## Principles of cloning, vectors and cloning strategies

#### **The DNA Molecule**



## **CLONING VECTORS**

 Cloning vectors are DNA molecules that are used to "transport" cloned sequences between biological hosts and the test tube.

**Cloning vectors share four common properties:** 

- 1. Ability to promote autonomous replication.
- 2. Contain a genetic marker (usually dominant) for selection.

3. Unique restriction sites to facilitate cloning of insert DNA.

4. Minimum amount of nonessential DNA to optimize cloning.

## PLASMIDS

- Bacterial cells may contain extrachromosomal DNA called plasmids.
- Plasmids are usually represented by small, circular DNA.
- Some plasmids are present in multiple copies in the cell





## **PLASMID VECTORS**

- Plasmid vectors are ≈1.2–
  3kb and contain:
- replication origin (ORI) sequence
- a gene that permits selection,
- Here the selective gene is ampr; it encodes the enzyme b-lactamase, which inactivates ampicillin.
- Exogenous DNA can be inserted into the bracketed region.



## SELECTIVE MARKER

- Selective marker is required for maintenance of plasmid in the cell.
- Because of the presence of the selective marker the plasmid becomes useful for the cell.
- Under the selective conditions, only cells that contain plasmids with selectable marker can survive
- Genes that confer resistance to various antibiotics are used.
- Genes that make cells resistant to ampicillin, neomycin, or chloramphenicol are used



## **ORIGIN OF REPLICATION**

- Origin of replication is a DNA segment recognized by the cellular DNA-replication enzymes.
- Without replication origin, DNA cannot be replicated in the cell.



## **MULTIPLE CLONING SITE**

- Many cloning vectors contain a multiple cloning site or polylinker: a DNA segment with several unique sites for restriction endo- nucleases located next to each other
- Restriction sites of the polylinker are not present anywhere else in the plasmid.
- Cutting plasmids with one of the restriction enzymes that recognize a site in the polylinker does not disrupt any of the essential features of the vector



## **MULTIPLE CLONING SITE**

 Gene to be cloned can be introduced into the cloning vector at one of the restriction sites present in the polylinker



(a) Sequence of polylinker



(b) Insertion of EcoRI restriction fragments





## **TYPES OF CLONING VECTORS**

## **CLONING VECTORS**

- Different types of cloning vectors are used for different types of cloning experiments.
- The vector is chosen according to the size and type of DNA to be cloned

## PLASMID VECTORS

- Plasmid vectors are used to clone DNA ranging in size from several base pairs to several thousands of base pairs (100bp -10kb).
- ColE1 based, pUC vehicles commercially available ones, eg pGEM3, pBlueScript



# Disadvantages using plasmids

- Cannot accept large fragments
- Sizes range from 0- 10 kb
- Standard methods of transformation are inefficient

## **BACTERIOPHAGE LAMBDA**

- Phage lambda is a bacteriophage or phage, i.e. bacterial virus, that uses *E. coli* as host.
- Its structure is that of a typical phage: head, tail, tail fibres.
- Lambda viral genome: 48.5 kb linear DNA with a 12 base ssDNA "sticky end" at both ends; these ends are complementary in sequence and can hybridize to each other (this is the cos site: cohesive ends).
- Infection: lambda tail fibres adsorb to a cell surface receptor, the tail contracts, and the DNA is injected.
- The DNA circularizes at the cos site, and lambda begins its life cycle in the *E. coli* host.

## **BACTERIOPHAGE LAMBDA**



## **COSMID VECTOR**

#### Purpose:

1. Clone large inserts of DNA: size ~ 45 kb

• Features:

Cosmids are Plasmids with one or two Lambda Cos sites.

 Presence of the Cos site permits *in vitro* packaging of cosmid DNA into Lambda particles



## **COSMID VECTOR**

- Thus, have some advantages of Lambda as Cloning Vehicle:
- Strong selection for cloning of large inserts
- Infection process rather than transformation for entry of chimeric DNA into *E. coli* host
- Maintain Cosmids as phage particles in solution
- But Cosmids are Plasmids: Thus do NOT form plaques but rather cloning proceeds via *E. coli* colony formation

### Yeast Artificial Chromosomes

Yeast artificial chromosomes are special linear DNA vectors that resemble normal yeast chromosome. YACs contain telomers that stabilize chromosome ends, centromer, that ensures chromosome partitioning between two daughter cells and a selective marker gene. YAC vectors can carry hundreds of thousands of base pairs of foreign DNA.



foreign DNA

## **Yeast Artificial Chromosomes**

#### Purpose:

- Cloning vehicles that propogate in eukaryotic cell hosts as eukaryotic Chromosomes
- Clone very large inserts of DNA: 100 kb 10 Mb
- Features:

YAC cloning vehicles are plasmids Final chimeric DNA is a linear DNA molecule with telomeric ends: **Artificial Chromosome** 

#### Additional features:

- Often have a selection for an insert
- YAC cloning vehicles often have a bacterial origin of DNA replication (ori) and a selection marker for propogation of the YAC through bacteria.
- The YAC can use both yeast and bacteria as a host

## **PACs and BACs**

- PACs P1-derived Artificial Chromosomes
- E. coli bacteriophage P1 is similar to phage lambda in that it can exist in E. coli in a prophage state.
- Exists in the *E. coli* cell as a plasmid, NOT integrated into the *E. coli* chromosome.
- P1 cloning vehicles have been constructed that permit cloning of large DNA fragments- few hundred kb of DNA
- Cloning and propogation of the chimeric DNA as a P1 plasmid inside *E. coli* cells

- BACs Bacterial Artificial Chromosomes
- These chimeric DNA molecules use a naturallyoccurring low-copy number bacterial plasmid origin of replication, such as that of F-plasmid in *E. coli*.
- Can be cloned as a plasmid in a bacterial host, and its natural stability generally permits cloning of large pieces of insert DNA, i.e. up to a few hundred kb of DNA.

## **RETROVIRAL VECTORS**

- Retroviral vectors are used to introduce new or altered genes into the genomes of human and animal cells.
- Retroviruses are RNA viruses.
- The viral RNA is converted into DNA by the viral reverse transcriptase and then is efficiently integrated into the host genome
- Any foreign or mutated host gene introduced into the retroviral genome will be integrated into the host chromosome and can reside there practically indefinitely.
- Retroviral vectors are widely used to study oncogenes and other human genes.

# Types of expression systems

- Bacterial: plasmids, phages
- Yeast: expression vectors: plasmids, yeast artifical chromosomes (YACs)
- Insect cells: baculovirus, plasmids
- Mammalian:
  - viral expression vectors (gene therapy):
    - SV40
    - vaccinia virus
    - adenovirus
    - retrovirus
  - Stable cell lines (CHO, HEK293)

## **EXPRESSION VECTORS**

- Allows a cloned segment of DNA to be translated into protein inside a bacterial or eukaryotic cell.
- Vectors will contain the ff:
- (a) in vivo promoter
- (b) Ampicillin selection
- (c) Sequencing primers

## **EXPRESSION VECTORS**

- Produces large amounts of a specific protein
- Permits studies of the structure and function of proteins
- Can be useful when proteins are rare cellular components or difficult to isolate

# Common problems with bacterial expression systems

- Low expression levels:
  - change promoter
  - change plasmid
  - change cell type
  - add rare tRNAs for rare codons on second plasmid
- Severe protein degradation:
  - use proteasome inhibitors and other protease inhibitors
  - try induction at lower temperature
- Missing post-translational modification: co-express with kinases etc.
- Glycosylation will not be carried out:
  - use yeast or mammalian expression system
- Misfolded protein (inclusion bodies):
  - co-express with GroEL, a chaperone
  - try refolding buffers

## **REPORTER GENE VECTORS**

- A gene that encodes a protein whose activity can be easily assayed in a cell in which it is not normally expressed
- These genes are linked to regulatory sequences whose function is being tested
- Changes in transcriptional activity from the regulatory sequences are detected by changes in the level of reporter gene expression

## **SHUTTLE VECTORS**

- Shuttle vectors can replicate in two different organisms, e.g. bacteria and yeast, or mammalian cells and bacteria.
- They have the appropriate origins of replication.
- Hence one can clone a gene in bacteria, maybe modify it or mutate it in bacteria, and test its function by introducing it into yeast or animal cells.

## **CLONING STRATEGY**

- Strategy depends on the starting information and desired endpoint.
- Starting Information or Resources:
  - Protein sequence
  - Positional cloning information
  - mRNA species / sequence
  - cDNA libraries
  - DNA sequence known or unknown
  - genomic DNA libraries
  - PCR product

How Are Genes Cloned Using Plasmids?

- To understand how genes are cloned, we need introduce three terms.
- Recombinant DNA- is mixed DNA
- Vector -it carries recombinant DNA into cells.
- Plasmids are tiny circular pieces of DNA that are commonly found in bacteria.

## Why Plasmids are Good Cloning Vectors

- small size (easy to manipulate and isolate)
- circular (more stable)
- replication independent of host cell
- several copies may be present (facilitates replication)
- frequently have antibody resistance (detection easy)

## How is foreign DNA Inserted into a Plasmid?

- To open up the DNA a **restriction enzyme** is used.
- Cut the DNA at a specific place called a restriction site.
- The result is a set of double-stranded DNA pieces with single-stranded ends
- These ends that jut out are not only "sticky" but they have gaps that can be now be filled with a piece of foreign DNA
- For DNA from an outside source to bond with an original fragment, one more enzyme is needed
- **DNA ligase** seals any breaks in the DNA molecule

## **RESTRICTION ENZYMES**

- Restriction enzymes: enzymes that cut DNA in specific places function:
- Inactivate foreign DNA
- Breaks only palindrome sequences, i.e. those exhibiting two-fold symmetry
- Important in DNA research, i.e. sequencing, hybridization
- Companies purify and market restriction enzymes





#### Restriction Enzyme Action of EcoRI



**Recombinant DNA** 

## **CLONING METHODOLOGY**

- **Cut** the cloning vector with R.E. of choice, eg *Eco* RI
- Cut DNA of interest with same R.E. or R.E. yielding same sticky ends, e.g. *Bam* HI and *Sau* 3A
- Mix the restricted cloning vector and DNA of interest together.
- Ligate fragments together using DNA ligase
- Insert ligated DNA into host of choice transformation of *E. coli*
- Grow host cells under restrictive conditions, grow on plates containing an antibiotic

## Subcloning an *Eco*RI fragment into a vector

Both plasmid and target DNAs are digested with *Eco*RI and gel purified.

The insert is ligated into the *Eco*RI site of the plasmid using T4 DNA ligase (requires ATP).

Three possible outcomes; insert can be in two possible orientations relative to the vector, or the vector has religated without the insert.




#### Lets see that one more time





### **BLUE/WHITE SCREENING**

- Colony Selection: finding the rare bacterium with recombinant DNA
- Only *E. coli* cells with resistant plasmids grow on antibiotic medium
- Only plasmids with functional *lacZ* gene can grow on Xgal
   *lacZ(+) => blue* colonies

  *lacZ* functional => polylinker intact => nothing inserted, no clone
  *lacZ(-) => white* colonies polylinker disrupted
  *successful insertion & recombination*!

## **α** -complementation

- The portion of the *lacZ* gene encoding the first 146 amino acids (the α -fragment) are on the plasmid
- The remainder of the *lacZ* gene is found on the chromosome of the host.
- If the α -fragment of the *lacZ* gene on the plasmid is intact (that is, you have a non-recombinant plasmid), these two fragments of the *lacZ* gene (one on the plasmid and the other on the chromosome) complement each other and will produce a functional β galactosidase enzyme.

#### **SCREENING RECOMBINANTS**



In the example shown above, the bgalactosidase gene is inactivated. The substrate "X-gal" turns blue if the gene is intact, ie. makes active enzyme. White colonies in X-gal imply the presence of recombinant DNA in the plasmid.

#### COMPLICATIONS

- *lacZ* gene not expressed constitutively
- X-gal does not activate gene expression
- must use IPTG as inducer
- (isopropyl-β-D-thio-galactoside)
- small inframe insertions may not inactivate α peptide
- still get blue colonies (often lighter less activity

### ENZYMES USED IN MOLECULAR BIOLOGY

Alkaline phosphatase	Removes phosphate groups from 5' ends of DNA (prevents unwanted re-ligation of cut DNA)	
DNA ligase	Joins compatible ends of DNA fragments (blunt/blunt or complementary cohesive ends). Uses ATP	
DNA polymerase I	Synthesises DNA complementary to a DNA template in the 5'-to-3'direction. Starts from an oligonucleotide primer with a 3' OH end	
Exonuclease III	Digests nucleotides progressiviely from a DNA strand in the 3' -to-5' direction	
Polynucleotide kinase	Adds a phosphate group to the 5' end of double- or single-stranded DNA or RNA. Uses ATP	
RNase A	Nuclease which digests RNA, not DNA	
Taq DNA polymerase	Heat-stable DNA polymerase isolated from a thermostable microbe ( <i>Thermus aquaticus</i> )	

#### ENZYMES USED IN MOLECULAR BIOLOGY



### **RESTRICTION ENZYMES**

The restriction enzymes most used in molecular biology labs cut within their recognition sites and generate one of three different types of ends.

#### **5' OVERHANGS**

5' overhangs: The enzyme cuts asymmetrically within the recognition site such that a short singlestranded segment extends from the 5' ends. *Bam* HI cuts in this manner.



## 3' OVERHANGS

 3' overhangs: Again, we see asymmetrical cutting within the recognition site, but the result is a singlestranded overhang from the two 3' ends. *Kpn*I cuts in this manner.



## **BLUNT ENDS**

 Blunts: Enzymes that cut at precisely opposite sites in the two strands of DNA generate blunt ends without overhangs. Smal is an example of an enzyme that generates blunt ends.



# Converting a 5' overhang to blunt end

- Both Klenow and T4 DNA polymerase can be used to fill in 5' protruding ends with dNTPs
- Used in joining DNA fragments with incompatible ends
- Once the ends have been blunted, ligation can proceed



# Converting a 3' overhang to a blunt end

- T4 DNA polymerase has a 3'-5' exonuclease activity
- In the presence of excess dNTPs will convert a 3' protruding end to a blunt end
- Ligation can know proceed



### **DIRECTIONAL CLONING**

- Often one desires to insert foreign DNA in a particular orientation
- This can be done by making two cleavages with two different restriction enzymes
- Construct foreign DNA with same two restriction enzymes
- Foreign DNA can only be inserted in one direction

- Good efficiency of ligation of foreign DNA into a vector can be achieved if both the vector and the insert DNA are cut with 2 different restriction enzymes which leave single stranded ends (cohesive ends).
- The DNA is ligated in only one direction, and there is only a low background of non-recombinant plasmids.
- If only one restriction enzyme is used to cut the vector and insert, then efficiency of ligation is lower, DNA can be inserted in two directions and tandem copies of inserts may be found.
- To avoid high background of non-recombinants, alkaline phosphatase is used to remove 5' phosphate groups from the cut vector to prevent selfligation.

### Alkaline phosphatase



Alkaline phosphatase removes 5' phosphate groups from DNA and RNA. It will also remove phosphates from nucleotides and proteins. These enzymes are most active at alkaline pH

### Alkaline phosphatase

- There are two primary uses for alkaline phosphatase in DNA manipulations:
- Removing 5' phosphates from plasmid and bacteriophage vectors that have been cut with a restriction enzyme. In subsequent ligation reactions, this treatment prevents self-ligation of the vector and thereby greatly facilitates ligation of other DNA fragments into the vector (e.g. subcloning).
- Removing 5' phosphates from fragments of DNA prior to labeling with radioactive phosphate. Polynucleotide kinase is much more effective in phosphorylating DNA if the 5' phosphate has previously been removed

#### DEPHOSPORYLATED VECTOR



#### R.E.S WITH COMPATIBLE ENDS



# Generating a new R.E. site at a blunt end

- Use linkers to generate a new R.E.
- Linkers are used to place sticky ends on to a blunt-ended molecule
- Short blunt ended synthetic ds DNA containing a R.E. site
- Experimental design:
  - (i) Blunt ended DNA + linker (T4 ligase)
  - (ii) Digest with appropriate R.E.
  - (iii) Ligate to vector



# Introducing a R.E. site by PCR

- R.E. site is designed into the 5' end of the PCR primer
- PCR fragment is digested with appropriate R.E., purified and ligated into plasmid vector







- It depends on the enzyme
- Some catalogues of enzymes provide anecdotal data on the efficiency of enzymes trying to work at the ends of DNA molecules.
- Generally, enzymes work better if they have a couple of extra nucleotides at the end they don't do very well if they are perched on the end of a molecule.

# Ligation of foreign DNA to plasmid vectors

Termini on foreign DNA	Requirements for cloning	Comments
Blunt -ended	High conc of DNAs and ligase	Large no. of non-recombinant clone
		R.E. sites may be eliminated Tandem copies of foreign DNA
Different protruding termini	Requires purification of plasmid after digestion	R.E. sites at junctions are conserved
		Low no. of non-recombinants Foreign DNA is inserted in only one orientation
Identical protruding termini	Phosphatase treatment of linear plasmid vector	R.E. sites at junctions are conserved
		Foreign DNA is inserted in either orientation
		Tandem copies of foreign DNA

#### IDENTIFICATION OF POSITIVE CLONES

- One of the first steps is to identify clones carrying the recombinant plasmid, with the desired DNA insert.
- This can be done by 'picking' clones choosing individual bacterial colonies in order to isolate the plasmid DNA from each of them.
- Single bacterial colonies are grown in culture broth containing the selection antibiotic in order to maintain the plasmid.
- The plasmid DNA is extracted by the standard minipreparation technique and then analysed by restriction digest.
- After digesting the DNA, different sized fragments are separated by agarose gel electrophoresis and the sizes determined by comparison with known DNA molecular weight marke

#### AGAROSE GEL ELECTROPHORESIS





### **RECOMBINANT DNA**

- R.E. are a useful tool for analysing Recombinant DNA
  - checking the size of the insert
  - checking the orientation of the insert
  - determining pattern of restriction sites within insert
- Sometimes it is important to determine the orientation of the DNA insert in relation to the vector sequence.
- This can be done simply by restriction digest using enzyme(s) which cut the vector sequence near to the insert and cut within the insert sequence (asymmetrically).

# APPLICATIONS

- Cloning DNA fragments
- Generating Libraries: essential step for genome mapping
- Positional cloning discovering disease genes
- Discovering genes from e.g. Protein sequence

#### **PCR cloning strategies**

- Cloning methods for PCR products are divided into three types:
  - (i) blunt-end cloning
  - (ii) sticky-end cloning
  - (iii) T-A cloning

#### **PCR Cloning Considerations**

- Nature of the Insert: not all PCR fragments will clone with the same efficiency into the same vector.
- Insert Size: The size of the fragment being cloned is a primary contributor to the overall cloning efficiency. Large fragments of DNA (≥ 5 kb) are amenable to cloning in high-copy number vectors, yet at a much lower efficiency.
- Vector-to-Insert Ratio:Optimization of molar concentration ratios of the vector to insert is critical to ensure efficient cloning. insert ratios: 1:1, 1:3,

### **T-A Cloning**

- When DNA fragments are generated *Taq* polymerase adds 1 or 2 extra adenines onto the end of 3' end of blunt ds DNA
- Several commercially available kits take advantage of this ability
- Use a plasmid vector with thymidine residues linked onto the 3' ends of linearised plasmid DNA



### TA CLONING VECTOR


#### **TA RECOMBINANT VECTOR**



# ANALYSIS OF CLONED DNA

- Is it the one you wanted?
- What are its molecular characteristics?

- Restriction mapping: determining the order of restriction sites in a cloned fragment:
- Gel electrophoresis: separates DNA fragments by molecular weight
- Southern Blot analysis: DNA is transferred ("blotted") to filter paper.Filter is exposed to a DNA probe. Binds specifically to target DNA immobilized on filter
- DNA sequencing: provides complete order of bases in a DNA fragment





### **DNA FORMS OF A PLASMID**

#### Uncut plasmid DNA can be in any of five forms:

- nicked
- circular
- Inear covalently closed
- supercoiled
- circular single-stranded.
- The exact distances between the bands of these different forms is influenced by:
  - percentage of agarose
  - time of electrophoresis
  - degree of supercoiling
  - the size of the DNA.
- Linear band = linear size of plasmid

Relative Positions of Different DNA Forms of a Plasmid on a Tris-Acetate Agarose Gel



### **CLONING INTO A PLASMID**





## **DNA CLONING**

- DNA cloning is a technique for reproducing DNA fragments.
- It can be achieved by two different approaches:
  - cell based
  - using polymerase chain reaction (PCR).
- a vector is required to carry the DNA fragment of interest into the host cell.

#### PLASMID CLONING STRATEGY

Involves five steps:

Enzyme restriction digest of DNA sample.

Enzyme restriction digest of DNA plasmid vector.

**Ligation of DNA** sample products and plasmid vector.

Transformation with the ligation products.

Growth on agar plates with selection for antibiotic resistance.

## **STEP 4. TRANSFORMATION OF LIGATION PRODUCTS**

- The process of transferring exogenous DNA into cells is call "transformation"
- There are basically two general methods for transforming bacteria. The first is a chemical method utilizing CaCl2 and heat shock to promote DNA entry into cells.
- A second method is called electroporation based on a short pulse of electric charge to facilitate DNA uptake.

#### CHEMICAL TRANSFORMATION WITH CALCIUM CHLORIDE





#### **STEP 5. GROWTH ON AGAR PLATES**



## STEP 5

 Blue colonies represent Ampicillin-resistant bacteria that contain pVector and express a functional alpha fragment from an intact LacZ alpha coding sequence.

White colonies represent Ampicillin-resistant bacteria that contain plnsert and do **not** produce LacZ alpha fragment

### **TERMS USED IN CLONING**

#### DNA recombination.

The DNA fragment to be cloned is inserted into a vector.

#### Transformation.

The recombinant DNA enters into the host cell and proliferates.

#### Selective amplification.

A specific antibiotic is added to kill *E. coli* without any protection. The transformed *E. coli* is protected by the antibiotic-resistance gene

Isolation of desired DNA clones