MODULE 1- LECTURE 3

TYPES, BIOLOGY AND SALIENT FEATURES OF VECTORS IN RECOMBINANT DNA TECHNOLOGY – PLASMID

1-3.1 Introduction:

DNA molecule used for carrying an exogenous DNA into a host organism and facilitates stable integration and replication inside the host system is termed as **Vector**. Molecular cloning involves series of sequential steps which includes restriction digestion of DNA fragments both target DNA and vector, ligation of the target DNA with the vector and introduction into a host organism for multiplication. Then the fragments resulted after digestion with restriction enzymes are ligated to other DNA molecules that serve as vectors.

In general, vectors should have following characteristics:

- Capable of replicating inside the host.
- Have compatible restriction site for insertion of DNA molecule (insert).
- Capable of autonomous replication inside the host (*ori* site).
- Smaller in size and able to incorporate larger insert size.
- Have a selectable marker for screening of recombinant organism.

1-3.2 Plasmids:

- Plasmids are naturally occurring extra chromosomal double-stranded circular DNA molecules which can autonomously replicate inside bacterial cells. Plasmids range in size from about 1.0 kb to over 250 kb.
- Plasmids encode only few proteins required for their own replication (replication proteins) and these proteins encoding genes are located very close to the *ori*. All the other proteins required for replication, e.g. DNA polymerases, DNA ligase, helicase, etc., are provided by the host cell.Thus, only a small region surrounding

the *ori* site is required for replication. Other parts of the plasmid can be deleted and foreign sequences can be added to the plasmid without compromising replication.

- The host range of a plasmid is determined by its *ori* region. Plasmids whose *ori* region is derived from plasmid Col E1 have a restricted host range. They only replicate in enteric bacteria, such as *E. coli*, *Salmonella,* etc. Plasmids of the RP4 type will replicate in most gram negative bacteria, to which they are readily transmitted by conjugation. Plasmids like RSF1010 are not conjugative but can be transformed into a wide range of gram -ve bacteria. Plasmids isolated from *Staphylococcus aureus* have a broad host range and can replicate in many other gram-positive bacteria.

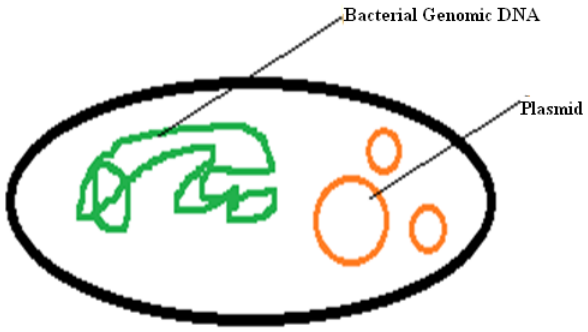

Fig 1-3.2: Bacterial Genomic DNA with plasmid

Some of the phenotypes which the naturally occurring plasmids confer on their host cells:

- Antibiotic resistance
- Antibiotic production
- Degradation of aromatic compounds
- Haemolysin production
- Sugar fermentation
- Enterotoxin production
- Heavy metal resistance
- Bacteriocin production
- Induction of plant tumors
- Hydrogen sulphide production

Most plasmids exist as double-stranded circular DNA molecules. However, the inter-conversion of super coiled, relaxed covalently closed circular DNA and open circular DNA is possible. Not all plasmids exist as circular molecules. Linear plasmids have been found in a variety of bacteria, e.g. *Streptomyces sp.* and *Borrelia burgdorferi.*

However, few types of plasmids are also able to replicate by integrating into bacterial chromosomal DNA; these are known as integrative plasmids or episomes. They are found mainly in prokaryotes but some eukaryotes are also found to harbour them. In prokaryotes they are found in *Escherichia coli***,** *Pseudomonas* species, *Agrobacterium* species etc. In eukaryotes they are mainly found in *Saccharomyces cerevisiae.*

1-3.2.1 Types of Plasmids

The plasmids are divided into 6 major classes as described below depending on the phenotype:

i) **Resistance or R plasmids** carry genes which give resistance to the bacteria from one or more chemical agents, such as antibacterial agents. R plasmids are very important in clinical microbiology as they can have profound consequences in the treatment of bacterial infections. Eg: RP4 plasmid, which is commonly found in *Pseudomonas* and in many other bacteria.

ii) **Fertility or F plasmids** are conjugative plasmid found in F^+ bacterium with higher frequency of conjugation. F plasmid carries transfer gene (*tra)* and has the ability to form Conjugation Bridge (F pilus) with F^- bacterium. Eg: F plasmid of *E. coli*.

iii) **Col plasmids** have genes that code for colicins, proteins that kill other bacteria. Eg: ColE1 of *E. coli*.

iv) **Degradative plasmids** allow the host bacterium to metabolize unusual molecules such as toluene and salicylic acid. Eg TOL of *Pseudomonas putida*.

v) **Virulence plasmids** confer pathogenicity on the host bacterium. Eg: Ti plasmids of *Agrobacterium tumefaciens*, which induce crown gall disease on dicotyledonous plants.

vi) **Cryptic Plasmids** do not have any apparent effect on the phenotype of the cell harboring them. They just code for enzymes required for their replication and maintenance in the host cell.

Based on the origin or source of plasmids, they have been divided into two major classes: such as natural and artificial.

i) **Natural plasmids**: They occur naturally in prokaryotes or eukaryotes. Example*:* ColE1.

ii) **Artificial plasmids**: They are constructed *in-vitro* by re-combining selected segments of two or more other plasmids (natural or artificial). Example: pBR322.

1-3.3 Natural Plasmids

Few examples of naturally occurring plasmids and their characteristics are listed in table below

Plasmid	**Size (kb)**	**Origin**	**Host range**	**Antibiotic resistance**	**Additional marker genes showing insertional inactivation**
RSF1010	8.6	*E.coli* (strain K-12)	Broad host range	Streptomycin and sulfonamides.	None
ColE1	6.6	*E.coli*	Narrow host range	None	Immunity to colicin E1
R100	94.2	*E.coli*	*E.coli* K-12, *Shigella flexneri* 2b	Streptomycin, chloramphenycol, tetracycline	Mercuric (ion) reductase, putative ethidium bromide (EtBr) resistant protein.

Table 1-3.3: Characteristics of natural plasmids

1-3.3.1 RSF1010

- It is a naturally occurring plasmid isolated from *E.coli K-12*.
- This plasmid has broad host range in gram negative bacteria.
- The size of plasmid is 8694bp.
- Antibiotic resistance genes for Streptomycin and sulfonamides have present.
- The replication of RSF1010 starts either bi- or uni-directionally from unique *ori-V* region (2347-2742).
- It cannot initiate transformation independently but can be transferred to host bacterium in presence of helper plasmid.
- Genebank accession no. M28829.

1-3.3.2 ColE1:

- It is a naturally occurring multicopy plasmid obtained from *E.coli (*copy number is around 40).
- The size of this natural plasmid is 6646bp.
- It forms the basis of many artificial vectors used in molecular cloning.
- The natural ColE1 plasmid has genes for colicin E1 production. Colicin is an antibacterial toxin produced under stressed condition. Cells harboring the plasmid will have resistance against the toxin.
- For using in molecular cloning experiment, colicin genes are replaced with selection marker (antibiotic resistant) gene.
- Genebank accession no. M33100.

1-3.4 Artificial Plasmids:

Naturally occurring plasmids has several limitations; for example, some are stringent and not relaxed (pSC101), some has poor marker genes (ColE1), and some are too large (RSF2124**).** To overcome the limitations of natural vectors, artificial plasmid are designed by combining different elements from diverse sources.

Artificial plasmids vectors are classified into two broad types based on their use:

1. Cloning vector
2. Expression vector

Apart from the following, there is another class of vectors known as **shuttle vector.** Shuttle vectors can be propagated in two or more different host species (both in prokaryotes and eukaryotes). Hence, inserted DNA can be manipulated and replicated in two different cellular systems.

Cloning vectors are designed for efficient transfer of foreign DNA into the host. Expression vectors have efficient machinery for cloning and expression of foreign gene in the host system. Selection of a vector depends upon various criteria decided by the experimental goal.

1-3.5 CloningVector:

A cloning vector is defined as a vector used for replication of a cloned DNA fragment in a host cell. These vectors are frequently engineered to contain “ori” – origin of replication sites particular to the host organism. Examples of commonly used cloning vectors are: pUC18, pUC19, pBluescript vectors etc.

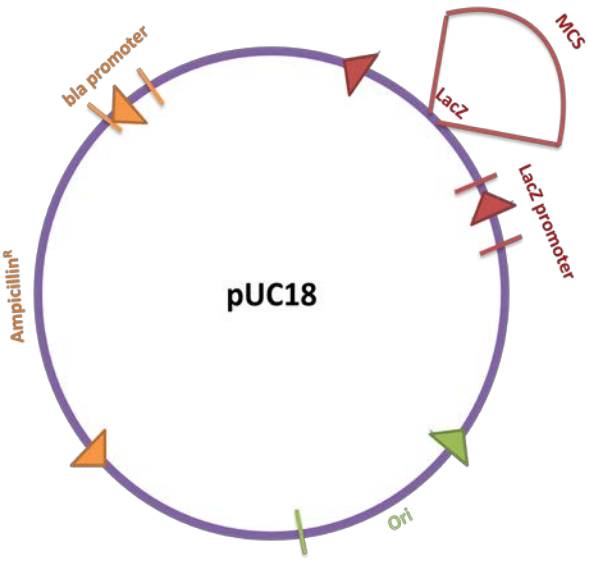

Fig 1-3.5: Cloning vector

Important features of a cloning vector used to carry DNA molecules are as follows:-

- Stability in host cell: Vectors should be stabile in host cell after introduction and should not get lost in subsequent generations. This permits replication of vectors producing large copies of gene of interest.
- Ability to control their own replication: This property enables them to multiply and exist in high copy number.
- Small size: Ideal vector should be less than or equal to 10kb. The small size is essential for easy introduction in cell by transformation, transduction and electroporation.
- Multiple cloning sites: This property permits the insertion of gene of interest and plasmid re-circularization.
- Should not be transferred by conjugation: This property of vector molecule prevents recombinant DNA to escape to natural population of bacteria.
- Selectable make gene: Vector molecules should have some detectable traits. These traits enable the transformed cells to be identified among the non-transformed ones. eg. antibiotic resistance gene.

1-3.6 Types of Cloning Vectors:

- Cloning vectors extensively used in molecular cloning experiments can be considered under following types: **plasmid**, **phage vector** and **cosmid.**
- Different vectors have different insert size and also vary in mode of replication inside the host.
- Mammalian genes are usually too large (~100 kb) and thus suffer from restrictions in complete inclusion with the conventional cloning vectors having limited insert size.
- Vectors engineered more recently, known as artificial chromosomes, have overcome this problem by mimicking the properties of host cell chromosomes. They have much larger insert size than other vectors.

Vector	Insert size	Source	Application
Plasmid	≤ 15 kb	Bacteria	Subcloning and downstream manipulation, cDNA cloning and expression assays
Phage	5-20 kb	Bacteriophage λ	Genomic DNA cloning, cDNA cloning and expression library
Cosmid	35-45 kb	Plasmid containing a bacteriophage λ *cos* site	Genomic library construction
BAC (bacterial artificial chromosome)	75-300 kb	Plasmid ocntaining *ori* from *E.coli* F- plasmid	Analysis of large genomes
YAC (yeast artificial chromosome)	100-1000 kb (1 Mb)	*Saccharomyces cerevisiae* centromere, telomere and autonomously replicating sequence	Analysis of large genome, YAC transgenic mice
MAC (mammalian artificial chromosome)	100 kb to > 1 Mb	Mammalian centromere, telomere and origin of replication	Under development for use in animal biotechnology and human gene therapy

Table 1-3.6: Different type of cloning vectors

1-3.7.1 Examples of Cloning Vector:

1-3.7.1.a pBR322

- pBR322 is a widely-used *E. coli* cloning vector. It was created in 1977 in the laboratory of Herbert Boyer at the University of California San Francisco. The *p* stands for "**plasmid**" and *BR* for "**Bolivar**" and "**Rodriguez**", researchers who constructed it.
- pBR322 is 4361 base pairs in length.
- pBR322 plasmid has the following elements:
  - "*rep*" replicon from plasmid pMB1 which is responsible for replication of the plasmid.
  - "*rop*" gene encoding Rop protein. Rop proteins are associated with stability of RNAI-RNAII complex and also decrease copy number. The source of "*rop*" gene is pMB1plasmid.
  - "*tet*" gene encoding tetracycline resistance derived from pSC101 plasmid.
  - "*bla*" gene encoding β lactamase which provide ampicillin resistance (source: transposon Tn3).

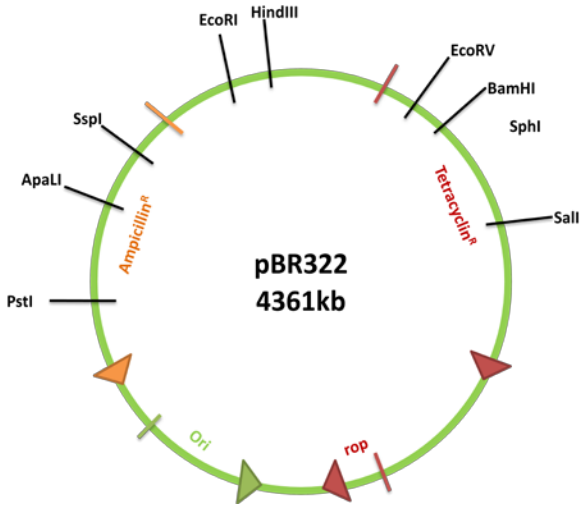

Fig 1-3.7.1.A: Plasmid pBR322.

1-3.7.1.b pUC plasmids:

- pUC plasmids are small, high copy number plasmids of size 2686bp.
- This series of cloning vectors were developed by Messing and co-workers in the University of California. The p in its name stands for plasmid and UC represents the University of California.
- pUC vectors contain a *lacZ* sequence and multiple cloning site (MCS) within *lacZ.* This helps in use of broad spectrum of restriction endonucleases and permits rapid visual detection of an insert.
- pUC18 and pUC19 vectors are identical apart from the fact that the MCS is arranged in opposite orientation.
- pUC vectors consists of following elements:
    - pMB1 "*rep*" replicon region derived from plasmid pBR322 with single point mutation (to increase copy number).
    - "*bla*" gene encoding β lactamase which provide ampicillin resistance which is derived from pBR322. This site is different from pBR322 by two point mutations.
    - *E.coli* lac operon system.
- "*rop*" gene is removed from this vector which leads to an increase in copy number.

An MCS is a short DNA sequence consisting of restriction sites for many different restriction endonucleases. MCS escalates the number of potential cloning strategies available by extending the range of enzymes that can be used to generate a restriction fragment suitable for cloning. By combining them within a MCS, the sites are made contiguous, so that any two sites within it can be cleaved simultaneously without excising vector sequences. The MCS is inserted into the *lacZ* sequence, which encodes the promoter and the α-peptide of β-galactosidase. Insertion of the MCS into the *lacZ* fragment does not affect the ability of the α-peptide to mediate complementation, while cloning DNA fragments into the MCS does. Therefore, recombinants can be detected by blue/white screening on growth medium containing X gal in presence of IPTG as an inducer.

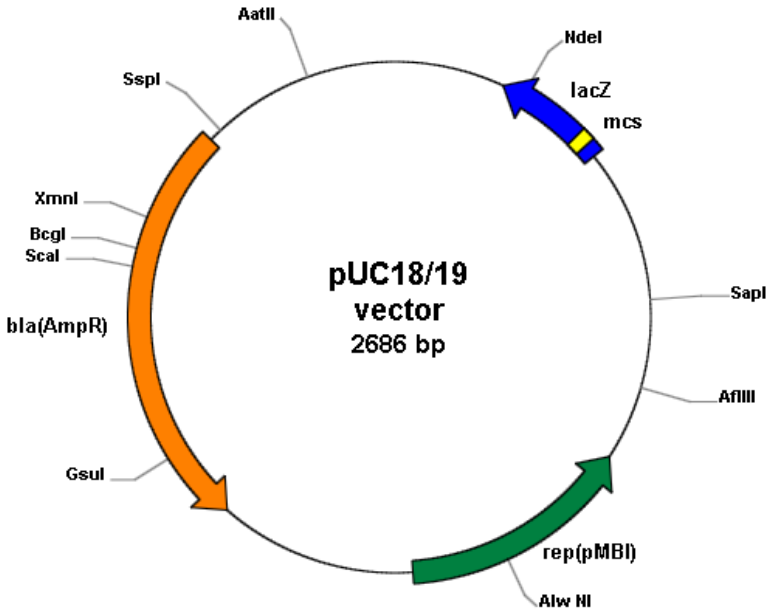

Fig 1-3.7.1.B: pUC plasmid

1-3.8 Expression Vector:

A vector used for expression of a cloned DNA fragment in a host cell is called as an expression vector. These vectors are frequently engineered to contain regulatory sequences that act as promoter and/or enhancer regions and lead to efficient transcription of the insert gene. Commonly used expression vector series are: pET vectors, pBAD vetors etc.

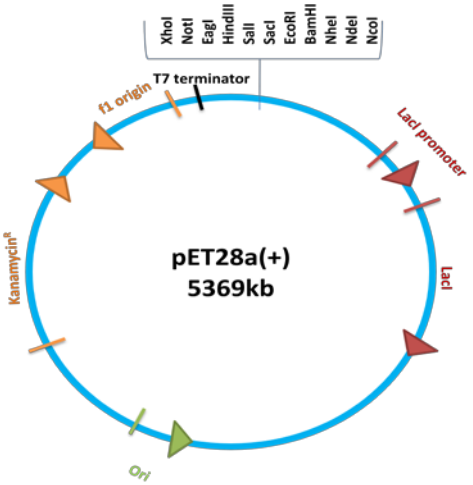

Fig 1-3.8: Expression vector

For an expression vector following features are essential:

- Promoter: Promoter is a sequence which is recognized by sigma subunit of RNA polymerase which is required for initiation of transcription of gene of interest.
- Terminator: It is a DNA element present at the end of a gene where transcription of gene ends. Terminator is short nucleotide sequences which can base pair with itself to form hair pin loop.
- Ribosome binding site: It is a short nucleotide sequence recognized by the ribosome as the point at which it should attach to the messenger molecule. The initiation codon of the gene is always a few nucleotides downstream of this.

1-3.8.2 Examples of Expression Vector:

1-3.8.2.a pET vector:

- pET vector system is a cloning and expression vector system for recombinant protein production in *E.coli*. This product is registered under trademark of Novagen Inc.
- The original pET vector system was constructed by Studier and colleagues. That plasmid is developed at Novagen with enhanced characteristics.
- Target genes are cloned under strong T7 bacteriophage promoter.
- The expression of the target protein is inducible by providing T7 RNA polymerase in the host cell as an inducing signal.
- Target gene is initially cloned to host cell that do not contain T7 RNA Polymerase, thus increasing plasmid stability.
- Once stabilized in a non-expression host, the recombinant plasmid is transferred to a expression host having T7 RNA Polymerase gene in the genome.
- Ampicillin and kannamycin resistance genes are available in pET vectors as selection marker.
- pET28 and pET32 are the most commonly used pET vectors.

1-3.9 Specialized Expression Vectors:

In molecular biology, vectors are generally designed for cloning a foreign gene into a host genome so that the host produces proteins which are normally not produced by host. But, apart from these applications, different specialized vectors have been constructed to achieve different application in genetic and molecular biology studies. The vectors constructed for thus specialized functions are termed as specialized vectors. Molecular and genetics study of a gene or protein can be aided by specialized vectors. Some of the applications of specialized vectors have been discussed below-

1-3.8.1: Promoter Probe Vectors:

Specialized vectors used for identification of efficient promoter region in a DNA segment are termed promoter probe vectors. Promoter-less reported genes (lacZ, GFP etc) are used for construction of promoter probe vectors. The expression of the reporter genes can be monitored and quantified easily using various biochemical or fluorescent techniques. Fusion of DNA fragment containing a promoter region upstream of the reporter gene drives the expression of the reported gene. However, there is no guarantee that the DNA sequence that behaves as promoter in recombinant host can behave in the same way in its native host (Pseudo- promoter). Further characterization is necessary to define a true novel promoter. Some of the widely used promoter probe vectors families are: pOT (eg. pRU1161, pRU1097 etc) and pJP2 (eg. pRU1156, pRU1157 etc). pOT vectors have higher copy number but lower stability as compared to pJP2 vectors.

1-3.9.2 Gene Fusion Vectors:

Fusion of one gene to another gene in order to produce a fusion protein is widely used in molecular biology studies. Fusion proteins are generated by cloning two or more target genes with a reporter gene (*His-tag, gfp, rfp, lacZ* etc) by using gene fusion vectors. Fusion proteins may provide improved properties like easy isolation and purification of target protein (His-tag), easy monitoring of gene expression level (GFP, RFP, *lacZ*), intracellular protein localization studies (GFP, RFP, LUC) etc. The target gene is cloned downstream of the promoter region present in the vector. Depending on the requirement, the target protein can be cloned either to the N-terminal or C-terminal

of the reporter protein. Different vectors have been commercially available to provide such flexibility in cloning site and reporter gene.

1-3.10 Viral Vectors:

In recombinant molecular biology, virus particles has been modified to use as a carrier of nucleic acid into a cell, termed as viral vectors. Viral vectors are highly efficient in transferring target DNA/RNA segment to the host cells with high specificity. Wild type virus are modified by deleting the non-essential genes and incorporating exogenous nucleic acid segments to construct a viral vector. Viral vectors have wide application in gene therapy and targeted drug delivery systems. Main advantages of viral vectors are-high transfer efficiency and high cell specificity. Although, there are certain safety issues associated with viral vectors and careful handing is essential during the experimental procedure.

Commonly used viral vectors are- Adenovirus, retrovirus, lentivirus, adeno associated virus (AAV), herpes simplex virus (HSV) etc. The properties of different viral vectors are summarized in the table below.

Viral vector	Insert type	Insert size	Immunogenicity	Host genome integration
Adeno virus	DNA	2-8 kb	Very high	Non integrating
Retro virus	RNA	2-8 kb	Low	Integrating
Lentivirus	RNA	7-18 kb	Low	Integrating
Adeno associated virus (AAV)	DNA	4.5 kb	Low	Non integrating
Herpes simplex virus (HSV)	DNA	>30kb	Low	Non integrating

Table 1-3.10: Types of viral vectors

1-3.10.1 Simian Virus 40 (SV40):

Simian virus 40 (SV40) was the first mammalian expression vector whose genome size is 5.2kb. It is a DNA virus which can infect human as well other mammalian cell lines. SV40 may integrate into the host genome, permitting stable transmission of insert DNA to daughter cells. Recombinant SV40 vectors (rSV40) display some unique features:

- SV40 is a well-known virus and nonreplicative vectors are easy-to-make
- SV40 can be produced in high titers (10^{12} IU/ml).
- They can infect both resting and dividing cells.
- Stable transgene expression can be achieved in a wide range of cell types.
  The major disadvantage of SV40 vector is the low packaging capacity with insert size of <5kb.

1-3.10.2 Baculoviral Vectors:

Baculovirus is a DNA virus with host range restricted to invertebrates, mostly insects. The baculovirus expression system has been used extensively for the expression of recombinant proteins in insect cells. Baculovirus is a group of insect virus and *Autographa californica* nucleopolyhedrovirus (AcMNPV) is the most extensively studied virus under this family.

The infection of AcMNPV is initiated by replication and transcription of of the DNA genome inside the nucleus and the assembly of the nucleocapsids. The nucleocapsids then bud off from the plasma membrane and initiate systemic infection.

Although, baculoviral vectors can transfect only insect cells, recombinant baculoviral vectors have been constructed containing mammalian cell specific promoters which can be used to infect mammalian cells as well.

Advantages of baculoviral vectors:

- Since insects cells are high eukaryotes, desired port-translational modification of complex protein can be achieved.
- They have higher packaging capacity of insert.
- Lower biosafety issue.
- High level of protein expression.

There are few drawbacks of baculoviral vectors such as-

- Foreign protein expression using an insect system is more complex and time consuming than a bacterial system.
- Sometime protein post-translational processing may be sub-optimal to compensate the secretory pathway of the protein.

1-3.11 Yeast Vector System:

Cloning and expression of a gene using yeast system has several advantages over E.coli system due to presence of eukaryotic post-translational modification machinery. Expression of complex proteins with proper modification and folding can be achieved by yeast eukaryotic system. These vectors have yeast origin of replication (ARS) for replication and maintenance in the yeast system and bacterial ori for maintaining inside a bacterial system.

Different types of yeast vector include YIp (yeast integrative plasmid), YEp (yeast episomal plasmid), YRp (yeast replicating plasmid), YCp (yeast centromere plasmid) etc.

YIp (yeast integrative plasmid) can integrate to the host genome by homologous recombination. This generally yields a single copy of recombinant vector DNA integrated to the host genome.

YEp (yeast episomal plasmid) can be maintained in the yeast system as an autonomously replicating episomal plasmid. This vector contains a part of 2μ plasmid which is essential for autonomous replication of the vector inside the yeast.

YRp (yeast replicating plasmid) are used to obtain a high copy number inside the host (upto 100 copy number).

YCp (yeast centromere plasmid) possess a centromeric region in addition to ARS which facilitates the mitotic segregation of the linear plasmid during replication. The copy number of this vector is essentially one per cell.

Bibliography:

- Cohen S.N., Chang A.C., Boyer H.W., Helling R.B. 1973. Construction of biologically functional bacterial plasmids *in vitro*. *Proc Natl Acad Sci USA* 70: 3240–3244.
- Esposito D, Garvey L.A, Chakiath C.S. 2009. Gateway Cloning for Protein Expression. *Met Mol Biol* 498: 31-54.
- Karunakaran R, Mauchline T.H, Hosie A.H, Poole P.S. 2005. A family of promoter probe vectors incorporating autofluorescent and chromogenic reporter proteins for studying gene expression in Gram-negative bacteria; *Microbiology*.151(10):3249-3256.
- Kay M. A, Glorioso J. C., Naldini L. 2001. Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics; *Nat. Med. Rev*. 7(1): 33-40.
- Kost T. A. and Condreay J. P. 2002. Recombinant baculoviruses as mammalian cell gene-delivery vectors; *Trends in Biotechnol* 20(4).
- Kroll J, Klinter S, Schneider C, Voss I, Steinbüchel A. 2010. Plasmid addiction systems: perspectives and applications in biotechnology. *Microb Biotechnol.* 3(6):634-57.
- Nilsson B., Abrahmsen L., Uhlen M. 1985. Immobilization and purification of enzymes with Staphylococcal protein A gene fusion vectors; *The EMBO Journal* 4(4):1075-1080.
- pET System Manual (10^{th} edition); Novagen Inc.
- Sikorski R. S., Hieter P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*; *Genetics* 122:19-27.

- Yu-chen Hu. 2005. Baculovirus as a highly efficient expression vector in insect and mammalian cells; *Acta Pharmacologica Sinica* 26 (4): 405–416.
- Ong Y, Ward J, Nagl S. 2008; Genome Database of Naturally Occurring Plasmids; URL: http://www.biochem.ucl.ac.uk/bsm/PLASMID/mainpage.htm
- Honda Y, Sakai H, Hiasa H, Tanaka K, Komano T, Bagdasarian M. 1991. Functional division and reconstruction of a plasmid replication origin: Molecular dissection of the *oriV* of the broad-host-range plasmid RSF1010; *Proc. Nati. Acad. Sci.* USA. 88:179-183.