TECHNIQUES IN GENETIC ENGINEERING

Işıl KURNAZ

CHAPTER 2

Tools of Genetic Engineering

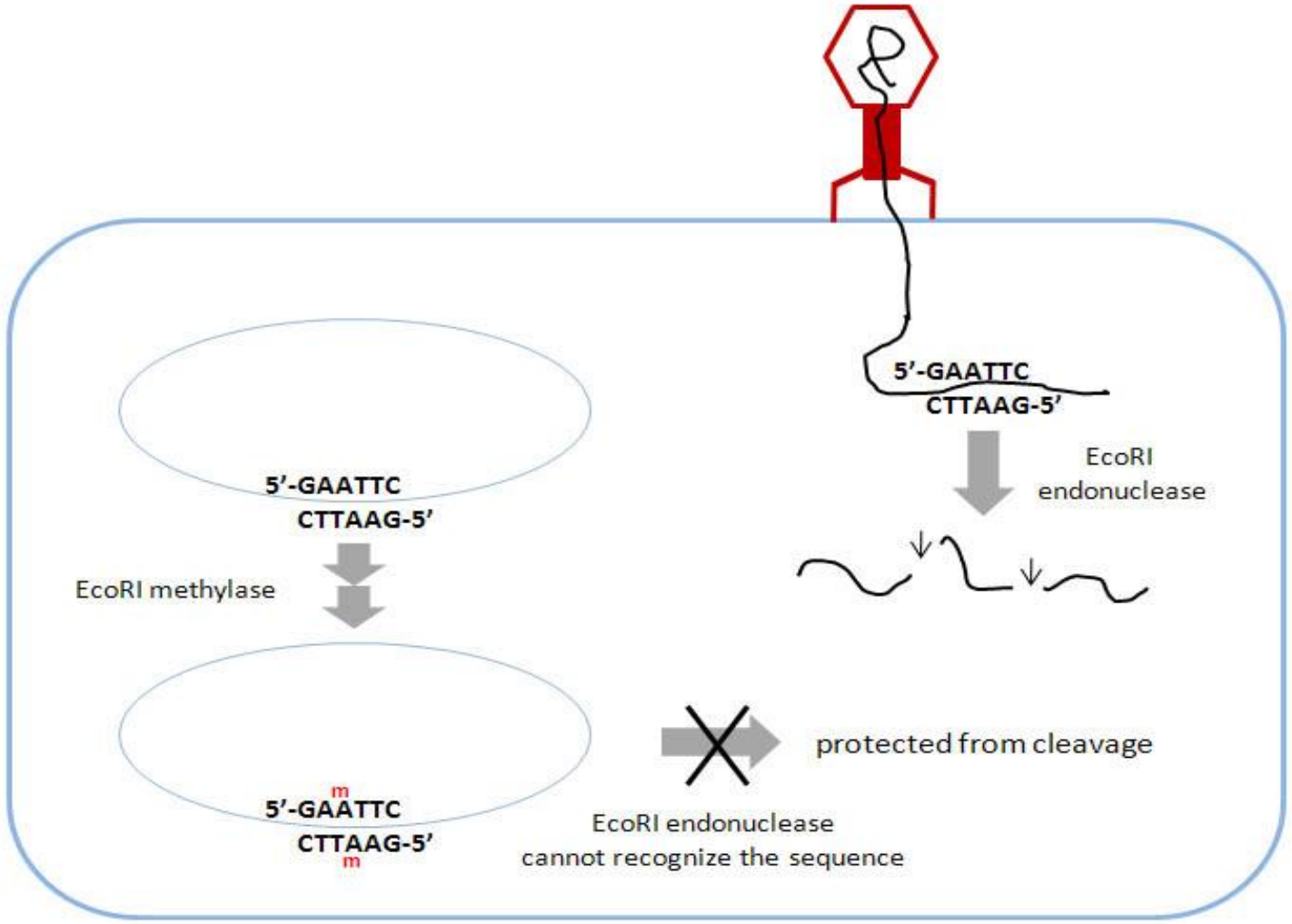
"Science is facts; just as houses are made of stones, so is science made of facts; but a pile of stones is not a house and a collection of facts is not necessarily science."

Henri Poincare (1854 – 1912)

Restriction Endonucleases

Restriction Endonucleases

Figure 2.1. A schematic summary of the restriction/modification system. In this example, the recognition motifs for EcoRI restriction endonuclease in the host genome are modified by the EcoRI methylase, which covalently adds a methyl group to the adenine nucleotide. This modification does not affect the structure of the host DNA, but simply disables the endonuclease from recognizing the motif, thus host genome is protected from cleavage. Phage DNA, on the other hand, has not been previously methylated, therefore EcoRI enzyme recognizes the cleavage site upon entry of phage DNA and cleaves it.



Type I endonucleases

These were the systems that were first characterized, and they consist of one enzyme with different subunits for recognition, cleavage and methylation (all-in-one). Their mechanism of cleavage relies on translocation of DNA until a mechanistic collision occurs (usually quite some distance away, up to 1000 bp), producing fairly random fragments. Thus, their lack of specific cleavage point makes them not suitable for specific gene cloning purposes.

Example : EcoDR2 TCANNNNNN GTCG

Type II endonucleases

These are the most common types of restriction enzymes, consisting of two subunits for restriction and modification. They usually bind DNA as homodimers, thus recognize symmetric sequences, and have fairly constant cut positions. The sites can either be *continuous*, as in the case of most 6-base-cutters (such as GAATTC for EcoRI), or *discontinuous*, where the two half-sites are separated by a number of random nucleotides (such as GCCNNNNNGGC for Bgl I).

The most common Type II enzymes which are used in the laboratory recognize 4 to 8 bases; but the recognition motif is palindromic, which means that not only each strand reads the same sequence 5' to 3', but also the nucleotides are "mirror image" through an axis that passes in the middle of sequence

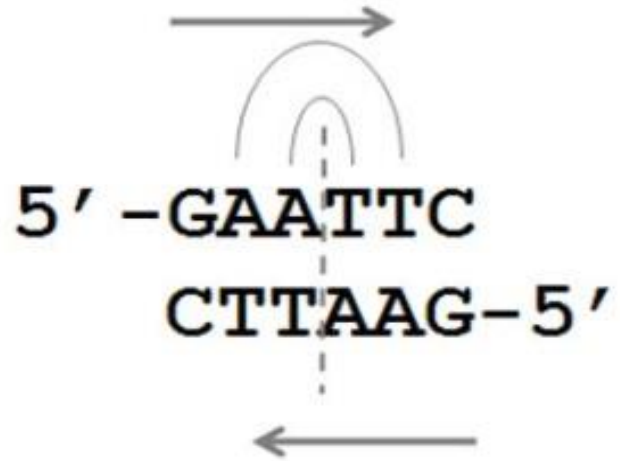


Figure 2.2. The palindromic recognition motif for EcoRI enzyme. The sequence reads the same in each orientation, and the nucleotides on either side of the midaxis are complementary.

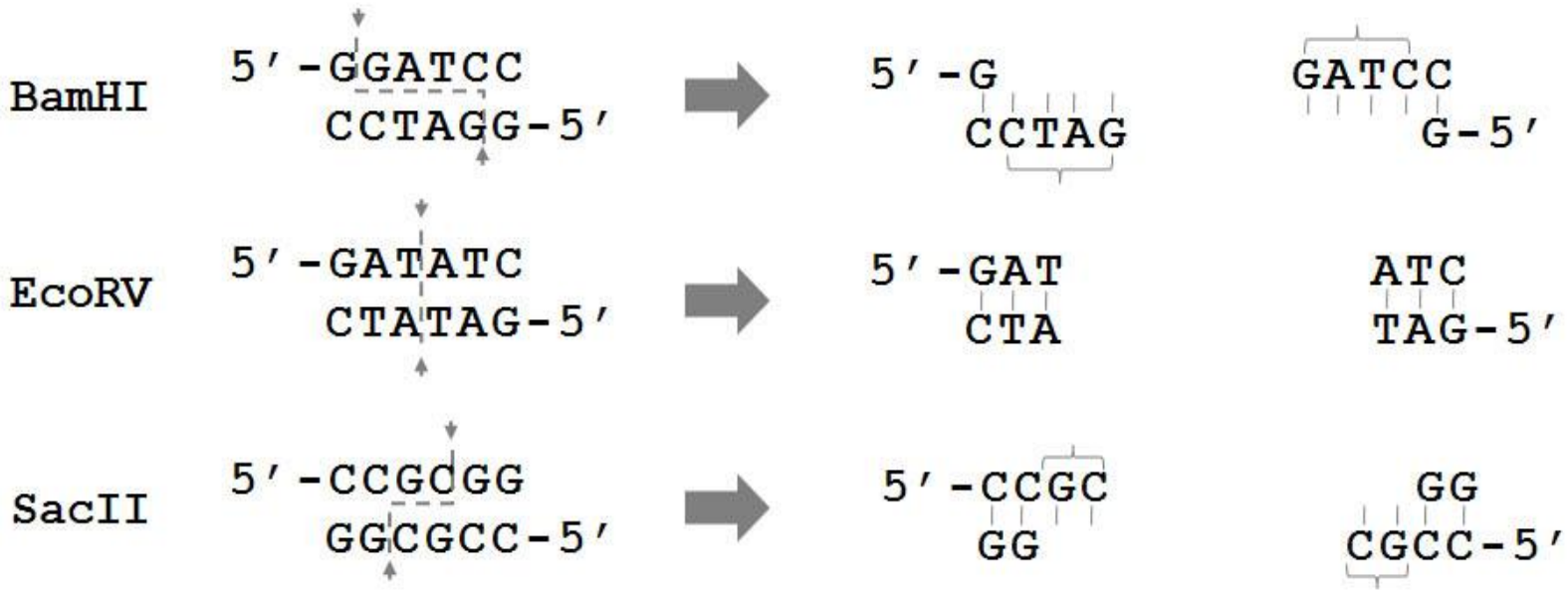


Figure 2.3. Different cut positions of restriction enzymes. BamHI cuts towards the 5’ end and generarets a sticky end with 3’ overhang , EcoRV cuts right in the middle of the sequence and generates a blunt end, and SacII cuts towards the 3’ end to generate a sticky end with 5’ overhang

Type IIs endonucleases

These group of enzymes actually work in pairs, and recognize asymmetric sequences. They cleave up to 20 bp away on one side of the recognition sequence. They are much more active on DNA containing multiple motifs. Their recognition site cannot be destroyed by blunting the ends of digestion products, which can be exploited in some cloning applications such as generating deletions along a DNA molecule.

Type III endonucleases

Enzymes of Type III consist of one complex of two subunits encoded by the *mod* (modification) and *res* (restriction) genes. The recognition sequence is a set of two copies of non-palindromic site in inverse orientation. The enzyme then cleaves at a specific distance (24 to 26 bp) away from one of the copies. Since the exact cut site is not pre-determined, these enzymes are non-suitable for cloning purposes.

Type IV endonucleases

These enzymes can recognize modified (even methylated) DNA. They are rather large proteins with two catalytic subunits, amd they cleave outside their recognition sites. They can be subgrouped among themselves: Enzymes that recognize *continuous* sequences (such as CTGAAG for Eco57I) cleave on one side only, whereas those that recognize *discontinuous* sequences (such as CGANNNNNNTGC for BcgI) cleave on both sides of the motif.

Isoschizomers and Neoschizomers

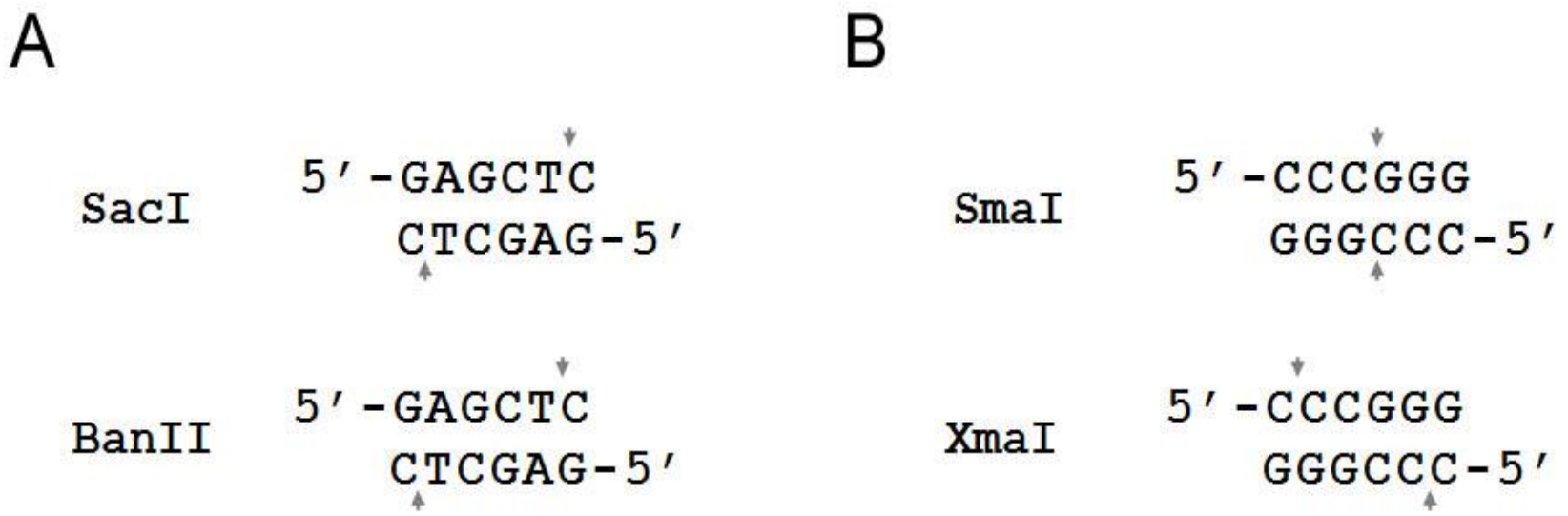


Figure 2.4. Isoschizomers and neoschizomers. (a) SacI and BanII in this example recognize exactly the same recognition motif and cut from exactly the same position, hence they are called isoschizomers of each other; (b) SmaI and XmaI in this example recognize the same motif but cut at different positions, hence they are neoschizomers.

Table 2.1. Isoschizomers of some enzymes and their recognition motifs as examples.

Enzyme	Isoschizomer(s)	Recognition Sequence
AclNI	SpeI	A/CTAGT
Bsp19I	NcoI	C/CATGG
Bsp106I	ClaI, Bsu15I	AT/CGAT
Eco32I	EcoRV	GAT/ATC
Sac II	SstII, KspI	CCGC/GG
XhoI	PaeR7I	C/TCGAG

Star Activity

Under non-standard conditions such as high pH, low ionic strength, high levels of organic solvents (for instance glycerol or DMSO), a non-standard ion in the reaction buffer (such as Mn^{2+} instead of Mg^{2+}) or even elongated incubation period, enzymes may exhibit non-specific recognition and cleavage, a phenomenon known as the **star activity**. In these cases, the enzymes will most likely cleave sequences which differ by one or two bases from the canonical recognition motifs (Fig.2.5).

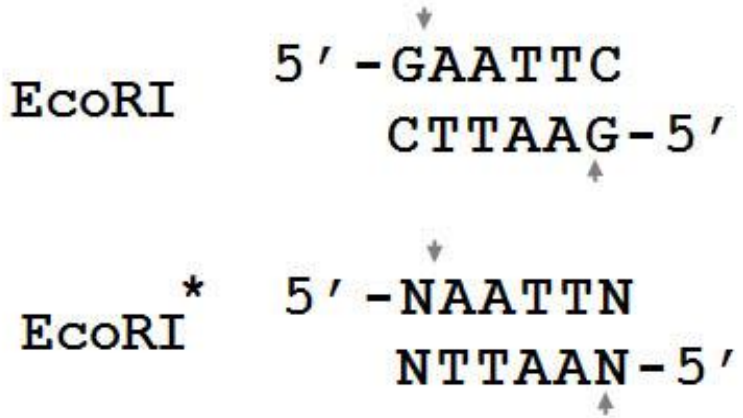
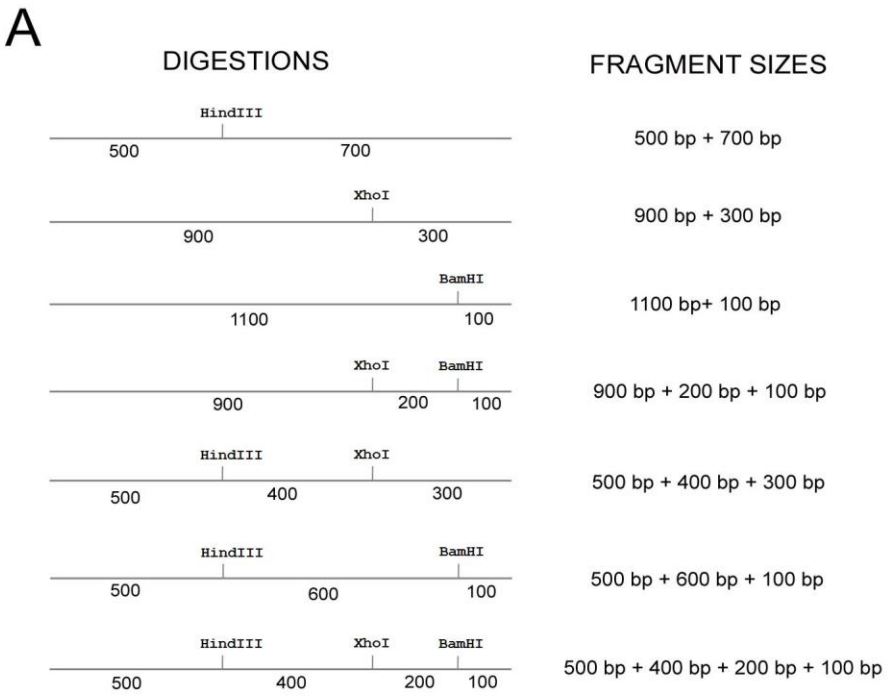


Figure 2.5. A schematic of star activity for EcoRI example. The recognition specificity is compromised during star activity and the enzyme will cleave sequences that are similar but not identical to its canonical motif

Restriction Mapping



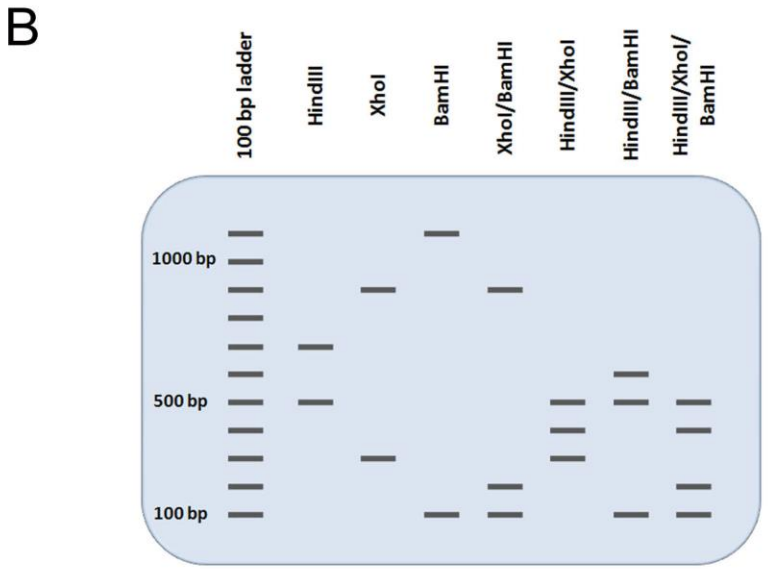


Figure 2.6. Restriction mapping principles. (a) a set of digestions conducted on the same DNA fragment; (b) how the digestion products would appear on the DNA agarose gel

APPENDIX I – DNA TECNHIQUES
i) DNA gel electrophoresis

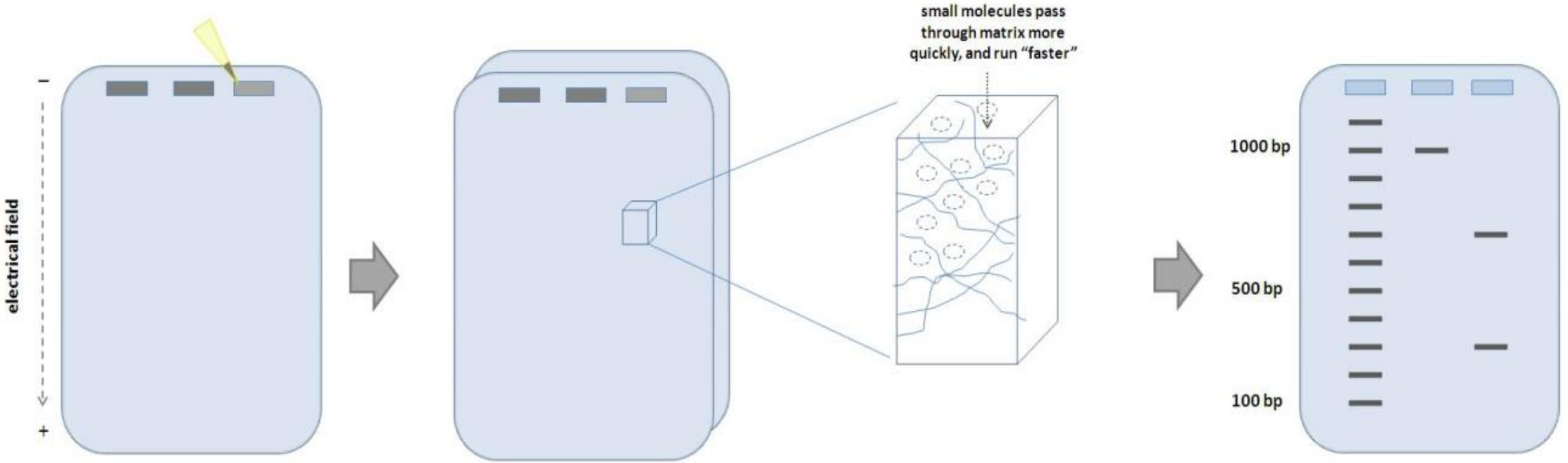


Figure A1.1. A schematic depiction of DNA agarose gel electrophoresis. The DNA samples are loaded into the wells (leftmost panel) and subjected to electrical field. DNA molecules travel through the agarose gel matrix that contains non-uniform pores (depicted as dashed circles); the shorter the DNA molecule, the faster it "runs" in the electrical field (middle panel). The size-separated DNA molecules are then visualized (by ethidium bromide or other methods) as bands (rightmost panel).

Restriction Fragment Length Polymorphism (RFLP)

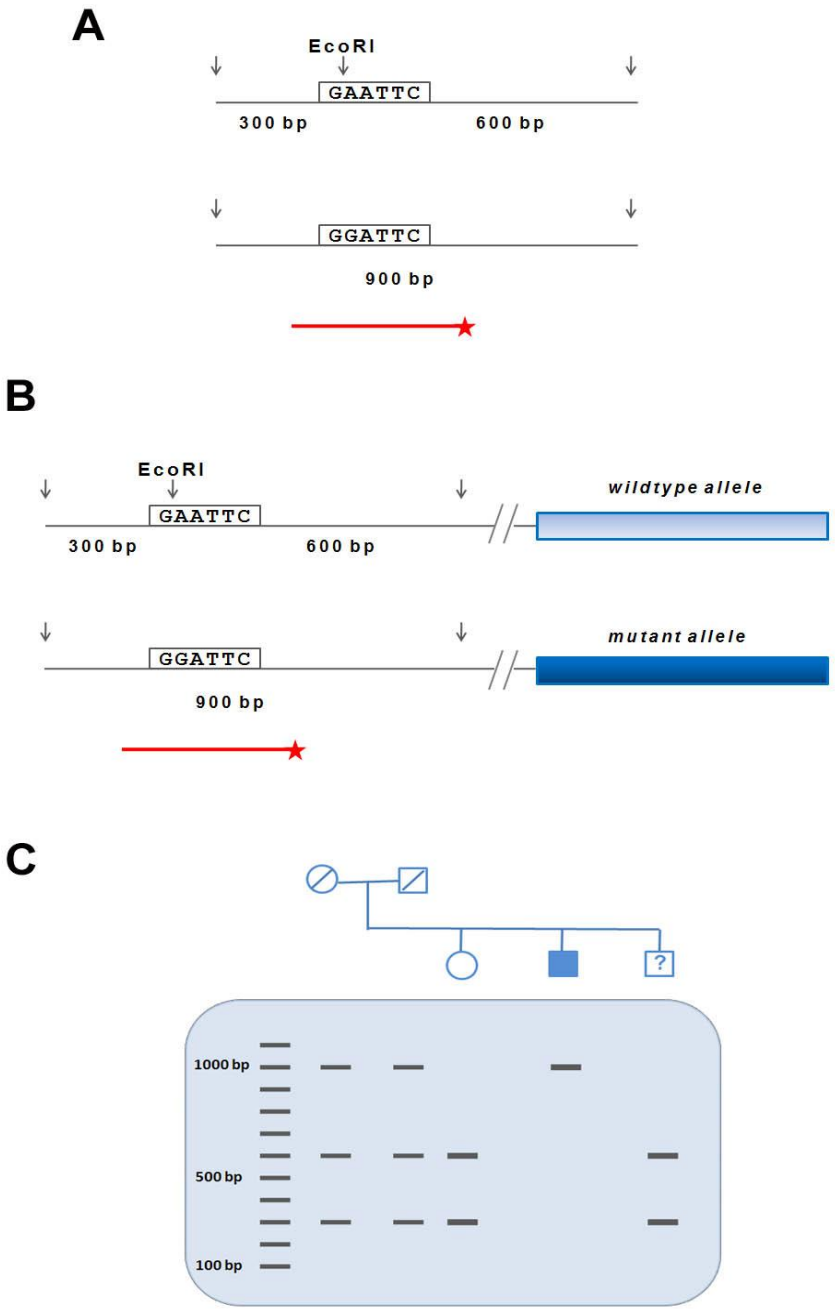


Figure 2.7. Restriction fragment length polymorphism (RFLP) principles. (a) a single nucleotide substitution changes the DNA sequence such that EcoRI recognition motif is abolished, which can be detected by a labeled probe after restriction digestion; (b) if this polymorphism was closely linked to a gene related to a disease in the population, then the polymorphism can be used to indirectly detect the presence of a wildtype or mutant allele of the gene; (c) the RFLP can be used to trace the mutant allele, thus the genetic disorder, within a family (question mark denotes the unborn child).

Restriction Fragment Length Polymorphism (RFLP)

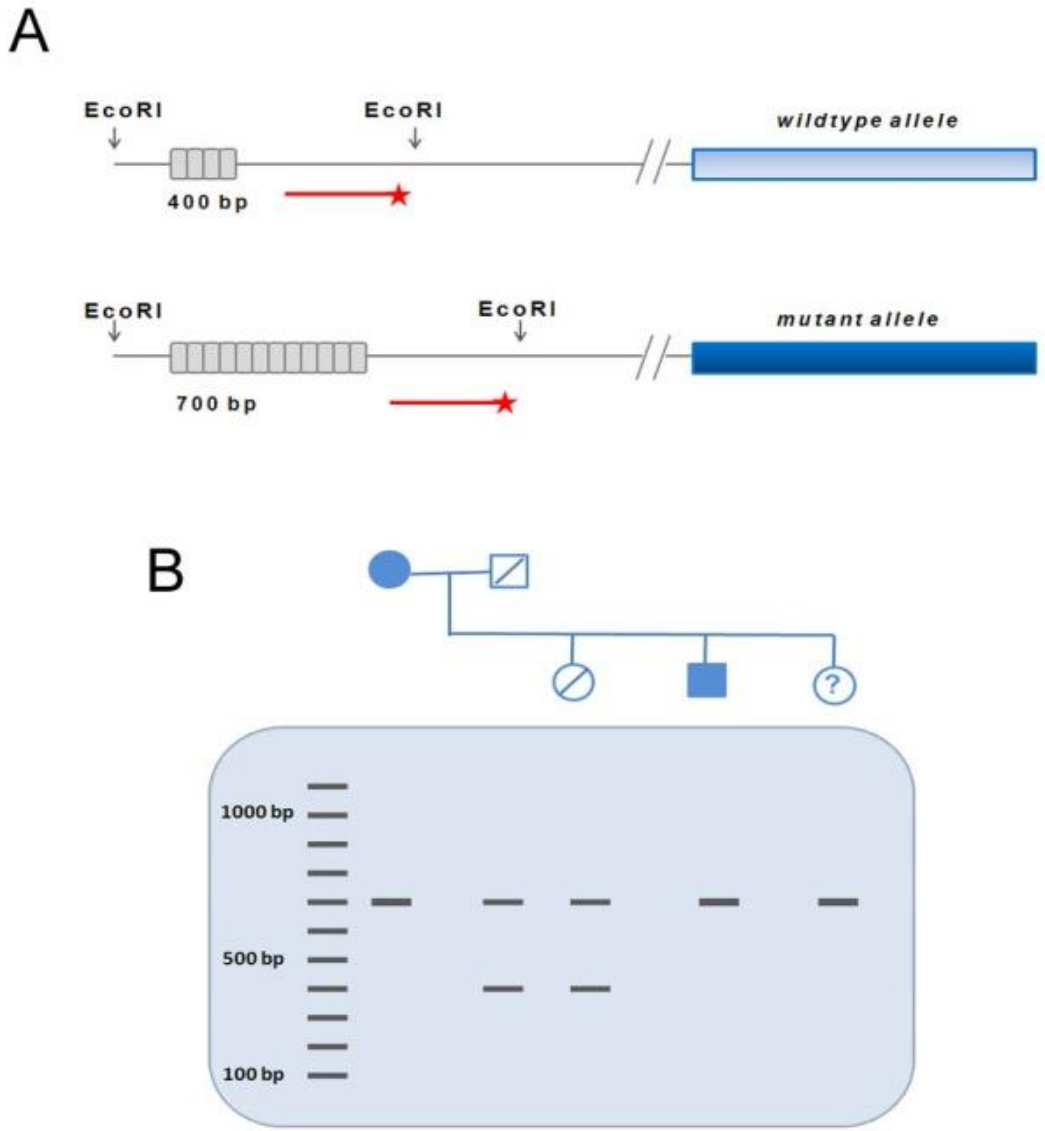


Figure 2.8. Restriction fragment length polymorphism (RFLP) due to variable number of tandem repeats. (a) different disease alleles could be linked to polymorphisms that result in different fragment length due to insertion of tandem repeats; (b) the fragment lengths can be analyzed by Southern blotting, using a labeled probe that can hybridize to both restriction fragments (question mark denotes the unborn child).

APPENDIX I – DNA TECNHIQUES
ii) Nucleic Acid Blotting (Southern blotting)

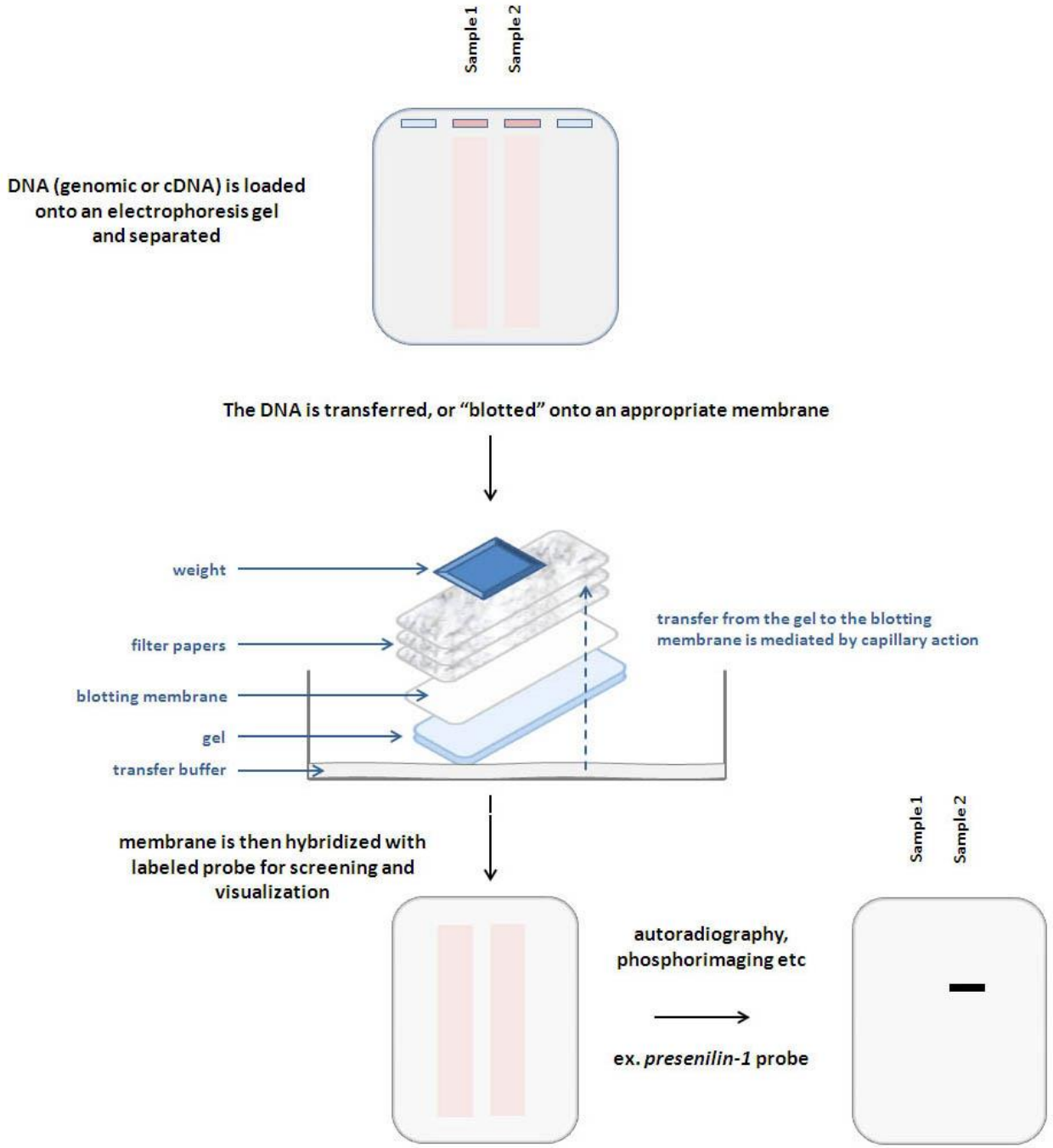


Figure A1.2. A schematic depiction of Southern blotting. The DNA samples are loaded into the wells, and then "blotted" onto a membrane. The membrane is then hybridized with a labeled probe against the target DNA (for example, presenilin-1 gene in the figure above). After autoradiography or other methods of visualization, any DNA that hybridizes to the probe is seen as a "band" on the membrane

Vectors

Vectors

Vectors are DNA molecules which can replicate autonomously and thus can be used to carry insert DNA into organisms and amplify this DNA *in vivo*. There are many types and many functions of vectors. The most commonly used vectors that will be covered in this section are (a) plasmids, (b) phage vectors, (c) cosmids, (d) bacterial artificial chromosomes, and (e) yeast artificial chromosomes. All of those vectors change in the size of insert they can carry, and the purpose for which they can be used.

Table 2.1. Various commonly used laboratory vectors and their key features

Vector type	Insert Size	Examples	Main Purposes
Plasmid	10-20 kb	pUC19, pCMV	DNA manipulation; protein expression; and many others
Phage (l, insertion)	around 10 kb	l gt11	cDNA libraries
Phage (l, replacement)	around 23 kb	EMBL4	genomic DNA libraries
cosmid	around 45 kb	pHM1; pJB8	genomic DNA libraries
phagemid	10-20 kb	pBlueScript	DNA manipulation; in vitro transcription; in vitro mutagenesis
BAC	130-150 kb	pBACe3.6	genomic DNA libraries
YAC	1000-2000 kb	pYAC4	genomic DNA libraries

Plasmids

Plasmids are perhaps the most commonly used vectors of all and are extra-chromosomal DNA molecules present in prokaryotes, serving a range of functions from production of conjugation pilli (F plasmids), conferring antibiotic resistance (R plasmids), sugar fermentation, heavy metal resistance etc depending on genes expressed on the plasmid. Plasmids are usually small, circular double stranded DNA molecules that have the capacity to replicate autonomously within bacteria, however still replication is coupled to host replication and can be found in two forms: *stringent*, which replicate once or twice per generation (**low copy number plasmids**), and *relaxed*, which replicate 10-200 copies in each generation (**high copy number plasmids**)

Table 2.2. Copy numbers of some key plasmids

Plasmid	Plasmid Size (approx., bp)	Copy number
pUC	2700	500-700
pBR322	2700	>25
ColE1	4500	>15

To be used in cloning purposes, all plasmids must contain the following DNA sequences at minimum:

1. **Origin of replication**, ***ori***, is required for autonomous replication within the host cell. If the host is a bacterium, then a bacterial *ori* is required; if the host is yeast, then a yeast *ori* is used.
2. **Selective marker** is required for selection of recombinant bacteria that contain the plasmid; the most common markers are antibiotic resistance genes.
3. **Multiple Cloning Site (MCS)** is an engineered sequence of DNA that contains multiple restriction sites that are unique (ie that are not found anywhere else in the plasmid), which will be used for inserting the foreign DNA.

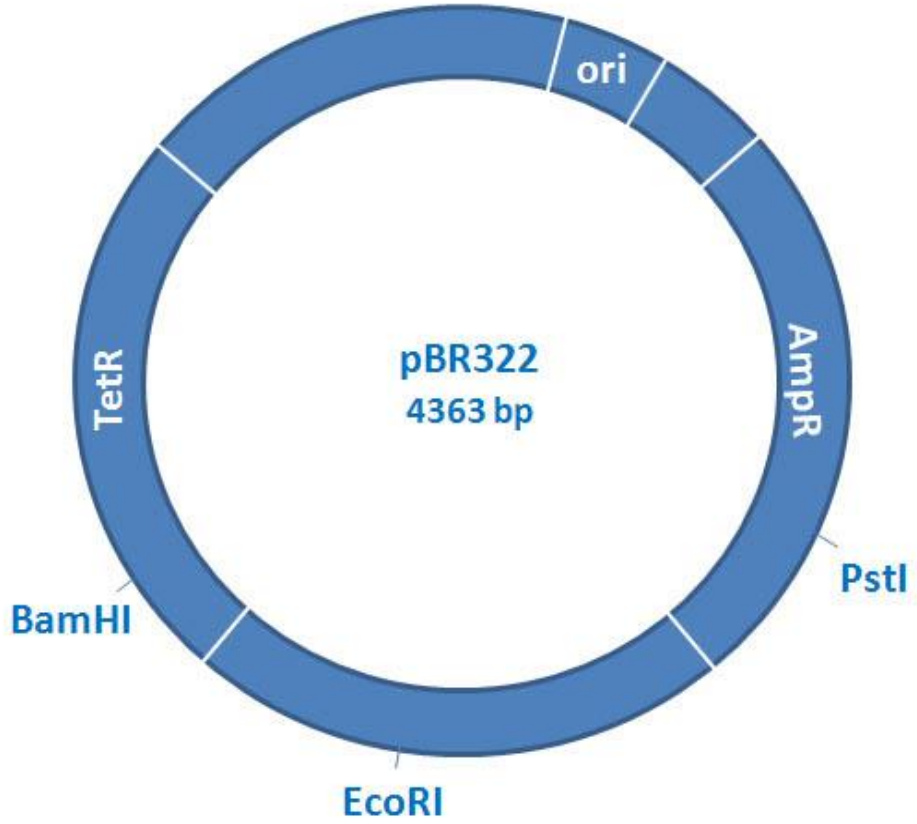


Figure 2.9. The simplified plasmid map of pBR322 vector. The *ori* sequence is derived from the ColE1 plasmid, and in addition the plasmid contains two antibiotic resistance genes that can be cleaved by two different restriction enzymes and used for cloning.

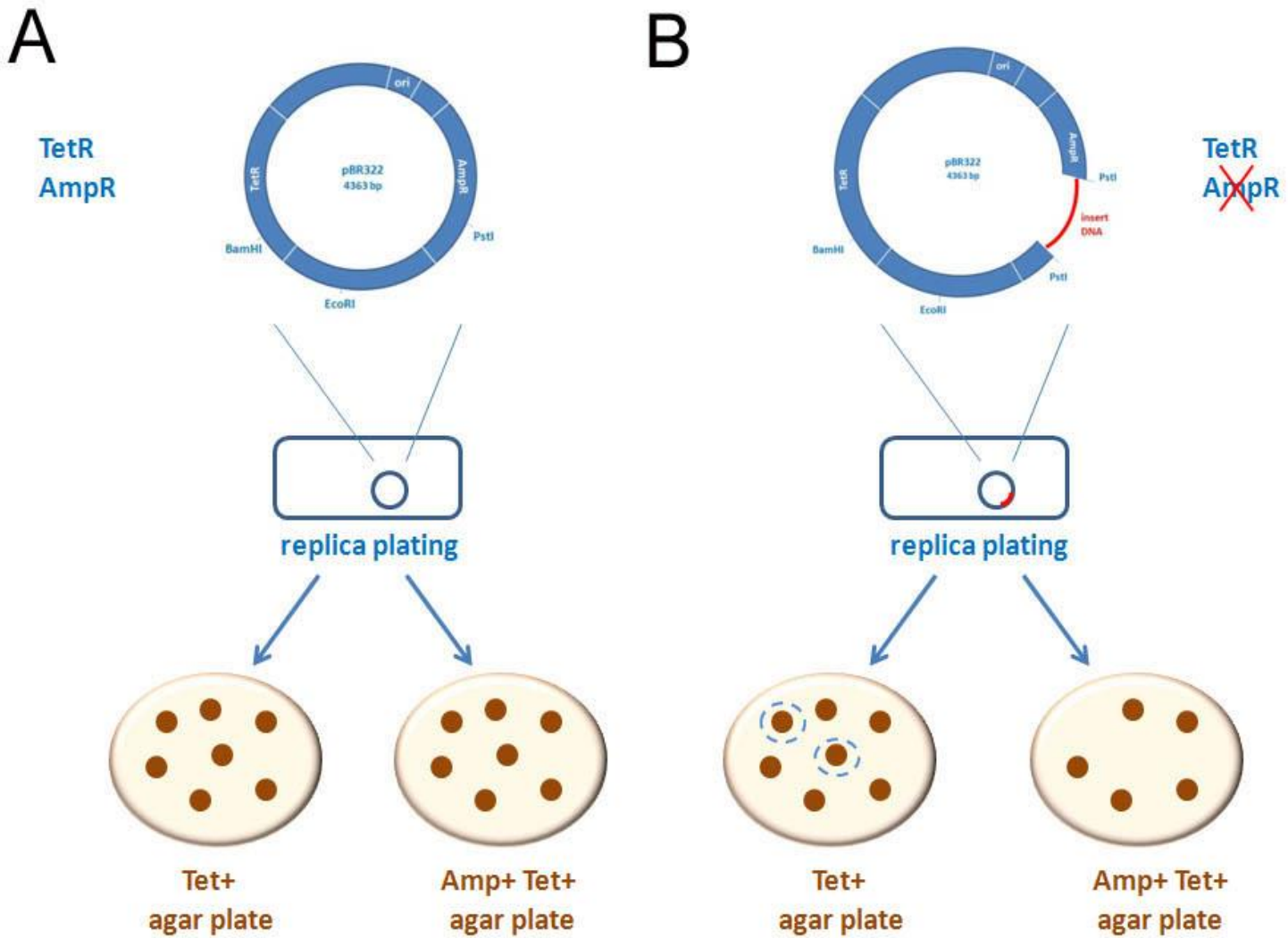


Figure 2.10. The schematic summary of cloning strategy using pBR322 vector. Essentially, transformed bacteria are replica plated onto two different agar plates – one containing ampicillin as a selection marker, and the other containing tetracycline. (a) If the plasmid vector contains no insert, then both resistance genes are intact, and bacteria will grow on both agar plates. (b) If the plasmid vector contains an insert DNA cloned into, for instance, the Ampicillin resistance gene, then the bacteria would grow colonies on the Tet+ agar plate, but not on the Amp+ plate

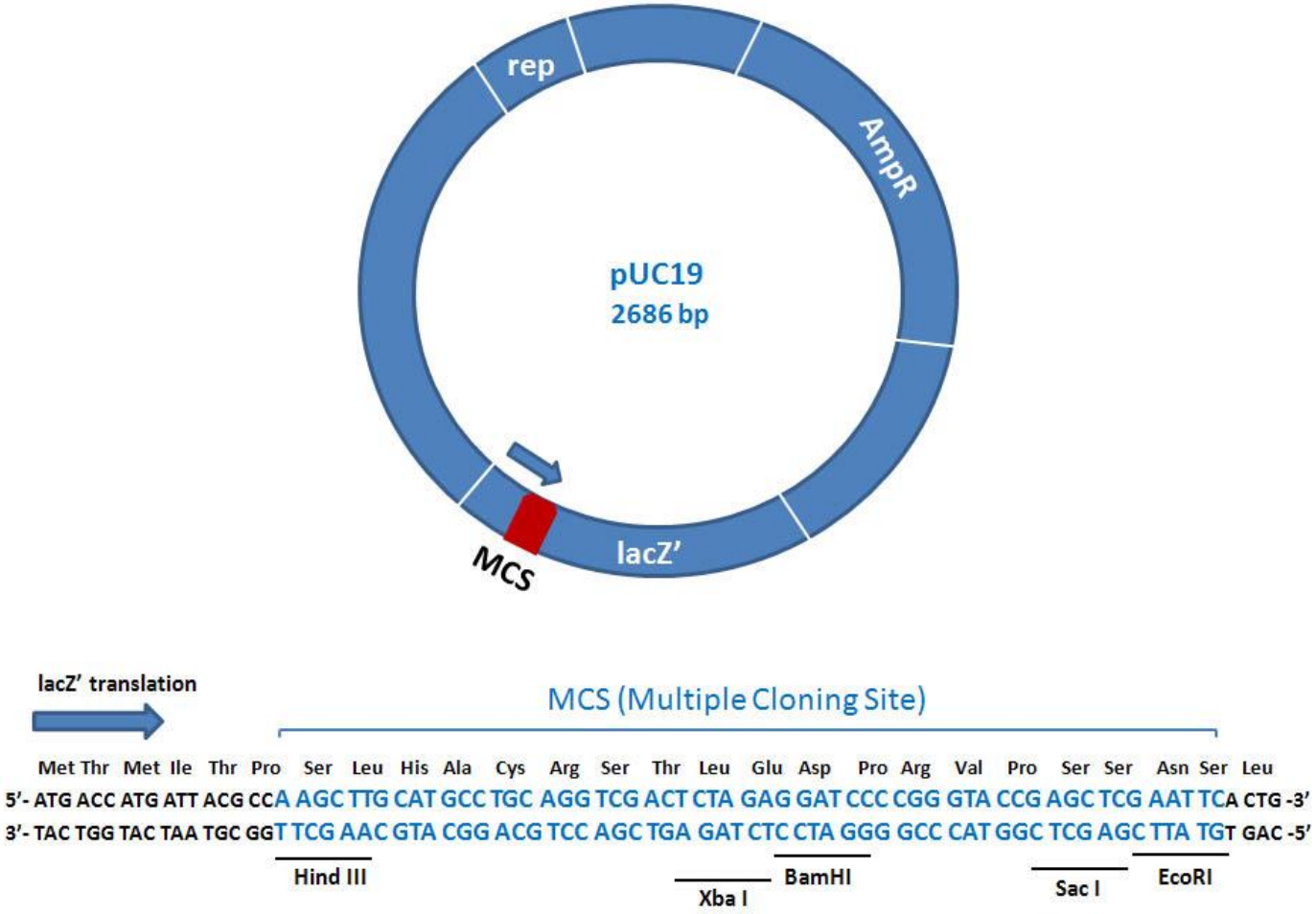


Figure 2.11. The simplified plasmid map of pUC19 vector. The *rep* sequence is derived from the pMB1 replicon and confers autonomous replication property to the plasmid; the AmpR gene allows for ampicillin selection; additionally the plasmid is engineered to harbor an MCS within the lacZ' coding sequence (only some of the recognition sequences within the MCS are shown).

Table 2.3. Some common selectable markers used for cloning and their modes of action.

Marker	Acts on :	Mode of action
ampicillin	prokaryotes (Gram negative bacteria)	Inhibits cell wall synthesis, thus bacteria cannot replicate in the presence of ampicillin
tetracycline	prokaryotes	Binds to the 30S ribosomal subunit and inhibits translocation of ribosomes
kanamycin	Prokaryotes and eukaryotes	Binds to ribosomal subunits and inhibits protein synthesis
puromycin	Prokaryotes and eukaryotes	Binds to ribosomal A site and causes premature chain termination
cycloheximide	eukaryotes	Targets the E site of 50S subunit of eukaryotic ribosomes

Phage vectors

Initial studies on phage life cycle and genome, which date back to 1950s and in particular to the work of Lwoff and his coworkers followed more recently by Ptashne and his group in 1990s, have shown that phages are convenient alternatives to plasmids due to the 50 kb genome that gets packaged into the head region, and that the central part of its genome is not necessary for packaging thus can be replaced with insert DNA.

Wildtype λ phage itself cannot be directly used for cloning purposes; first of all, it contains very few unique restriction sites that can be exploited for cloning, and secondly there is a maximum size limit (between 78% and 105% of wildtype DNA length, or 37-53 kb) to the DNA that can be packaged to the phage heads. Therefore, phage genome had to be engineered before it could be used as a cloning vector.

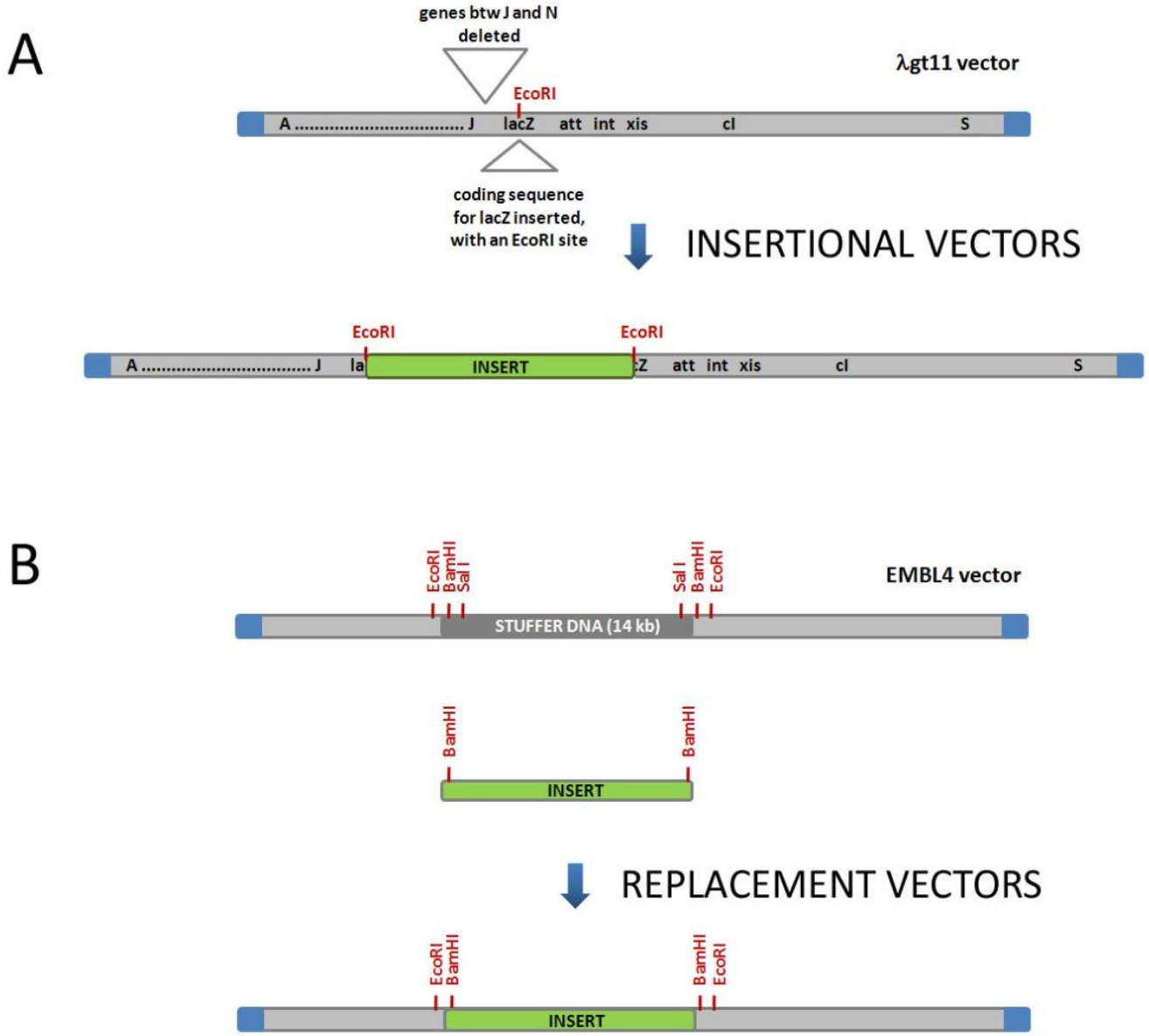


Figure 2.14. Examples of (a) an insertional vector, and (b) a replacement vector. (a) in insertional vectors, the insert DNA is cloned into the MCS, which in this case resides within the lacZ coding sequence; (b) in replacement vectors there is a central stuffer for proper packaging of the vector, which will be "swapped" with the insert DNA

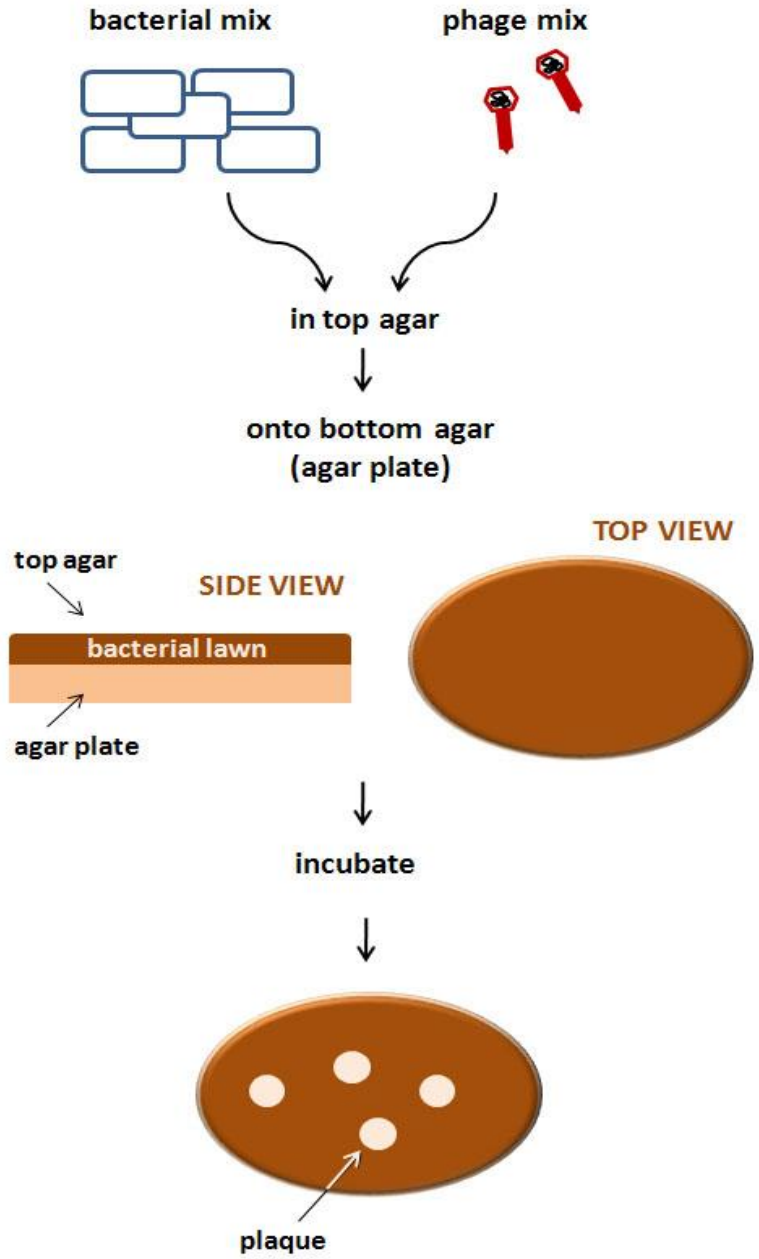


Figure 2.15. The plaque assay. Bacteria and phages are mixed together in top agar, and then poured on top of an agar plate (the "bottom" agar), where the bacteria grows into a "bacterial lawn", and after incubation and growth of phage, the lysed bacteria will yield a clear plaque

Cosmids and Phagemids

Cosmids are essentially plasmid vectors that contain λ phage *cos* site, developed by Collins and Bruning (1978). Since they are based on plasmids and have an origin of replication, they can replicate in the cell like a plasmid, but because they have *cos* sites they can be packaged like a phage, leading a "dual life". And since they can be packaged like a phage particle, they may even at times carry up to 45 kb inserts, much larger capacity than a common plasmid, and even larger than a typical λ phage vector that can carry only up to 25 kb. Just like the phage vectors, cosmids are also not as commonplace as they used to be; more advanced "specialist" vectors are now in use.

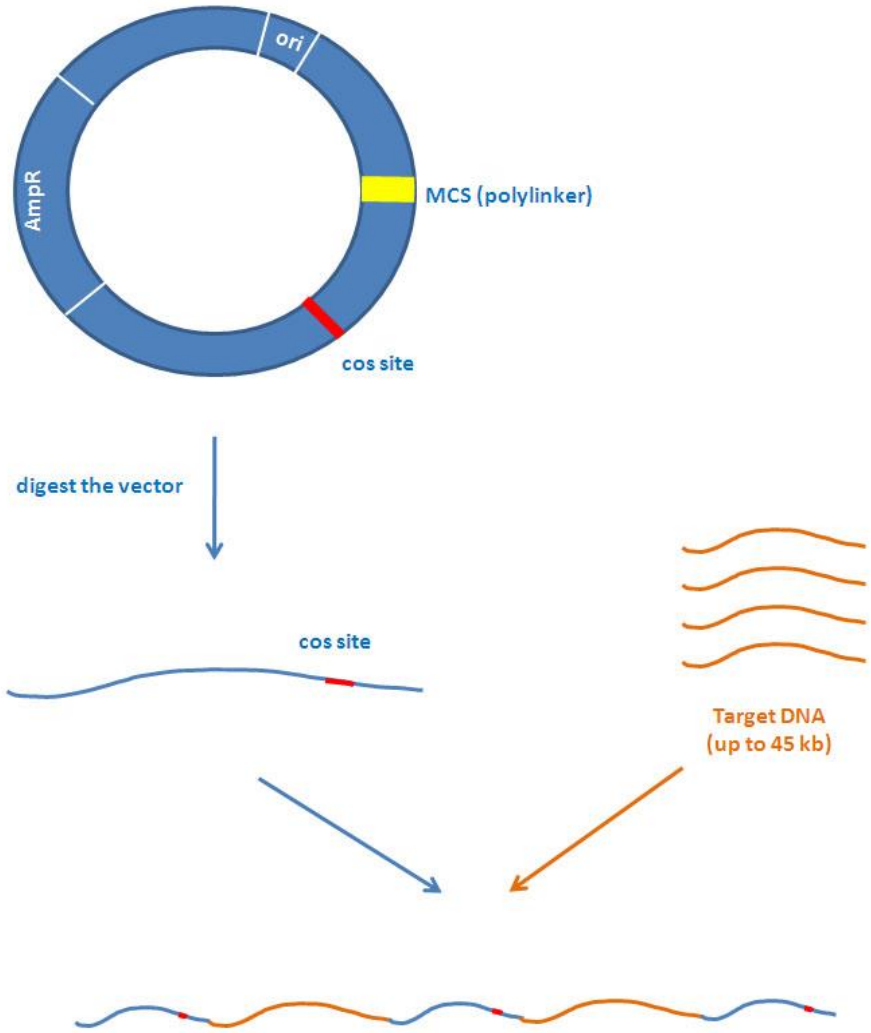


Figure 2.17. A schematic of a typical cosmid vector, showing the major elements: an origin of replication (ori), a selectable marker (such as Ampicillin resistance gene), a polylinker or MCS, and a cos site. Upon digestion of the vector within the polylinker, target DNA is cloned into the linearized vector, which can then be packaged into phage particles.

Cosmids and Phagemids

Phagemids are essentially plasmids that contain origin of replication for single-stranded phages (such as f1), thus bacteria which are transformed with this plasmid and infected with a helper phage (such as M13 or f1) can produce single-stranded copies of plasmids, which in turn can be packaged into phage heads. In the absence of a helper phage, the DNA would be propagated like a normal plasmid in bacteria. Since this vector is a hybrid of both plasmids and single-stranded phage vectors, it can be used to generate single stranded DNA to be used in sequencing reactions.

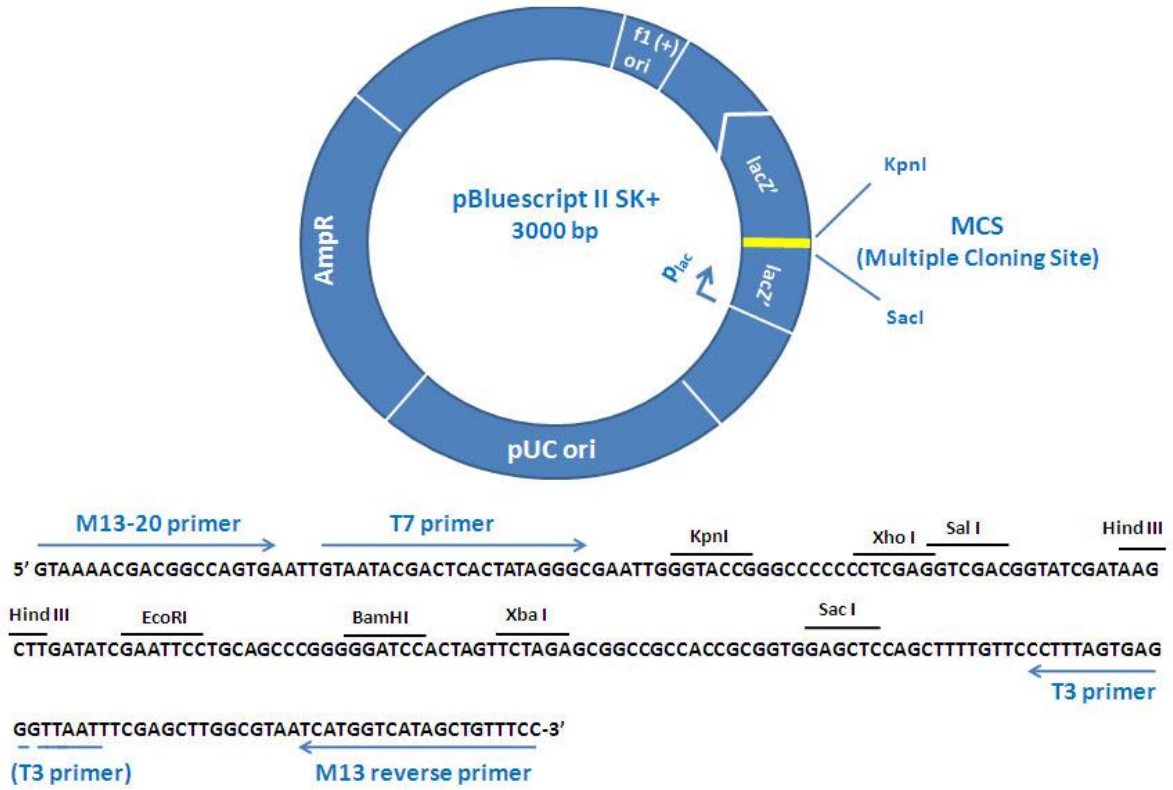


Figure 2.18. A schematic diagram of the pBluescript II SK+ phagemid, with ampicillin resistance gene, multiple cloning site (MCS), f1 origin (in + orientation, hence SK+), and pUC ori sequence. The base sequence of MCS is given below the map, with main restriction sites indicated above the sequence.

Bacterial Artificial Chromosomes (BACs)

BACs are based on the Fertility (F) plasmid that is designed to carry large DNA sequences of usually 150 to 350 kb, mainly for genome library construction purposes. They are based on F plasmids, because the *par* genes ascertain the even distribution of large sized recombinant plasmids to the next generation.

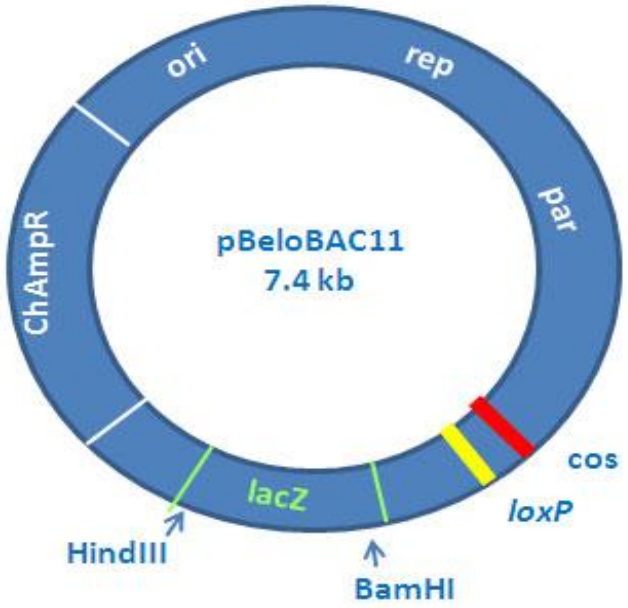


Figure 2.19. A schematic diagram of the pBeloBAC11 vector. The vector contains the chloramphenicol resistance gene as a selectable marker, and ori sequence, rep and par sequences, loxP site and cos site, in addition to a lacZ gene that includes two restriction recognition motifs for cloning (Hind III and BamHI).

Yeast Artificial Chromosomes (YACs)

YACs can mimic normal chromosome structure (hence the name "artificial chromosomes", as in BACs), as they consist of two "arms" between which large DNA fragments of up to 100 kb can be cloned. Each arm contains a telomere at the end for stabilization like a normal chromosome, and one of the arms contains an autonomous replication sequence (ARS) required for yeast chromosome replication, a centromere (CEN), and a marker for selection of recombinant yeast (usually a gene for the synthesis of an amino acid, such as *trp1* or *his3*).

Table 2.4. Similarities and differences between BAC and YAC vectors

	BAC vector	YAC vector
mode	circular	Linear
copy number	1-2 per cell	1 per cell
capacity	100 – 350 kb	almost limitless
stability of insert	Stable	unstable

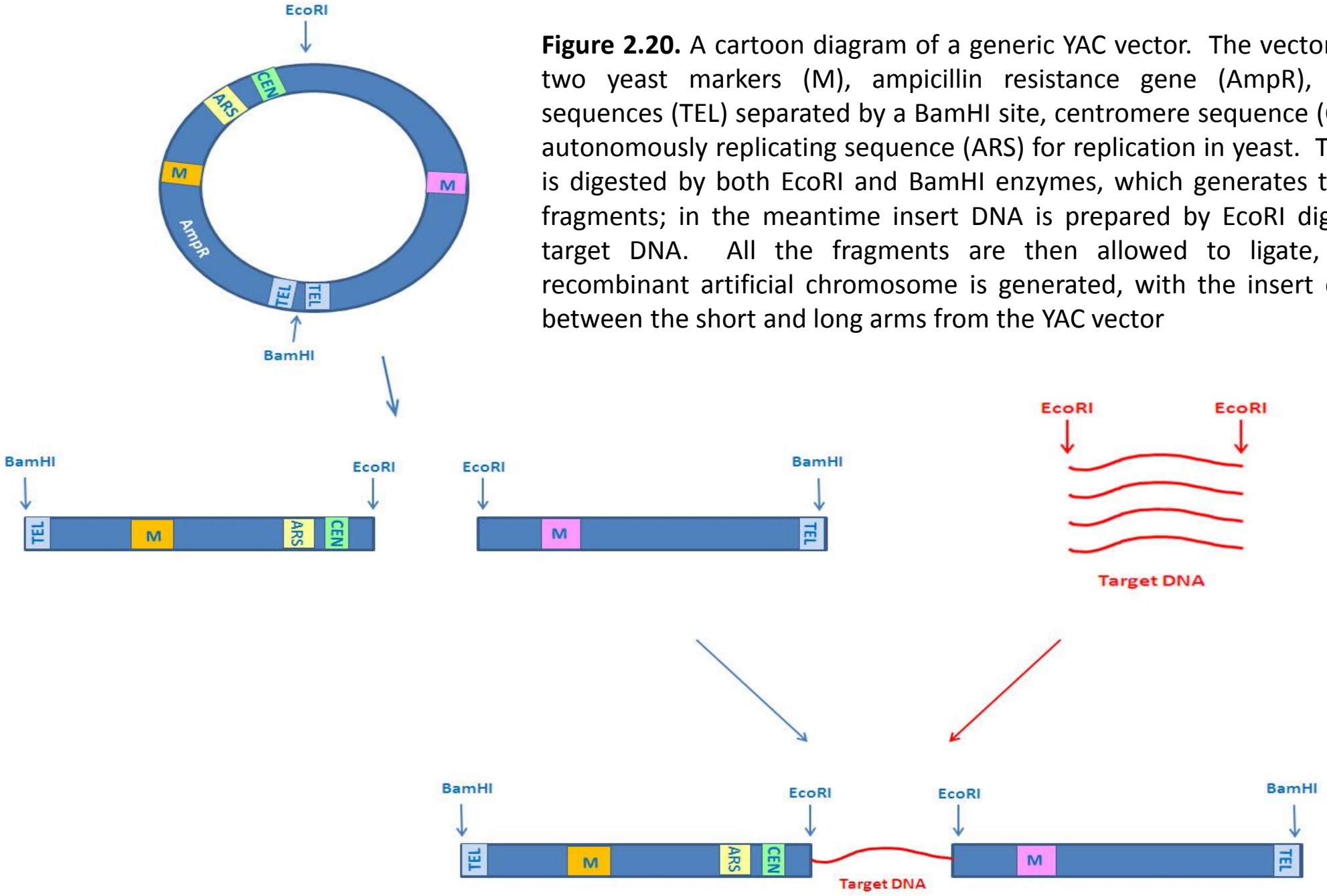


Figure 2.20. A cartoon diagram of a generic YAC vector. The vector contains two yeast markers (M), ampicillin resistance gene (AmpR), telomere sequences (TEL) separated by a BamHI site, centromere sequence (CEN), and autonomously replicating sequence (ARS) for replication in yeast. The vector is digested by both EcoRI and BamHI enzymes, which generates two linear fragments; in the meantime insert DNA is prepared by EcoRI digestion of target DNA. All the fragments are then allowed to ligate, and the recombinant artificial chromosome is generated, with the insert carried in between the short and long arms from the YAC vector

Expression vectors

If one needs to synthesize specific proteins for further analysis, there are several expression vectors that one can choose from, depending on which cell type from which organism will be used for gene expression.

pBluescript vectors (see phagemids, page 36) can be used for *in vitro* transcription and translation, due to the presence of T7 and T3 phage promoters on either side of the polylinker (Fig.2.18). However, if you would like to express your protein in a bacterial cell, a plant cell or a mammalian cell you need to use a plasmid that contains either a bacterial, plant or mammalian promoter, respectively, upstream of the gene of interest.

pGEX series of vectors are commonly used for bacterial expression of proteins in large quantities. These vectors express the gene of interest as a fusion to the glutathione-S-transferase (GST) gene driven by the *tac* promoter, a hybrid of *trp* and *lac* promoters (Fig.2.21). The promoter is upstream of the GST gene, followed by a polylinker (or, MCS) where the gene of interest will be cloned **in frame** with the GST, and this GST fusion will be used for a variety of purposes

pCMV vectors for mammalian expression, however, typically rely on the expression of the gene of interest from a promoter that strongly activates transcription in mammalian cells – since viruses are very good hijackers of transcriptional apparatus of their host cells, strong promoters from viruses that infect mammalian cells, such as cytomegalovirus (CMV) or simian virus 40 (SV40) are usually found in these vectors. In the case of pCMV, it is the CMV promoter (Fig.2.21). In the particular example given (pCMV-HA), the promoter is upstream of a hemagglutunin (HA) tag followed by an MCS (Fig.2.21). The HA tag can be used for a variety of purposes, which will be discussed in the following chapters

Figure 2.21. Cartoon diagrams of pGEX-4T-2 and pCMV-HA vectors.

(a) pGEX vectors typically contain an antibiotic resistance gene, ori sequence, a bacterial promoter, in this case the hybrid *tac* promoter, driving the expression of a glutathione-S-transferase (GST) gene followed by a polylinker or MCS sequence, where gene of interest will be cloned for bacterial expression;

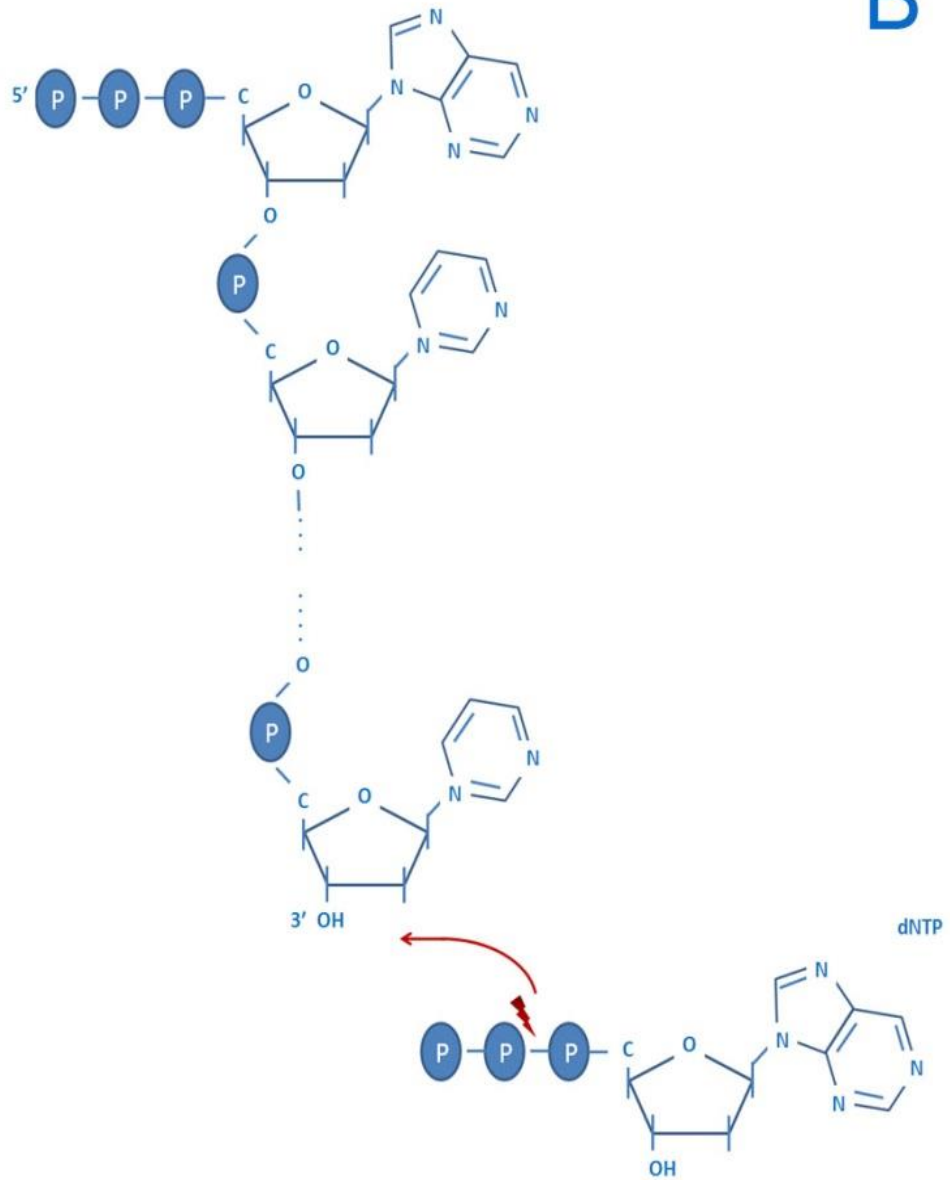
(b) pCMV vectors typically contain an ori sequence, antibiotic selection for cloning, and a cytomegalovirus (CMV) promoter upstream of an HA tag followed by an MCS

Modifying enzymes

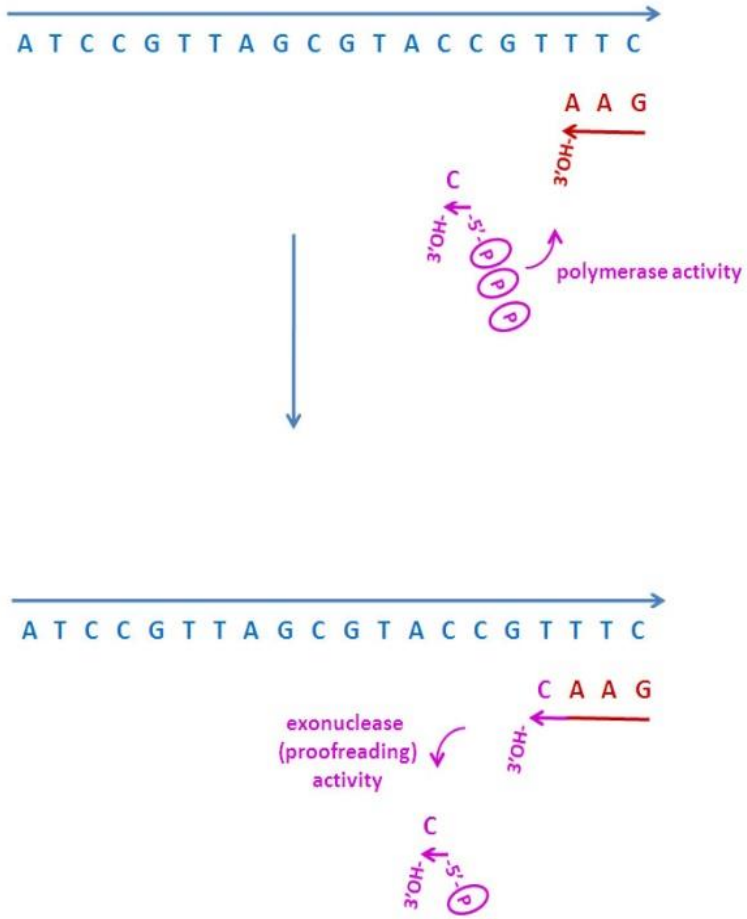
Polymerases

DNA polymerases are DNA-dependent DNA polymerases, ie they use a DNA strand as a template and synthesize a complementary DNA molecule in 5'-to-3' direction, also called the 5'-3'-polymerase activity. DNA polymerases, however, cannot start *de* novo synthesis, ie all DNA polymerases require an existing free 3'-OH group, routinely provided by **primer** sequences in molecular biology applications, to start a synthesis reaction, due to their proofreading activities (also called 3'-to-5'-exonuclease activity; Fig.2.22).

A

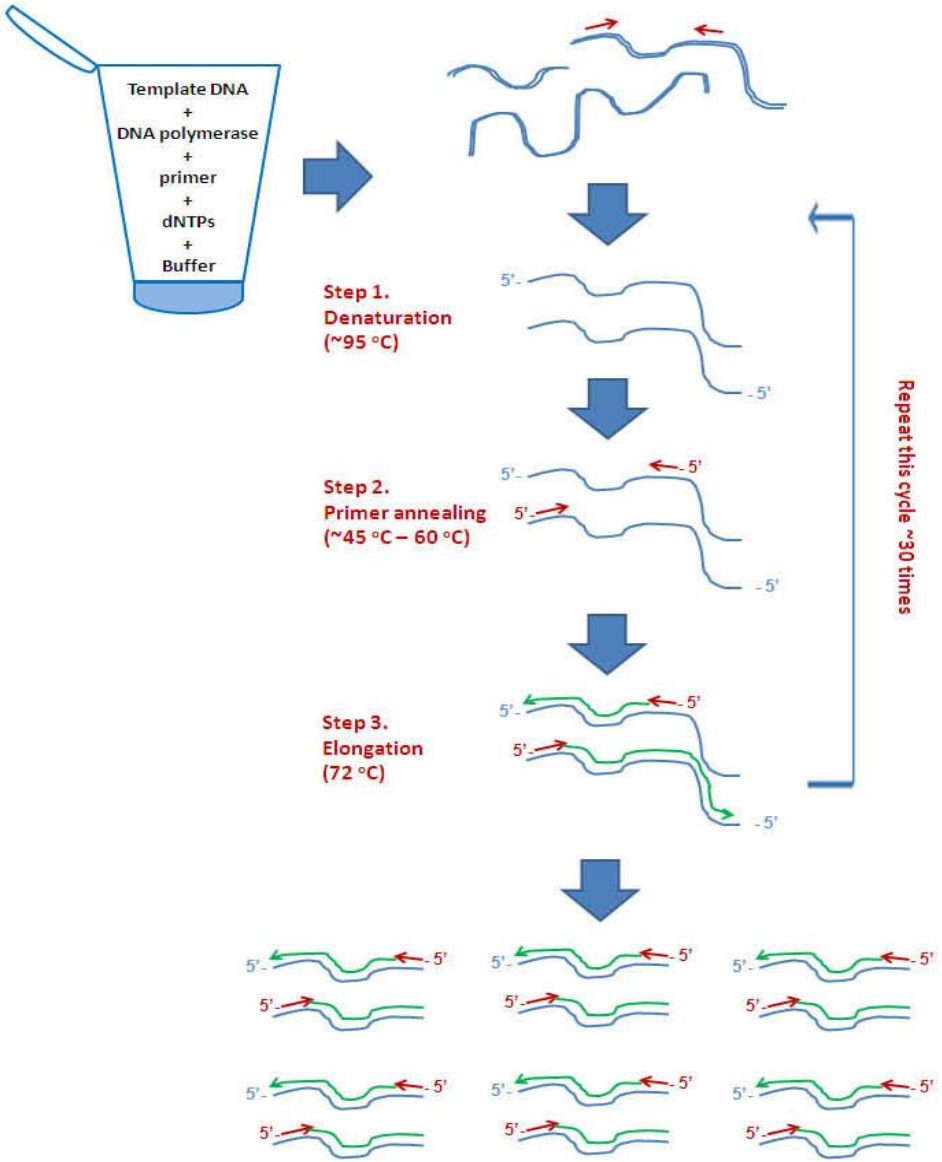

5'
P
P
P
C
O
N
N
N
N
O
P
C
O
N
N
O
O
P
C
O
N
N
3' OH
dNTP
P
P
P
C
O
N
N
N
N
OH

B


A T C C G T T A G C G T A C C G T T T C
A A G
3'OH-
C
3'OH-
5'-
P
P
P
polymerase activity
A T C C G T T A G C G T A C C G T T T C
C A A G
exonuclease
(proofreading)
activity
3'OH-
C
3'OH-
5'-
P

APPENDIX I – DNA TECNHIQUES: iii) Polymerase Chain Reaction (PCR)

A



B

5'- ATCGCTACGCTACTCGACTATCGACAGCATCGACTC-3'
CGTAGCTGAG-5'

5'-ATCGCTACGC
3'-TAGCGATGCGATGAGCTGATAGCTGTCGTAGCTGAG-5'

C

Human : 5'- ATCGCTACGCTACTCGACTATCGACAGCATCGACTC-3'
Mouse : 5'- ATAGCTTAGCTAGTCAACTAACGACAGGATCGACTC-3'
Rat : 5'- ATAGCTTCGCTACTCGACTAACGACAGCATCGACTC-3'

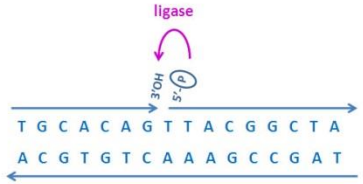
Degenerate primer: 5'-AT(C/A)GCT(A/T)(C/A)GC

Protein sequence : Ala – Met – Ile – Leu – Arg – His – Val – Ser – Cys – Tyr .....

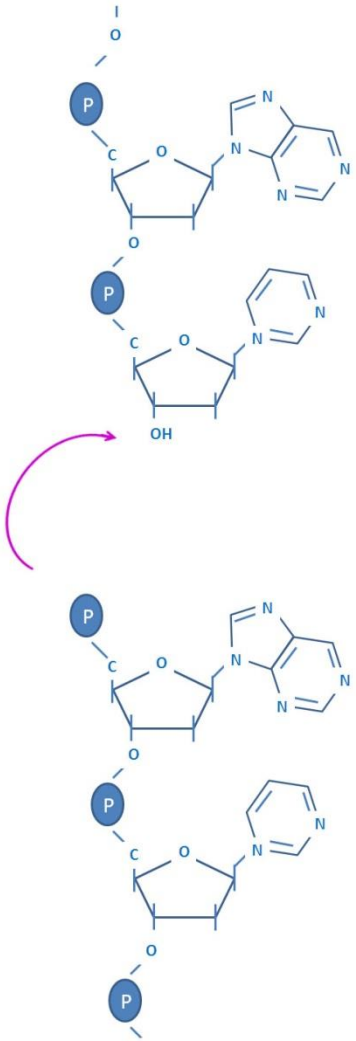
Predicted DNA sequence : GCn – ATG – ATt – CTn – CGn – CAt – GTn – TCn – TGt – TAt . .....
c TTa AGa CAc AGt c c
a TTg AGg AGc

Degenerate primer : 5'-GCnATGAT(t/c/a)CTnCG

A



B



Ligases

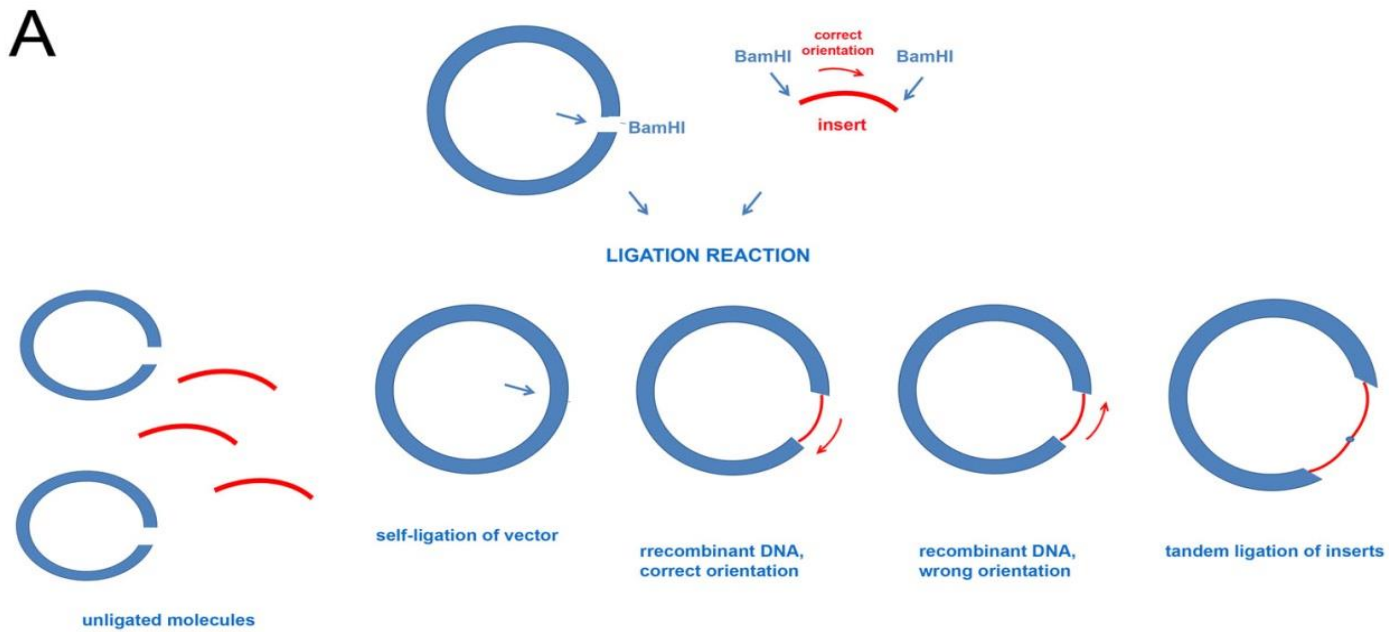
DNA ligases catalyze the covalent bond formation between a 5’-phosphate group and a 3’-hydroxyl group, however unlike DNA polymerases they cannot join a free deoxynucleotide to an existing primer sequence; instead they act as DNA repair enzymes and “seal” the “nicks” of double-strand breaks in nucleic acids .

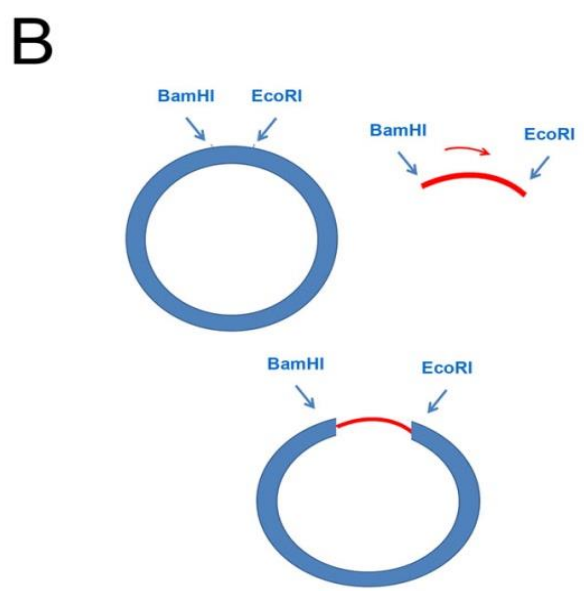
The DNA strands to be joined by DNA ligases usually result from hydrolysis of an existing sugar-phosphate backbone, therefore the 5’-end of one of the strands has only one phosphate group, instead of three phosphates in dNTPs. Therefore, ligases require external energy source so as to carry out the catalysis of phosphdiester bond formation: The *E. coli* DNA ligase requires NADH, and the T4 DNA ligase uses ATP as the source of energy required for the ligation reaction.

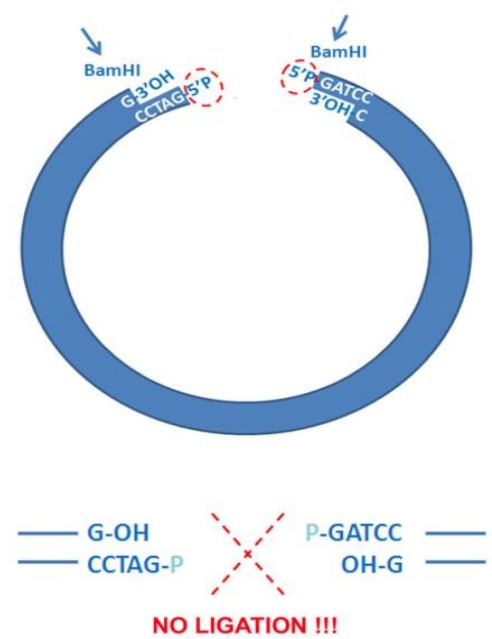
The ligation reaction will include a mixture of non-ligated fragments, correctly-ligated fragments, incorrectly-ligated fragments, self-annealing of the plasmid vector, and sometimes even tandem ligations (Fig.2.24a).

Using two different restriction enzymes for cloning (RE1 on one end of the fragment, and RE2 on the other end of the fragment) usually circumvents one of the ligation mistakes, ie ligation of the insert to the vector DNA in the wrong orientation (Fig.2.24b). This is particularly important if the insert DNA will be used for expression of proteins (where the orientation will affect promoter-driven expression).

In order to avoid self-ligation or self-annealing of the vector, it is customary to treat the digested vector with phosphatases (see Alkaline Phosphatases, below) prior to ligation – this means that the vector cannot be ligated since it lacks 5'-phosphates on both strands (Fig.2.24c); there is still the problem with vector-insert ligation reaction, however the bacterial DNA repair mechanism will correct this "mistake" of single-strand nicks upon transformation of the bacteria.


A
BamHI
correct orientation
BamHI
BamHI
insert
LIGATION REACTION
unligated molecules
self-ligation of vector
rrecombinant DNA, correct orientation
recombinant DNA, wrong orientation
tandem ligation of inserts


B
BamHI
EcoRI
BamHI
EcoRI
BamHI
EcoRI


C
BamHI
BamHI
G 3'OH
CCTAG 5'P
5'P GATCC
3'OH C
G-OH
CCTAG-P
P-GATCC
OH-G
NO LIGATION !!!

Alkaline Phosphatases

Alkaline phosphatases (AP) modify nucleic acids by removing the 5’ phosphate groups (Fig.2.25), and are mostly active at alkaline pH. The most commonly used alkaline phosphatases in cloning laboratories are the shrimp alkaline phosphatase, derived from cold water shrimp, and calf intestinal alkaline phosphatase. These enzymes are preferred in genetic engineering experiments essentially for the relative ease of inactivation – both these enzymes can easily be inactivated by heat treatment (65^{o}C for shrimp and 75^{o}C for calf intestinal AP).

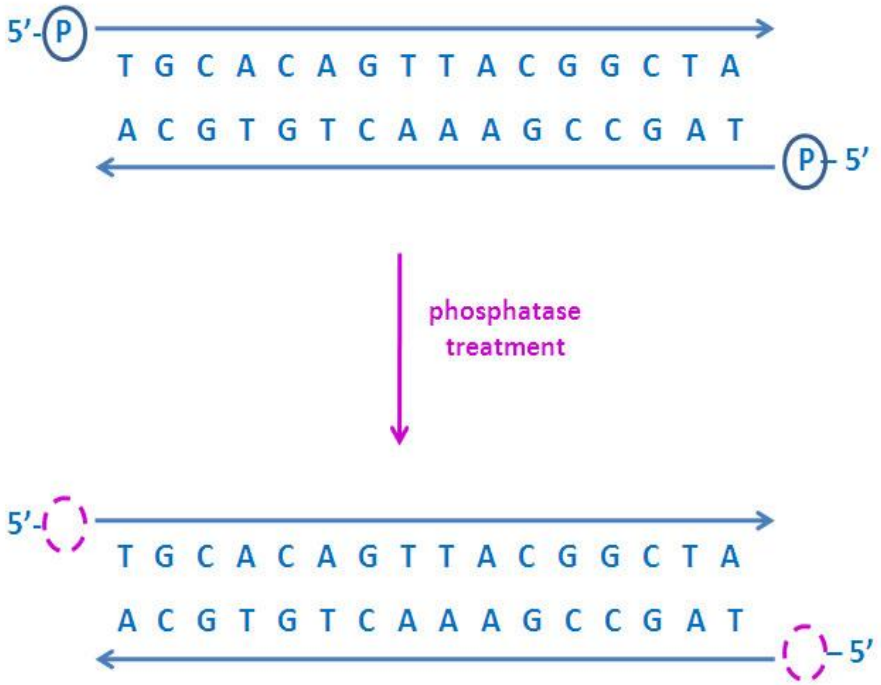


Figure 2.25 Phosphatases remove the 5’ phosphate groups from nucleic acids

Recombinases

Site-specific recombination (which will be covered in Chapter 7.c.3) could also be employed for cloning, which in the long term is an effective way to sub-clone DNA sequences from one vector to many others. The typical example is the GATEWAY™ series of vectors, which is based on bacteriophage lambda recombination mechanism, consisting of the *att* sites (*attB* site on *E. coli* and the *attP* site on the phage) and the recombinase enzyme (see Phage vectors, Fig. 2.14.a).

Basic Principles of Cloning:

Bacterial Transformation

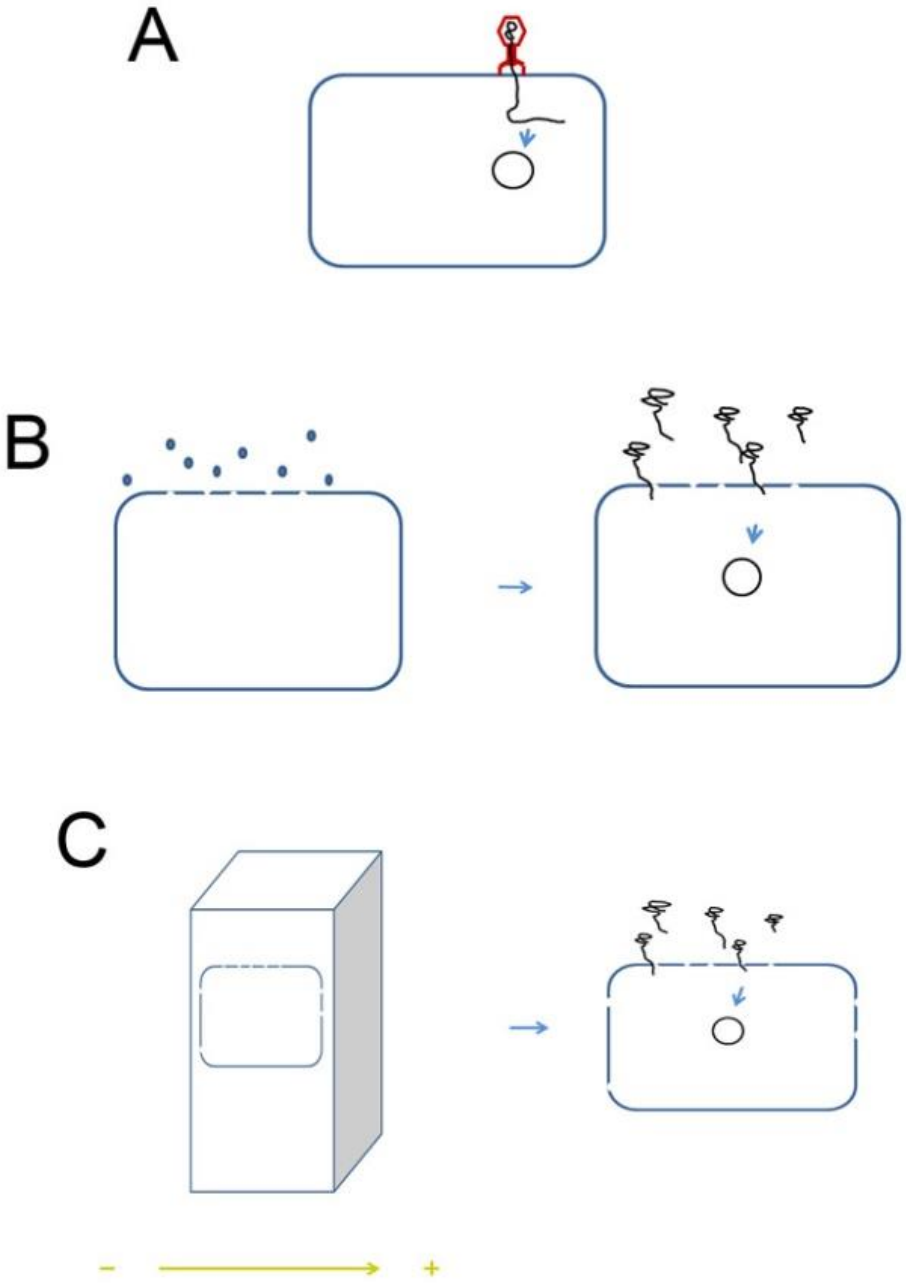
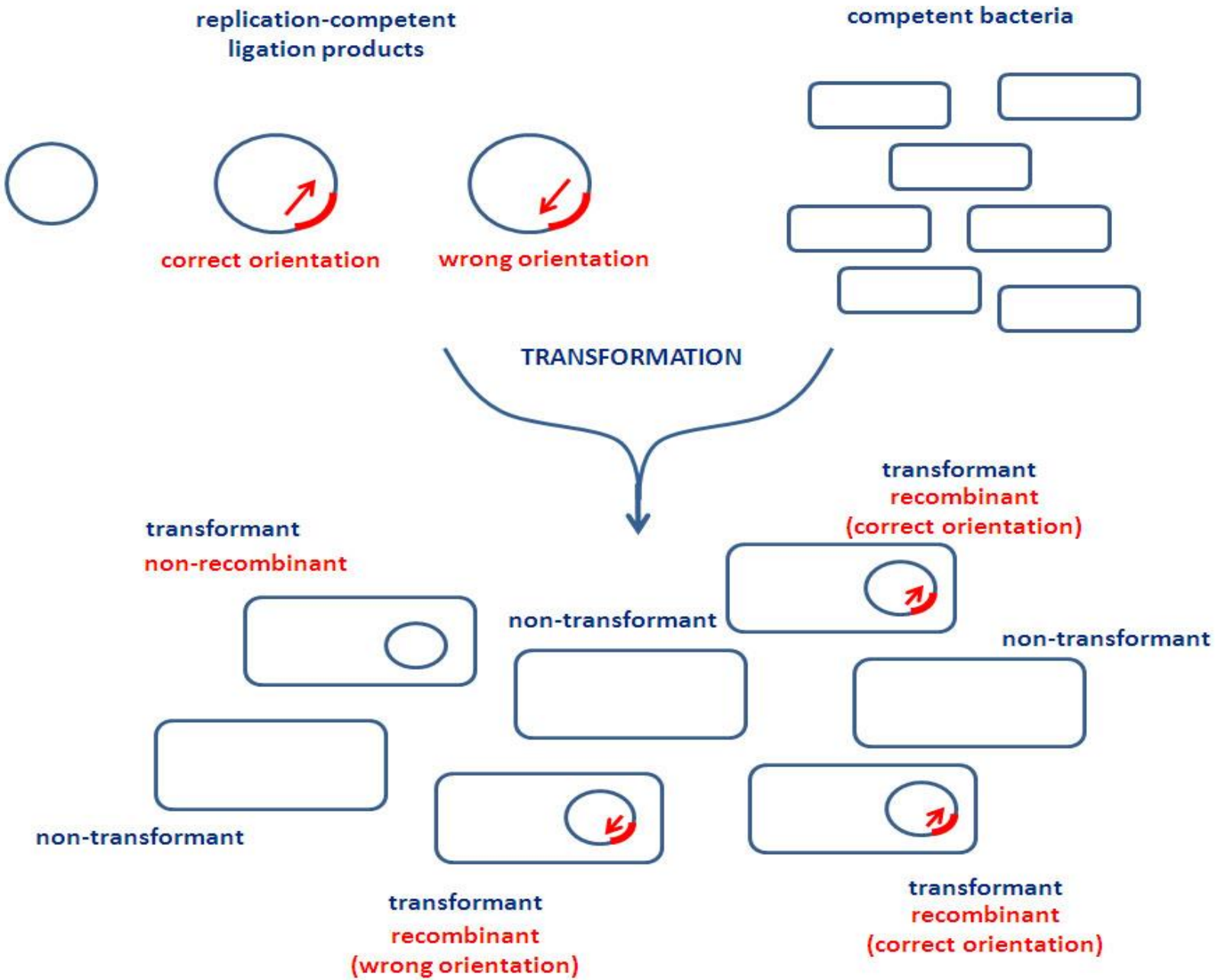


Figure 2.26 Transforming bacteria. Three of the possible methods to transfer recombinant DNA to bacterial cells are schematized. (a) Bacteria can be transformed using phage vectors; (b) bacteria can be made competent with chemicals such as calcium chloride, which creates temporary pores in the cell wall and membrane, leaving the bacteria ready to uptake foreign DNA; or (c) bacteria can be electroporated, where the cell wall and membrane are temporarily disrupted, and foreign DNA can be transferred (the electrical field is indicated in khaki).

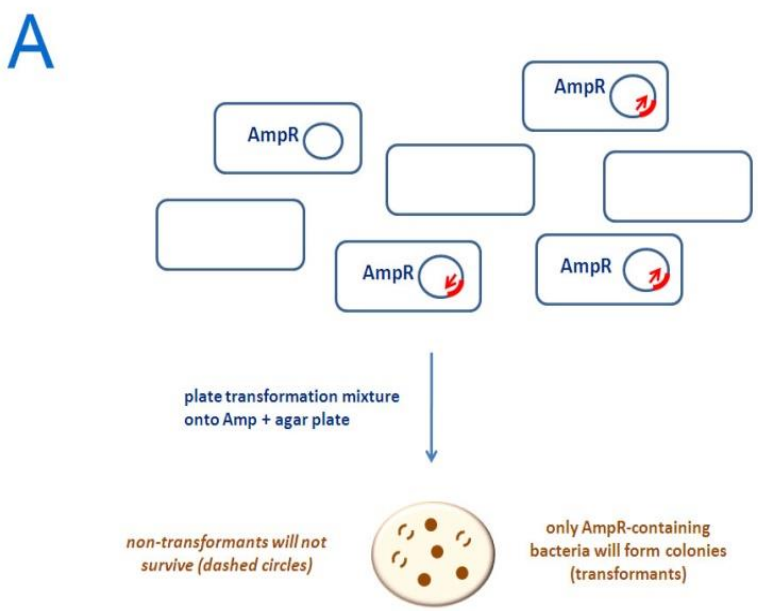
$$transformation\ efficiency = \frac{\#\ of\ recombinant\ colonies}{\#\ of\ total\ competent\ cells\ used}$$

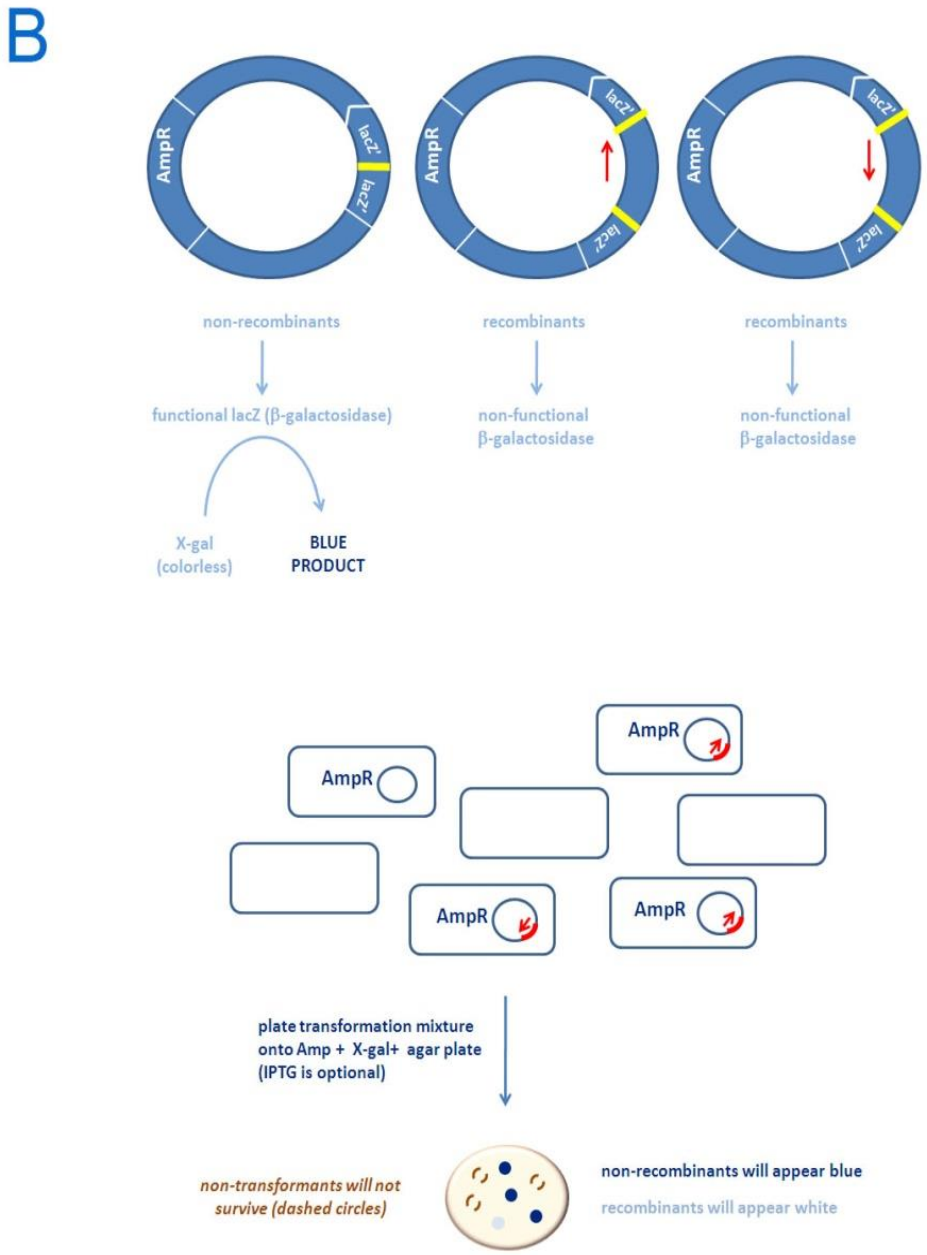
Figure 2.27 Transformation products. When competent cells are transformed with the ligation mixture, the resulting bacterial culture will contain both non-transformants, which will not grow on antibiotic selection, and transformants, which will produce colonies in antibiotic selection. The transformants could be either non-recombinant, containing self-ligated vector, or recombinant, containing insert-ligated vector with either correct or wrong orientation of the insert.



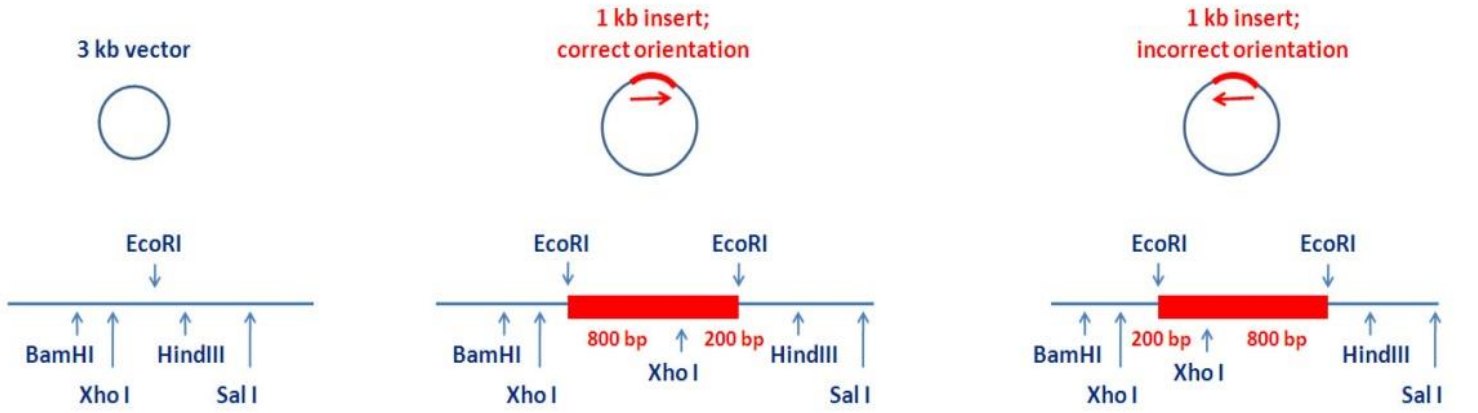
Screening for recombinants

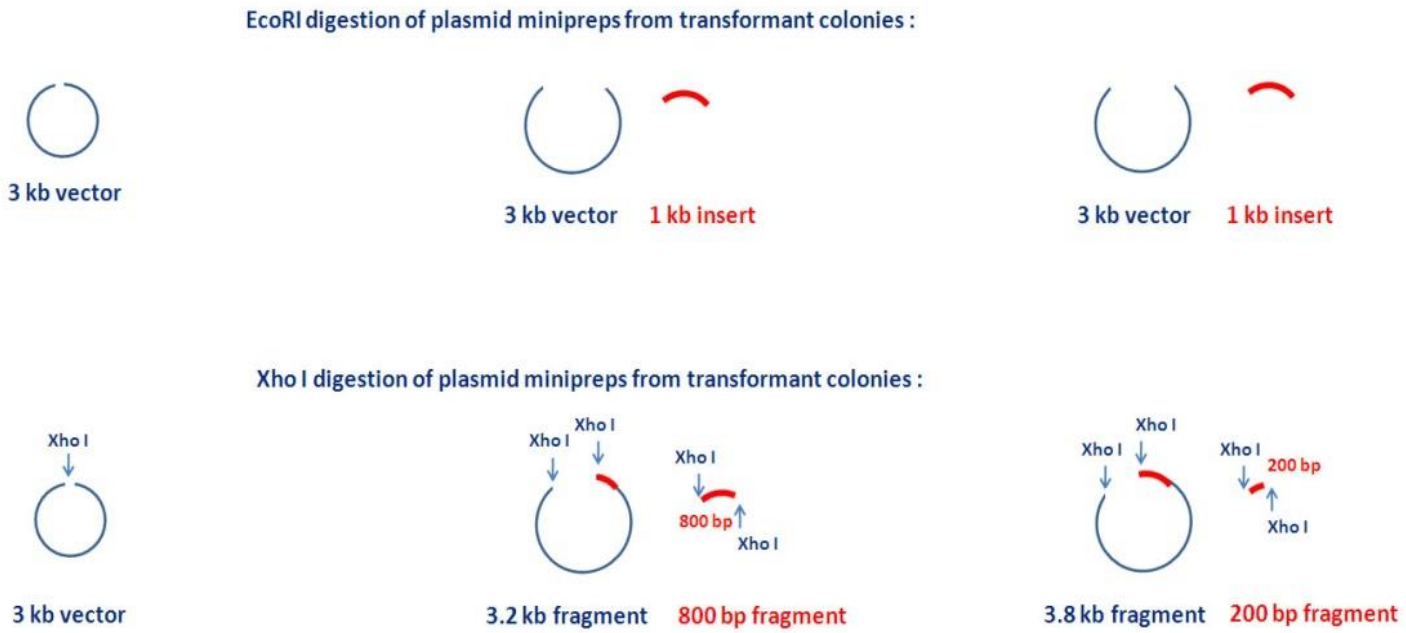
Figure 2.28 Screening for transformant colonies. (a) **Antibiotic selection** is a fast and economic method for screening for transformants, since non-transformant bacteria do not contain the plasmid bearing an antibiotic resistance gene and eventually die under antibiotic selection.





(b) **Blue-white screening** can distinguish not only between transformant and non-transformant bacteria, but also between non-recombinant vs recombinant plasmids, since non-recombinant plasmids code for intact, functional b-galactosidase and convert the colorless X-gal substrate into a blue product, whereas recombinant plasmids cannot produce an intact enzyme due to interruption of the coding sequence by insert





Restriction enzyme screening used to identify correct orientation of inserts relies on the presence of a recognition motif for an enzyme within the MCS also within the insert, different than the enzyme used in cloning the insert. This is only required if the insert and vector were digested each by a single enzyme, such as EcoRI (see Fig.2.29 to the left).

Using the same enzyme for screening will be completely uninformative, as this would simply drop out the entire insert, without giving us any clue as to which orientation the insert was cloned to begin with. However, if there is a unique recognition motif within the insert that would generate asymmetric fragments of different sizes, this motif could be used to identify the orientation of the insert.

In the example to the left, there is an Xho I site within the insert that cleaves the fragment into two fragments of 800 and 200 bp.