DNA7 8/22/05 10:17 AM Page 132

Chapter 7 Cloning Vectors for Eukaryotes

Vectors for yeast and other fungi, 132
Cloning vectors for higher plants, 141
Cloning vectors for animals, 151

Most cloning experiments are carried out with *E. coli* as the host, and the widest variety of cloning vectors are available for this organism. *E. coli* is particularly popular when the aim of the cloning experiment is to study the basic features of molecular biology such as gene structure and function. However, under some circumstances it may be desirable to use a different host for a gene cloning experiment. This is especially true in biotechnology (Chapter 13), where the aim may not be to study a gene, but to use cloning to control or improve synthesis of an important metabolic product (e.g. a hormone such as insulin), or to change the properties of the organism (e.g. to introduce herbicide resistance into a crop plant). We must therefore consider cloning vectors for organisms other than *E. coli*.

7.1 Vectors for yeast and other fungi

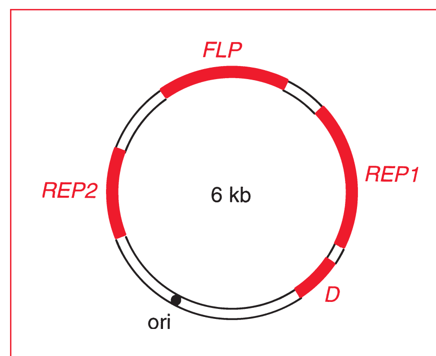
The yeast *Saccharomyces cerevisiae* is one of the most important organisms in biotechnology. As well as its role in brewing and breadmaking, yeast has been used as a host organism for the production of important pharmaceuticals from cloned genes (p. 292). Development of cloning vectors for yeast has been stimulated greatly by the discovery of a plasmid that is present in most strains of *S. cerevisiae* (Figure 7.1). The 2 μm plasmid, as it is called, is one of only a very limited number of plasmids found in eukaryotic cells.

7.1.1 Selectable markers for the 2 μm plasmid

The 2 μm plasmid is an excellent basis for a cloning vector. It is 6 kb in size, which is ideal for a vector, and exists in the yeast cell at a copy number of between 70 and 200. Replication makes use of a plasmid origin, several enzymes provided by the host cell, and the proteins coded by the *REP1* and *REP2* genes carried by the plasmid.

DNA7 8/22/05 10:17 AM Page 133

Figure 7.1 The yeast 2 μm plasmid. *REP1* and *REP2* are involved in replication of the plasmid, and *FLP* codes for a protein that can convert the A form of the plasmid (shown here) to the B form, in which the gene order has been rearranged by intramolecular recombination. The function of *D* is not exactly known.



However, all is not perfectly straightforward in using the 2 μm plasmid as a cloning vector. First, there is the question of a selectable marker. Some yeast cloning vectors carry genes conferring resistance to inhibitors such as methotrexate and copper, but most of the popular yeast vectors make use of a radically different type of selection system. In practice a normal yeast gene is used, generally one that codes for an enzyme involved in amino acid biosynthesis. An example is the gene *LEU2*, which codes for β-isopropyl-malate dehydrogenase, one of the enzymes involved in the conversion of pyruvic acid to leucine.

In order to use *LEU2* as a selectable marker, a special kind of host organism is needed. The host must be an **auxotrophic** mutant that has a non-functional *LEU2* gene. Such a *leu2*$^-$ yeast is unable to synthesize leucine and can survive only if this amino acid is supplied as a nutrient in the growth medium (Figure 7.2(a)). Selection is possible because transformants contain a plasmid-borne copy of the *LEU2* gene, and so are able to grow in the absence of the amino acid. In a cloning experiment, cells are plated out onto **minimal medium**, which contains no added amino acids. Only transformed cells are able to survive and form colonies (Figure 7.2(b)).

7.1.2 Vectors based on the 2 μm plasmid – yeast episomal plasmids

Vectors derived from the 2 μm plasmid are called **yeast episomal plasmids (YEps)**. Some YEps contain the entire 2 μm plasmid, others include just the 2 μm origin of replication. An example of the latter type is YEp13 (Figure 7.3).

YEp13 illustrates several general features of yeast cloning vectors. First, it is a **shuttle vector**. As well as the 2 μm origin of replication and the selectable *LEU2* gene, YEp13 also includes the entire pBR322 sequence, and can therefore replicate and be selected for in both yeast and *E. coli*. There are several lines of reasoning behind the use of shuttle vectors. One is that it may be difficult to recover the recombinant DNA molecule from a transformed yeast colony. This is not such a problem with YEps, which are present in yeast cells primarily as plasmids, but with other yeast vectors, which may integrate into

DNA7 8/22/05 10:17 AM Page 134

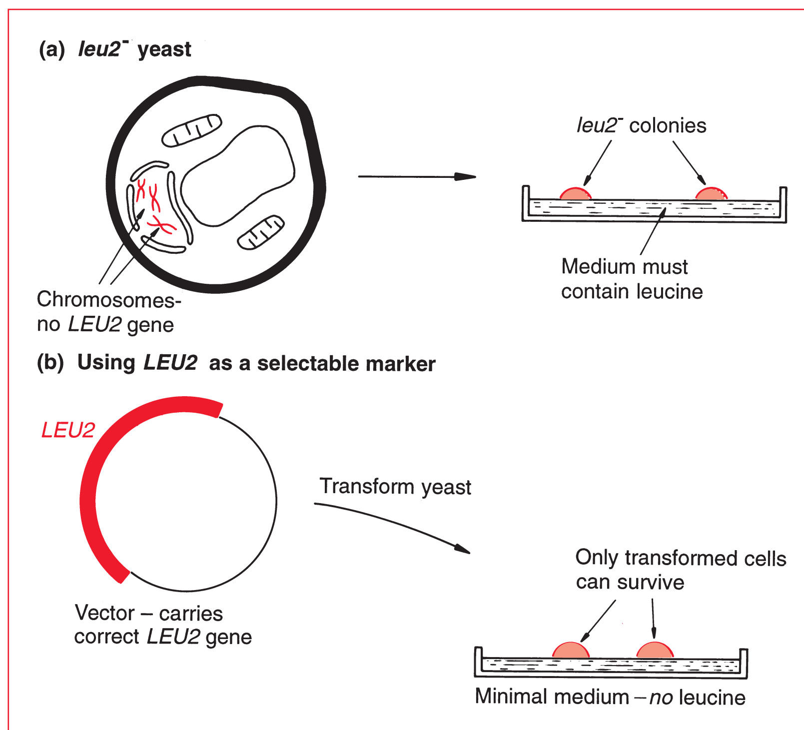


Figure 7.2 Using the *LEU2* gene as a selectable marker in a yeast cloning experiment.

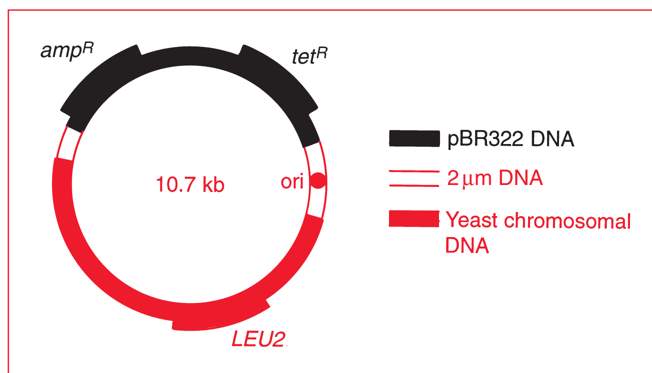


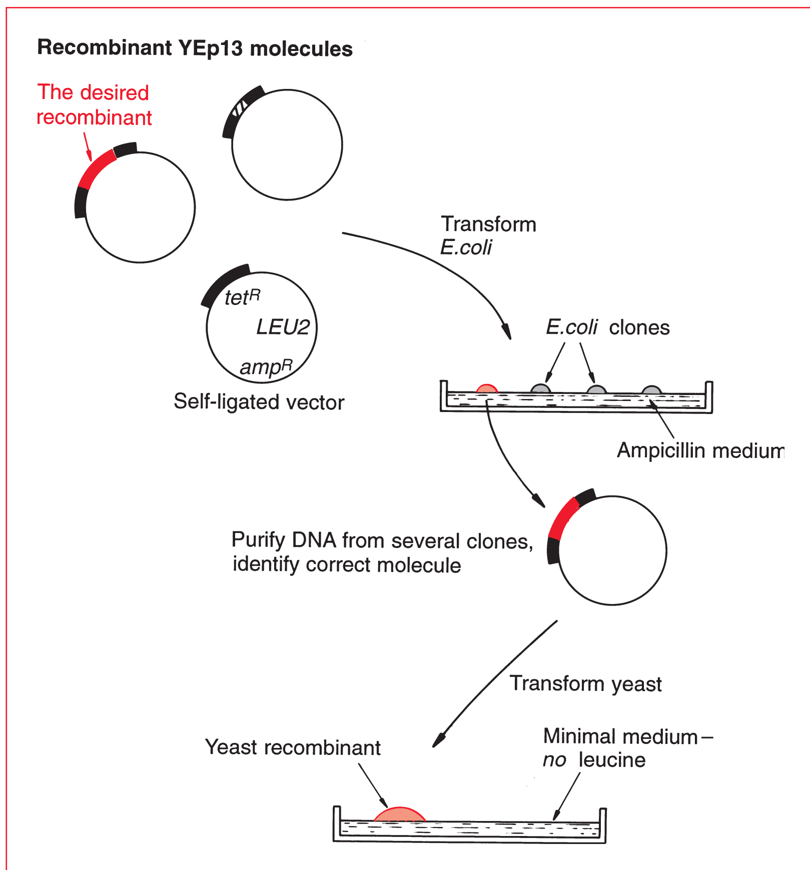
Figure 7.3 A yeast episomal plasmid, YEp13.

one of the yeast chromosomes (p. 135), purification may be impossible. This is a disadvantage because in many cloning experiments purification of recombinant DNA is essential in order for the correct construct to be identified by, for example, DNA sequencing.

The standard procedure when cloning in yeast is therefore to perform the initial cloning experiment with *E. coli*, and to select recombinants in this organism. Recombinant plasmids can then be purified, characterized, and the correct molecule introduced into yeast (Figure 7.4).

DNA7 8/22/05 10:17 AM Page 135

Figure 7.4 Cloning with an *E. coli*–yeast shuttle vector such as YEp13.



7.1.3 A YEp may insert into yeast chromosomal DNA

The word 'episomal' indicates that a YEp can replicate as an independent plasmid, but also implies that integration into one of the yeast chromosomes can occur (see the definition of 'episome' on p. 16). Integration occurs because the gene carried on the vector as a selectable marker is very similar to the mutant version of the gene present in the yeast chromosomal DNA. With YEp13, for example, **homologous recombination** can occur between the plasmid *LEU2* gene and the yeast mutant *LEU2* gene, resulting in insertion of the entire plasmid into one of the yeast chromosomes (Figure 7.5). The plasmid may remain integrated, or a later recombination event may result in it being excised again.

7.1.4 Other types of yeast cloning vector

In addition to YEps, there are several other types of cloning vector for use with *S. cerevisiae*. Two important ones are as follows:

(1) **Yeast integrative plasmids (YIps)** are basically bacterial plasmids carrying a yeast gene. An example is YIp5, which is pBR322 with an inserted

DNA7 8/22/05 10:17 AM Page 136

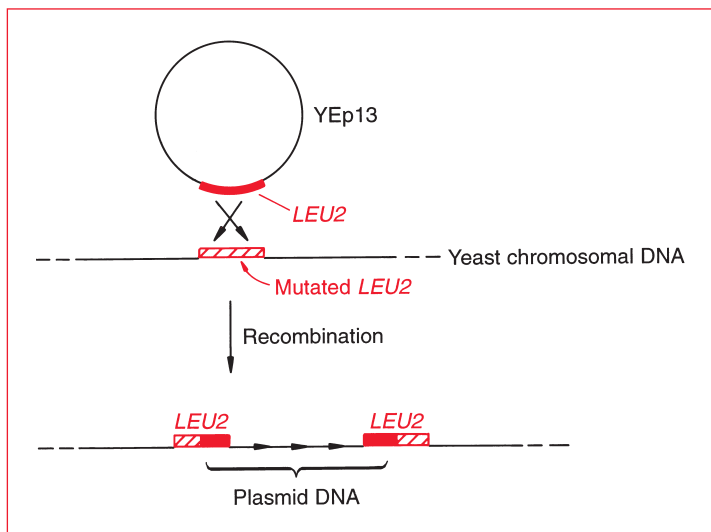


Figure 7.5 Recombination between plasmid and chromosomal *LEU2* genes can integrate YEp13 into yeast chromosomal DNA. After integration there are two copies of the *LEU2* gene; usually one is functional, and the other mutated.

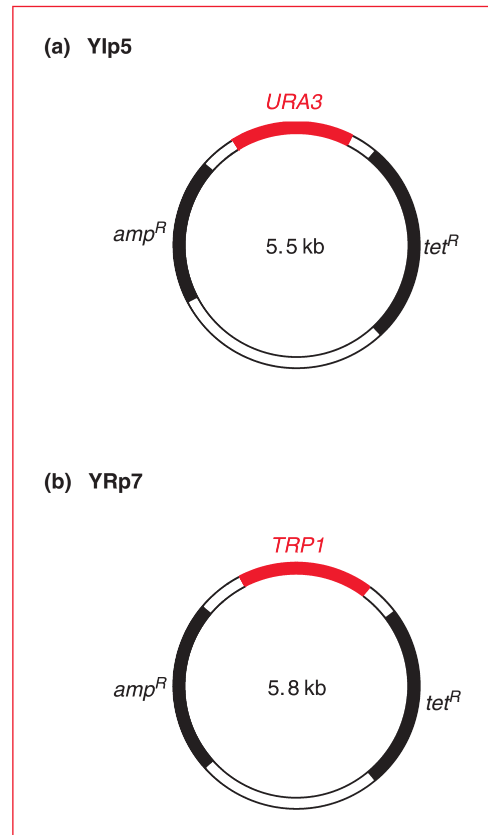
URA3 gene (Figure 7.6(a)). This gene codes for orotidine-5′-phosphate decarboxylase (an enzyme that catalyses one of the steps in the biosynthesis pathway for pyrimidine nucleotides) and is used as a selectable marker in exactly the same way as *LEU2*. A YIp cannot replicate as a plasmid as it does not contain any parts of the 2 μm plasmid, and instead depends for its survival on integration into yeast chromosomal DNA. Integration occurs just as described for a YEp (Figure 7.5).

(2) **Yeast replicative plasmids (YRps)** are able to multiply as independent plasmids because they carry a chromosomal DNA sequence that includes an origin of replication. Replication origins are known to be located very close to several yeast genes, including one or two which can be used as selectable markers. YRp7 (Figure 7.6(b)) is an example of a replicative plasmid. It is made up of pBR322 plus the yeast gene *TRP1*. This gene, which is involved in tryptophan biosynthesis, is located adjacent to a chromosomal origin of replication. The yeast DNA fragment present in YRp7 contains both *TRP1* and the origin.

Three factors come into play when deciding which type of yeast vector is most suitable for a particular cloning experiment. The first of these is **transformation frequency**, a measure of the number of transformants that can be obtained per microgram of plasmid DNA. A high transformation frequency is necessary if a large number of recombinants are needed, or if the starting DNA is in short supply. YEps have the highest transformation frequency, providing between 10 000 and 100 000 transformed cells per μg. YRps are also quite productive, giving between 1000 and 10 000 transformants per μg, but a YIp yields less than 1000 transformants per μg, and only 1–10 unless special procedures are used. The low transformation frequency of a YIp reflects the

DNA7 8/22/05 10:17 AM Page 137

Figure 7.6 A YIp and a YRp.



fact that the rather rare chromosomal integration event is necessary before the vector can be retained in a yeast cell.

Furthermore, YEps and YRps also have the highest copy numbers: 20–50 and 5–100, respectively. In contrast, a YIp is usually present at just one copy per cell. These figures are important if the objective is to obtain protein from the cloned gene, as the more copies there are of the gene the greater the expected yield of the protein product.

So why would one ever wish to use a YIp? The answer is because YIps produce very stable recombinants, as loss of a YIp that has become integrated into a chromosome occurs at only a very low frequency. On the other hand, YRp recombinants are extremely unstable, the plasmids tending to congregate in the mother cell when a daughter cell buds off, so the daughter cell is non-recombinant. YEp recombinants suffer from similar problems, though an improved understanding of the biology of the 2 µm plasmid has enabled more stable YEps to be developed in recent years. Nevertheless, a YIp is the vector of choice if the needs of the experiment dictate that the recombinant yeast cells must retain the cloned gene for long periods in culture.

DNA7 8/22/05 10:17 AM Page 138

7.1.5 Artificial chromosomes can be used to clone long pieces of DNA in yeast

The final type of yeast cloning vector to consider is the **yeast artificial chromosome (YAC)**, which presents a totally different approach to gene cloning. The development of YACs has been a spin-off from fundamental research into the structure of eukaryotic chromosomes, work that has identified the key components of a chromosome as being (Figure 7.7):

(1) The centromere, which is required for the chromosome to be distributed correctly to daughter cells during cell division.
(2) Two telomeres, the structures at the ends of a chromosome, which are needed in order for the ends to be replicated correctly and which also prevent the chromosome from being nibbled away by exonucleases.
(3) The origins of replication, which are the positions along the chromosome at which DNA replication initiates, similar to the origin of replication of a plasmid.

Once chromosome structure had been defined in this way the possibility arose that the individual components might be isolated by recombinant DNA techniques and then joined together again in the test tube, creating an artificial chromosome. As the DNA molecules present in natural yeast chromosomes are several hundred kilobases in length, it might be possible with an artificial chromosome to clone long pieces of DNA.

The structure and use of a YAC vector

Several YAC vectors have been developed but each one is constructed along the same lines, with pYAC3 being a typical example (Figure 7.8(a)). At first glance pYAC3 does not look much like an artificial chromosome, but on closer examination its unique features become apparent. pYAC3 is essentially a pBR322 plasmid into which a number of yeast genes have been inserted. Two of these genes, *URA3* and *TRP1*, have been encountered already as the selectable markers for YIp5 and YRp7, respectively. As in YRp7, the DNA fragment that carries *TRP1* also contains an origin of replication, but in pYAC3 this fragment is extended even further to include the sequence called *CEN4*,

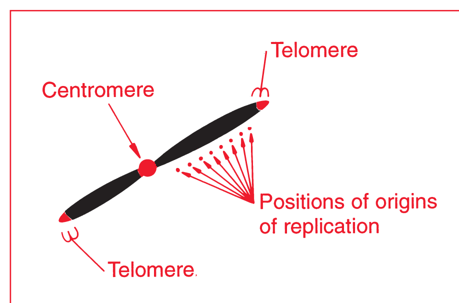


Figure 7.7 Chromosome structure.

DNA7 8/22/05 10:17 AM Page 139

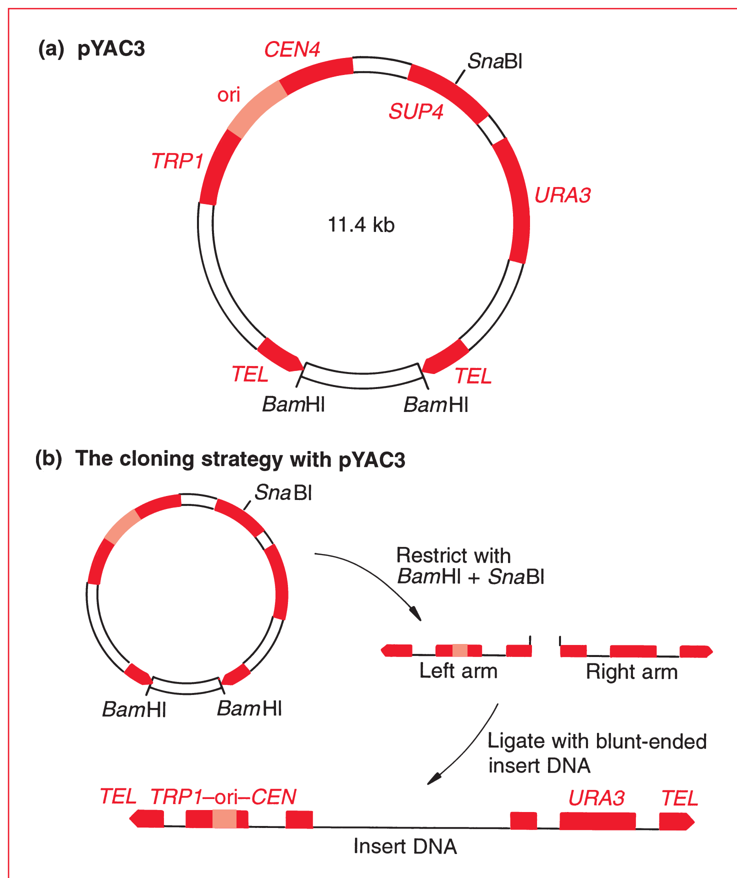


Figure 7.8 A YAC vector and the way it is used to clone large pieces of DNA.

which is the DNA from the centromere region of chromosome 4. The *TRP1*–origin–*CEN4* fragment therefore contains two of the three components of the artificial chromosome.

The third component, the telomeres, is provided by the two sequences called *TEL*. These are not themselves complete telomere sequences, but once inside the yeast nucleus they act as seeding sequences onto which telomeres will be built. This just leaves one other part of pYAC3 that has not been mentioned: *SUP4*, which is the selectable marker into which new DNA is inserted during the cloning experiment.

The cloning strategy with pYAC3 is as follows (Figure 7.8(b)). The vector is first restricted with a combination of *Bam*HI and *Sna*BI, cutting the molecule into three fragments. The fragment flanked by *Bam*HI sites is discarded, leaving two arms, each bounded by one *TEL* sequence and one *Sna*BI site. The DNA to be cloned, which must have blunt ends (*Sna*BI is a blunt end cutter, recognizing the sequence TACGTA), is ligated between the two arms, producing the artificial chromosome. Protoplast transformation (p. 105) is then used to introduce the artificial chromosome into *S. cerevisiae*. The yeast strain

DNA7 8/22/05 10:17 AM Page 140

that is used is a double auxotrophic mutant, *trp1$^-$ ura3$^-$*, which is converted to *trp1$^+$ ura3$^+$* by the two markers on the artificial chromosome. Transformants are therefore selected by plating onto minimal medium, on which only cells containing a correctly constructed artificial chromosome are able to grow. Any cell transformed with an incorrect artificial chromosome, containing two left or two right arms rather than one of each, is not able to grow on minimal medium as one of the markers is absent. The presence of the insert DNA in the vector can be checked by testing for insertional inactivation of *SUP4*, which is carried out by a simple colour test: white colonies are recombinants, red colonies are not.

Applications for YAC vectors

The initial stimulus in designing artificial chromosomes came from yeast geneticists who wanted to use them to study various aspects of chromosome structure and behaviour, for instance to examine the segregation of chromosomes during meiosis. These experiments established that artificial chromosomes are stable during propagation in yeast cells and raised the possibility that they might be used as vehicles for genes that are too long to be cloned as a single fragment in an *E. coli* vector. Several important mammalian genes are greater than 100 kb in length (e.g. the human cystic fibrosis gene is 250 kb), beyond the capacity of all but the most sophisticated *E. coli* cloning systems (p. 129), but well within the range of a YAC vector. Yeast artificial chromosomes therefore opened the way to studies of the functions and modes of expression of genes that had previously been intractable to analysis by recombinant DNA techniques. A new dimension to these experiments was provided by the discovery that under some circumstances YACs can be propagated in mammalian cells, enabling the functional analysis to be carried out in the organism in which the gene normally resides.

Yeast artificial chromosomes are equally important in the production of gene libraries. Recall that with fragments of 300 kb, the maximum insert size for the highest capacity *E. coli* vector, some 30 000 clones are needed for a human gene library (p. 129). However, YAC vectors are routinely used to clone 600 kb fragments, and special types are able to handle DNA up to 1400 kb in length, the latter bringing the size of a human gene library down to just 6500 clones. Unfortunately these 'mega-YACs' have run into problems with insert stability, the cloned DNA sometimes becoming rearranged by intramolecular recombination. Nevertheless, YACs have been of immense value in providing long pieces of cloned DNA for use in large scale DNA sequencing projects.

7.1.6 Vectors for other yeasts and fungi

Cloning vectors for other species of yeast and fungi are needed for basic studies of the molecular biology of these organisms and to extend the possible uses of yeasts and fungi in biotechnology. Episomal plasmids based on the *S. cerevisiae* 2 μm plasmid are able to replicate in a few other types of yeast,

DNA7 8/22/05 10:17 AM Page 141

but the range of species is not broad enough for 2 μm vectors to be of general value. In any case, the requirements of biotechnology are better served by integrative plasmids, equivalent to YIps, as these provide stable recombinants that can be grown for long periods in fermenters (p. 277). Efficient integrative vectors are now available for a number of species, including yeasts such as *Pichia pastoris* and *Kluveromyces lactis*, and the filamentous fungi *Aspergillus nidulans* and *Neurospora crassa*.

7.2 Cloning vectors for higher plants

Cloning vectors for higher plants were developed in the 1980s and their use has led to the **genetically modified (GM) crops** that are in the headlines today. We will examine the genetic modification of crops and other plants in Chapter 15. Here we look at the cloning vectors and how they are used.

Three types of cloning system have been used with varying degrees of success with higher plants:

(1) Vectors based on naturally occurring plasmids of *Agrobacterium*.
(2) Direct gene transfer using various types of plasmid DNA.
(3) Vectors based on plant viruses.

7.2.1 *Agrobacterium tumefaciens* – nature's smallest genetic engineer

Although no naturally occurring plasmids are known in higher plants, one bacterial plasmid, the Ti plasmid of *Agrobacterium tumefaciens*, is of great importance.

A. tumefaciens is a soil microorganism that causes crown gall disease in many species of dicotyledonous plants. Crown gall occurs when a wound on the stem allows *A. tumefaciens* bacteria to invade the plant. After infection the bacteria cause a cancerous proliferation of the stem tissue in the region of the crown (Figure 7.9).

The ability to cause crown gall disease is associated with the presence of the Ti (**t**umour **i**nducing) plasmid within the bacterial cell. This is a large (greater than 200 kb) plasmid that carries numerous genes involved in the infective process (Figure 7.10(a)). A remarkable feature of the Ti plasmid is that, after infection, part of the molecule is integrated into the plant chromosomal DNA (Figure 7.10(b)). This segment, called the **T-DNA**, is between 15 and 30 kb in size, depending on the strain. It is maintained in a stable form in the plant cell and is passed on to daughter cells as an integral part of the chromosomes. But the most remarkable feature of the Ti plasmid is that the T-DNA contains eight or so genes that are expressed in the plant cell and are responsible for the cancerous properties of the transformed cells. These genes also direct synthesis of unusual compounds, called opines, that the bacteria use

DNA7 8/22/05 10:17 AM Page 142

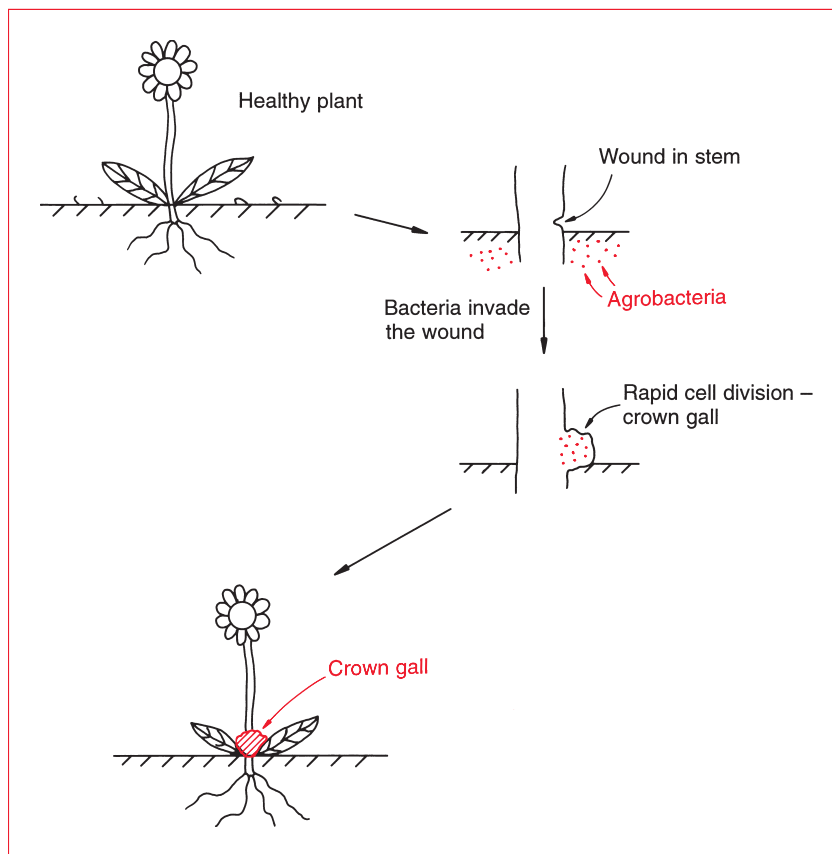


Figure 7.9 Crown gall disease.

as nutrients (Figure 7.10(c)). In short, *A. tumefaciens* genetically engineers the plant cell for its own purposes.

Using the Ti plasmid to introduce new genes into a plant cell

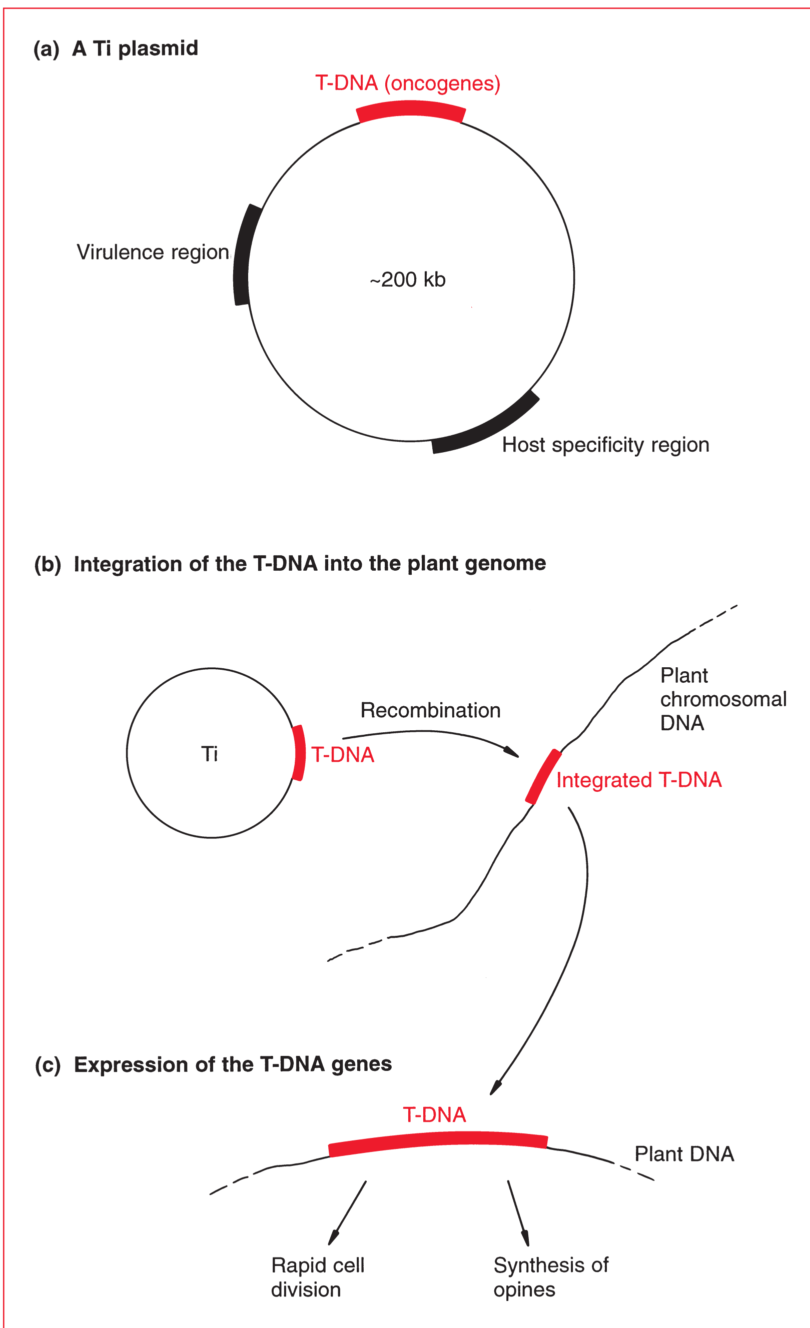
It was realized very quickly that the Ti plasmid could be used to transport new genes into plant cells. All that would be necessary would be to insert the new genes into the T-DNA and then the bacterium could do the hard work of integrating them into the plant chromosomal DNA. In practice this has proved quite a tricky proposition, mainly because the large size of the Ti plasmid makes manipulation of the molecule very difficult.

The main problem is of course that a unique restriction site is an impossibility with a plasmid 200 kb in size. Novel strategies have to be developed for inserting new DNA into the plasmid. Two are in general use:

(1) The binary vector strategy (Figure 7.11) is based on the observation that the T-DNA does not need to be physically attached to the rest of the Ti plasmid. A two plasmid system, with the T-DNA on a relatively small

DNA7 8/22/05 10:17 AM Page 143

Figure 7.10 The Ti plasmid and its integration into the plant chromosomal DNA after *A. tumefaciens* infection.



molecule, and the rest of the plasmid in normal form, is just as effective at transforming plant cells. In fact, some strains of *A. tumefaciens*, and related agrobacteria, have natural binary plasmid systems. The T-DNA plasmid is small enough to have a unique restriction site and to be manipulated using standard techniques.

DNA7 8/22/05 10:17 AM Page 144

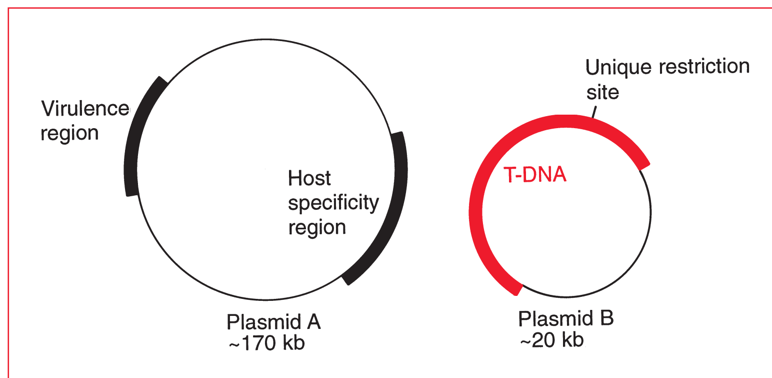


Figure 7.11 The binary vector strategy. Plasmids A and B complement each other when present together in the same *A. tumefaciens* cell. The T-DNA carried by plasmid B is transferred to the plant chromosomal DNA by proteins coded by genes carried by plasmid A.

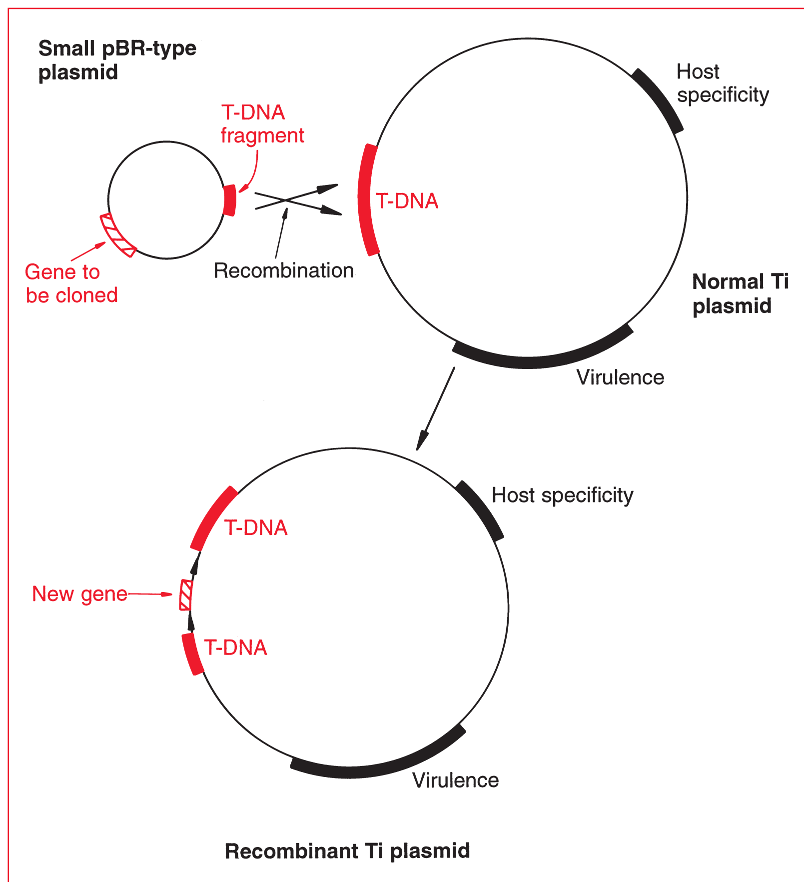


Figure 7.12 The cointegration strategy.

(2) The cointegration strategy (Figure 7.12) uses an entirely new plasmid, based on pBR322 or a similar *E. coli* vector, but carrying a small portion of the T-DNA. The homology between the new molecule and the Ti plasmid means that if both are present in the same *A. tumefaciens* cell, recombination can integrate the pBR plasmid into the T-DNA region. The

DNA7 8/22/05 10:17 AM Page 145

gene to be cloned is therefore inserted into a unique restriction site on the small pBR plasmid, introduced into *A. tumefaciens* cells carrying a Ti plasmid, and the natural recombination process left to integrate the new gene into the T-DNA. Infection of the plant leads to insertion of the new gene, along with the rest of the T-DNA, into the plant chromosomes.

Production of transformed plants with the Ti plasmid

If *A. tumefaciens* bacteria that contain an engineered Ti plasmid are introduced into a plant in the natural way, by infection of a wound in the stem, then only the cells in the resulting crown gall will possess the cloned gene (Figure 7.13(a)). This is obviously of little value to the biotechnologist. Instead a way of introducing the new gene into every cell in the plant is needed.

There are several solutions, the simplest being to infect not the mature plant but a culture of plant cells or protoplasts (p. 105) in liquid medium (Figure 7.13(b)). Plant cells and protoplasts whose cell walls have re-formed can be treated in the same way as microorganisms: for example, they can be plated onto a selective medium in order to isolate transformants. A mature plant regenerated from transformed cells will contain the cloned gene in every cell and will pass the cloned gene to its offspring.

However, regeneration of a transformed plant can occur only if the Ti vector has been '**disarmed**' so that the transformed cells do not display cancerous properties. Disarming is possible because the cancer genes, all of which lie in the T-DNA, are not needed for the infection process; infectivity is controlled mainly by the virulence region of the Ti plasmid. In fact, the only parts of the T-DNA that are involved in infection are two 25 bp repeat sequences found at the left and right borders of the region integrated into the plant DNA. Any DNA placed between these two repeat sequences will be treated as 'T-DNA' and transferred to the plant. It is therefore possible to remove all the cancer genes from the normal T-DNA, and replace them with an entirely new set of genes, without disturbing the infection process.

A number of disarmed Ti cloning vectors are now available, a typical example being the binary vector pBIN19 (Figure 7.14). The left and right T-DNA borders present in this vector flank a copy of the *lacZ′* gene, containing a number of cloning sites, and a kanamycin resistance gene that functions after integration of the vector sequences into the plant chromosome. As with a yeast shuttle vector (p. 133), the initial manipulations that result in insertion of the gene to be cloned into pBIN19 are carried out in *E. coli*, the correct recombinant pBIN19 molecule then being transferred to *A. tumefaciens* and thence into the plant. Transformed plant cells are selected by plating onto agar medium containing kanamycin.

The Ri plasmid

Over the years there has also been interest in developing plant cloning vectors based on the **Ri plasmid** of *Agrobacterium rhizogenes*. Ri and Ti plasmids are

DNA7 8/22/05 10:17 AM Page 146

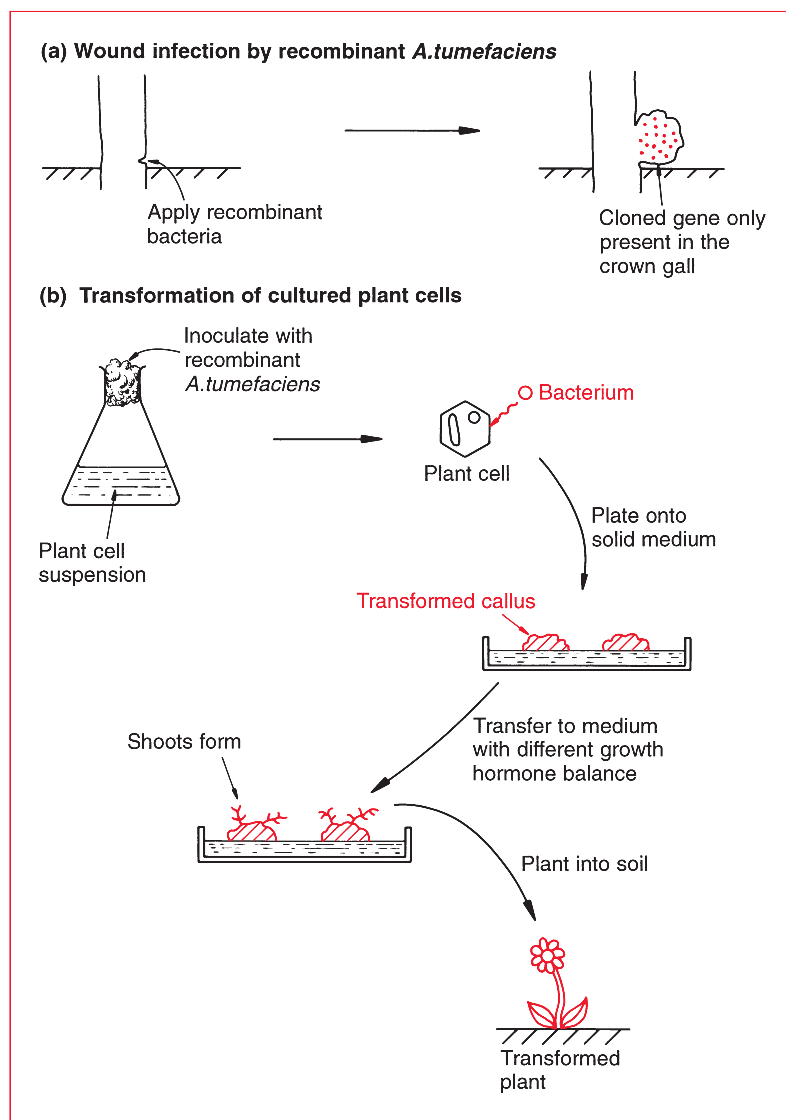


Figure 7.13 Transformation of plant cells by recombinant *A. tumefaciens*. (a) Infection of a wound: transformed plant cells are present only in the crown gall. (b) Transformation of a cell suspension: all the cells in the resulting plant are transformed.

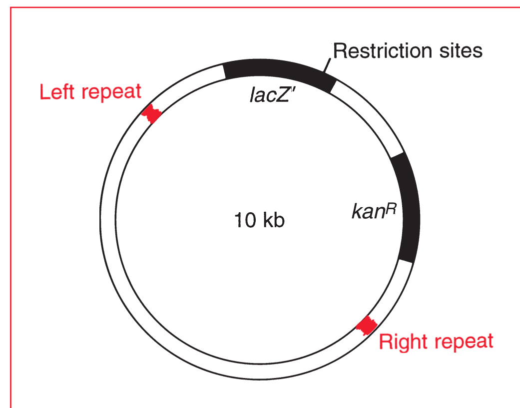
very similar, the main difference being that transfer of the T-DNA from an Ri plasmid to a plant results not in a crown gall but in hairy root disease, typified by a massive proliferation of a highly branched root system. The possibility of growing transformed roots at high density in liquid culture has been explored by biotechnologists as a potential means of obtaining large amounts of protein from genes cloned in plants (p. 298).

Limitations of cloning with *Agrobacterium* plasmids

Higher plants are divided into two broad categories, the monocots and the dicots. Several factors have combined to make it much easier to clone genes

DNA7 8/22/05 10:17 AM Page 147

Figure 7.14 The binary Ti vector pBIN19. *kan*^{R} = kanamycin resistance gene.



in dicots such as tomato, tobacco, potato, peas and beans, but much more difficult to obtain the same results with monocots. This has been frustrating because monocots include wheat, barley, rice and maize, which are the most important crop plants and hence the most desirable targets for genetic engineering projects.

The main difficulty stems from the fact that in nature *A. tumefaciens* and *A. rhizogenes* infect only dicotyledonous plants; monocots are outside of the normal host range. For some time it was thought that this natural barrier was insurmountable and that monocots were totally resistant to transformation with Ti and Ri vectors, but eventually artificial techniques for achieving T-DNA transfer were devised. However, this was not the end of the story. Transformation with an *Agrobacterium* vector normally involves regeneration of an intact plant from a transformed protoplast, cell or callus culture. The ease with which a plant can be regenerated depends very much on the particular species involved and, once again, the most difficult plants are the monocots. Attempts to circumvent this problem have centred on the use of biolistics – bombardment with microprojectiles (p. 105) – to introduce plasmid DNA directly into plant embryos. Although this is a fairly violent transformation procedure it does not appear to be too damaging for the embryos, which still continue their normal development programme to produce mature plants. The approach has been successful with maize and several other important monocots.

7.2.2 Cloning genes in plants by direct gene transfer

Biolistics circumvents the need to use *Agrobacterium* as the means of transferring DNA into the plant cells. **Direct gene transfer** takes the process one step further and dispenses with the Ti plasmid altogether.

Direct gene transfer into the nucleus

Direct gene transfer is based on the observation, first made in 1984, that a supercoiled bacterial plasmid, although unable to replicate in a plant cell on

DNA7 8/22/05 10:17 AM Page 148

its own, can become integrated by recombination into one of the plant chromosomes. The recombination event is poorly understood but is almost certainly distinct from the processes responsible for T-DNA integration. It is also distinct from the chromosomal integration of a yeast vector (p. 135), as there is no requirement for a region of similarity between the bacterial plasmid and the plant DNA. In fact, integration appears to occur randomly at any position in any of the plant chromosomes (Figure 7.15).

Direct gene transfer therefore makes use of supercoiled plasmid DNA, possibly a simple bacterial plasmid, into which an appropriate selectable marker (e.g. a kanamycin resistance gene) and the gene to be cloned have been inserted. Biolistics is frequently used to introduce the plasmid DNA into plant embryos, but if the species being engineered can be regenerated from protoplasts or single cells, then other strategies, possibly more efficient than biolistics, are possible.

One method involves resuspending protoplasts in a viscous solution of polyethylene glycol, a polymeric, negatively charged compound that is thought to precipitate DNA onto the surfaces of the protoplasts and to induce uptake

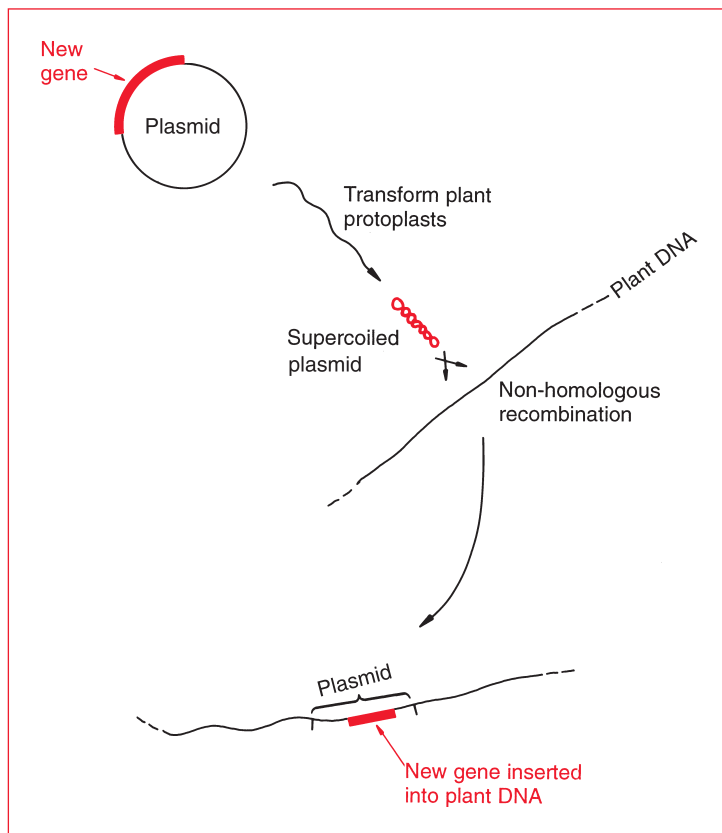


Figure 7.15 Direct gene transfer.

DNA7 8/22/05 10:17 AM Page 149

by endocytosis (Figure 7.16). Protoplasts can also be fused with DNA-containing liposomes (as shown for animal cells in Figure 5.14(b)), or intact cells can be vigorously shaken with DNA-coated silica needles which penetrate the cell wall and transfer the DNA into the interior.

After treatment, protoplasts are left for a few days in a solution that encourages regeneration of the cell walls. The cells are then spread onto selective medium to identify transformants and to provide callus cultures from which intact plants can be grown (exactly as described for the *Agrobacterium* system, Figure 7.13(b)).

Transfer of genes into the chloroplast genome

If biolistics is used to introduce DNA into a plant embryo, then some particles may penetrate one or more of the chloroplasts present in the cells. Chloroplasts contain their own genomes, distinct from (and much shorter than) the DNA molecules in the nucleus, and under some circumstances plasmid DNA can become integrated into this chloroplast genome. Unlike the integration of DNA into plant chromosomes, integration into the chloroplast genome will not occur randomly. Instead the DNA to be cloned must be flanked by sequences similar to the region of the chloroplast genome into which the DNA is to be inserted, so that insertion can take place by homologous recombination (see p. 135). Each of these flanking sequences must be 500 bp or so in length.

A plant cell contains tens of chloroplasts, and probably only one per cell becomes transformed, so the inserted DNA must carry a selectable marker such as the kanamycin resistance gene, and the embryos must be treated with the antibiotic for a considerable period to ensure that the transformed genomes propagate within the cell. Although this means that chloroplast transformation is a difficult method to carry out successfully, it is becoming important as an adjunct to the more traditional methods for obtaining GM crops. As each cell has many chloroplasts, but only one nucleus, a gene inserted into the chloroplast genome is likely to be expressed at a higher level than one placed in the nucleus. This is particularly important when the engineered plants are to be used for production of important proteins (Chapter 13).

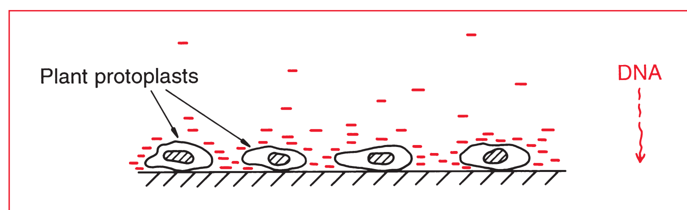


Figure 7.16 Direct gene transfer by precipitation of DNA onto the surfaces of protoplasts.

DNA7 8/22/05 10:17 AM Page 150

7.2.3 Attempts to use plant viruses as cloning vectors

Modified versions of λ and M13 bacteriophages are important cloning vectors for *E. coli* (Chapter 6). Most plants are subject to viral infection, so could viruses be used to clone genes in plants? If they could they would be much more convenient to use than other types of vector, because with many viruses transformation can be achieved simply by rubbing the virus DNA onto the surface of a leaf. The natural infection process then spreads the virus throughout the plant.

The potential of plant viruses as cloning vectors has been explored for several years but without great success. One problem is that the vast majority of plant viruses have genomes not of DNA but of RNA. RNA viruses are not so useful as potential cloning vectors because manipulations with RNA are rather 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 for gene cloning.

Caulimovirus vectors

Although one of the first successful plant genetic engineering experiments, back in 1984, used a caulimovirus vector to clone a new gene into turnip plants, two general difficulties with these viruses have limited their usefulness.

The first is that the total size of a caulimovirus genome is, like that of λ, constrained by the need to package it into its protein coat. Even after deletion of non-essential sections of the virus genome the capacity for carrying inserted DNA is still very limited. Recent research has shown that it might be possible to circumvent this problem by adopting a helper virus strategy, similar to that used with phagemids (p. 120). In this strategy, the cloning vector is a **cauliflower mosaic virus (CaMV)** genome that lacks several of the essential genes, which means that it can carry a large inserted gene but cannot by itself direct infection. Plants are inoculated with the vector DNA along with a normal CaMV genome. The normal viral genome provides the genes needed for the cloning vector to be packaged into virus proteins and spread through the plant.

This approach has considerable potential, but does not solve the second problem, which is the extremely narrow host range of caulimoviruses. This restricts cloning experiments to just a few plants, mainly brassicas such as turnips, cabbages and cauliflowers. Caulimoviruses have, however, been important in genetic engineering as the source of highly active promoters that work in all plants and that are used to obtain expression of genes introduced by Ti plasmid cloning or direct gene transfer.

Geminivirus vectors

What of the geminiviruses? These are particularly interesting because their natural hosts include plants such as maize and wheat, and they could there-

DNA7 8/22/05 10:17 AM Page 151

fore be potential vectors for these and other monocots. But geminiviruses have presented their own set of difficulties, one problem being that during the infection cycle the genomes of some geminiviruses undergo rearrangements and deletions, which would scramble up any additional DNA that has been inserted, an obvious disadvantage for a cloning vector. Research over the years has addressed these problems, but it seems increasingly unlikely that geminiviruses will ever find applications in plant cloning.

7.3 Cloning vectors for animals

Considerable effort has been put into the development of vector systems for cloning genes in animal cells. These vectors are needed in biotechnology for the synthesis of **recombinant protein** from genes that are not expressed correctly when cloned in *E. coli* or yeast (Chapter 13), and methods for cloning in humans are being sought by clinical molecular biologists attempting to devise techniques for **gene therapy** (p. 319), in which a disease is treated by introduction of a cloned gene into the patient.

The clinical aspect has meant that most attention has been directed at cloning systems for mammals, but important progress has also been made with insects. Cloning in insects is interesting because it makes use of a novel type of vector that we have not met so far. We will therefore examine insect vectors before concluding the chapter with an overview of the cloning methods used with mammals.

7.3.1 Cloning vectors for insects

The fruit fly, *Drosophila melanogaster*, has been and still is one of the most important model organisms used by biologists. Its potential was first recognized by the famous geneticist Thomas Hunt Morgan, who in 1910 started to carry out genetic crosses between fruit flies with different eye colours, body shapes and other inherited characteristics. These experiments led to the techniques still used today for gene mapping in insects and other animals. More recently, the discovery that the homeotic selector genes of *Drosophila* – the genes that control the overall body plan of the fly – are closely related to equivalent genes in mammals, has led to *D. melanogaster* being used as a model for the study of human developmental processes. The importance of the fruit fly in modern biology makes it imperative that vectors for cloning genes in this organism are available.

P elements as cloning vectors for *Drosophila*

The development of cloning vectors for *Drosophila* has taken a different route to that followed with bacteria, yeast, plants and mammals. No plasmids are known in *Drosophila* and although fruit flies are, like all organisms, suscepti-

ble to infection with viruses, these have not been used as the basis for cloning vectors. Instead, cloning in *Drosophila* makes use of a **transposon** called the **P element**.

Transposons are common in all types of organisms. They are short pieces of DNA (usually less than 10kb in length) that can move from one position to another in the chromosomes of a cell. P elements, which are one of several types of transposon in *Drosophila*, are 2.9kb in length and contain three genes flanked by short inverted repeat sequences at either end of the element (Figure 7.17(a)). The genes code for transposase, the enzyme that carries out the transposition process, and the inverted repeats form the recognition sequences that enable the enzyme to identify the two ends of the inserted transposon.

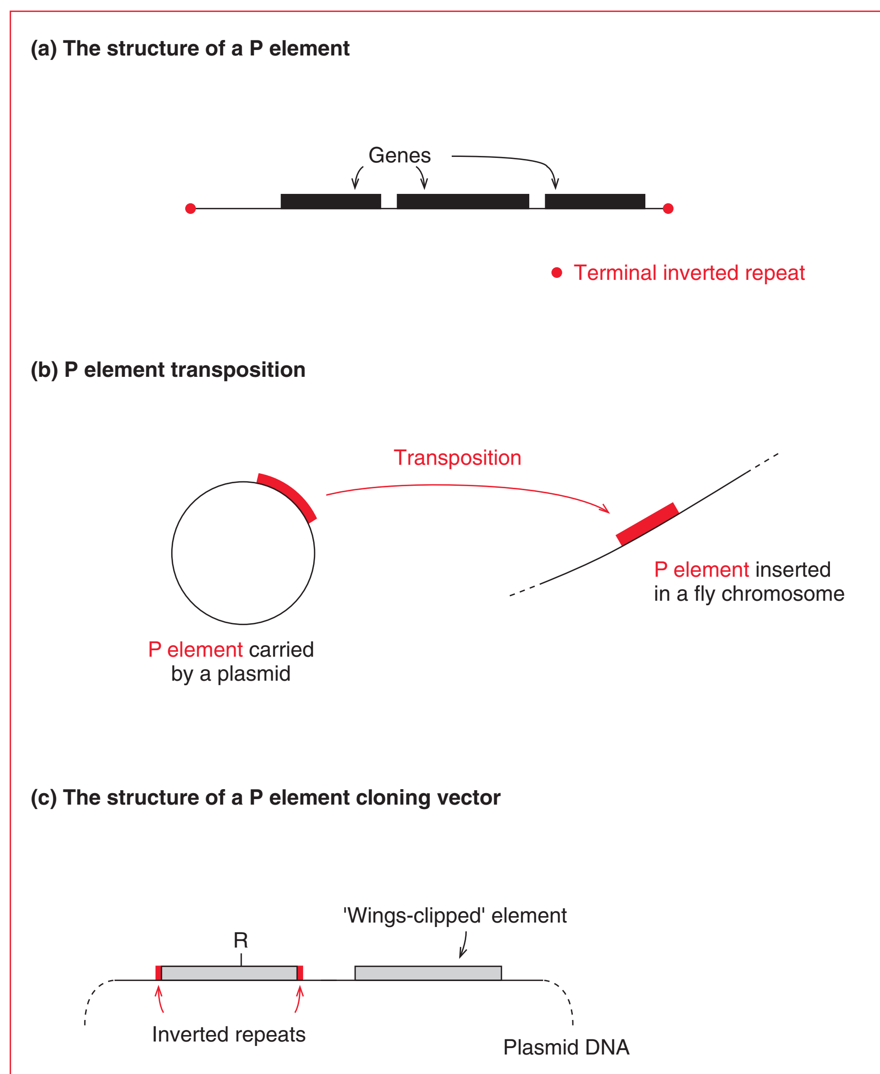


Figure 7.17 Cloning in *Drosophila* with a P element vector. (a) The structure of a P element. (b) Transposition of a P element from a plasmid to a fly chromosome. (c) The structure of a P element cloning vector. The left-hand P element contains a cloning site (R) that disrupts its transposase gene. The right-hand P element has an intact transposase gene but cannot itself transpose because it is 'wings-clipped' – it lacks terminal inverted repeats.

DNA7 8/22/05 10:17 AM Page 153

As well as moving from one site to another within a single chromosome, P elements can also jump between chromosomes, or between a plasmid carrying a P element and one of the fly's chromosomes (Figure 7.17(b)). The latter is the key to the use of P elements as cloning vectors. The vector is a plasmid that carries two P elements, one of which contains the insertion site for the DNA that will be cloned. Insertion of the new DNA into this P element results in disruption of its transposase gene, so this element is inactive. The second P element carried by the plasmid is therefore one that has an intact version of the transposase gene. Ideally this second element should not itself be transferred to the *Drosophila* chromosomes, so it has its 'wings clipped': its inverted repeats are removed so that the transposase does not recognize it as being a real P element (Figure 7.17(c)). Once the gene to be cloned has been inserted into the vector, the plasmid DNA is microinjected into fruit fly embryos. The transposase from the wings-clipped P element directs transfer of the engineered P element into one of the fruit fly chromosomes. If this happens within a germline nucleus then the adult fly that develops from the embryo will carry copies of the cloned gene in all its cells. P element cloning was first developed in the 1980s and has made a number of important contributions to *Drosophila* genetics.

Cloning vectors based on insect viruses

Although virus vectors have not been developed for cloning genes in *Drosophila*, one type of virus, the **baculovirus**, has played an important role in gene cloning with other insects. The main use of baculovirus vectors is in the production of recombinant protein, and we will return to them when we consider this topic in Chapter 13.

7.3.2 Cloning in mammals

At present, gene cloning in mammals is carried out for one of three reasons:

(1) To achieve a **gene knockout**, which is an important technique used to help determine the function of an unidentified gene (p. 268). These experiments are usually carried out with rodents such as mice.

(2) For production of recombinant protein in a mammalian cell culture, and in the related technique of **pharming**, which involves genetic engineering of a farm animal so that it synthesizes an important protein such as a pharmaceutical, often in its milk (p. 296).

(3) In **gene therapy**, in which human cells are engineered in order to treat a disease (p. 319).

Cloning vectors for mammals

For many years it was thought that viruses would prove to be the key to cloning in mammals. This expectation has only partially been realized. The first cloning experiment involving mammalian cells was carried out in 1979 with a vector based on simian virus 40 (SV40). This virus is capable of

DNA7 8/22/05 10:17 AM Page 154

infecting several mammalian species, following a lytic cycle in some hosts and a lysogenic cycle in others. The genome is 5.2 kb in size (Figure 7.18(a)) and contains two sets of genes, the 'early' genes, expressed early in the infection cycle and coding for proteins involved in viral DNA replication, and the 'late' genes, coding for viral capsid proteins. SV40 suffers from the same problem as λ and the plant caulimoviruses, in that packaging constraints limit the amount of new DNA that can be inserted into the genome. Cloning with SV40 therefore involves replacing one or more of the existing genes with the DNA to be cloned. In the original experiment a segment of the late gene region was replaced (Figure 7.18(b)), but early gene replacement is also an option.

In the years since 1979 a number of other types of virus have been used to clone genes in mammals. **Adenoviruses** enable DNA fragments of up to 8 kb to be cloned, longer than is possible with an SV40 vector, though adenoviruses are more difficult to handle because their genomes are bigger. **Papillomaviruses**, which also have a relatively high capacity for inserted DNA,

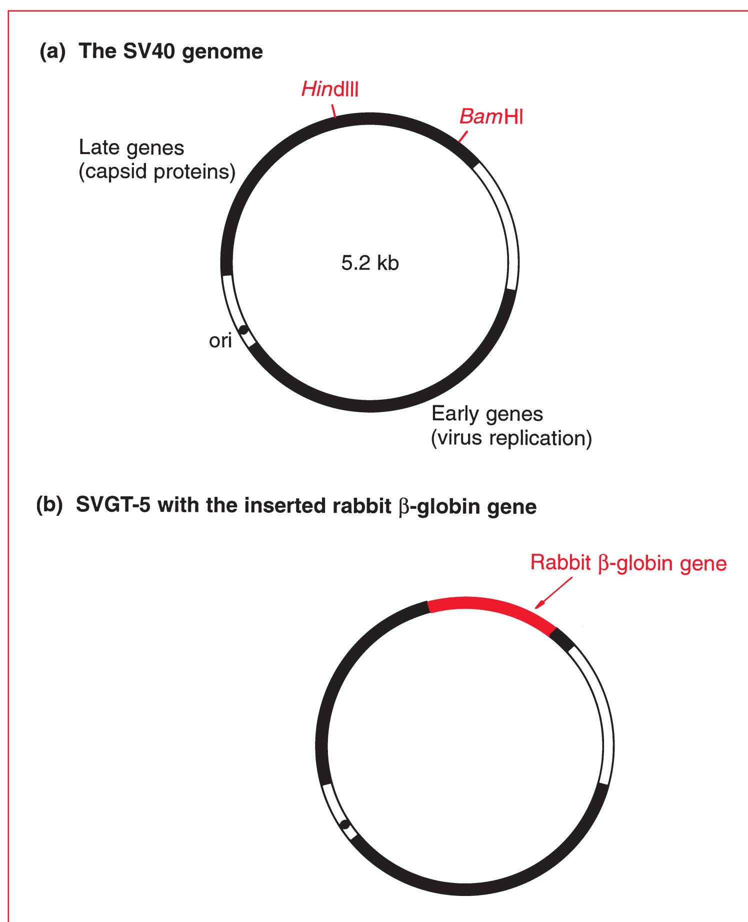


Figure 7.18 SV40 and an example of its use as a cloning vector. To clone the rabbit β-globin gene the *Hin*dIII to *Bam*HI restriction fragment was deleted (resulting in SVGT-5) and replaced with the rabbit gene.

DNA7 8/22/05 10:17 AM Page 155

have the important advantage of enabling a stable transformed cell line to be obtained. **Adeno-associated virus (AAV)** is unrelated to adenovirus but often found in the same infected tissues, because AAV makes use of some of the proteins synthesized by adenovirus in order to complete its replication cycle. In the absence of a helper virus, the AAV genome inserts into its host's DNA. With most integrative viruses, this is a random event, but AAV has the unusual property of always inserting at the same position, within human chromosome 19. Knowing exactly where the cloned gene will be in the host genome is important if the outcome of the cloning experiment must be checked rigorously, as is the case in applications such as gene therapy. AAV vectors are therefore looked on as having major potential in this area. However, at present, **retroviruses** are the most commonly used vectors for gene therapy. Although they insert at random positions the resulting integrants are very stable, which is a major advantage as it means that the therapeutic effects of the cloned gene will persist for some time. We will return to gene therapy in Chapter 14.

Many mammalian viruses kill their host cells soon after infection, so special tricks are needed if these are to be used for anything other than short-term transformation experiments. Bovine papillomavirus (BPV), which causes warts on cattle, is particularly attractive because it has an unusual infection cycle in mouse cells, taking the form of a multicopy plasmid with about 100 molecules present per cell. It does not cause the death of the mouse cell, and BPV molecules are passed to daughter cells on cell division. Shuttle vectors consisting of BPV and pBR322 sequences, and capable of replication in both mouse and bacterial cells, have been used for the production of recombinant proteins in mouse cell lines.

Gene cloning without a vector

One of the reasons why virus vectors have not become widespread in mammalian gene cloning is because it was discovered in the early 1990s that the most effective way of transferring new genes into mammalian cells is by microinjection. Although a difficult procedure to carry out, microinjection of bacterial plasmids, or linear DNA copies of genes, into mammalian nuclei results in the DNA being inserted into the chromosomes, possibly as multiple copies in a tandem, head-to-tail arrangement (Figure 7.19). This procedure is generally looked on as more satisfactory than the use of a viral vector because it avoids the possibility that viral DNA will infect the cells and cause defects of one kind or another.

A knockout mouse (p. 269), which has copies of the cloned gene in all its cells, can be generated by microinjection of a fertilized egg cell which is subsequently cultured *in vitro* for several cell divisions and then implanted into a foster mother. Alternatively an **embryonic stem (ES) cell** can be used. These are obtained from within an early embryo and, unlike most mammalian cells, are totipotent, meaning that their developmental pattern is not pre-set and

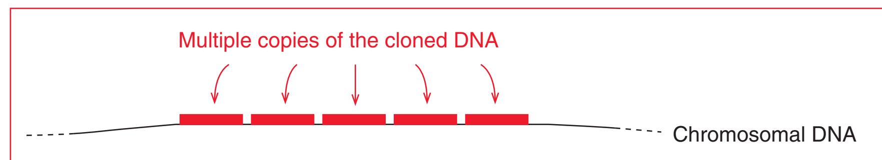


Figure 7.19 Multiple copies of cloned DNA molecules inserted as a tandem array in a chromosomal DNA molecule.

cells descended from them can form many different structures in the adult mouse. After microinjection, the ES cell is placed back in an embryo which is implanted into the foster mother. The resulting mouse is a **chimera**, comprising a mixture of engineered and non-engineered cells, because the embryo that receives the ES cell also contains a number of ordinary cells that contribute, along with the ES cell, to the make-up of the adult mouse. Non-chimeric mice, which contain the cloned gene in all their cells, are obtained by allowing the chimera to reproduce, as some of the offspring will be derived from egg cells that contain the cloned gene.

Further reading

Bevan, M. (1984) Binary *Agrobacterium* vectors for plant transformation. *Nucleic Acids Research*, **12**, 8711–21.

Brisson, N., Paszkowski, J., Penswick, J.R. *et al*. (1984) Expression of a bacterial gene in plants by using a viral vector. *Nature*, **310**, 511–14. [The first cloning experiment with a caulimovirus.]

Broach, J.R. (1982) The yeast 2 μm circle. *Cell*, **28**, 203–4.

Burke, D.T., Carle, G.F. & Olson, M.V. (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science*, **236**, 806–12.

Chilton, M.D. (1983) A vector for introducing new genes into plants. *Scientific American*, **248** (June), 50–9. [The Ti plasmid.]

Colosimo, A., Goncz, K.K., Holmes, A.R. *et al*. (2000) Transfer and expression of foreign genes in mammalian cells. *Biotechniques*, **29**, 314–21.

Evans, M.J., Carlton, M.B.L. & Russ, A.P. (1997) Gene trapping and functional genomics. *Trends in Genetics*, **13**, 370–4. [Includes a description of the use of ES cells.]

Graham, F.L. (1990) Adenoviruses as expression vectors and recombinant vaccines. *Trends in Biotechnology*, **8**, 20–5.

Hamer, D.H. & Leder, P. (1979) Expression of the chromosomal mouse β-maj-globin gene cloned in SV40. *Nature*, **281**, 35–40.

Hansen, G. & Wright, M.S. (1999) Recent advances in the transformation of plants. *Trends in Plant Science*, **4**, 226–31.

Nadolska-Orczyk, A., Orczyk, W. & Przetakiewicz, A. (2000) *Agrobacterium*-mediated transformation of cereals – from technique development to its application. *Acta Physiologiae Plantarum*, **22**, 77–8.

Parent, S.A. *et al.* (1985) Vector systems for the expression, analysis and cloning of DNA sequences in *S. cerevisiae*. *Yeast*, **1**, 83–138. [Details of different yeast cloning vectors.]

Paszkowski, J., Shillito, R.D., Saul, M. *et al*. (1984) Direct gene transfer to plants. *EMBO Journal*, **3**, 2717–22.

DNA7 8/22/05 10:17 AM Page 157

Rubin, G.M. & Spradling, A.C. (1982) Genetic transformation of *Drosophila* with transposable element vectors. *Science*, **218**, 348–53. [Cloning with P elements.]

Timmermans, M.C.P., Das, O.P. & Messing, J. (1994) Geminiviruses and their uses as extrachromosomal replicons. *Annual Review of Plant Physiology and Plant Molecular Biology*, **45**, 79–112.

Viaplana, R., Turner, D.S. & Covey, S.N. (2001) Transient expression of a GUS reporter gene from cauliflower mosaic virus replacement vectors in the presence and absence of helper virus. *Journal of General Virology*, **82**, 59–65. [The helper virus approach to cloning with CaMV.]