Unit 6a
Gene Cloning
Vectors
Molecular Biotechnology (Ch 4)
Principles of Gene Manipulation (Ch 3 & 4)
Analysis of Genes and Genomes (Ch 5)
Plasmids

Naturally occurring plasmids

- occur widely in bacteria

- are covalently closed circular dsDNA

- are replicons, stably inherited as extra-chromosomal DNA

- can be 1 kbp to 500 kbp in size (compared to 4000 kbp chromosome)

- bacteria can contain several different types of plasmid simultaneously

- many naturally occurring plasmids carry genes for restriction enzymes, antibiotic resistance, or other genes
Bacterial Vectors

All vectors:
1. must replicate autonomously in their specific host even when joined to foreign DNA

2. should be easily separated from host chromosomal DNA

   • *E. coli* chromosomal DNA: ~ 4 million bp
   • typical plasmid vector: ~ 3 to 10 kb

ori - origin of replication
sequence at which DNA polymerase initiates replication

Most modern cloning vectors in *E. coli* are derived from naturally-occurring plasmid col E1. Most of col E1 was deleted except for an origin of replication and an antibiotic resistance gene.
Vectors

Types

**plasmids** - can occur naturally in as circular dsDNA in bacteria (up to 15 kb)

**bacteriophage** - viruses of bacteria (~10-50 kb) used in the construction of cDNA and genomic libraries

**BAC** - bacterial artificial chromosome (130-150 kb)

**YAC** - Yeast artificial chromosome (1000-2000 kb)

Each type of vector has specific applications but primary function is to carry foreign DNA (foreign to bacteria) and have it replicated by the bacteria
Introduction of foreign DNA into E. coli

Bacterial cells with plasmid

Lyse cells and recover plasmids

Isolate plasmid vectors

Cut with a restriction enzyme to linearize without fragmenting

Add foreign DNA with compatible ends and ligate

Recombinant plasmid vectors

“Competent” bacterial cells

Plasmid transformed bacteria

Plasmid vector transform

Lyse cells and recover plasmids
Plasmids

Plasmid vectors (engineered) should

1. **have naturally-occurring copy number control sequences deleted**
   
   Result: higher plasmid copy number in the bacteria
   
   (eg. 20-200 copies/cell)
   
   eg. pUC19

2. **be small in size** (3-5 kbp) to facilitate separation from chromosomal DNA
   
   Transfer efficiency declines for plasmids >15 kbp

3. **carry selectable markers** to allow selection of:
   
   i) plasmid transformed bacteria
   
   ii) transformed bacteria that carry the recombinant vector
3. Selectable markers - explanation

i) selection of plasmid-transformed bacteria

Achieved by engineering a plasmid to carry and express a gene for antibiotic detoxification which will allow any bacterium carrying that plasmid (and gene) to be resistant to the toxic effects of the antibiotic.

Following transformation of bacterial cells (antibiotic sensitive without this plasmid), cells are plated on an antibiotic containing media.

Commonly used antibiotics: Ampicillin, streptomycin, kanamycin, hygromycin

Removes non-transformed bacteria which will be present in the transformation mixture
Bacterial transformation

Plasmid with antibiotic resistance gene

Competent bacterial cells sensitive to antibiotic cmpd

Petri plate with bacterial agar media containing an appropriate antibiotic

Bacterial colony derived from a single cell

A significant % of cells are not transformed
3. Selectable markers

Antibiotic selection - kills all cells that did not get the vector

However - there is no guarantee that all vectors got an insert.

Therefore, many cloning vectors let us screen for those with inserts based on insertional inactivation of a gene.

Example: X-gal assay for the LacZ gene
Expression of LacZ gene

*this is what happens if we don’t insert a fragment*

Expression of gene

in bacterial cell

β-galactosidase

galactose

Released indolyl cmpd dimerizes after oxidation to yield an insoluble indigo dye (blue color)

NOTE: X-gal is taken up by cells but the product becomes insoluble inside the cells

Colonies expressing LacZ gene

Agar plate

Amp

X-gal
Selection for bacteria carrying recombinant plasmids

- **Plasmid**
  - Ori
  - Amp<sup>r</sup>
  - LacZ gene

- **Restriction enzyme**
  - Foreign DNA
  - DNA Ligase

- **Transform bacteria**
  - Plate on media with antibiotic and X-gal

- **Amp<sup>r</sup>**
  - Will die in presence of antibiotic
  - Will grow and produce blue colonies

- **Non-recombinant plasmid**
  - Will grow and produce white colonies

- **Recombinant plasmid**
  - Will grow and produce blue colonies
  - Plate on media with antibiotic and X-gal
• X-Gal structure
Plasmids

Plasmid vectors (engineered) should: (cont’d)

4. **Have a multiple cloning site (MCS) region** A constructed ds sequence containing a series of consecutive, unique (i.e., only one site in the plasmid) restriction enzyme recognition sites. The sequence is a multiple of 3 bp long and is inserted into the coding region of the selectable marker (i.e., LacZ gene) gene.

Sequence is relatively short (54 bp in pUC19)
As multiple of 3, the insertion does not change the reading frame of coding region of the gene, adds 18 amino acids to protein - but doesn’t affect product activity.

**Value:** Allows different restriction enzymes to be used to open plasmid to match the restriction enzyme used to create the DNA being inserted.
Multiple cloning site (MCS) from pUC 18/19 plasmids

Paired vectors have oppositely oriented MCS
In-frame insertion of MCS adds 12-18 amino acids

Translation of mRNA into protein

Transcription into mRNA

End is generally away from catalytic site
Example: Insertion at an Eco RI site

Insertion of foreign DNA

EcoRI ends into vector
Example: Insertion at a Bam HI site

**Diagram:**

- **EcoRI**
- **K**
- **B**
- **X**
- **P**
- **HindIII**

**LacZ - MCS**
Fragments can be inserted using either 1 or 2 restriction sites

EcoRI  K  B  X  P  HindIII

LacZ - MCS

Etc, etc
Example: One enzyme, two possible orientations

Goal: Express Red gene by Green promoter in plasmid

Possibilities?
Example: One enzyme, two possible orientations

Goal: Express Red gene by Green promoter in plasmid
Example: One enzyme, two possible orientations

Plasmid

Promoter

Cut with E

Isolate red gene

Cut with E

Ligase

Goal: Express Red gene by Green promoter in plasmid

E EcoRI
K KpnI
X Xbal
S SalI
H HindIII
When the same restriction enzyme cuts on both ends, the insert can be ligated to the vector in either of two possible orientations.
Example: Two enzymes, one possible orientation
This is called directional cloning

Orientation is defined with reference to the direction of transcription
Example: Two enzymes, one possible orientation
This is called directional cloning

Goal: Express Red gene by Green promoter in plasmid

Plasmid

Promoter

E  EcoRI
K  KpnI
S  SalI
H  HindIII
X  XbaI
Example: Two enzymes, one possible orientation
This is called directional cloning

Goal: Express Red gene by Green promoter in plasmid
Example: Two enzymes, one possible orientation
This is called directional cloning

Goal: Express Red gene by Green promoter in plasmid
Bacteriophage Derived Vectors – 2nd type of vector

Common phage (virus) of E. coli is lambda

Viruses typically consist of a DNA (RNA) based genome and a protein coat

A virus uses host cells to replicate and express their genome

Features of natural lambda

- ds DNA, ~49 kbp, encodes 67 genes
- Outside its host, phage exists as a phage particle (capsid particle)
- Linear dsDNA wrapped in a protein coat
- ds except for both 5’ termini which are 12 nt extensions
  - The extensions are complementary allowing circularization
- The 12 nt extension is called a COS site
Bacteriophage Lambda

- Complementary 12 nt COS sequence allows circularization of the genome inside a host cell.
- ~49 kb dsDNA inside the capsid particle.

Diagram showing the head, tail, and capsid particle of Bacteriophage Lambda.
Bacteriophage Lambda

The cos site in (multimer) Lambda DNA looks like this:

---GGGGCGGCGACCT---
---CCCCGCGCTGGA---

terminase cuts like this:

---G/GGGCGGCGACCTC---
---CCCCGCGCTGGA/G---

resulting in these cos ends in the linear Lambda DNA:

---G
---CCCCGCGCGCTGGA

GGCGGCGACCTC---
G---

from
http://www.bioinformatics.nl/molbi/SCLRResources/lambda.htm
Bacteriophage lambda (\(\lambda\)) infection of \textit{E. coli}

Bacterial cell

Host chromosome

 Adsorption to host cell and insertion of phage DNA

\(\lambda\) DNA

Circularization of \(\lambda\) DNA

Integration of \(\lambda\) DNA into host chromosome

Expression of \(\lambda\) genes

Prophage

Lysogenic pathway
Infection (cont’d)

Integration of λ DNA into host chromosome

Expression of λ genes

• Replication of λ DNA
• Synthesis of head and tail proteins
• Packaging in particles

Induction event

Lysogenic pathway

• Host cell lysis
• Release of infectious particles

Lytic pathway
Only one gene is expressed in the lysogenic pathway - cl repressor protein which prevent the first stage in the sequential expression of the lambda genome genes.

During lytic pathway, different sets of genes are expressed early (DNA replication) mid (synthesis of head and tail structures and assemble of capsid particle) and late (for lysis of host cell) in the pathway

The capsid coat (head and tail structures) are very important to the phage as it allows the phage to get its DNA into healthy E. coli cells.
Lytic Infectious process

Bacteria spread uniformly over surface

Bacteria with very low level of lambda-infected cells spread over plate (1/10,000)

Time

~20 h

Zone of lysed bacteria
Plaque
Bacterial lawn
Plaque - results from a repeated infection, phage replication and lysis with release of phage particles. Lysed cell zone contains free phage particles and proteins expressed in the *E. coli* prior to lysis.
Plaque - results from a repeated infection, phage replication and lysis with release of phage particles. Lysed cell zone contains free phage particles and proteins expressed in the *E. coli* prior to lysis.
Capsid head structure

The head structure is a defined size

To be a viable phage particle the head structure must contain between 38-51 kbp of linear dsDNA which is terminated by 5’ COS sequences
Lambda vectors - features

Must create space in the lambda genome for insertion of extra DNA

1. Lambda vectors have had part of the genome deleted to allow 1-8 kb of foreign DNA to be inserted. Deleted part is non-essential to function as a vector. Eg the vectors are 41 – 47 kbp long
Lambda vectors - features

2. Selectable marker to distinguish recombinant from non-recombinant vectors. (No antibiotics are required as uninfected cells are required.)

If LacZ gene is used as a cloning site:

- no insert: blue plaque
- insert in LacZ: clear plaque
Lambda vectors - features

2. Selectable marker to distinguish recombinant from non-recombinant vectors. (No antibiotics are required as uninfected cells are required.)

If lambda repressor gene (cl) is used as a cloning site:
• no insert: repressor prevents phage replication, no plaques
• insert in cl: clear plaque

ie. Only phage with inserts will give you a plaque
Lambda vectors - features

3. One or more unique restriction sites within the selectable marker to allow for insertion of foreign DNA.

Because the vector is large (~43 kbp) getting multiple restriction sites which are unique (occur only once in the vector) is more difficult but has been achieved with specific vectors.

4. The ability to repackage recombinant phage DNA into a capsid particle.

For phage this process is normally done inside the bacterial cell but for recombinant DNA work it must be done in vitro if this is to be useful vector.

Recombinant Lambda DNA can be packaged in-vitro using bacterial cell extracts prepared from cells expressing the mid stage of the lytic infection genes. The extract contains the enzymes and precursors for the head and tail structure and for packaging of the recombinant DNA into capsid particles. Addition of the extract to ligated phage/foreign DNA will result in packaging any DNA which is 38-51 kbp long and terminated by cos sequences.
Construction of recombinant lambda 

- capsid

Linear phage DNA with unique EcoRI site in marker gene

EcoRI ends

Cos ends

Concatemaric DNA

Ligation

Foreign DNA

In vitro packaging based on recognition of COS and distance between cos sites

Can form, but cannot be packaged, not replicate

Unique EcoRI site within marker gene

Alkaline phosphatase

E

AP

EcoRI

(millions of particles)
Alkaline phosphatase - prevents ligation of lambda arms to each other
Foreign DNA alone cannot be packaged as an infectious particle as packaging requires recognition of a cos site at each end of the sequence (and 38-51 kb of DNA between the cos sites).

Molar ratio of vector to foreign DNA arranged to favor foreign DNA insertion vs self-ligation of foreign DNA (higher molar amount of the vector)

The resulting phage particles are used to infect *E. coli* cells and plated as described previously.

If insertion site is in the λ repressor (*cl* gene), then all plaques will be due to recombinant phage.

If insertion site is in LacZ marker gene, then all clear plaques will indicate recombinant phage and blue plaques will be due to non-recombinant phage.
Lambda phage is most commonly used for cDNA cloning

Lambda vector

(many copies)

In-vitro packaging of DNA with phage proteins

ds-cDNA with adaptors for restriction sites

Infection

Plating

Each plaque from phage with different insert
BAC (Bacterial Artificial Chromosome) vector

Designed for **cloning of large pieces of genomic DNA** (50-250 kbp)

BAC vectors are used to find contiguous pieces of genomic DNA so that the arrangement of the sequences in a chromosome can be determined.

BAC vectors are ds circular DNA which contains an ORI with tight replication control so that ideally only a single copy of the BAC with inserted DNA is retained in any cell. (Ori is derived from F’ plasmid of E coli)

Multiple copies of large DNA inserts leads to a low (but unacceptable) frequency of recombination.
BAC vector

from
http://escience.ws/b572/L18/L18.htm
BAC vectors

**Sequences for maintenance in E.coli**
- ori - E.coli rep. derived from high-copy plasmid pUC9
- CM(R) - chloramphenicol resistance gene; typical plasmid vectors use ampicillin resistance as a selectable marker, so it's better to have a different gene for BACs
- cloning site - site for insert

**Selection against clones with no inserts**
- PUCLINK stuffer fragment interrupts the sacB gene
- sacB encodes levanosucrase, which converts sucrose to levan, which is toxic to E. coli.
- PUCLINK stuffer fragment can be excised with restriction enzymes NotI (5'GC^GGCCGC3') , BamHI (5'G^GATCC3') or EcoRI (5'G^AATTC3').
- If the plasmid is recircularized with itself, the sacB promoter will now be directly upstream from the sacB coding sequence, and sacB will be expressed
- If the plasmid ligates with an insert, then sacB will not be expressed, and cells will survive.

**Working with large inserts**
- NotI site is best, because the recognition is 8 bp (5'GC^GGCCGC3'), rather than 6. (Remember, $4^6=4096; 4^8=65536$). Because NotI cuts, on the average, once every 65536 bp, most inserts can be excised as a single fragment, or perhaps several large fragments. 6-cutter enzymes would produce many smaller fragments.
- For physical reasons, circular BACs are much less susceptible to shearing than linear YACs
BAC - Selection for clones with inserts

Transformants are grown on levanosucrose.

If the vector recircularizes with itself, sacB is expressed, and cells are killed.

If the vector contains an insert, sacB is not expressed. Thus, only clones with inserts will be selected.
Summary of Vectors

• **Plasmids** - designed for cloning small pieces (0-8 kb) for foreign DNA
  - high copy number, high amplification
  - chemically treated cells (competent cells) allows plasmids to leak through weaken plasma membrane of bacteria (size limitations)
  - Transformation \( \sim 10^8 \) transformants/\( \mu \)g DNA

• **Bacteriophage lambda** - primarily used for cloning small cDNAs (0-8 kb) (single genes)
  » Valuable because of high transfection efficiency \( \sim 10^9/\mu \)g DNA and ability to segregate large numbers of plaques/plate

• **BAC vectors** - designed for cloning very large pieces of genomic DNA (50-250 kb) requires special transformation system to get large DNA into *E.coli* cells \( \sim 10^7/\mu \)g DNA