk J<sup>5</sup>, Kok E<sup>5</sup>, De Giacomo M<sup>6</sup>, Einspanier R<sup>7</sup>, Onori R<sup>6</sup>, Brera C<sup>6</sup>, Bikker P<sup>8</sup>, van der Meulen J<sup>8</sup>, Kleter G<sup>5</sup>.

### Author information

#### Abstract

Within the frame of the EU-funded MARLON project, background data were reviewed to explore the possibility of measuring health indicators during post-market monitoring for potential effects of feeds, particularly genetically modified (GM) feeds, on livestock animal health, if applicable. Four case studies (CSs) of potential health effects on livestock were framed and the current knowledge of a possible effect of GM feed was reviewed. Concerning allergenicity (CS-1), there are no case-reports of allergic reactions or immunotoxic effects resulting from GM feed consumption as compared with non-GM feed. The likelihood of horizontal gene transfer (HGT; CS-2) of GMO-related DNA to different species is not different from that for other DNA and is unlikely to raise health concerns. Concerning mycotoxins (CS-3), insect-resistant GM maize may reduce fumonisins contamination as a health benefit, yet other Fusarium toxins and aflatoxins show inconclusive results. For nutritionally altered crops (CS-4), the genetic modifications applied lead to compositional changes which require special considerations of their nutritional impacts. No health indicators were thus identified except for possible beneficial impacts of reduced mycotoxins and nutritional enhancement. More generally, veterinary health data should ideally be linked with animal exposure information so as to be able to establish cause-effect relationships.