I. Restriction endonucleases

Prokaryotic enzymes that make double-stranded cuts in DNA.
Name = *1st letter of Genus* + *two letters of species* (+ strain) + roman numeral:

*Eco*RI	*Escherichia coli* strain R, enzyme I
*Kpn*I	*Klebsiella pneumoniae*, enzyme I
*Hin*dIII	*Haemophilus influenzae* strain d, enzyme III

Recognize a specific sequence, usually a **palindrome**, 4-8 bp. Length (and composition) of sequence determines frequency of cutting.

*Alu*I	5'-AGCT-3' 3'-TCGA-5'	4 bp, cuts on average every 256 bp
*Eco*RI	5'-GAATTC-3' 3'-CTTAAG-5'	6 bp, cuts on average every 4096 bp
*Not*I	5'-GCGGCCGC-3' 3'-CGCCGGCG-5'	8 bp, cuts on average every 65,536 bp less frequently in AT-rich regions

Cut in a specific place relative to the recognition sequence (usually within). Can generate single-stranded overhangs that are complementary (**sticky ends**) or **blunt ends**. 5' end has phosphate, 3' end has hydroxyl.

*Alu*I	5'-AG|CT-3' 3'-TC|GA-5'	5'-AG-OH P CT-3' 3'-TC-P HO-GA-5'	blunt ends
*Eco*RI	5'-G|AATTC-3' 3'-CTTAA|G-5'	5'-G-OH P AATTC-3' 3'-CTTAA-P HO-G-5'	5' overhang
*Kpn*I	5'-GGTAC|C-3' 5'-C|CTAGG-3'	5'-GGTAC-OH P C-3' 3'-C-P HO-CTAGG-5'	3' overhang

Fragments that have complementary overhangs, even if from different sources, can be joined together by using **DNA ligase**.

II. Other enzymes used in cloning

DNA polymerases: 5' overhangs can be filled in to make them blunt-ended:

*Eco*RI **5'-G** → (polymerase + dTTP + dATP) → **GAATT**
3'-CTTAA **CTTAA**

3' overhangs can be trimmed off to make them blunt-ended:

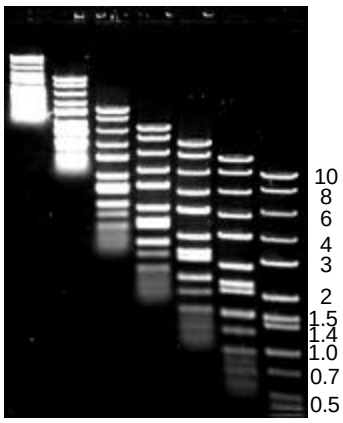
*Kpn*I 5'-GGTAC → (polymerase with 3'→5' exonuclease) → 5'-G
5'-C 3'-C

phophatase: removes the 5' phosphates to prevent ligation

III. Gel electrophoresis

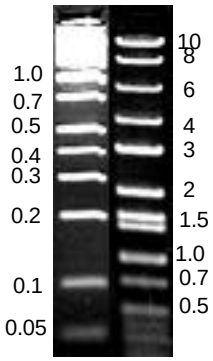
DNA fragments can be separated by size by agarose gel electrophoresis. Phosphate groups give DNA a negative charge, so fragments move in an electric field, from negative to positive. Agarose provides a sieve – smaller fragments move through pores more easily, quickly.

Higher concentrations of agarose separate lower size fragments better.



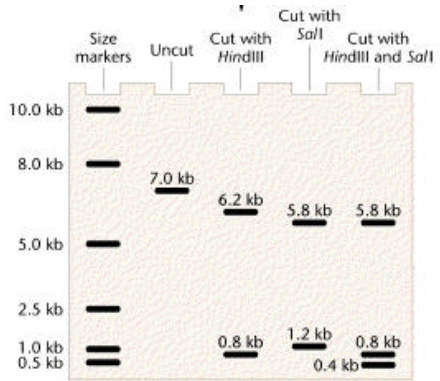
These figures show separation of DNA fragments on agarose gels. The gel on the left was loaded at 15-minute intervals, so the lane on the left ran for 15 min, the next for 30 min, and the lane on the right for 1 hr 45 min.

The gel on the right shows how different concentrations of agarose separate different ranges of sizes. The lane on the left is 3% agarose – the smaller fragments are resolved; the lane on the right is 0.6% agarose – larger fragments are resolved.

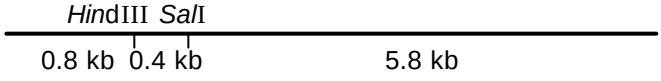


IV. Restriction mapping

We can "map" DNA with unknown sequence by digestion with single and multiple enzymes:



1. uncut = 7.0 kb
2. *Hin*dIII gives 6.2 kb + 0.8 kb;
   therefore, one *Hin*dIII site, 0.8 kb from one end.
3. *Sal*I gives 5.8 kb + 1.2 kb;
   therefore, one *Sal*I site, 1.2 kb from one end.
4. *Hin*dIII + *Sal*I gives 5.8 kb + 0.8 kb + 0.4 kb;
   the 1.2 kb *Sal*I band was cut by *Hin*dIII, so:



V. Features of cloning vectors

Ability to maintain itself in the host cell:
origin of replication that regulates copy number
single copy, low copy (<100/cell), high copy (200-1000/cell)
selectable marker, usually an antibiotic resistance gene

Ease of inserting fragments:
multiple restriction enzyme sites = **polylinker** or **multiple cloning site**
system for selecting for insertion of a DNA fragment

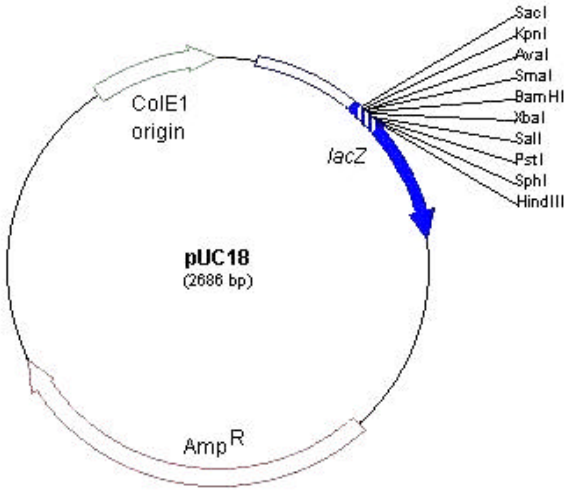
Ease of introduction into and recovery from host cells

VI. Plasmid vectors

pUC18: ColE1 origin system keeps copy number at 200-500 per cell.
Confers resistance to ampicillin by β-lactamase gene (*bla*).
Polylinker is <u>within</u> coding region of *lacZ* gene; insertion of DNA into polylinker disrupts β-gal production. When grown in the presence of X-gal, colonies expressing β-gal turn blue, while other colonies remain white = **blue-white selection** for colonies containing inserted DNA.

Insertion into cell is via transformation; recovery from cell is easy because plasmids stay in solution, but chromosomal DNA sticks to cell membrane.

Plasmid vectors typically hold up to about 20 kb of insert easily.



SacI KpnI SmaI BamHI XbaI SalI PstI SphI HindIII
ATGACCATGATTACGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCACTG
TACTGGTACTAATGCTTAAGCTCGAGCCATGGGCCCCTAGGAGATCTCAGCTGGACGTCCGTACGTTCGAACCGTGAC

VII. Vectors for cloning larger fragments

Other vectors have been developed in part to allow cloning of larger fragments.

l **replacement vectors**. All the genes essential for lytic growth are found in the first 20 kb and the last 10 kb of λ. One can replace the central region, which has only lysogeny genes, with any DNA. DNA is ligated into the left and right "arms" and packaged into phage heads *in vitro*. Only clones carrying an insert of the right size (10-20 kb, enough to make one headful) will be packaged. Packaged particles are used to infect *E. coli*. Lytic growth occurs, producing plaques (if grown on plates) or lysis (if grown in liquid culture). This makes more DNA and more phage particles to propagate the clone.

Cosmids are plasmid vectors that carry λ *cos* sequences, which allows them to be packaged in λ particles *in vitro*. One can then introduce them into host cells by infection rather than transformation, after which they behave just like plasmids. Cosmid vectors typically carry inserts of about 40 kb.

Bacterial artificial chromosomes (**BAC**s) are plasmids based on the F plasmid. They can carry inserts up to several hundred kb. They are often introduced into the host cell by **electroporation**, an improved transformation technology, and generally maintained at a single copy per cell, which avoids toxicity of too much foreign DNA.

VIII. Specialized vectors

Expression vectors carry a bacterial or phage promoter, allowing one to express protein encoded by the cloned DNA. Example: human insulin produced by *E. coli* replaced insulin purified from pig pancreases.

T-DNA is used to introduce DNA into plant genomes. Derived from **Ti** (tumor-inducing) plasmids of *Agrobacterium tumifaciens*, which produces crown gall tumors. Engineered versions have sequences allowing growth in *E. coli* (*ori*, kan^R), and the genes necessary for integration into the plant genome.

Drosophila P element vectors. Somewhat similar to a λ replacement vector. The transposase gene is cut out of a P element, and any DNA you desire is put in it's place. This construct has the *cis* sequences for transposition (the inverted repeats and nearby sequences at the ends), but not the *trans* sequences (the transposase); it is therefore non-autonomous. For transformation, the P element construct is injected into the germlines of Drosophila embryos along with transposase. Transformants are detected by expression of a marker gene on the P construct (usually $white^+$).

Gene therapy vectors. Some genetic diseases can be treated by providing a wild-type copy of the defective gene. Vector must provide a way to get the gene into cells where it must be expressed.

1. Retroviral vectors. Similar to λ replacement vectors and *P* element vectors – replace viral sequences with gene of interest, use virus to carry gene into target cells. Problems: random integration, attacks only proliferating cells.
2. Adenovirus vectors. Normally infects respiratory epithelia (non-dividing cells); genome does not integrate into chromosome, but persists inside cell. Used in gene therapy for cystic fibrosis.
3. HACs (human artificial chromosomes)? Not yet, but some day…