DNA Cloning Strategies

RossMcManus
Institute of Molecular Medicine, TCD
DMMC

What is Cloning?

- *Process by which a single individual is replicated to give rise to many essentially identical individuals*

- *Clone: A group of essentially identical individuals at the genetic level, derived from a single progenitor*
  - *Asexual Reproduction*
  - *DNA molecules,Viruses, Bacteria, Higher Organisms*

Cloning - General Principles

Cloning - General Principles

- *<u>Essentially</u> Identical*
  - *No Recombination, But Mutation will Occur*

DNA Cloning

- **Rationale**:
  - Massive amplification of DNA sequences
  - Stable propagation of DNA sequences

- A single DNA molecule can be amplified allowing it to be
  - Studied - Sequenced
  - Manipulated - Mutagenised or Engineered
  - Expressed - Generation of Protein

Whats it used for?

- Manipulating DNA sequences
  - Transgenic or knockout animals
  - Genetically Modified Organisms
  - Protein expression
    - Vaccines
    - Biologicals - infliximab
  - DNA perpetuation and analysis

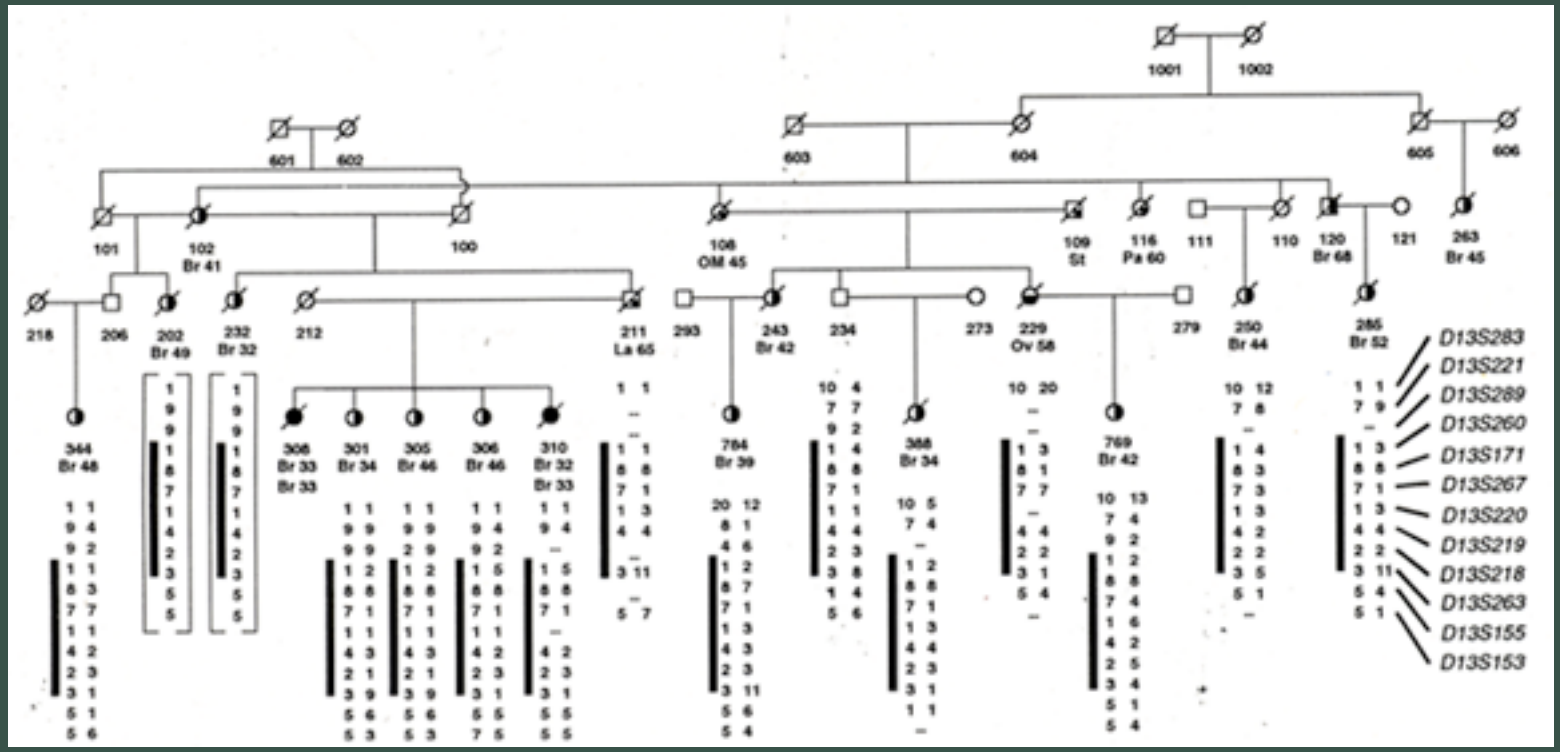
Maturity-onset obesity
in transgenic mice
expressing CMV-ADAR2
medschool1.mc.vanderbilt.edu/

Cloning Strategy

- **Strategy** depends on the starting information and desired endpoint.

- **Starting Information or Resources:**
  - Positional cloning information
  - Protein sequence
  - mRNA species / sequence
    - cDNA libraries
    - Subtractive hybridisation / genechip
  - DNA sequence known or unknown
    - genomic DNA libraries

Positional Cloning: *BRCA2*



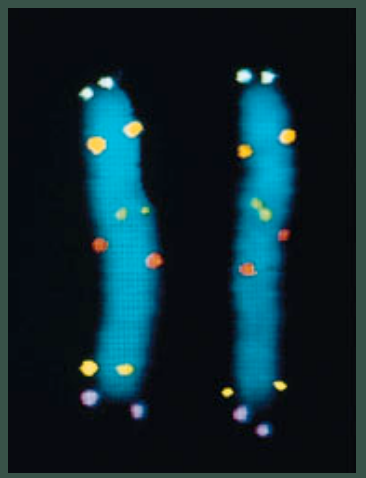
Wooster et al., 1994. **Science** 265: 2088-2090

Methodology

- Ability to **cut**, **modify** and **rejoin** DNA sequences
  - Restriction enzymes
    - Specifically recognise and cleave DNA sequences
  - DNA Polymerases
    - Klenow, Taq, Vent
    - In vitro generation of DNA
  - Phosphatases, Kinases
    - Removal and addition of Phosphates
  - Ligases, Topoisomerases
    - Reseal phosphodiester bonds

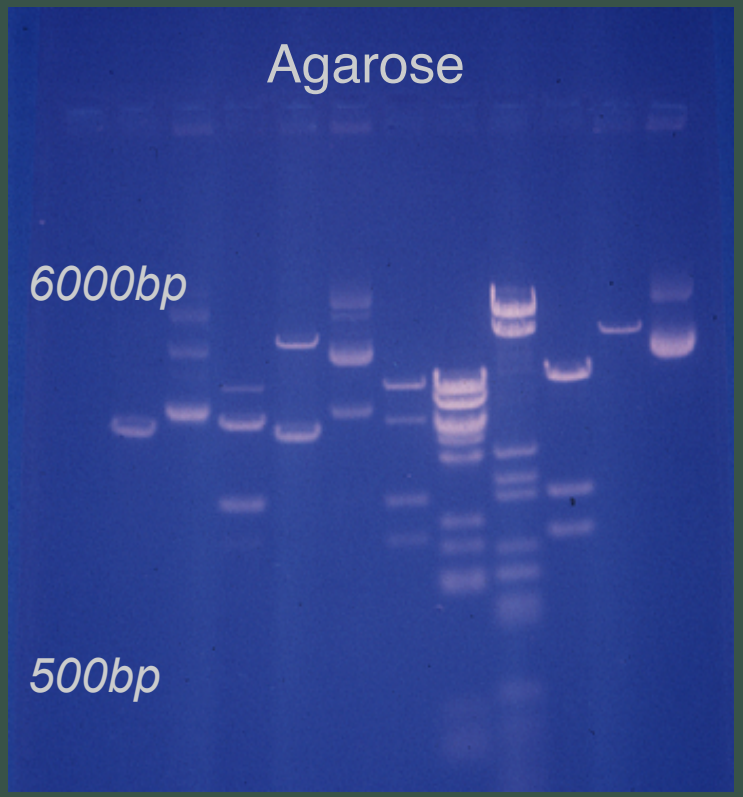
Methodology

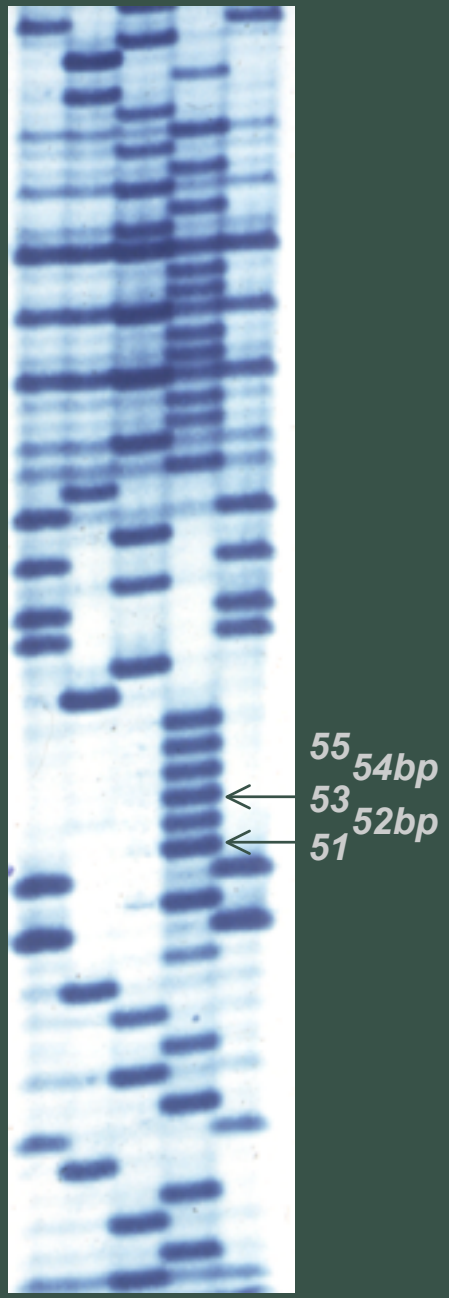
- Ability to **identify** DNA sequences
  - Sequencing
  - Hybridisation
  - Antibody detection



- Ability to **propagate** DNA
  - Cloning vectors (many varieties)
  - PCR
  - Transformation / Transfection of host cells

- Ability to **separate** DNA fragments
  - Electrophoresis

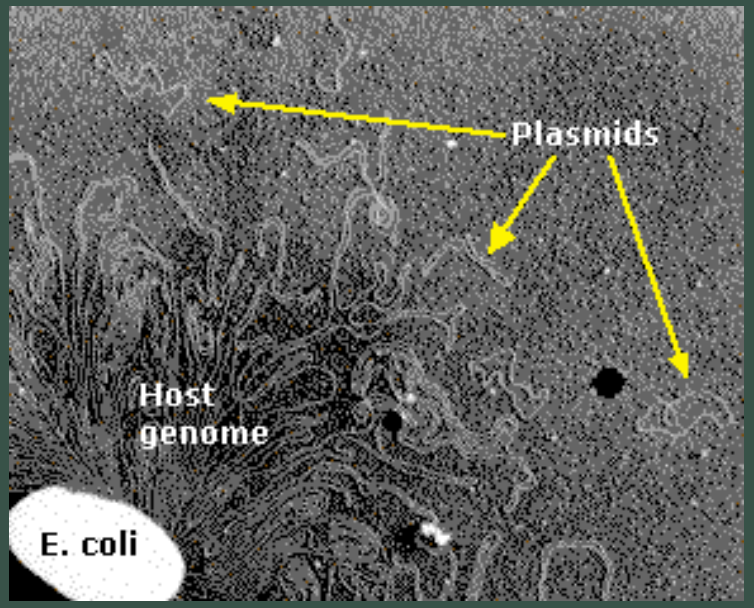




Polyacrylamide

Cloning Vectors:
Basic Components

- **Plasmids:** naturally occurring bacterial 'mini-chromosomes'
- **Vectors:** Engineered derivatives of plasmids used for cloning



- **Origin of replication** (Ori)
  - ColE1, p15a plasmids (copy number 1 - >50)
  - f1, M13 - filamentous phage
  - Eukaryotic- SV40

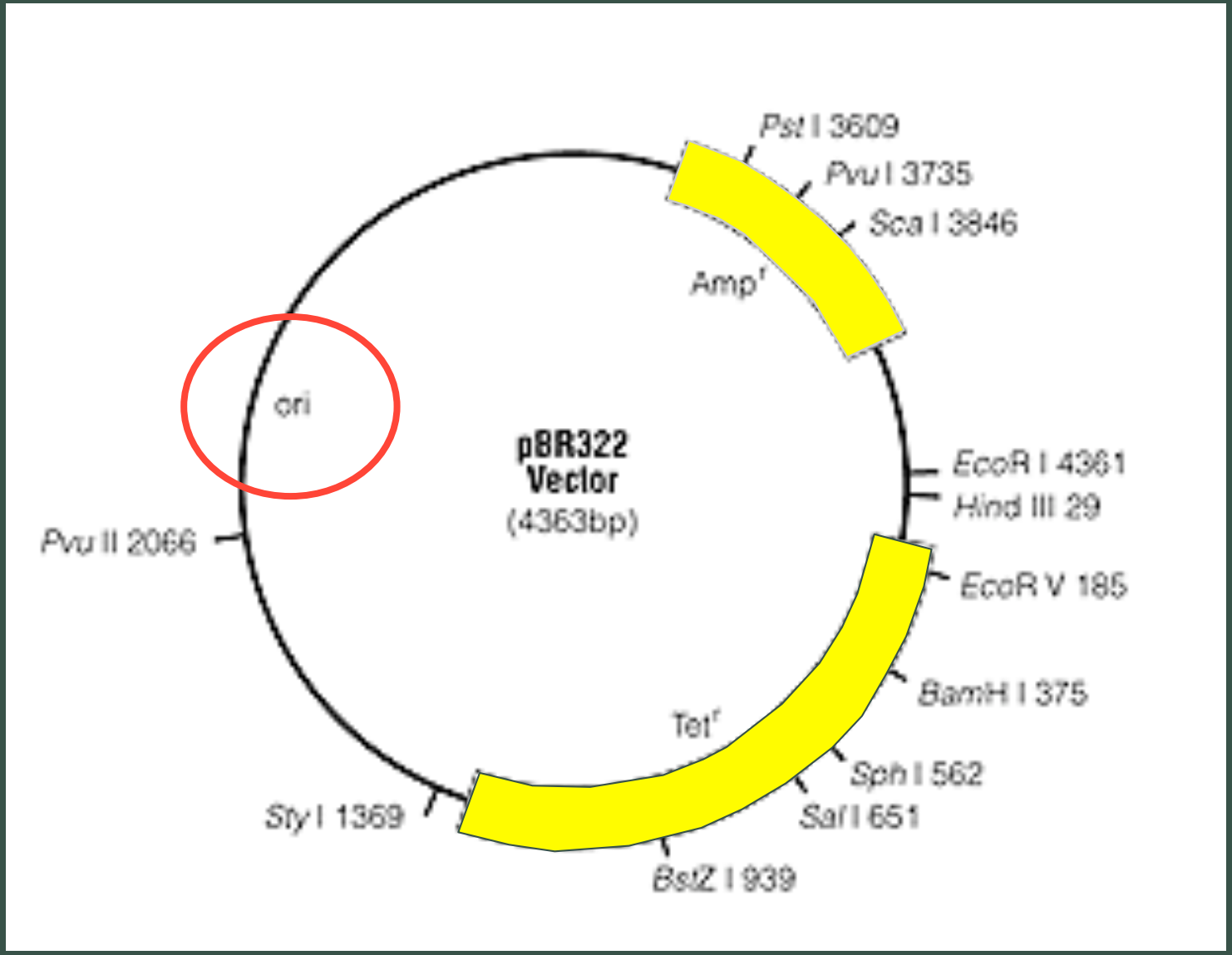
Mechanism of **plasmid selection**

- Bacterial Cells:
  - Ampicillian — Cell wall synthesis
  - Chloramphenicol — 50S ribosomal S/U
  - Kanamycin — Inhibits translocation
  - Tetracycline — AA-tRNA / ribosome binding
- Eukaryotic Cells:
  - Neomycin
  - G418

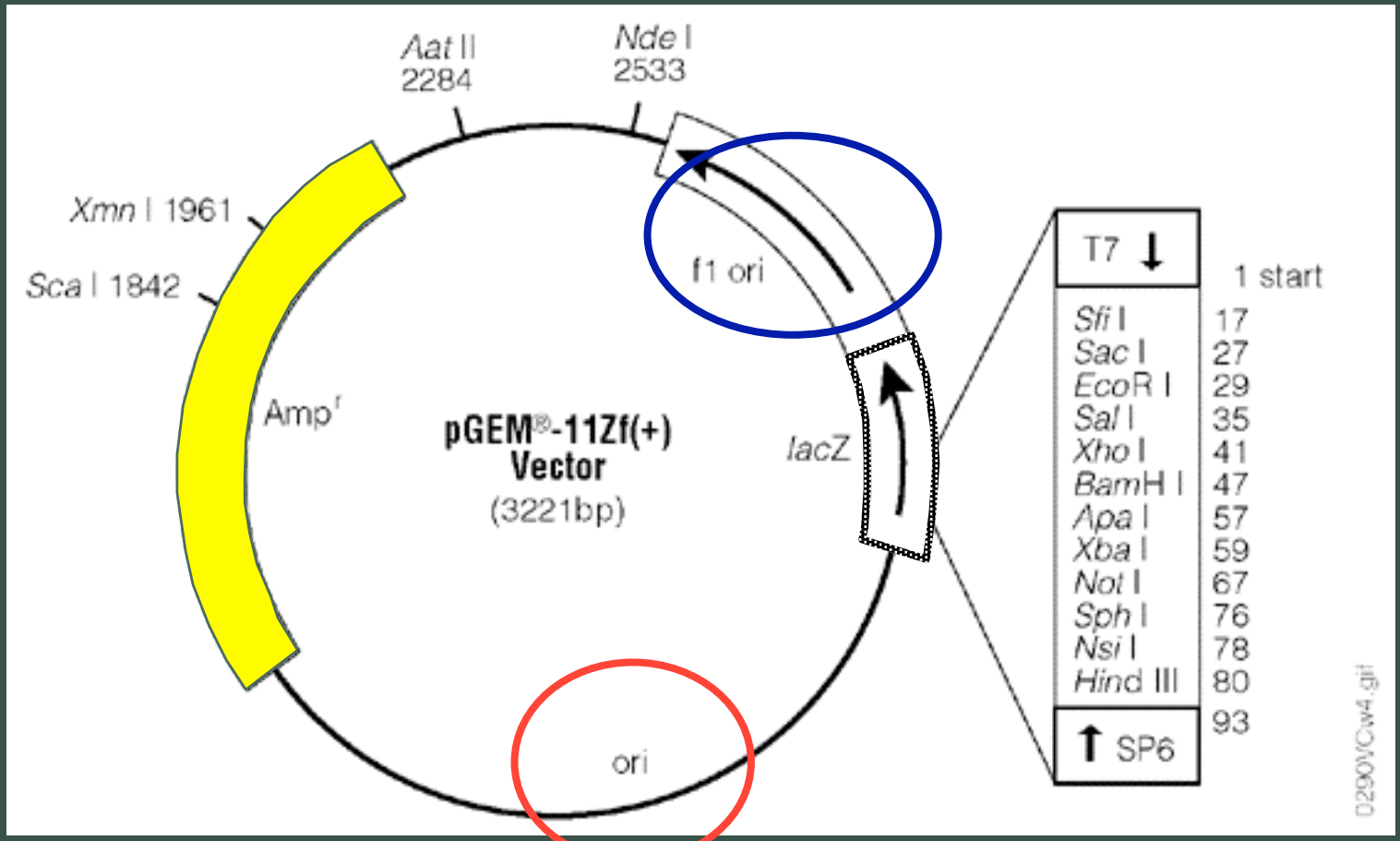
Cloning Vectors: Basic Components

- **Multiple cloning site** (**MCS**)
  - Selection of unique restriction enzyme sites
- Mechanism of **insert detection**
  - LacZ α peptide (blue / white colour)
  - Antibiotic inactivation
- **Other features**
  - T7 / SP6 promoters (? bidirectional)
  - Turning on and off expression in vivo (Tet on / off)
  - Sequencing primer binding sites (M13, T7)

pBR322


Pst I 3609
Pvu I 3735
Sca I 3846
Amp'
ori
pBR322
Vector
(4363bp)
EcoR I 4361
Hind III 29
EcoR V 185
BamH I 375
Tet'
Sph I 562
Sal I 651
Sty I 1369
BstZ I 939
Pvu II 2066

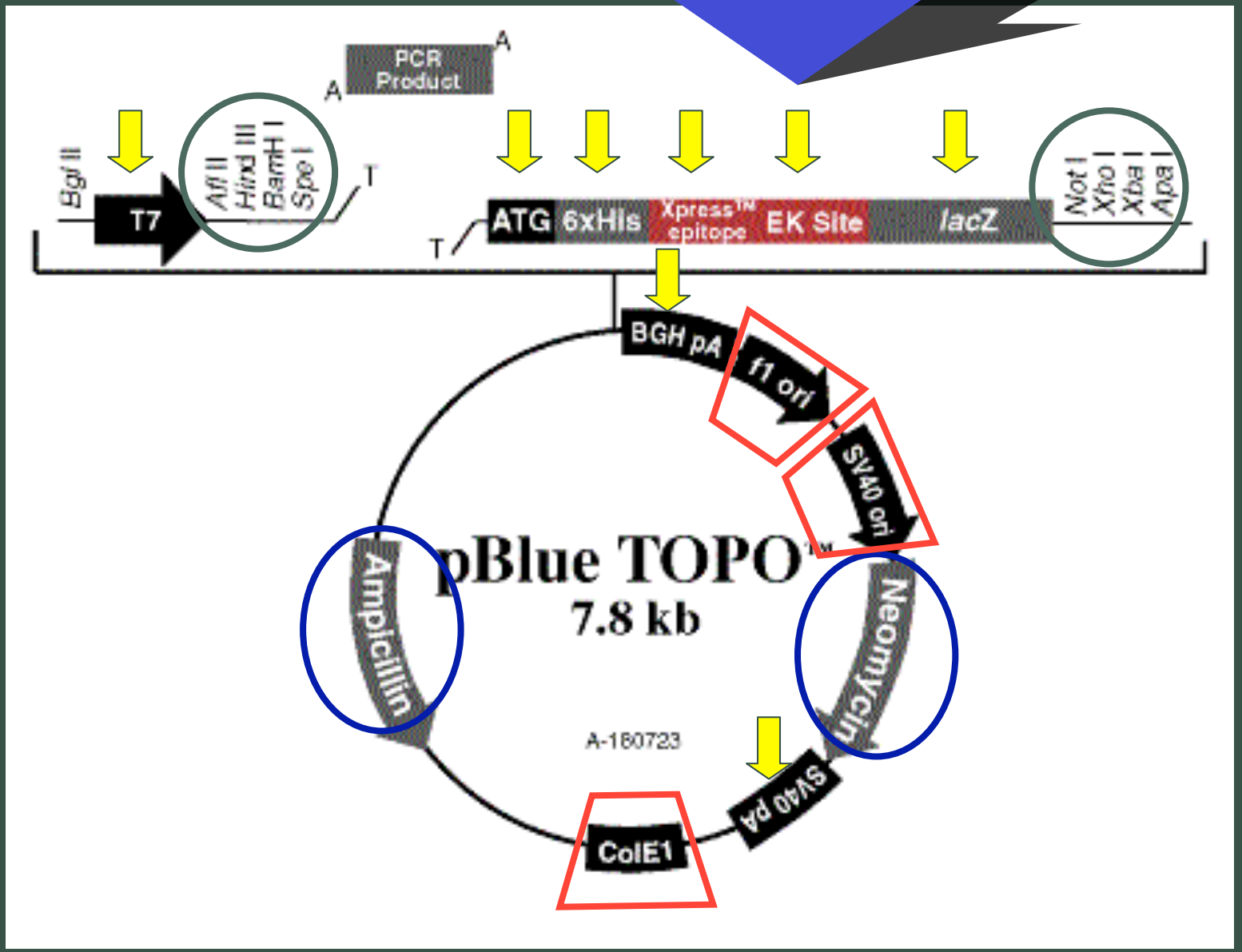
pGEM


Aat II
2284
Nde I
2533
Xmn I 1961
Sca I 1842
Amp^r
f1 ori
pGEM®-11Zf(+)
Vector
(3221bp)
lacZ
ori
T7
1 start
Sfi I 17
Sac I 27
EcoR I 29
Sal I 35
Xho I 41
BamH I 47
Apa I 57
Xba I 59
Not I 67
Sph I 76
Nsi I 78
Hind III 80
93
SP6
D290VCw4.gif

Cloning Vectors: Specialised features

- **Specialised vectors** exist to perform different functions
- **Expression**
  - Expression cassettes - protein expression;
  - Require promoter, initiation & termination codons, RBS (Kozak,polyA tail [3'UTR])
  - Fusion proteins - purification systems - his tag, GST, Maltose bp
- **Reporter genes**
  - analysing promoters, enhancers, translational processes
- **Sequencing**
  - Production of ssDNA

Expression Vector

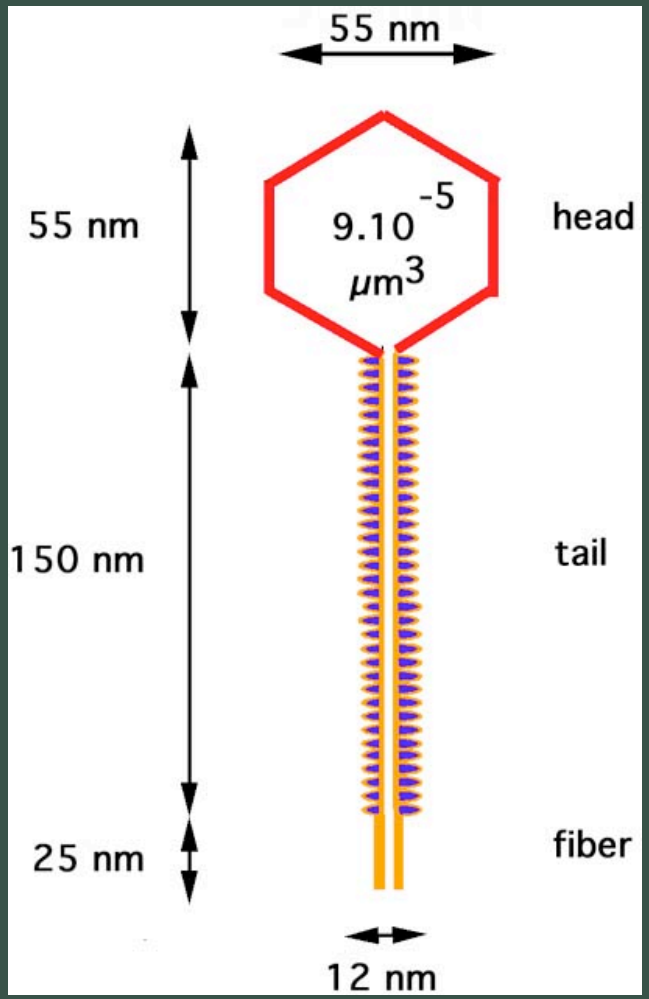


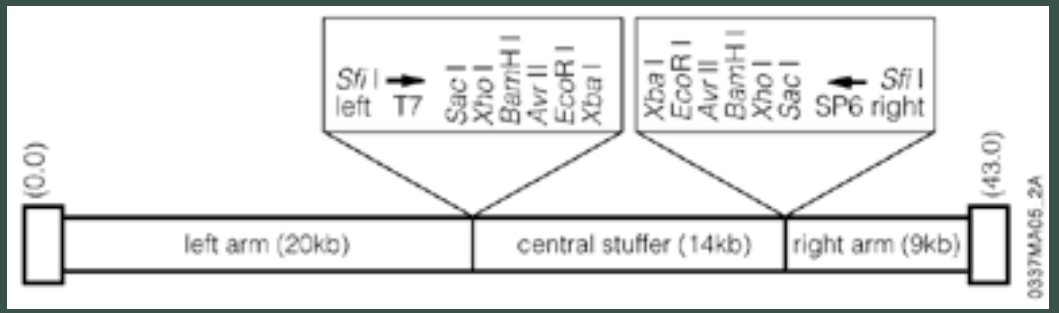
Cloning Vectors:

more **Specialised features**

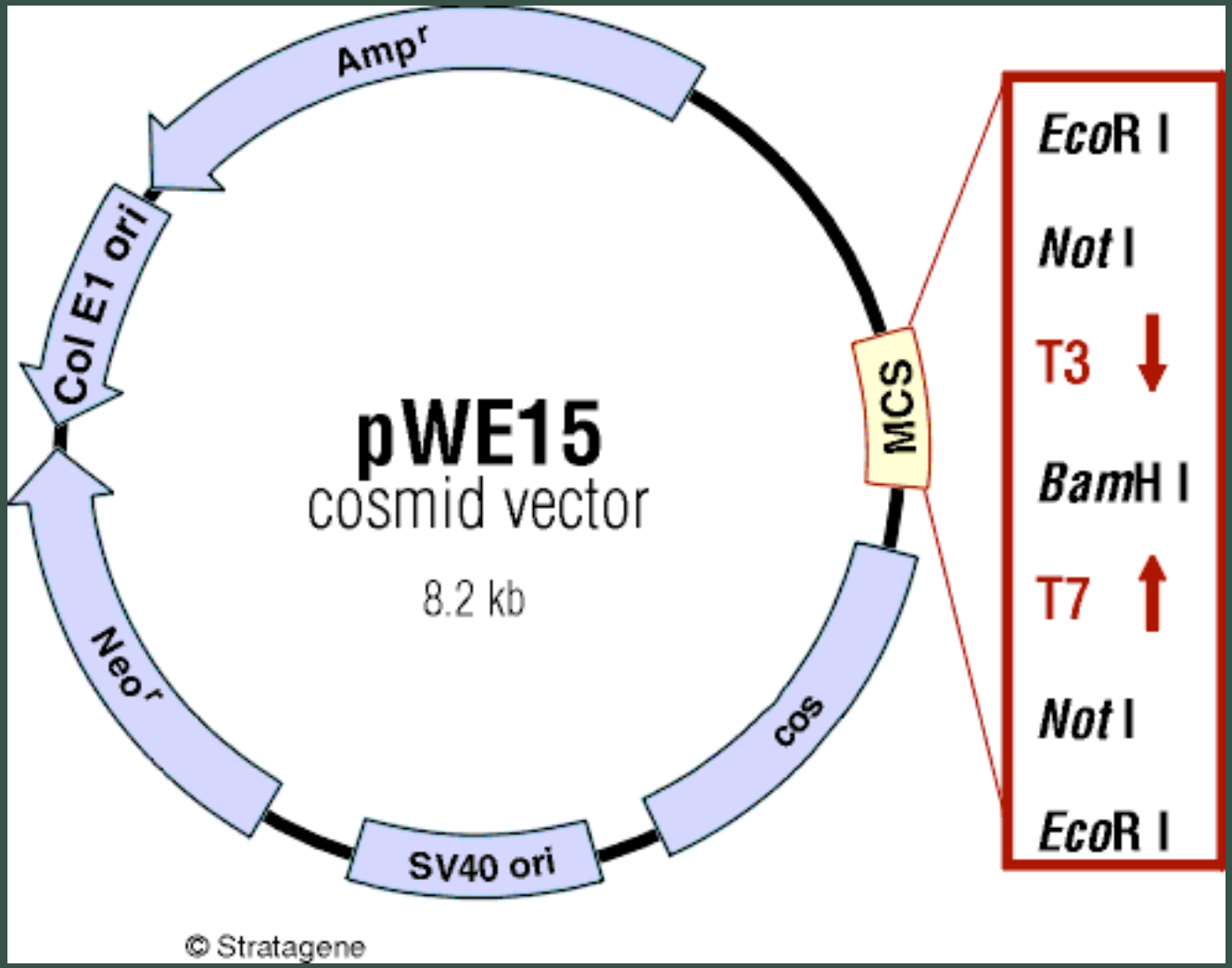
- **Cloning** (large scale)
  - High efficiency transfer: Lambda, P1
  - Large insert size: PAC, BAC, YAC, HAC
- **Mutagenesis**
  - Sequential alterations depend on strand selection
- **Combination vectors**
  - Shuttle vectors
  - Cosmids
  - Phagemids

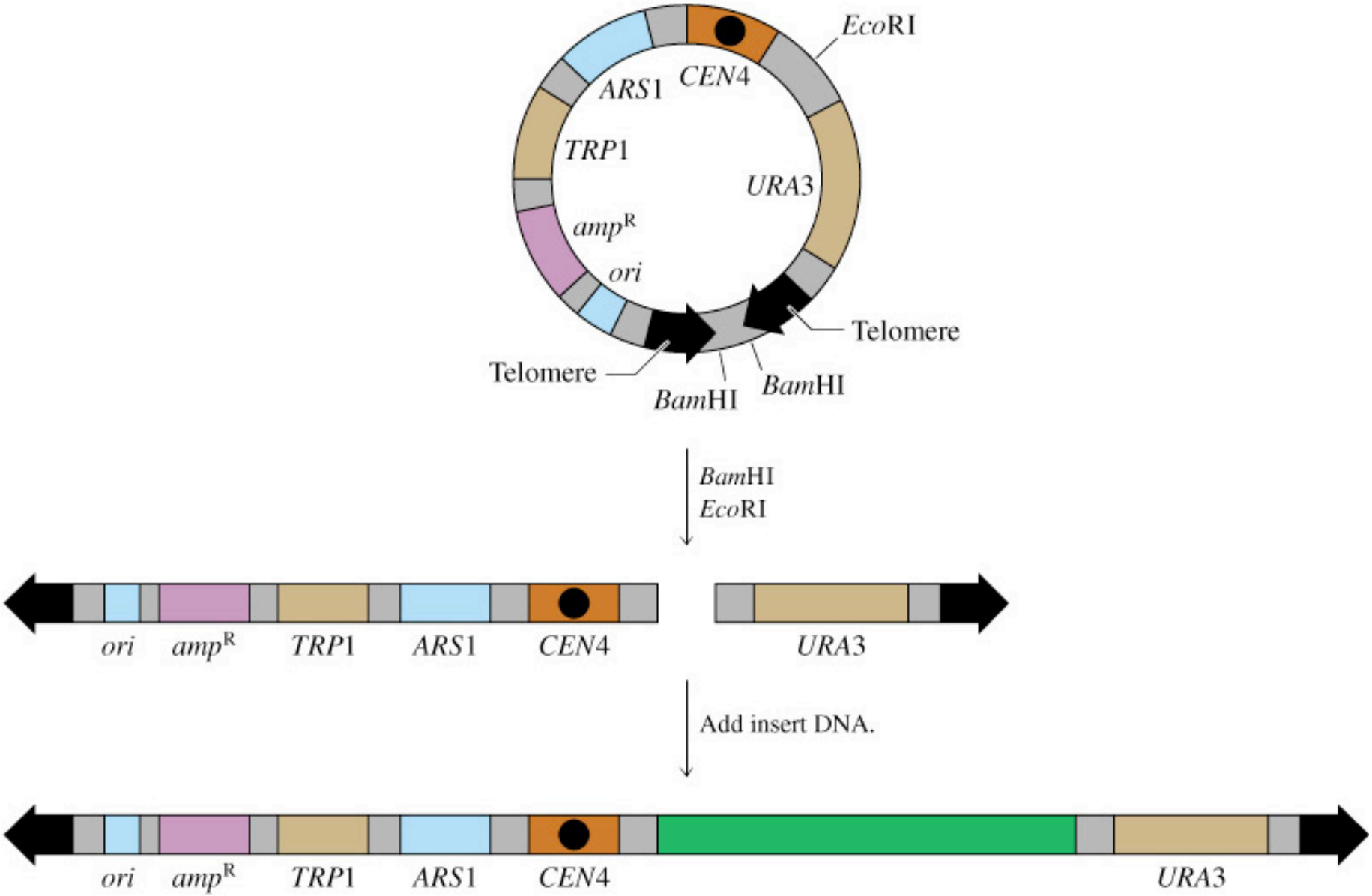
Bacteriophage Lambda-λ

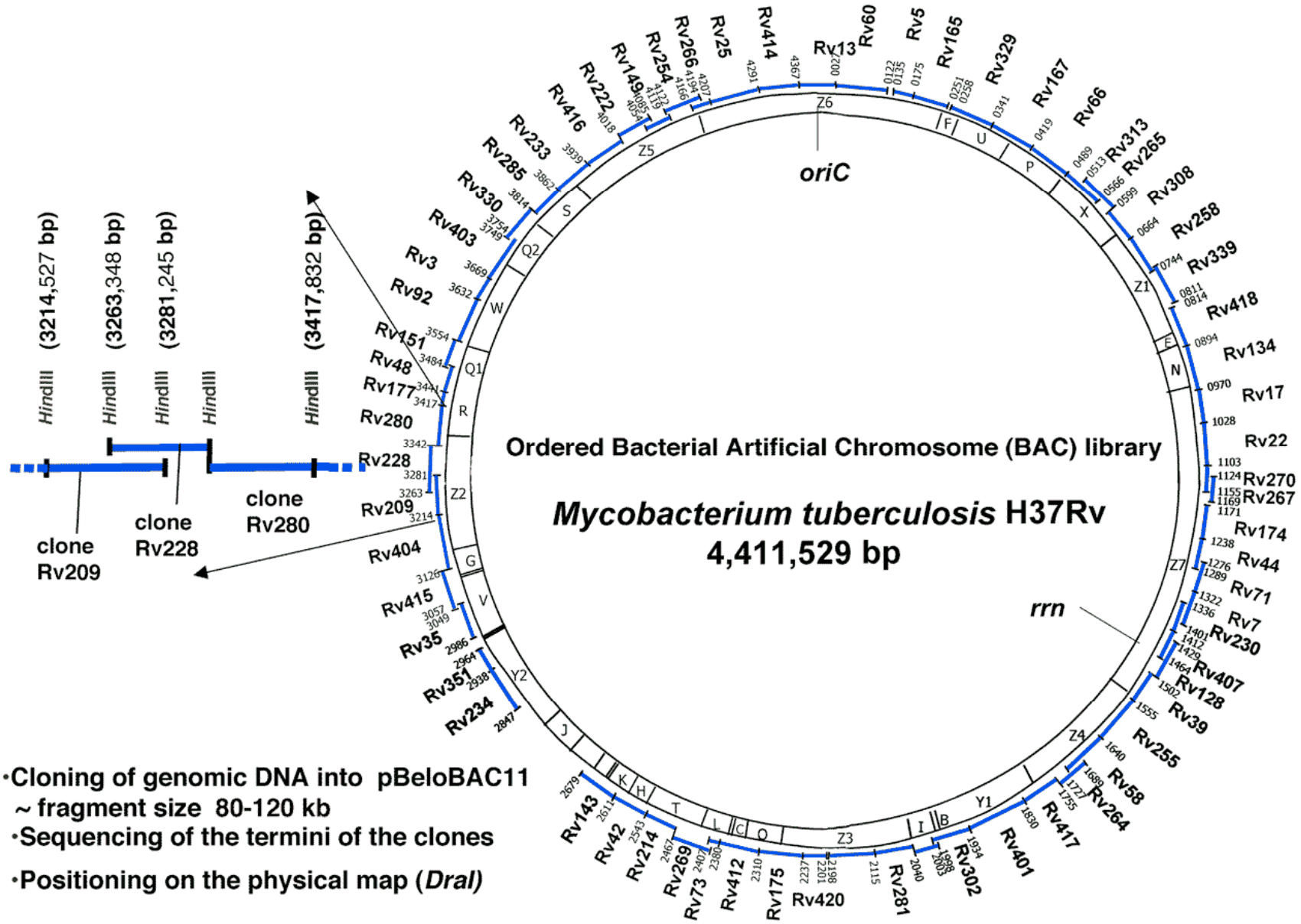




Cosmids - Half Plasmid, Half Phage



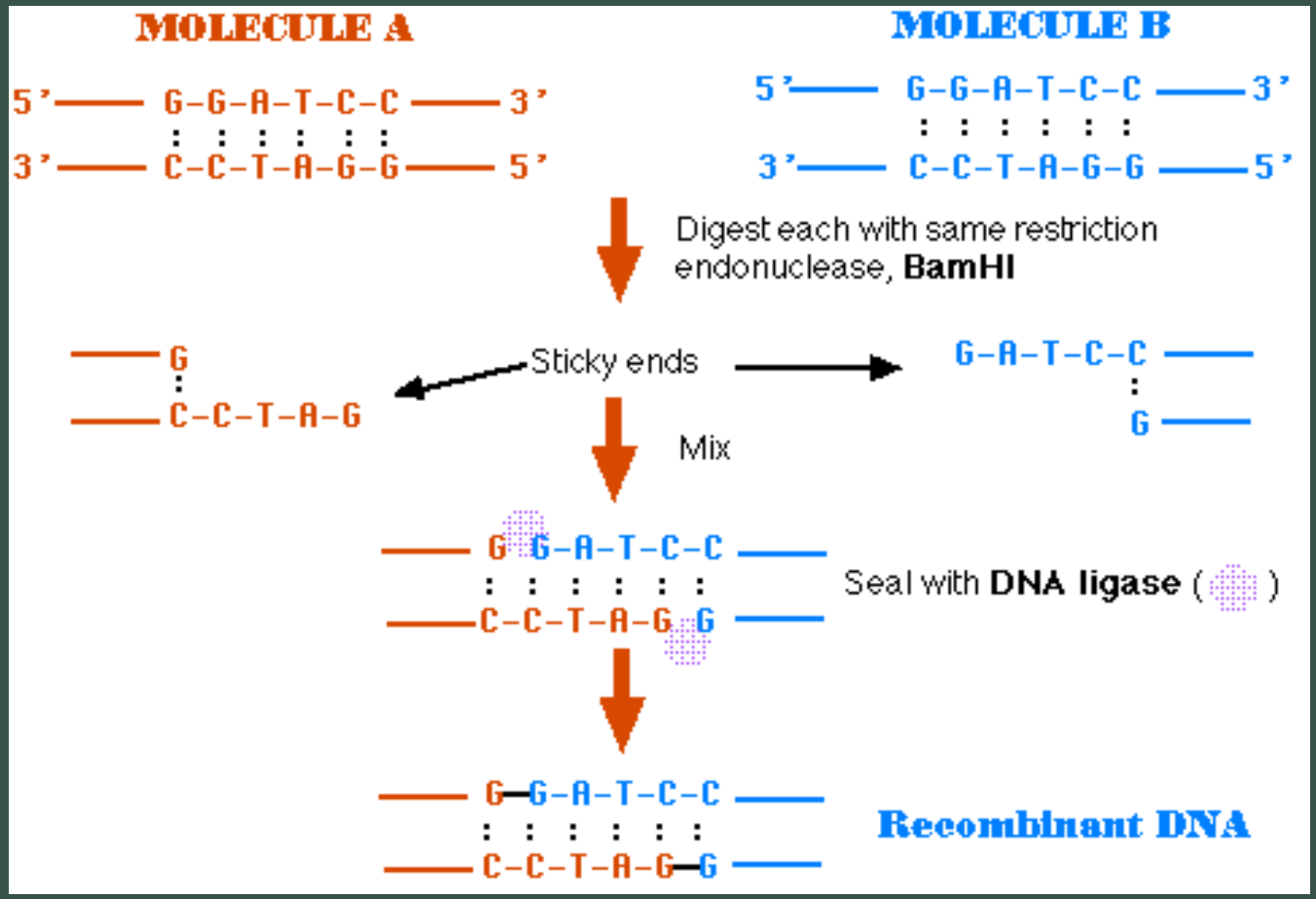

EcoRI
CEN4
ARS1
TRP1
URA3
ampR
ori
Telomere
Telomere
BamHI
BamHI
BamHI
EcoRI
ori
ampR
TRP1
ARS1
CEN4
URA3
Add insert DNA.
ori
ampR
TRP1
ARS1
CEN4
URA3



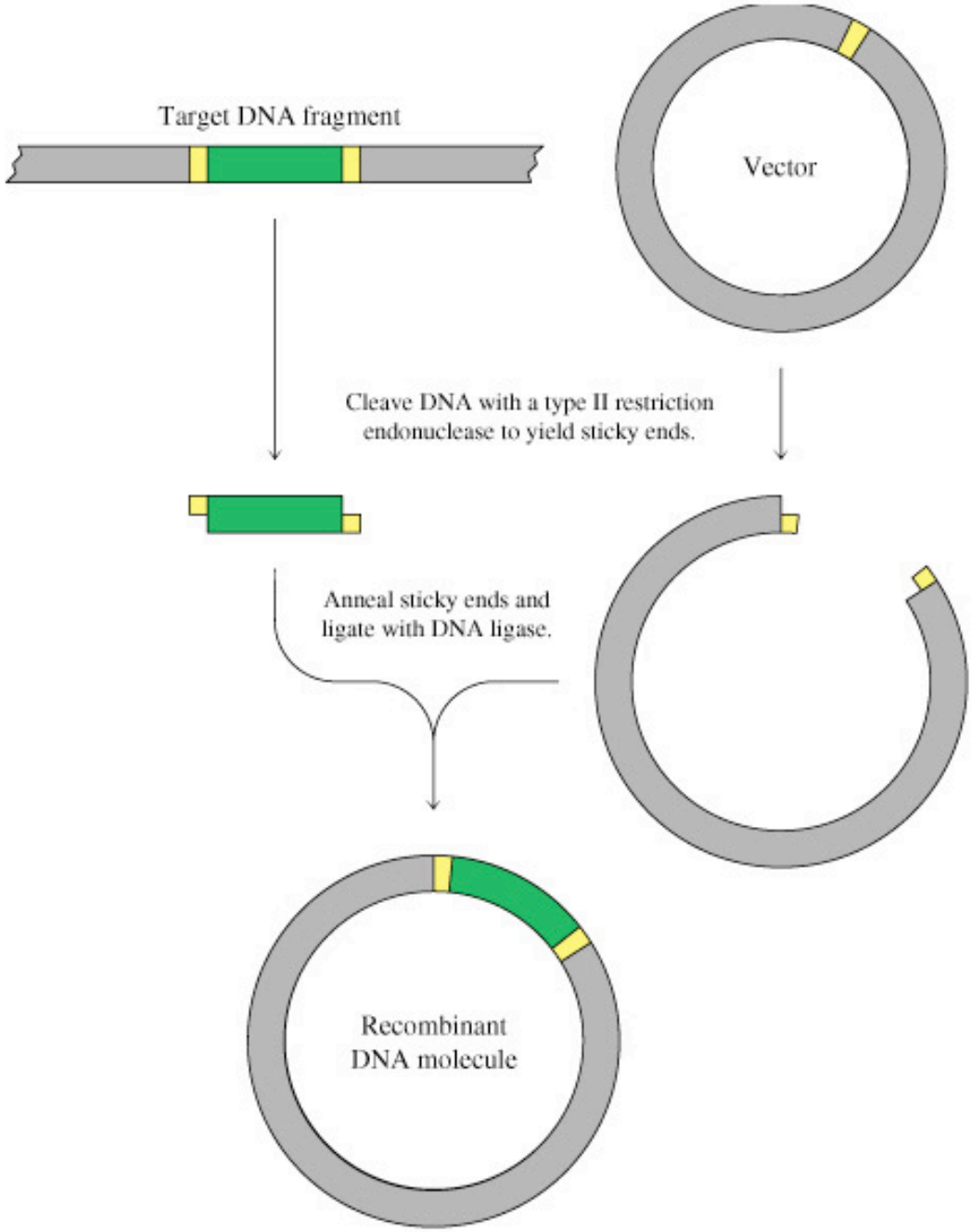
- Cloning of genomic DNA into pBeloBAC11
  ~ fragment size 80-120 kb
- Sequencing of the termini of the clones
- Positioning on the physical map (*DraI*)

Restriction Endonucleases

- Discovered early 70's
- Allows DNA to be **cut** specifically and reproducibly

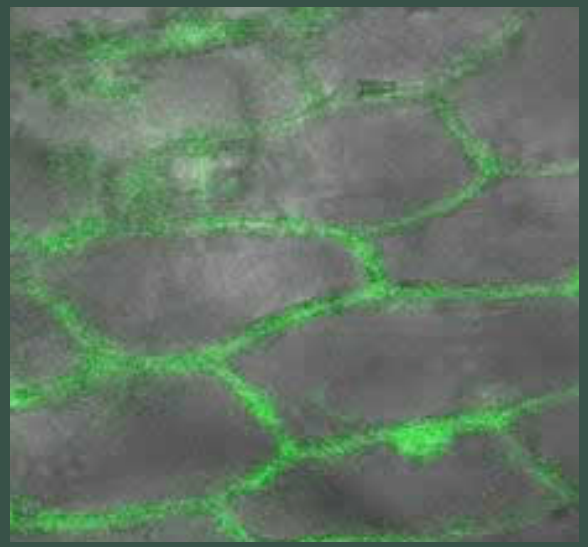


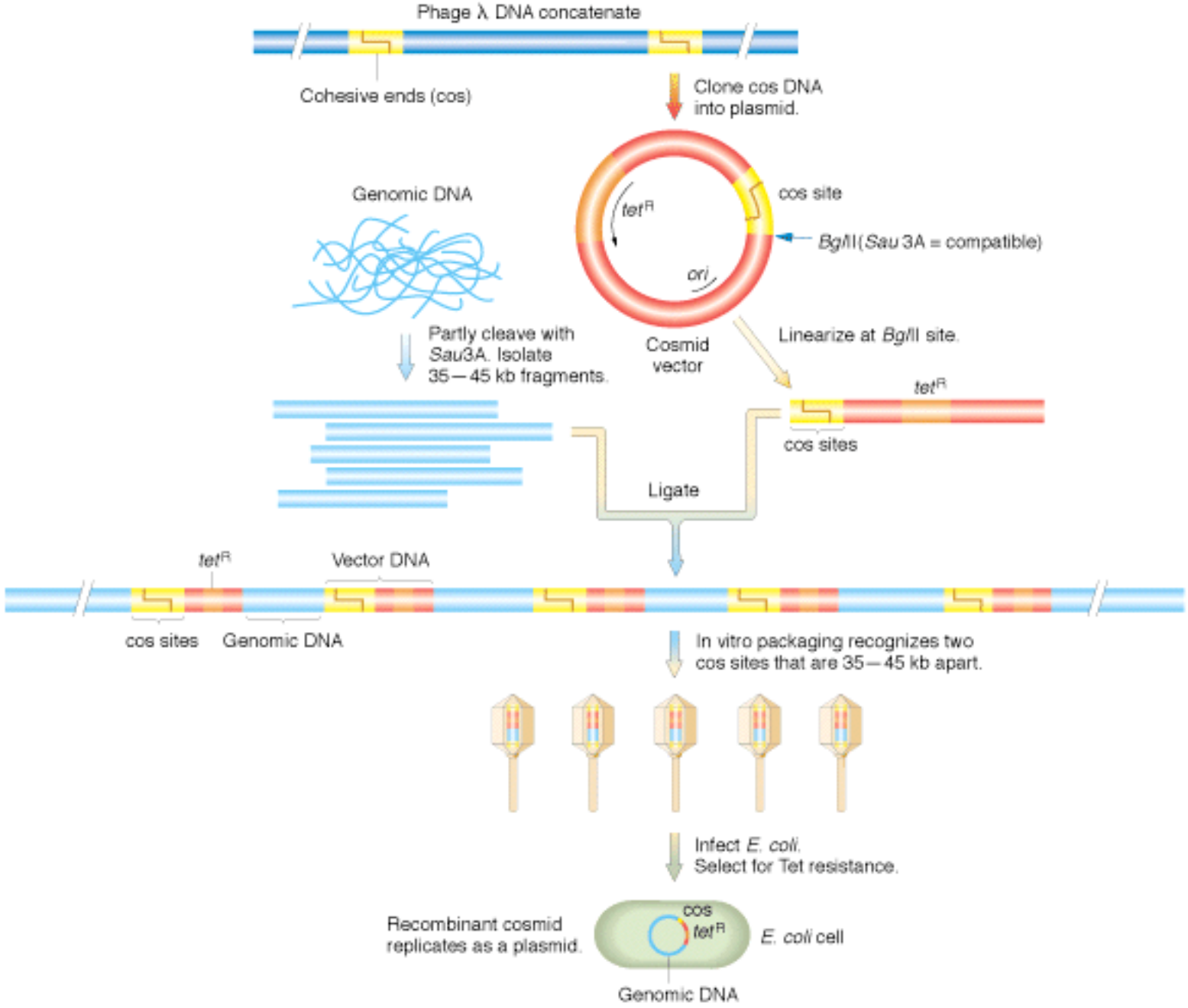
- Loads of enzymes with different site specificities
  - Flexibility in cloning
  - Incorporation of sites during PCR
- Allowed first detailed genetic mapping
  - RFLPs (SNPs)
- Vector construction

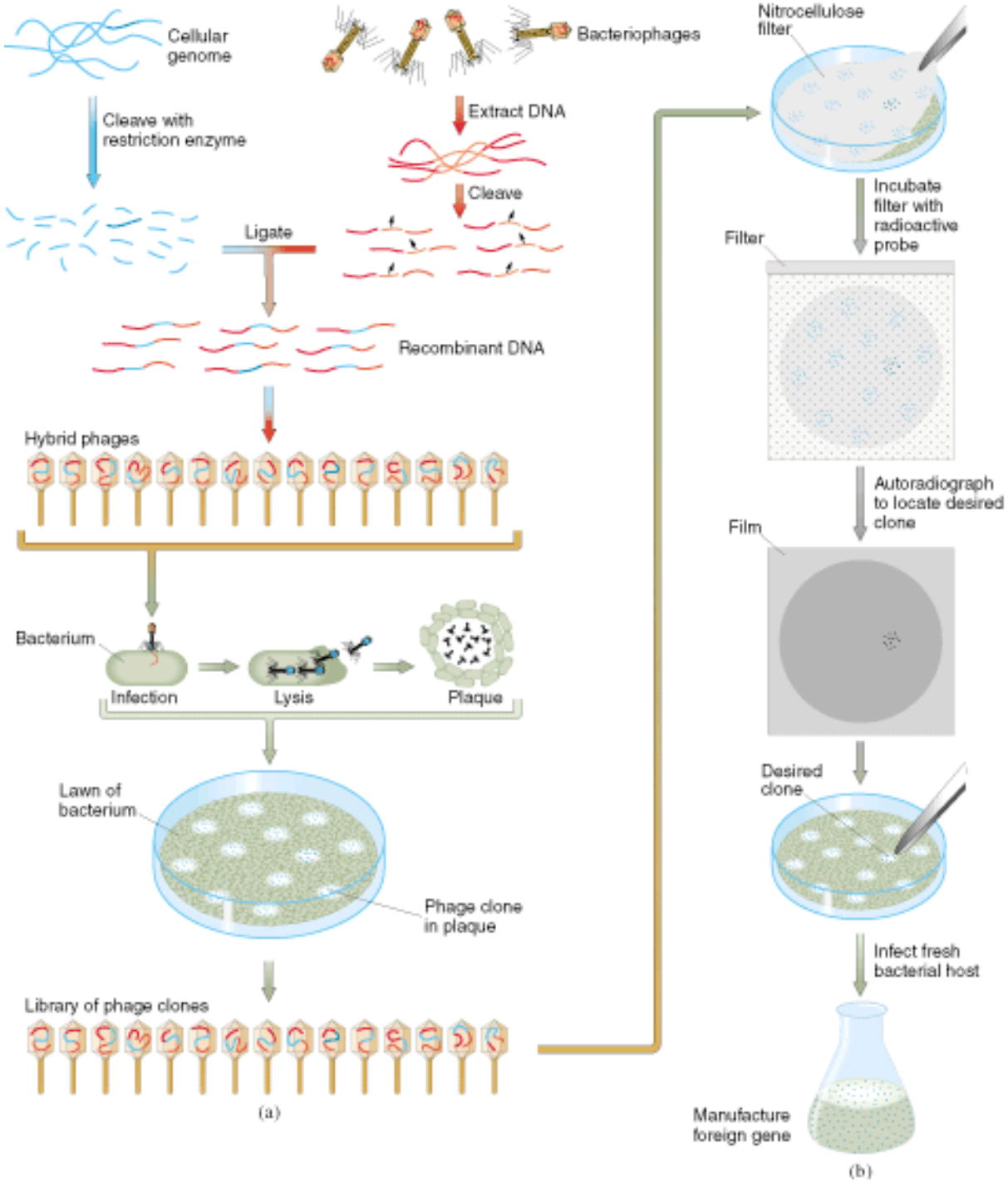

Target DNA fragment
Vector
Cleave DNA with a type II restriction endonuclease to yield sticky ends.
Anneal sticky ends and ligate with DNA ligase.
Recombinant DNA molecule

Cloning: Applications

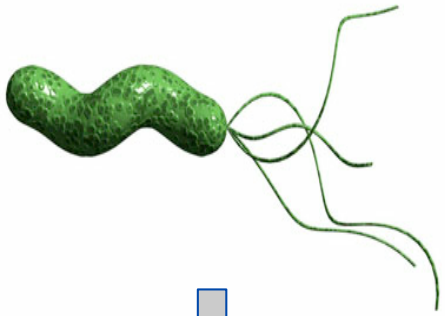
- Cloning DNA fragments
- Generating Libraries
  - Essential step for Genome Mapping
  - Shotgun Cloning
- Positional cloning
  - Discovering Disease Genes
- Discovering genes from e.g. Protein sequence
- Fusion proteins - GFP
- Gene regulation - Luciferase reporters




Phage λ DNA concatenate
Cohesive ends (cos)
Clone cos DNA into plasmid.
Genomic DNA
tetR
cos site
BglII (Sau 3A = compatible)
ori
Cosmid vector
Partly cleave with Sau3A. Isolate 35—45 kb fragments.
Linearize at BglII site.
tetR
cos sites
Ligate
tetR
Vector DNA
cos sites
Genomic DNA
In vitro packaging recognizes two cos sites that are 35—45 kb apart.
Infect E. coli. Select for Tet resistance.
Recombinant cosmid replicates as a plasmid.
cos
tetR
E. coli cell
Genomic DNA


Cellular genome
Bacteriophages
Cleave with restriction enzyme
Extract DNA
Cleave
Ligate
Recombinant DNA
Hybrid phages
Bacterium
Infection
Lysis
Plaque
Lawn of bacterium
Phage clone in plaque
Library of phage clones
(a)
Nitrocellulose filter
Incubate filter with radioactive probe
Filter
Autoradiograph to locate desired clone
Film
Desired clone
Infect fresh bacterial host
Manufacture foreign gene
(b)

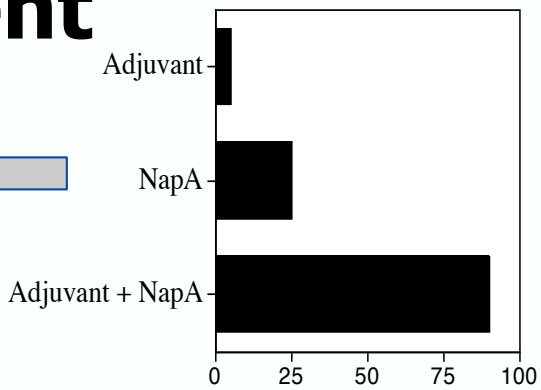
H. pylori: vaccine development

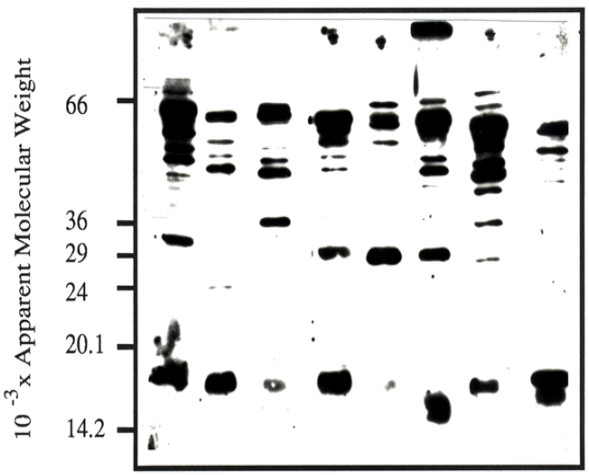


NapA in Phase I clinical trials as vaccine

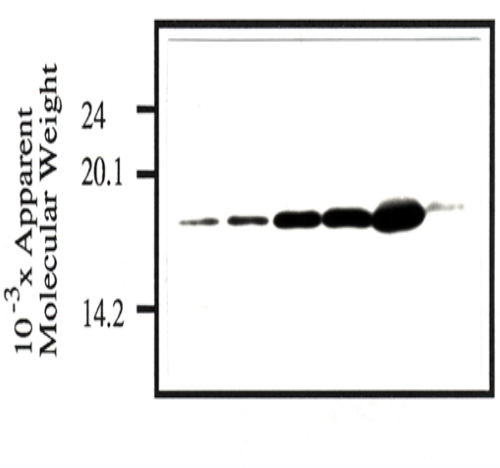
J. Exp. Med. (2000)

Vaccination

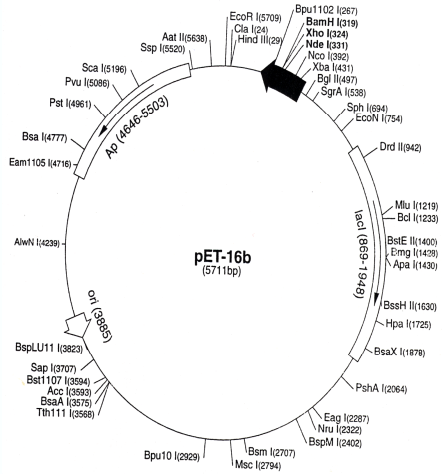




Identification



Purification



Cloning

H Windle, D Kelleher, R McManus

Libraries *II*

- **Efficiency** of reaction
  - important since all of the genome needs to be cloned
- E.coli 4,600,000 base pairs
  - 4600/8 = 575 **independent, non-overlapping** clones of 8 kb required to cover entire genome
- In practice many more required to cover breakpoints and provide **redundancy**

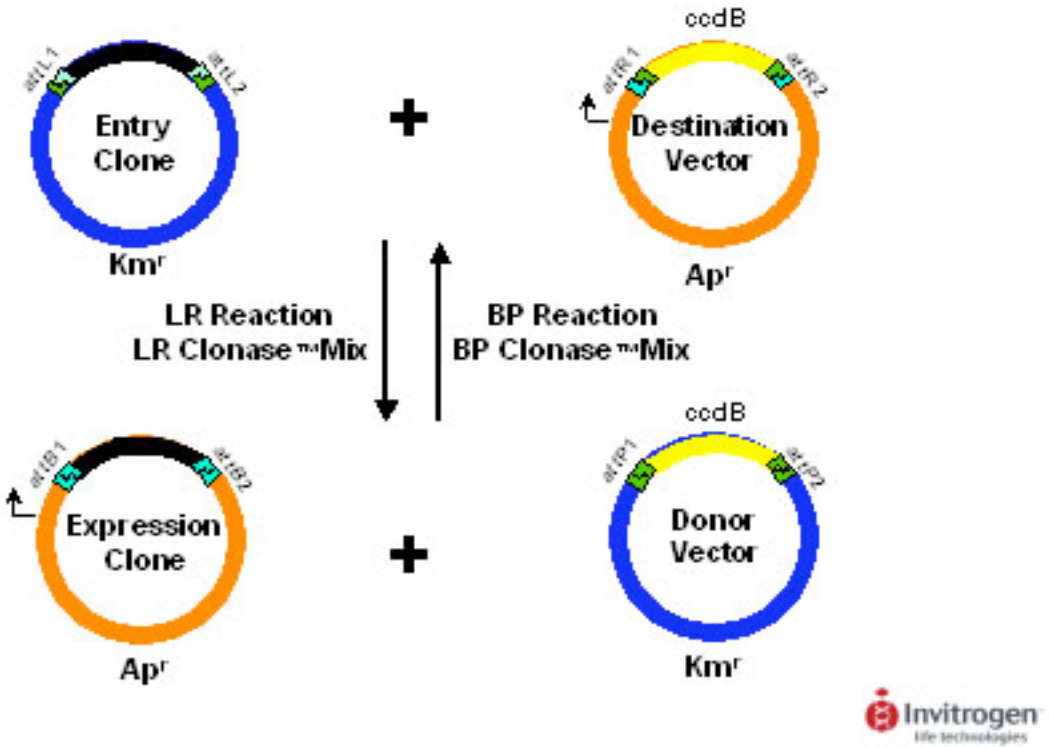
Screening

- Must be able to screen a library for recombinants to detect gene or sequence of interest
  - **Probe** e.g. labelled DNA sequence
  - This will hydrogen bond (**hybridise**) with similar sequences in library
- Clones derived from transformation of host cell must be screened in order to detect sequences of interest.
  - Plaque hybridisation
  - Colony hybridisation
  - Antibody screening

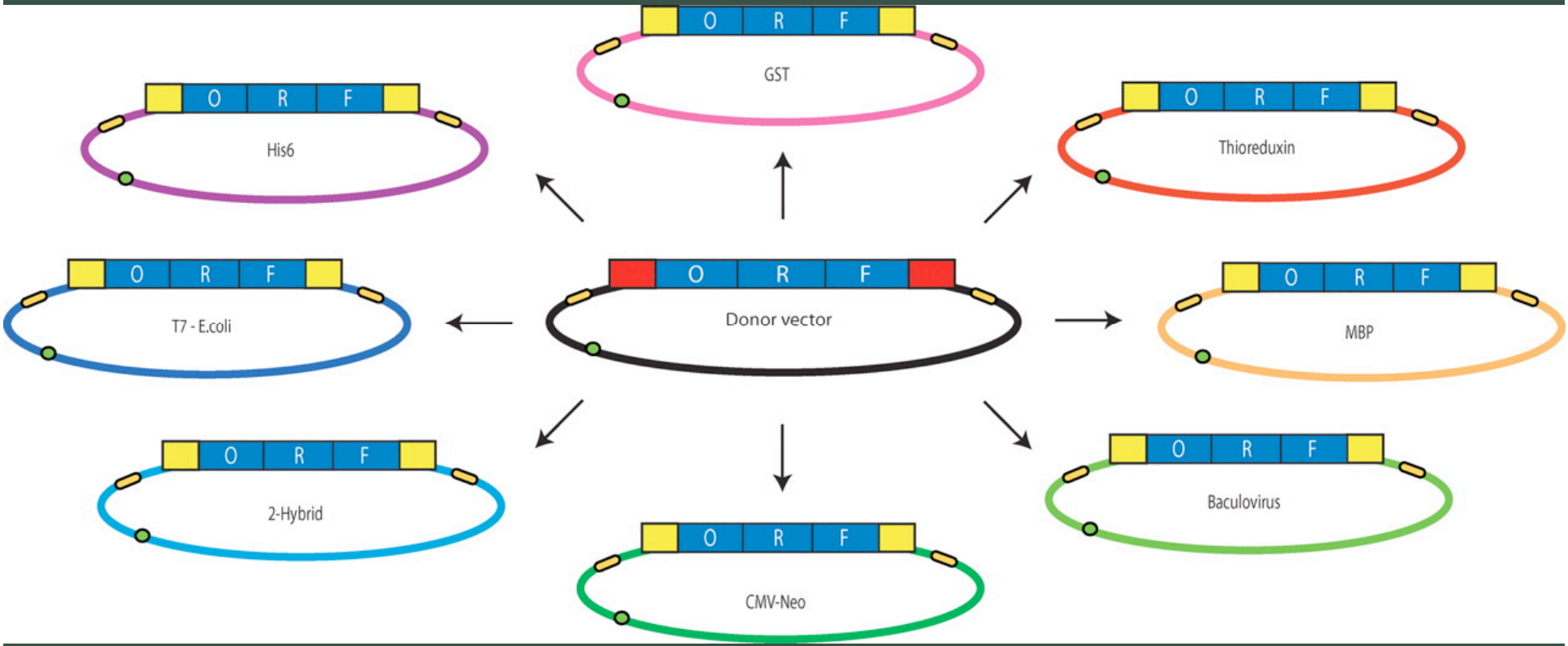
cDNA library

- **Isolate** mRNA
- **Reverse transcribe** mRNA into cDNA
- **Insert** cDNA into vector of choice
  - Amplify library
- **Screen** for sequence of interest using a known DNA probe (hybridisation)
    - Find related sequences
- Screen for differentially expressed mRNAs (or deleted DNA)
  - **Subtractive hybridisation**

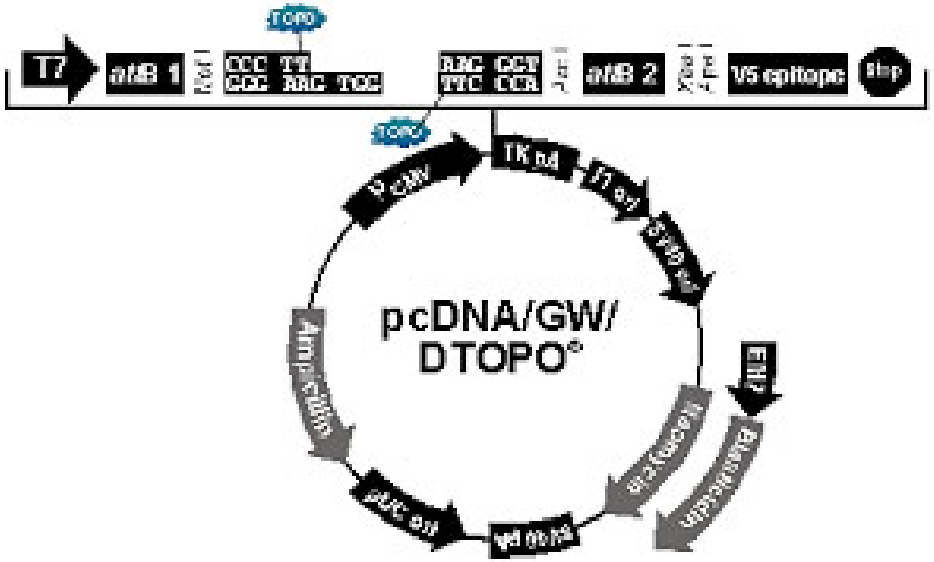
Gateway® Technology



GATEWAY cloning System

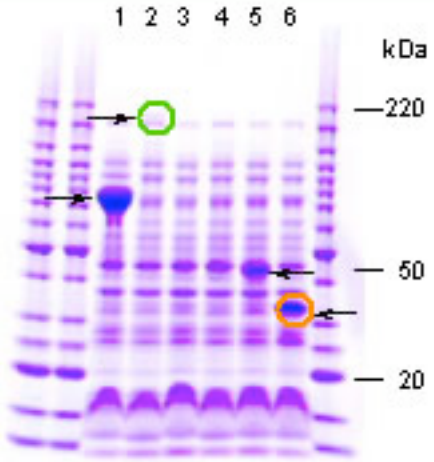


pcDNA/GW/ D-TOPO® Vectors



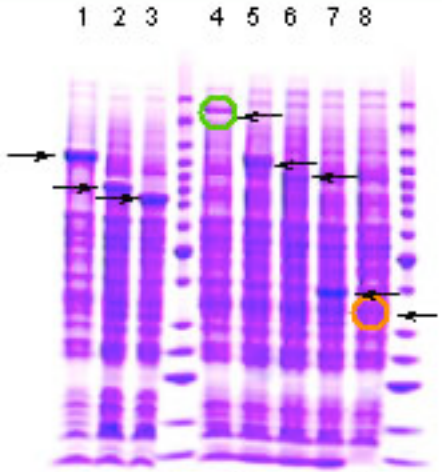
Invitrogen
life technologies

Expression of Full-Length Human ORFs



***E. coli* strain BL21 SI**

Lane 1: 6xHis-GUS
Lane 2: 6xHis-MAP4
Lane 3: 6xHis-β-Adaptin
Lane 4: 6xHis-Transferrin Receptor
Lane 5: 6xHis-Tyr Kinase
Lane 6: 6xHis-EIF4e



Baculovirus/*Sf*9 Insect Cells

Lane 1: GST-GUS
Lane 2: 6xHis-GUS
Lane 3: GUS
Lane 4: MAP4
Lane 5: β-Adaptin
Lane 6: Transferrin Receptor
Lane 7: Tyr Kinase
Lane 8: EIF4e

Invitrogen
life technologies

Sources of Clones

- NEDO (Japan)
  - >18,500
- MGC (mammalian gene collection: US)
  - >10,000
- DKFZ (Germany)
  - >10,000
- Full length and coding region only available
  - (Gateway & similar vectors)

Commercial Sources

- IMAGE consortium
- Genecopia - ORF express
- Invitrogen - Ultimate ORF collection
- Open Biosystems - Freedom ORF collection / Incyte gene collection
- Origene - True clone collection
- RZPD - Full ORF shuttle