Department of BSBE
Indian Institute Of Technology Guwahati

BT 207

Dr. Sanjukta Patra

Vectors

Requirements for Cloning

- Enzymes
- Cloning Vectors
- Selection methods
- Expression Host
- Methods for introducing DNA into a host cell

V**ector -** a DNA molecule used as a vehicle to artificially carry foreign genetic material into another cell, where it can be replicated and/or expressed.

The four major types of vectors are plasmids, viral vectors, cosmids, and artificial chromosomes.

Of these, the most commonly used vectors are plasmids.

Vectors have an origin of replication, a multiple cloning site, and a selectable marker.

Types of vectors:

1. Cloning
2. Expression
3. Shuttle vectors

Vectors

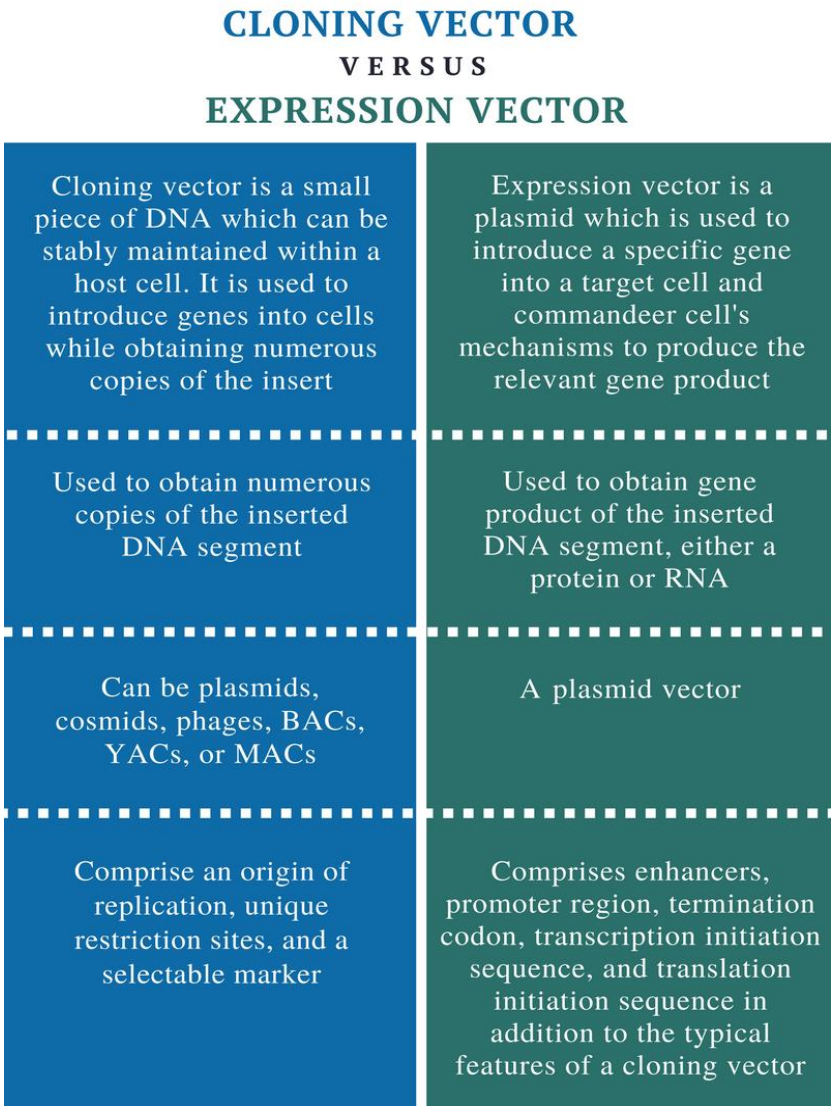
Type:

1. Cloning
2. Expression
3. Shuttle vectors

Difference between cloning and expression vector:

Cloning vector is used to carry foreign DNA segments into any host cell.

Expression vector is a type of plasmid, which contains suitable expression signals to maximize target gene expression.



Cloning vector

- **Cloning vector** - a small piece of DNA that can be stably maintained in an organism, and into which a foreign DNA fragment can be inserted for cloning purposes.
- The vector contains features that allow for the convenient insertion or removal of a DNA fragment to or from the vector.
- Treating the vector and the foreign DNA with a restriction enzyme that cuts the DNA.
- DNA fragments thus generated contain either blunt ends or overhangs known as sticky ends.
- Vector DNA and foreign DNA with compatible ends can then be joined together by molecular ligation.
- There are many types of cloning vectors, but the most commonly used ones are genetically engineered plasmids.
- Others are:
- Bacteriophages (such as phage λ)
- Cosmids
- Fosmids
- Bacterial artificial chromosomes (BACs)
- Yeast artificial chromosomes (YACs)

Should have: Origin of replication, unique restriction site, selectable marker

What determines the choice vector?

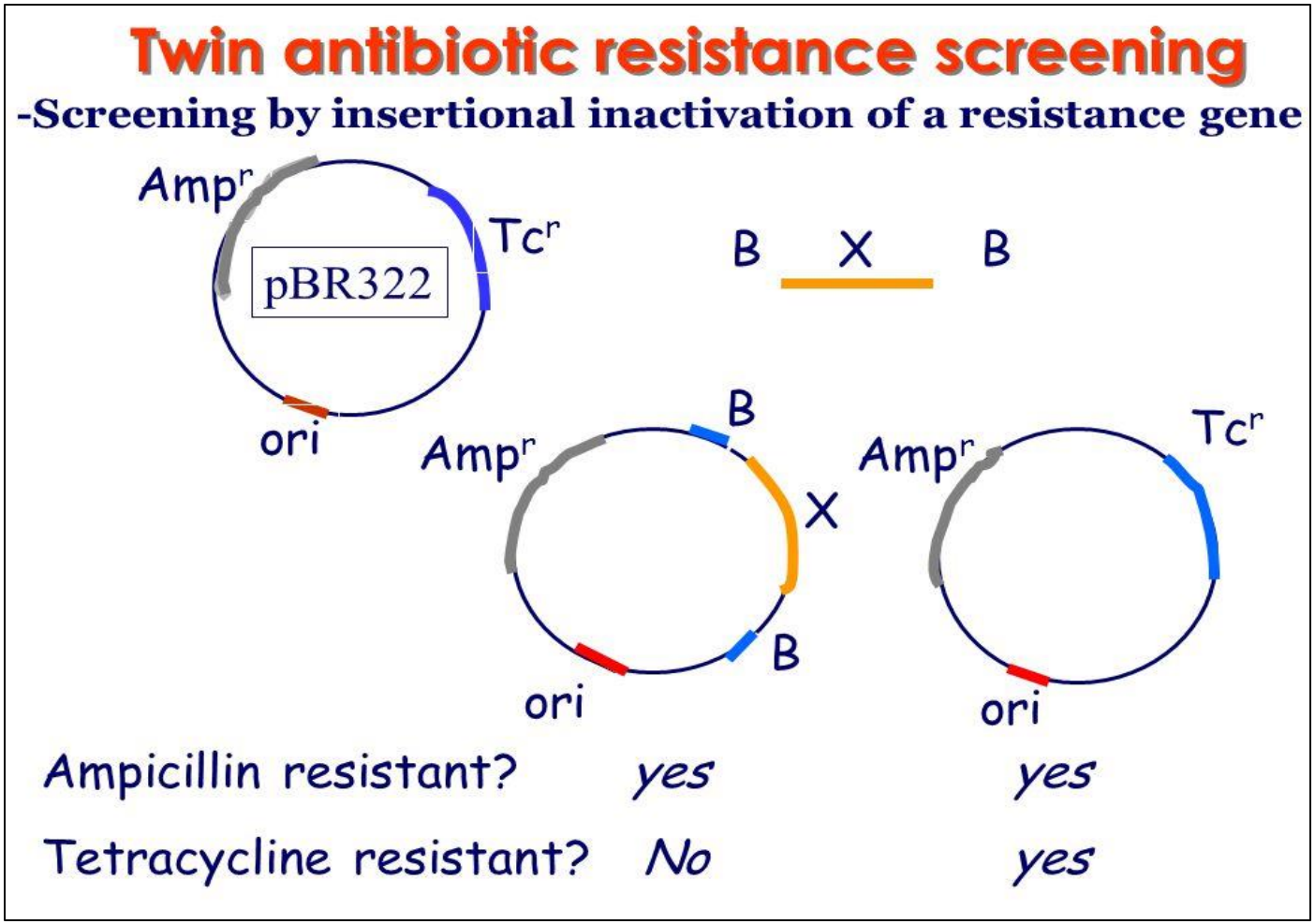
- Insert size
- Vector size
- Restriction sites
- Regulatory sequences (as promoters)
- Copy number
- Ability to screen for inserts
- Cloning efficiency

- **What down-stream experiments do you plan?**

Table 5.1 Maximum DNA insert possible with different cloning vectors. YACs are discussed on p. 159.

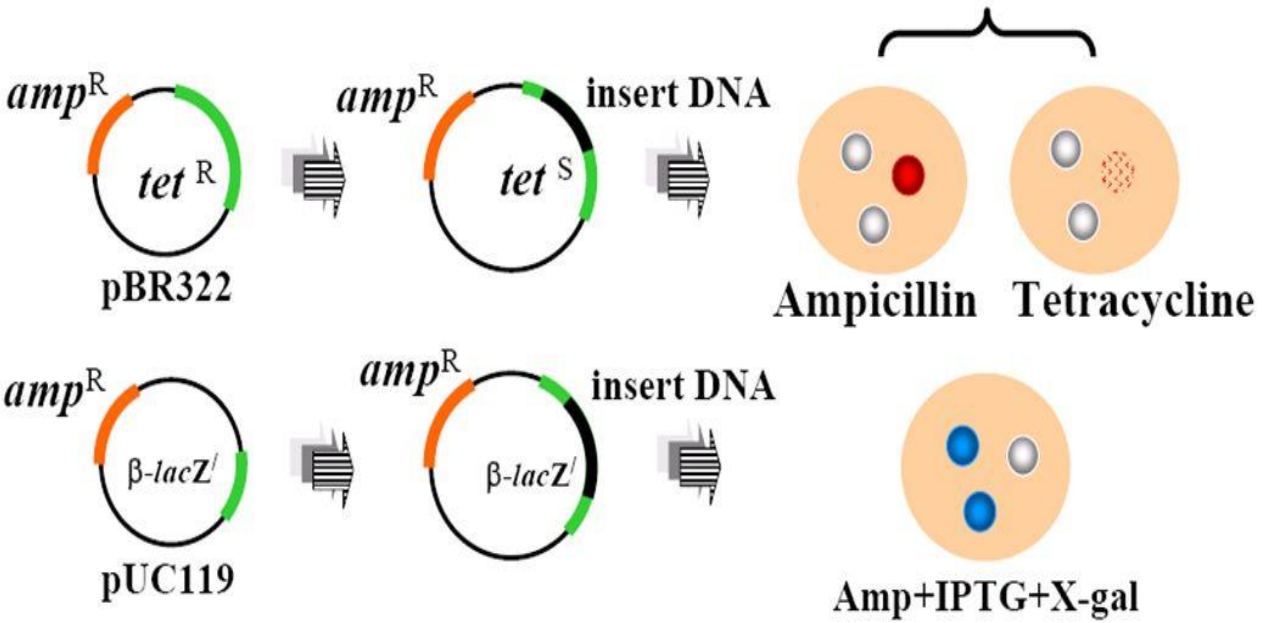
Vector	Host	Insert size
λ phage	*E. coli*	5–25 kb
λ cosmids	*E. coli*	35–45 kb
P1 phage	*E. coli*	70–100 kb
PACs	*E. coli*	100–300 kb
BACs	*E coli*	≤ 300 kb
YACs	*Saccharomyces cerevisiae*	200–2000 kb

SCREENING

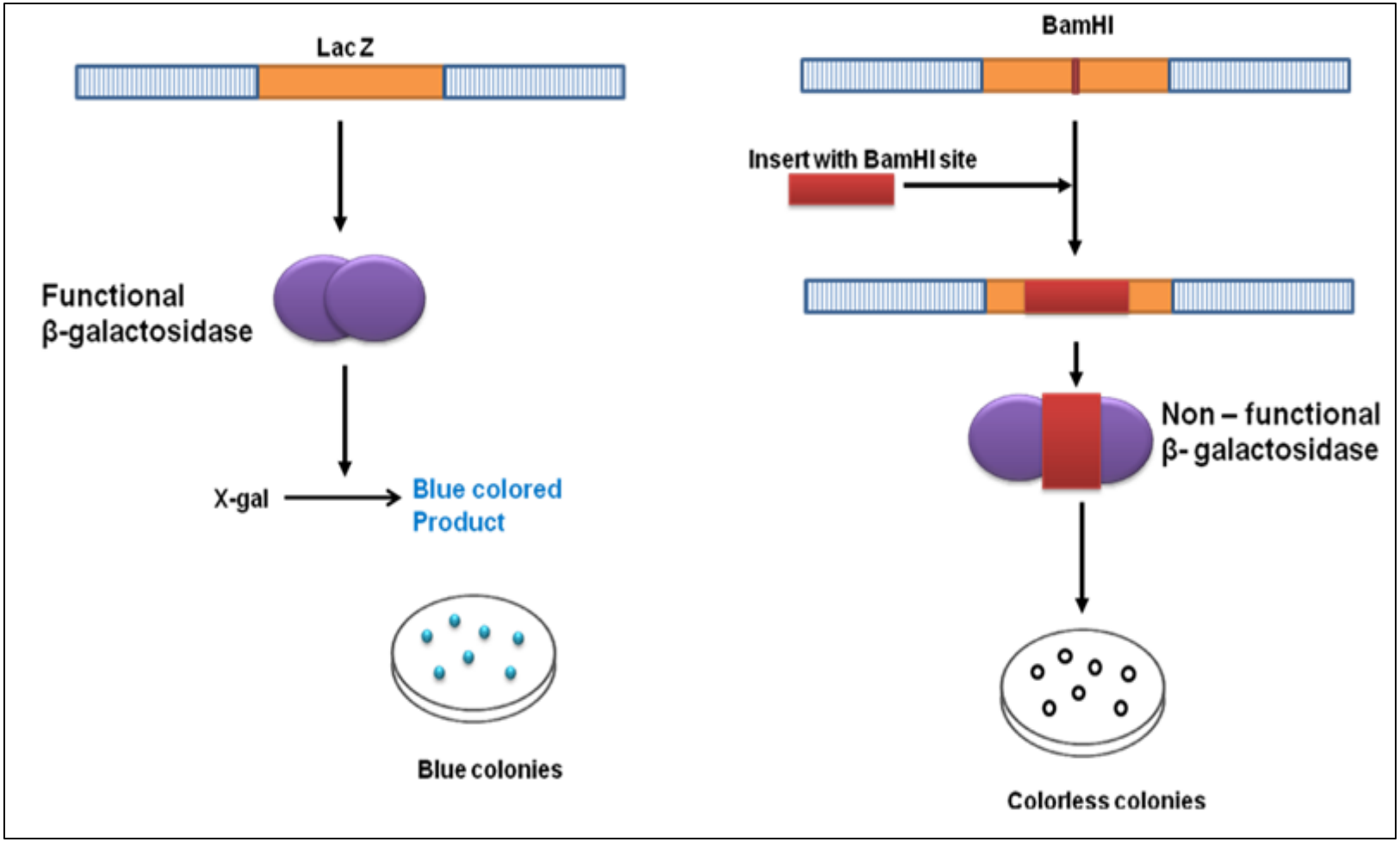

Twin antibiotic resistance screening
-Screening by insertional inactivation of a resistance gene
Amp^r
Tc^r
pBR322
ori
B X B
B
Amp^r
X
B
ori
Amp^r
Tc^r
ori
Ampicillin resistant? yes yes
Tetracycline resistant? No yes

H1-2 Twin antibiotic resistance

A vector with two antibiotic resistance genes can be used to screen for recombinants if the target fragment is inserted into one of the genes, thus insertionally inactivating it.



Blue white screening



Cloning vectors

Allowing the exogenous DNA to be inserted, stored, and manipulated mainly at DNA level.

1. **Plasmid vectors**
2. **Bacteriophage vectors**
3. **Cosmids**
4. **Fosmids**
5. **BACs & YACs**
6. **MACs**

What determines the choice vector?

- Insert size
- Vector size
- Restriction sites
- Regulatory sequences (as promoters)
- Copy number
- Cloning efficiency
- Ability to screen for inserts

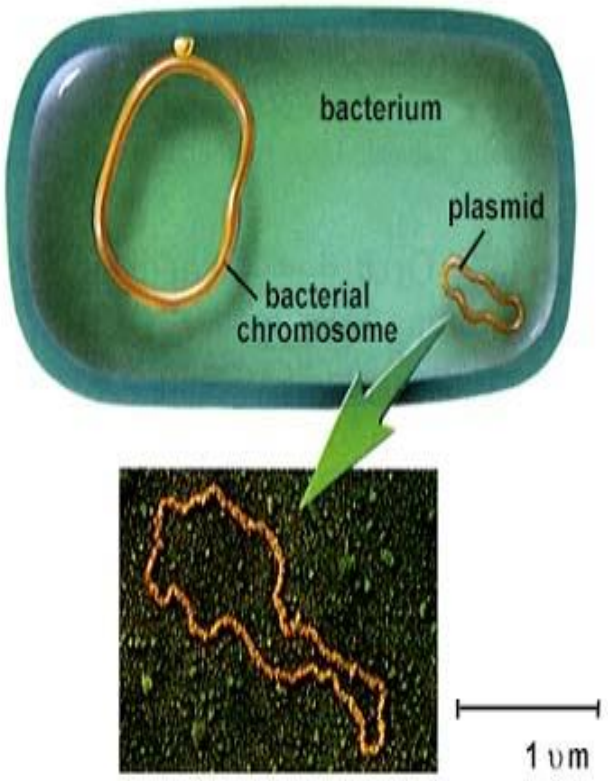
- **What down-stream experiments do you plan?**

Table 5.1 Maximum DNA insert possible with different cloning vectors. YACs are discussed on p. 159.

Vector	Host	Insert size
λ phage	*E. coli*	5–25 kb
λ cosmids	*E. coli*	35–45 kb
P1 phage	*E. coli*	70–100 kb
PACs	*E. coli*	100–300 kb
BACs	*E coli*	≤ 300 kb
YACs	*Saccharomyces cerevisiae*	200–2000 kb

Plasmid vectors

- Double-stranded, circular, self-replicating, extra-chromosomal DNA molecules.
- Circular DNA molecules present in the cytoplasm of the bacteria
- Capable of autonomous replication
- Can transfer genes from one cell to other
- Act as vectors in genetic engineering.
- Can also be present in Yeasts



Advantages:

Small, easy to handle

Useful for cloning small DNA fragments (< 10kbp)

Disadvantages:

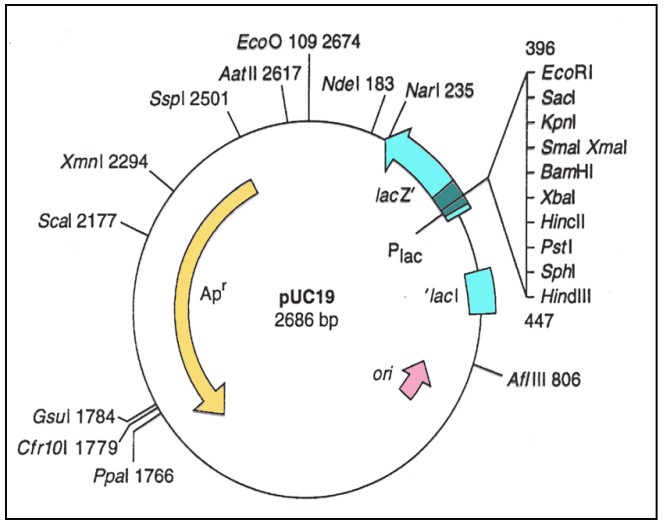
Less useful for cloning large DNA fragments (> 10kbp)

Plasmid vector:

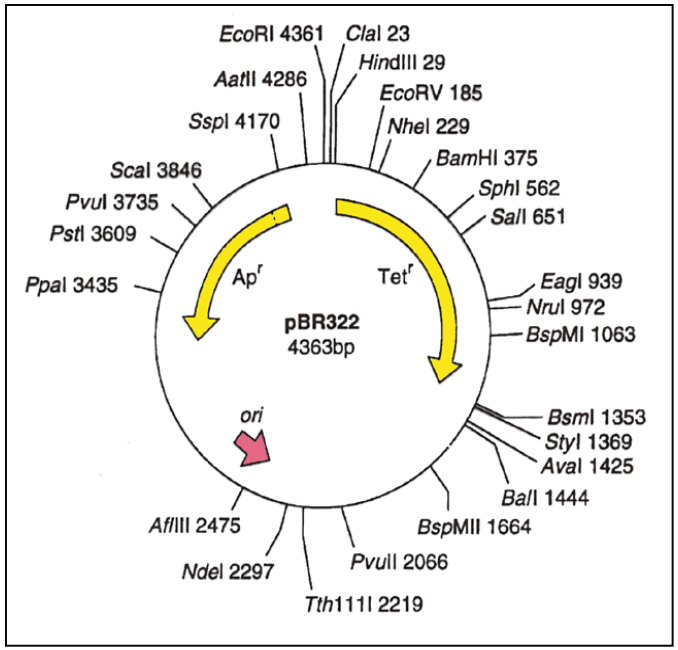
Plasmid vectors may encode genetic information for properties:

1. Resitance to Antibiotics
2. Bacteriocin production
3. Enterotoxin production
4. Enhanced pathogenicity
5. Reduced Sensitivity to mutagens
6. Degrade complex organic molecules

- Contains an origin of replication, allowing for replication independent of host's genome.
- Contains Selective markers: Selection of cells containing a plasmid
  - twin antibiotic resistance
  - blue-white screening
- Contains a multiple cloning site (MCS)
- Easy to be isolated from the host cell.



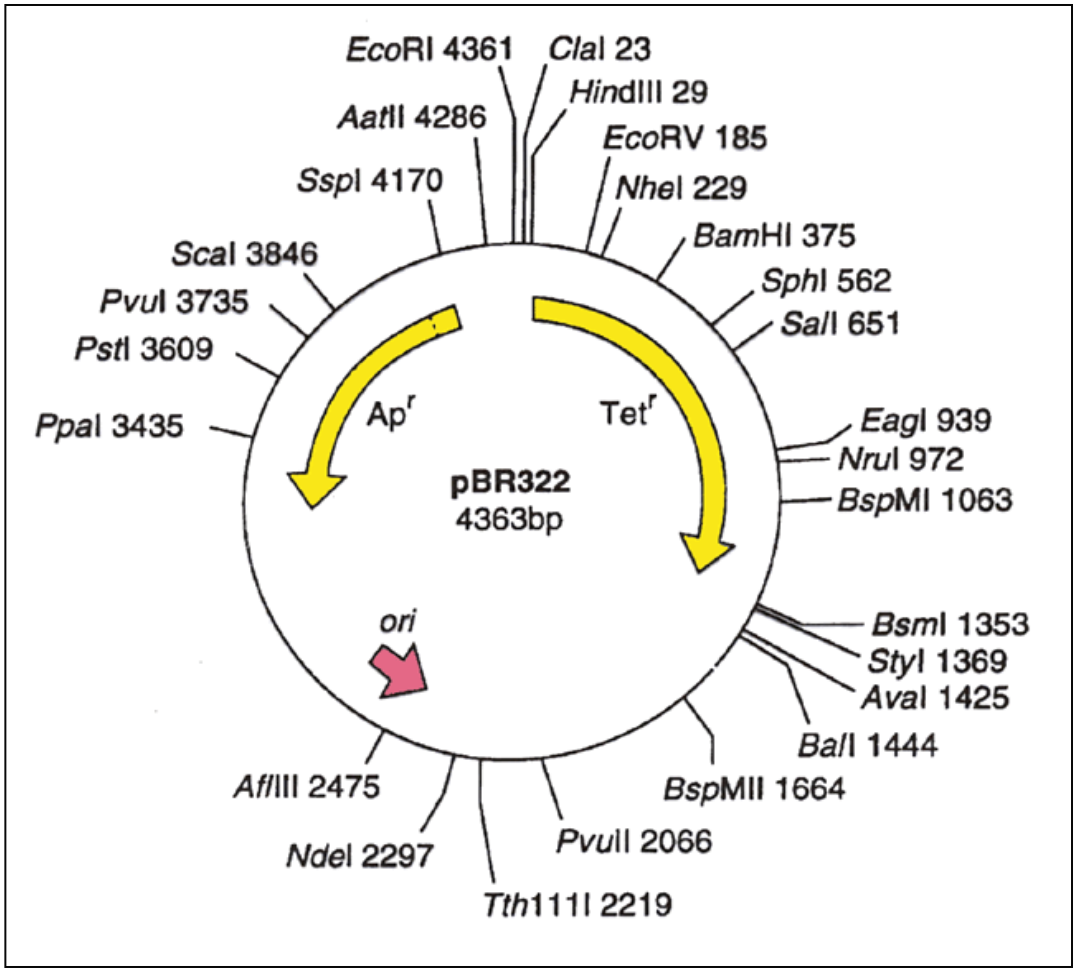
Vector Map



- **pBR322** is a plasmid and was one of the first widely used *E. coli* cloning vectors. Created in 1977 in the laboratory of Herbert Boyer at the University of California, San Francisco.

- Named after Francisco Bolivar Zapat - p stands for "plasmid," and BR for "Bolivar" and "Rodriguez.“

- pBR322 is 4361 base pairs in length and has two antibiotic resistance genes – the gene *bla* encoding the ampicillin resistance (Amp^R) protein, and the gene *tetA* encoding the tetracycline resistance (Tet^R) protein.
- It contains the origin of replication of pMB1, and the *rop* gene, which encodes a restrictor of plasmid copy number.
- The plasmid has unique restriction sites for more than forty restriction enzymes. Eleven of these forty sites lie within the Tet^R gene. There are two sites for restriction enzymes HindIII and ClaI within the promoter of the Tet^R gene.
- There are six key restriction sites inside the Amp^R gene.
- The circular sequence is numbered such that after the unique EcoRI site, the count increases through the Tet^R gene.
- The Amp^R gene is penicillin beta-lactamase.
- Promoters P1 and P3 are for the beta-lactamase gene. P3 is the natural promoter, and P1 is artificially created by the ligation of two different DNA fragments to create pBR322. P2 is in the same region as P1, but it is on the opposite strand and initiates transcription in the direction of the tetracycline resistance gene.

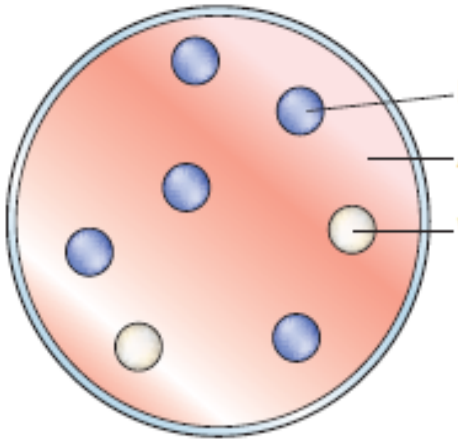

EcoRI 4361
ClaI 23
HindIII 29
AatII 4286
EcoRV 185
SspI 4170
NheI 229
BamHI 375
ScaI 3846
SphI 562
PvuI 3735
SalI 651
PstI 3609
Ap^r
Tet^r
PpaI 3435
EagI 939
NruI 972
pBR322
4363bp
BspMI 1063
ori
BsmI 1353
StyI 1369
AvaI 1425
BalI 1444
AflIII 2475
BspMII 1664
NdeI 2297
PvuII 2066
Tth111I 2219

pBR322

pUC series

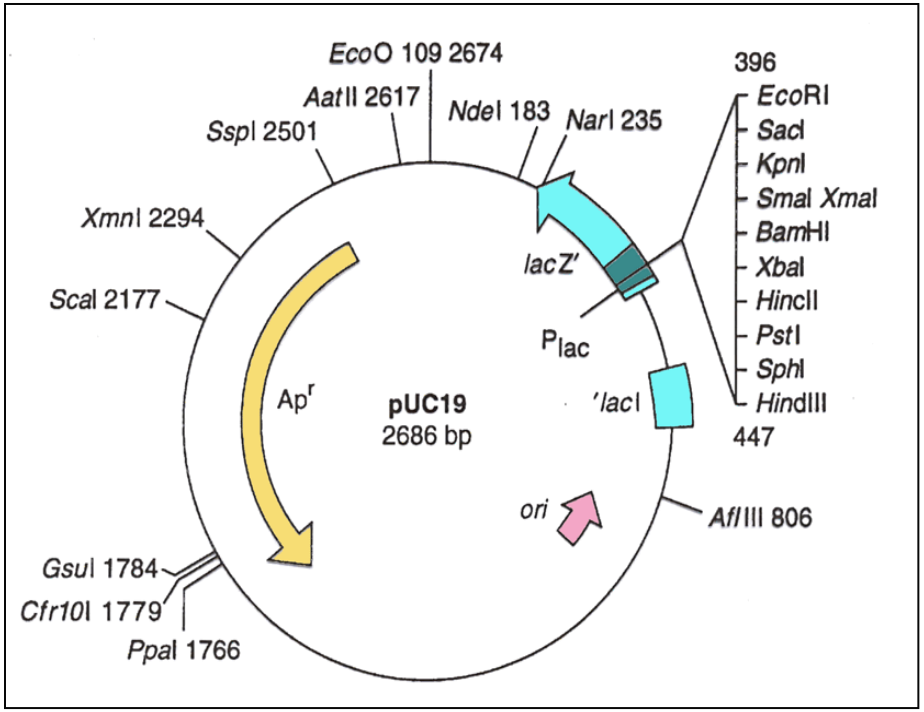
- The features of pBR322 have been enhanced by the construction of a series of plasmids termed pUC (produced at the University of California).
- There is an antibiotic resistance gene for tetracycline and origin of replication for *E. coli*. There is presence of multiple restriction sites termed the multiple cloning site (MCS).
- It additionally codes for a portion of a polypeptide called β-galactosidase facilitating blue white screening.
- When the pUC plasmid has been used to transform the host cell *E. coli* the gene may be switched on by adding the inducer IPTG (isopropyl-β-D-thiogalactopyranoside).
- Its presence causes the enzyme β-galactosidase to be produced. The functional enzyme is able to hydrolyse a colourless substance called X-gal (5-bromo-4-chloro-3-indolylb-galactopyranoside) into a blue insoluble material (5,50-dibromo-4,40–dichloro indigo).
- However if the gene is disrupted by the insertion of a foreign fragment of DNA, a non-functional enzyme results which is unable to carry out hydrolysis of X-gal.
- This makes the recombinant pUC plasmid to be easily detected as it is white or colourless in the presence of X-gal.
- An intact non-recombinant pUC plasmid will be blue since its gene is fully functional and not disrupted. This elegant system, termed blue/white selection.

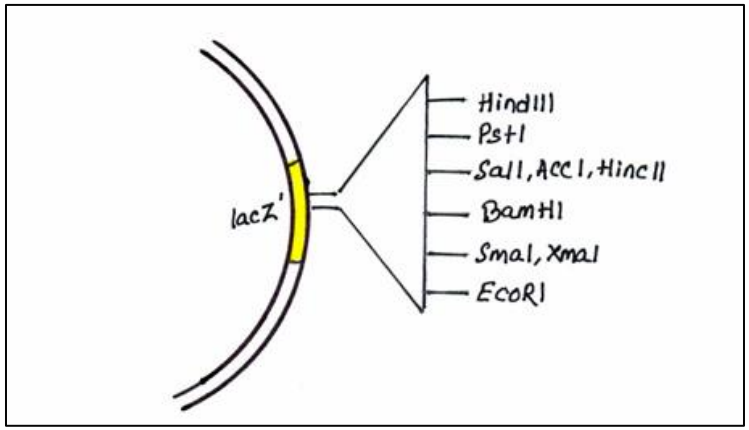
Screening for pUC8 recombinants



Blue colonies = β-galactosidase synthesized
X-gal → blue product
White colonies = β-galactosidase not synthesized
X-gal → no blue product

Recombinants are screened by plating onto agar containing X-gal and IPTG.


EcoO 109 2674
AatII 2617
NdeI 183
NarI 235
SspI 2501
XmnI 2294
ScaI 2177
GsuI 1784
Cfr10I 1779
PpaI 1766
AflIII 806
396
EcoRI
SacI
KpnI
SmaI XmaI
BamHI
XbaI
HincII
PstI
SphI
HindIII
447
lacZ'
Plac
'lacI
ori
Ap^r
pUC19
2686 bp


lacZ'
HindIII
PstI
SalI, ACCI, HincII
BamHI
SmaI, XmaI
EcoRI

pUC19

- **pUC19** is one of a series of plasmid cloning vectors created by Joachim Messing and co-workers.
- The designation "pUC" is derived from the classical "p" prefix (denoting "plasmid") and the abbreviation for the University of California, where early work on the plasmid series had been conducted.
- It is a circular double stranded DNA and has 2686 base pairs.
- pUC19 – blue white screening.
- pUC18 is similar to pUC19, but the MCS region is reversed. Notably, it has a N-terminal fragment of β-galactosidase (*lacZ*) gene of *E. coli*.
- The multiple cloning site (MCS) region is split into codons 6-7 of the lacZ gene, providing for many restriction endonucleases restriction sites.
- In addition to β-galactosidase, pUC19 also encodes for an ampicillin resistance gene (amp^R), via a β-lactamase enzyme that functions by degrading ampicillin and reducing its toxicity to the host.
- The *ori* site, or origin of replication, is derived from the plasmid pMB1.
- pUC19 is small but has a high copy number. The high copy number is a result of the lack of the *rop* gene and a single point mutation in the ori of pMB1.
- The *lacZ* gene codes for β-galactosidase. The recognition sites for HindIII, SphI, PstI, SalI, XbaI, BamHI, SmaI, KpnI, SacI and EcoRI restric tion enzymes have been derived from the vector M13mp19

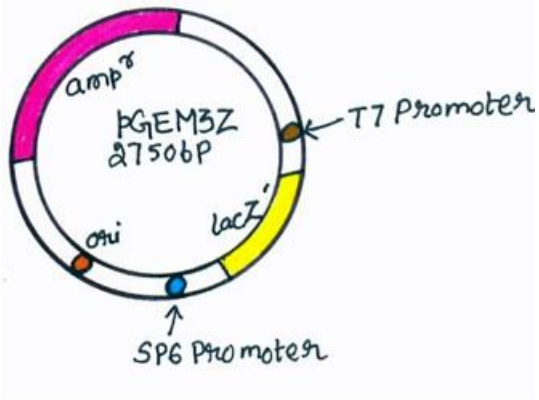
pUC vectors carry different combinations of restriction sites and show greater flexibility in the types of DNA fragment that can be cloned.

Clustering of the restriction sites allows a DNA fragment with two different sticky ends to be cloned without involving linker attachment.

DNA cloned into a member of the pUC series can be transferred directly to its M13mp counterpart, because of the same restriction site clusters and it can be analysed by DNA sequencing or in vitro mutagenesis.

PGEM3Z plasmid

- pGEM3Z is having similarity with a pUC vector because of the presence of the amp^R genes and *lacZ'* genes containing a cluster of restriction sites.
- The difference 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.
- These two promoter sequences lie on either side of the cluster of restriction site used for introduction of new DNA into the pGEM3Z molecule.
- **Advantages:**
- High copy number
- Identification can be achieved by single step by plating on agar medium containing ampicillin and X gal.
- Time consumption is less.



Expression vector

Elements of expression vectors

- an origin of replication
- a selectable marker
- and a suitable site for the insertion of a gene like the multiple cloning site

- Usually a plasmid or virus is designed for gene expression in cells.
- The vector is used to introduce a specific gene into a target cell, and can commandeer the cell's mechanism for protein synthesis to produce the protein encoded by the gene.
- The vector is engineered to contain regulatory sequences that act as enhancer and promoter regions and lead to efficient transcription of the gene carried on the expression vector.
- The goal of a well-designed expression vector is the efficient production of protein, achieved by the production of significant amount of stable messenger RNA, which can then be translated into protein.
- The expression of a protein may be tightly controlled, and the protein is only produced in significant quantity when necessary through the use of an inducer, in some systems however the protein may be expressed constitutively.
- *Escherichia coli* is commonly used as the host for protein production, but other cell types may also be used.

Prokaryotic expression vector

- Promoter - commonly use inducible promoters derived from *lac* operon and the T7 promoter. Other strong promoters used include Trp promoter and Tac-Promoter, which are a hybrid of both the Trp and Lac Operon promoters.
- Ribosome binding site (RBS) - follows the promoter, and promotes efficient translation of the protein of interest.
- Translation initiation site - Shine-Dalgarno sequence enclosed in the RBS, 8 base-pairs upstream of the AUG start codon.

Eukaryotes expression vector

Eukaryote expression vectors require sequences that encode for:

- Polyadenylation tail: Creates a polyadenylation tail at the end of the transcribed pre-mRNA that protects the mRNA from exonucleases and ensures transcriptional and translational termination: stabilizes mRNA production.
- Minimal UTR length: UTRs contain specific characteristics that may impede transcription or translation, and thus the shortest UTRs or none at all are encoded for in optimal expression vectors.
- Kozak sequence: Vectors should encode for a Kozak sequence in the mRNA, which assembles the ribosome for translation of the mRNA.

Examples of mammalian **expression vectors** include

- the adenoviral **vectors**
- the pSV
- the pCMV series of plasmid **vectors**
- vaccinia and retroviral **vectors**
- baculovirus

The promoters for cytomegalovirus (CMV) and SV40 are commonly used in mammalian **expression vectors** to drive gene **expression**.

A **shuttle vector** - a plasmid that can propagate in two different host species.
DNA inserted into a shuttle vector can be tested or manipulated in two different cell types.
Advantage - they can be manipulated in *E. coli*, then used in a system which is more difficult or slower to use (e.g. yeast).

1. Shuttle vectors include plasmids that can propagate in eukaryotes and prokaryotes (e.g. both *Saccharomyces cerevisiae* and *Escherichia coli*.

2. In different species of bacteria example - both *E. coli* and *Rhodococcus erythropolis*.

3. Adenovirus shuttle vectors, which can propagate in *E. coli* and mammals.

Shuttle vectors are frequently used to quickly make multiple copies of the gene in *E. coli* (amplification).

Example - yeast shuttle vector. Yeast shuttle vectors have components that allow for replication and selection in both *E. coli* cells and yeast cells.

The *E. coli* component of a yeast shuttle vector includes an origin of replication and a selectable marker, e.g. antibiotic resistance, beta lactamase, beta galactosidase.

The yeast component of a yeast shuttle vector includes an autonomously replicating sequence (ARS), a yeast centromere (CEN), and a yeast selectable marker (e.g. URA3, a gene that encodes an enzyme for uracil synthesis).

Questions