VirginiaTech
Invent the Future

Institute for
Critical Technology and Applied Science

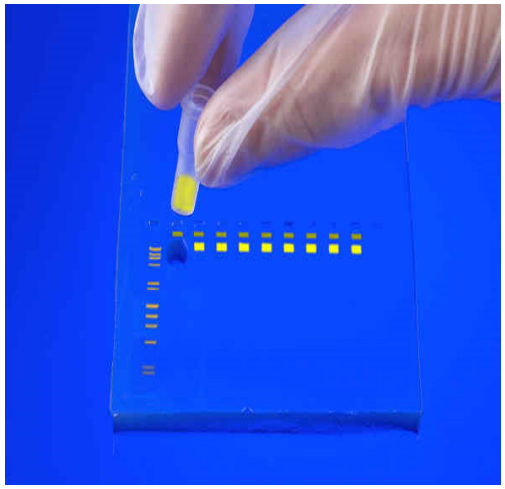
Restriction endonucleases and ligases

Mohamed N. Seleem

Big picture: Gene cloning

- From Gene to Protein
- Generally use bacteria as the “factory”

What we have done so far?



- Decide which gene to clone
- Designing primers
- PCR amplification (xx billions copies)
- Gel electrophoresis
- Cutting our gene from gel

What is next

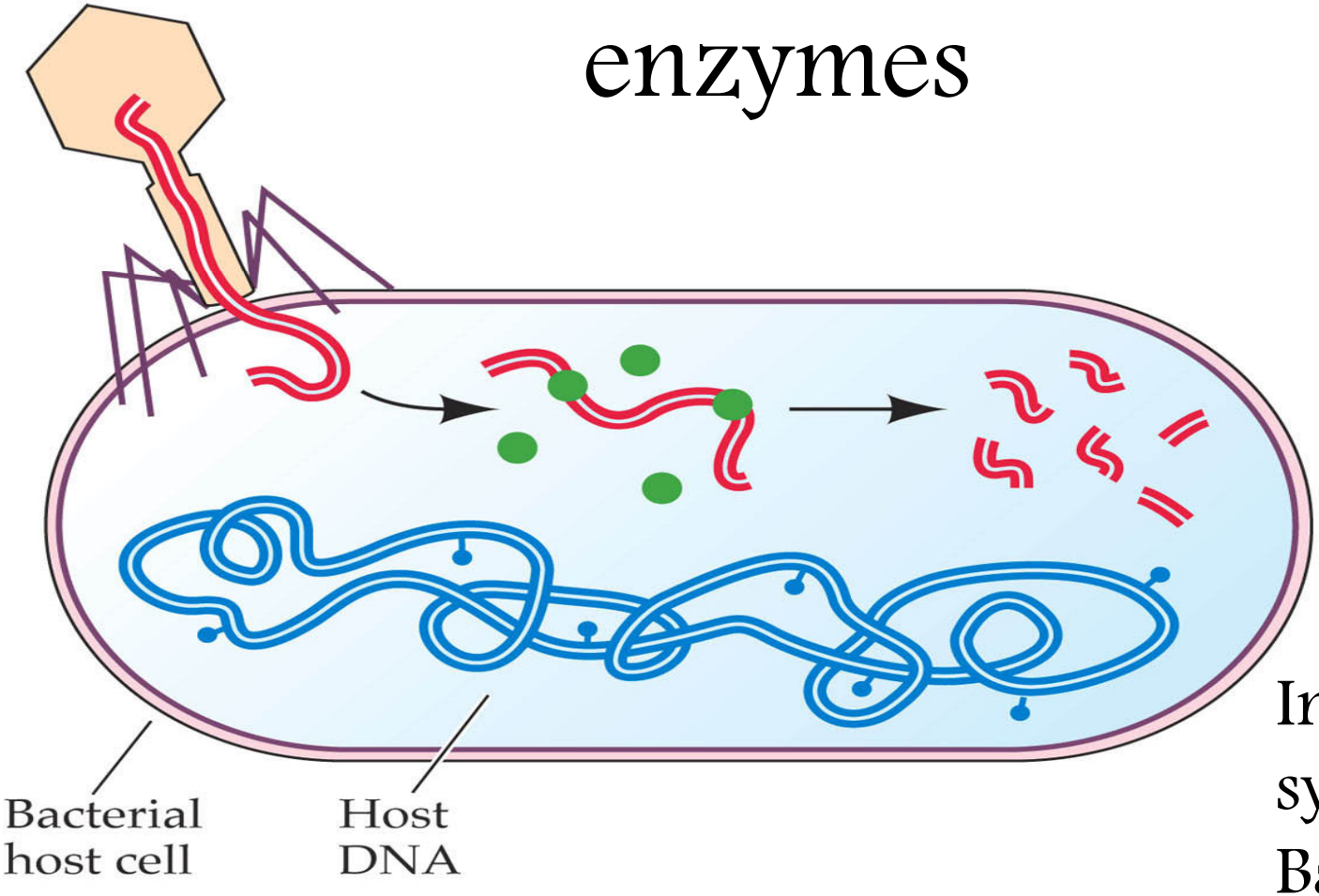
- Move my gene to the bacteria to purify protein as a final product.

Problem: Gene by itself does not go inside bacteria, it needs a carrier (vector or plasmid)

Cloning Tools

- Restriction endonucleases
- Ligase
- Vectors
- Host
- Methods for introducing DNA into a host cell

Restriction endonucleases enzymes



Immune system of Bacteria

LIFE: THE SCIENCE OF BIOLOGY, Seventh Edition, Figure 16.1 Bacteria Fight Invading Viruses with Restriction Enzymes

VirginiaTech
Invent the Future

Haemophilus influenzae

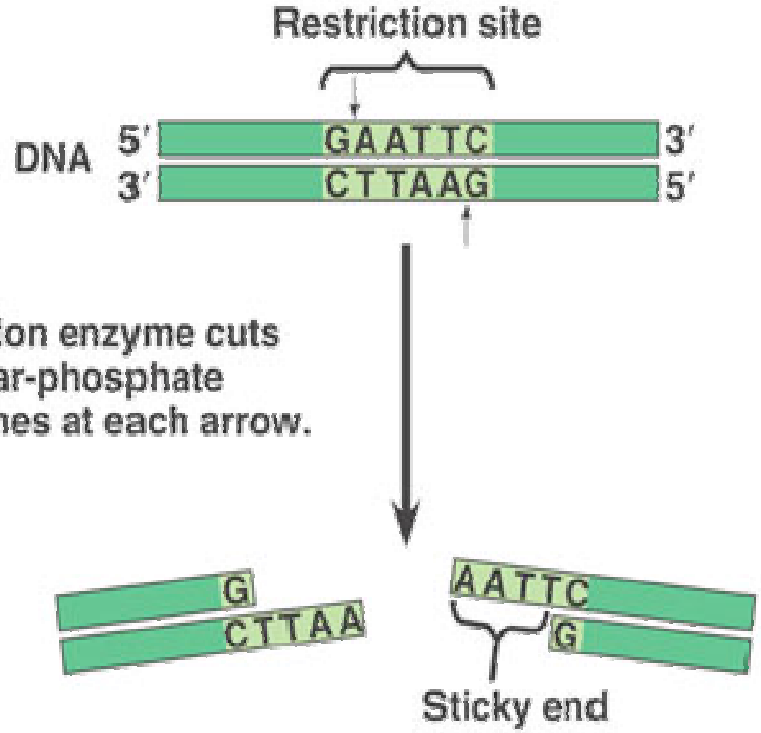
To date well over 3000 diff REs to choose from, which were found from screening >10000 bacteria

Hamilton Smith
1973 Nobel Prize

Cutting DNA

- Restriction endonucleases (restriction enzymes)
  - sticky ends
  - blunt ends

- Nomenclature
  - *Eco*RI
  - *E* = genus (*Escherichia*)
  - *co* = species (*coli*)
  - R = strain
  - I = # of enzyme



VirginiaTech
Invent the Future

Some restriction enzymes

Table 8.1 Properties of Some Restriction Enzymes

Enzyme	Bacterial Source	Restriction Site*
*Bam*HI	*Bacillus amyloliquefaciens* H	G↓GATCC CCTAG↑G
*Eco*RI	*Escherichia coli* RY13	G↓AATTC CTTAA↑G
*Eco*RII	*E. coli* R245	CC↓GG GG↑CC
*Hin*dII	*Haemophilus influenzae* Rd	GTPy↓PuAC CAPu↑PyTG
*Hin*dIII	*H. influenzae* Rd	A↓AGCTT TTCGA↑A
*Hin*fI	*H. influenzae* Rf	G↓ANTC CTNA↑G
*Hpa*I	*H. parainfluenzae*	GTT↓AAC CAA↑TTG
*Msp*I	*Moraxella* sp.	CC↓GG GG↑CC
*Sma*I	*Serratia marcescens*	CCC↓GGG GGG↑CCC

*Arrows indicate sites of cleavage; Py = pyrimidine (either T or C); Pu = purine (either A or G); N = any nucleotide (A, T, G, or C).

- 100's of restriction enzymes (RE's) are commercially available
- Artificial RE sites can be inserted at ends of any gene
- RE's and ligase allow precise cutting and pasting of any DNA sequences

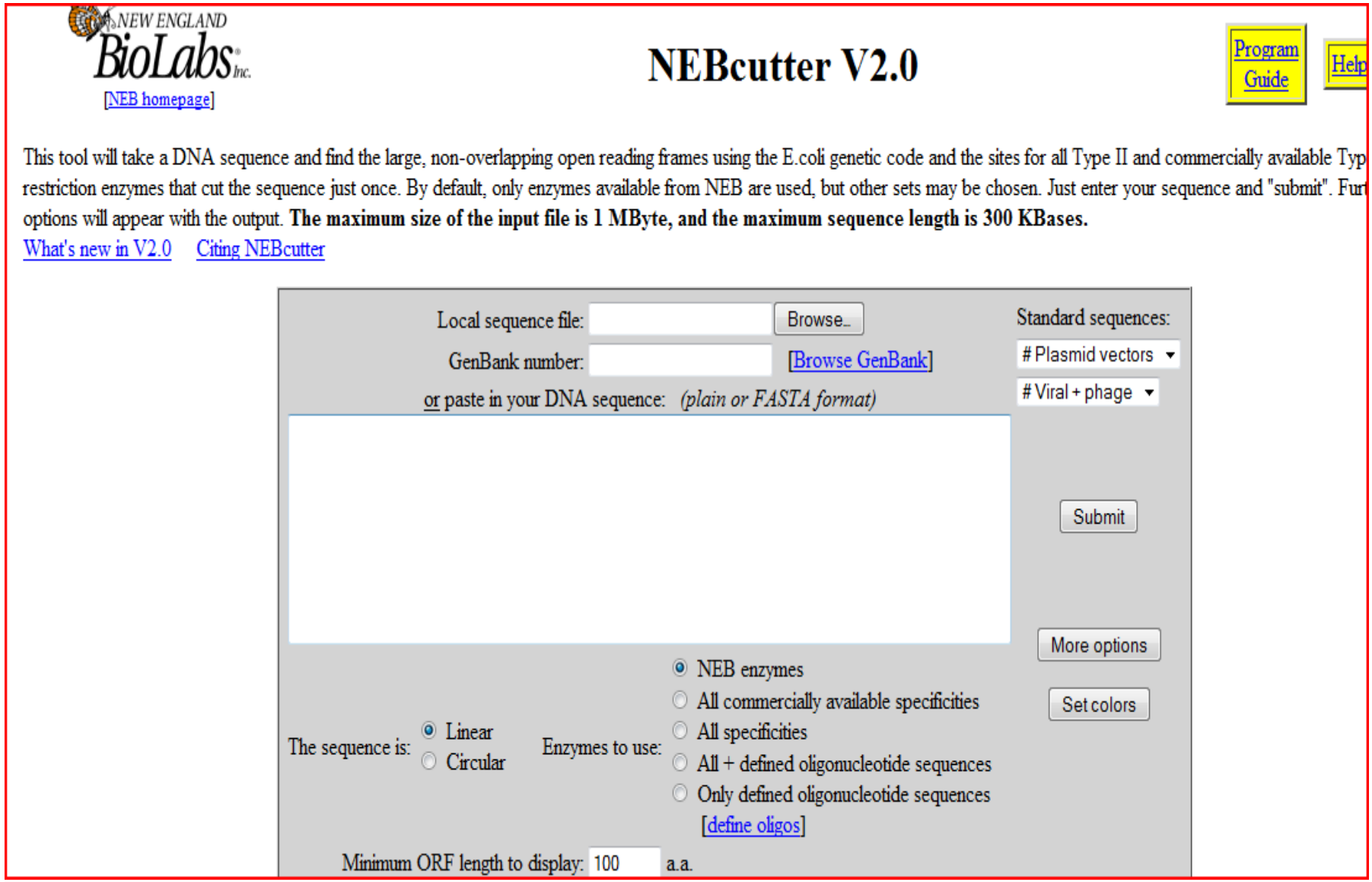
Which enzymes I should use

Pubmed

http://www.ncbi.nlm.nih.gov

Omp25 *Brucella*

```
     gene            1..642
                     /gene="omp25"
     CDS             1..642
                     /gene="omp25"
                     /codon_start=1
                     /product="outer membrane protein 25"
                     /protein_id="CAQ68393.1"
                     /db_xref="GI:206596726"
                     /db_xref="GOA:B5U6Y0"
                     /db_xref="InterPro:IPR011250"
                     /db_xref="UniProtKB/TrEMBL:B5U6Y0"
                     /translation="MRTLKSLVIVSAALLPFSATAFAADAIQEQPPVPAPVEVAPQYS
                     WAGGYTGLYLGYGWNKAKTSTVGSIKPDDWKAGAFAGWNFQQDQIVYGVEGDAGYSWA
                     KKSKDGLEVKQGFEGSLRARVGYDLNPVMPYLTAGIAGSQIKLNNGLDDESKFRVGWT
                     AGAGLEAKLTDNILGRVEYRYTQYGNKNYDLAGTNVRNKLDTQDIRVGIGYKF"
ORIGIN
        1 atgcgcactc ttaagtctct cgtaatcgtc tcggctgcgc tgctgccgtt ctctgcgacc
       61 gcttttgctg ccgacgccat ccaggaacag cctccggttc cggctccggt tgaagtagct
      121 ccccagtata gctgggctgg tggctatacc ggtctttacc tcggctacgg ctggaacaag
      181 gccaagacca gcaccgttgg cagcatcaag cctgacgatt ggaaggctgg cgcttttgct
      241 ggctggaact tccagcagga ccagatcgta tatggtgttg aaggtgatgc aggttattcc
      301 tgggccaaga agtccaagga cggcctggaa gtcaagcagg gctttgaagg ctcgctgcgt
      361 gcccgcgtcg gctacgacct gaacccggtt atgccgtacc tcacggctgg tattgccggt
      421 tcgcagatca agcttaacaa cggcttggac gacgaaagca agttccgcgt gggttggacg
      481 gctggtgccg gtctcgaagc caagctgacg gacaacatcc tcggccgcgt tgagtaccgt
      541 tacacccagt acggcaacaa gaactacgat ctggccggta cgaatgtccg caacaagctg
      601 gacacgcagg atatccgcgt cggcatcggc tacaagttct aa
//
```


NEW ENGLAND
BioLabs Inc.
[NEB homepage]
NEBcutter V2.0
Program Guide
Help
This tool will take a DNA sequence and find the large, non-overlapping open reading frames using the E.coli genetic code and the sites for all Type II and commercially available Typ
restriction enzymes that cut the sequence just once. By default, only enzymes available from NEB are used, but other sets may be chosen. Just enter your sequence and "submit". Furt
options will appear with the output. The maximum size of the input file is 1 MByte, and the maximum sequence length is 300 KBases.
What's new in V2.0
Citing NEBcutter
Local sequence file:
Browse...
GenBank number:
[Browse GenBank]
or paste in your DNA sequence: (plain or FASTA format)
Standard sequences:
Plasmid vectors
Viral + phage
Submit
More options
Set colors
The sequence is:
Linear
Circular
Enzymes to use:
NEB enzymes
All commercially available specificities
All specificities
All + defined oligonucleotide sequences
Only defined oligonucleotide sequences
[define oligos]
Minimum ORF length to display: 100 a.a.

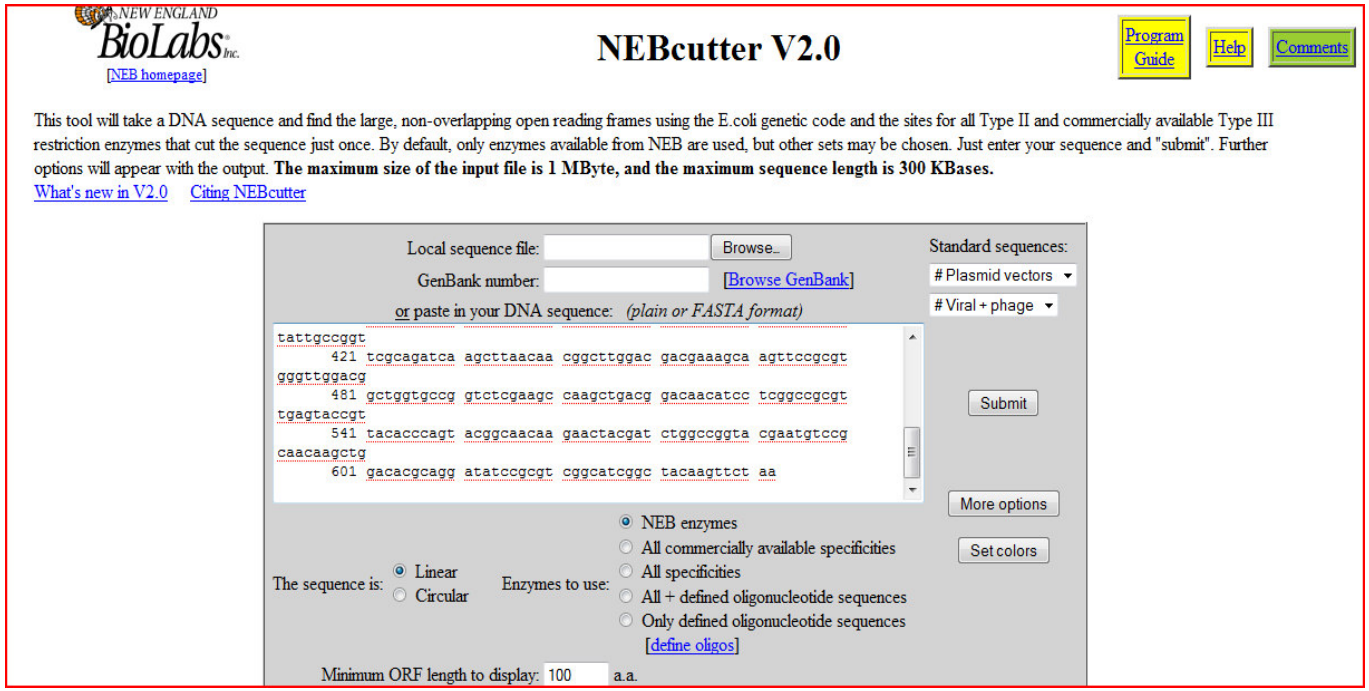
VirginiaTech
Invent the Future

Copy and paste sequence

copy

paste

```
                     /db_xref="GOA:B5U6Y0"
                     /db_xref="InterPro:IPR011250"
                     /db_xref="UniProtKB/TrEMBL:B5U6Y0"
                     /translation="MRTLKSLVIVSAALLPFSATAFAADAIQEQPPVPAPVEVAPQYS
                     WAGGYTGLYLGYGWNKAKTSTVGSIKPDDWKAGAFAGWNFQQDQIVYGVEGDAGYSWA
                     KKSKDGLEVKQGFEGSLRARVGYDLNPVMPYLTAGIAGSQIKLNNGLDDESKFRVGWT
                     AGAGLEAKLTDNILGRVEYRYTQYGNKNYDLAGTNVRNKLDTQDIRVGIGYKF"
ORIGIN
        1 atgcgcactc ttaagtctct cgtaatcgtc tcggctgcgc tgctgccgtt ctctgcgacc
       61 gcttttgctg ccgacgccat ccaggaacag cctccggttc cggctccggt tgaagtagct
      121 ccccagtata gctgggctgg tggctatacc ggtctttacc tcggctacgg ctggaacaag
      181 gccaagacca gcaccgttgg cagcatcaag cctgacgatt ggaaggctgg cgcttttgct
      241 ggctggaact tccagcagga ccagatcgta tatggtgttg aaggtgatgc aggttattcc
      301 tgggccaaga agtccaagga cggcctggaa gtcaagcagg gctttgaagg ctcgctgcgt
      361 gcccgcgtcg gctacgacct gaacccggtt atgccgtacc tcacggctgg tattgccggt
      421 tcgcagatca agcttaacaa cggcttggac gacgaaagca agttccgcgt gggttggacg
      481 gctggtgccg gtctcgaagc caagctgacg gacaacatcc tcggccgcgt tgagtaccgt
      541 tacacccagt acggcaacaa gaactacgat ctggccggta cgaatgtccg caacaagctg
      601 gacacgcagg atatccgcgt cggcatcggc tacaagttct aa
//
```

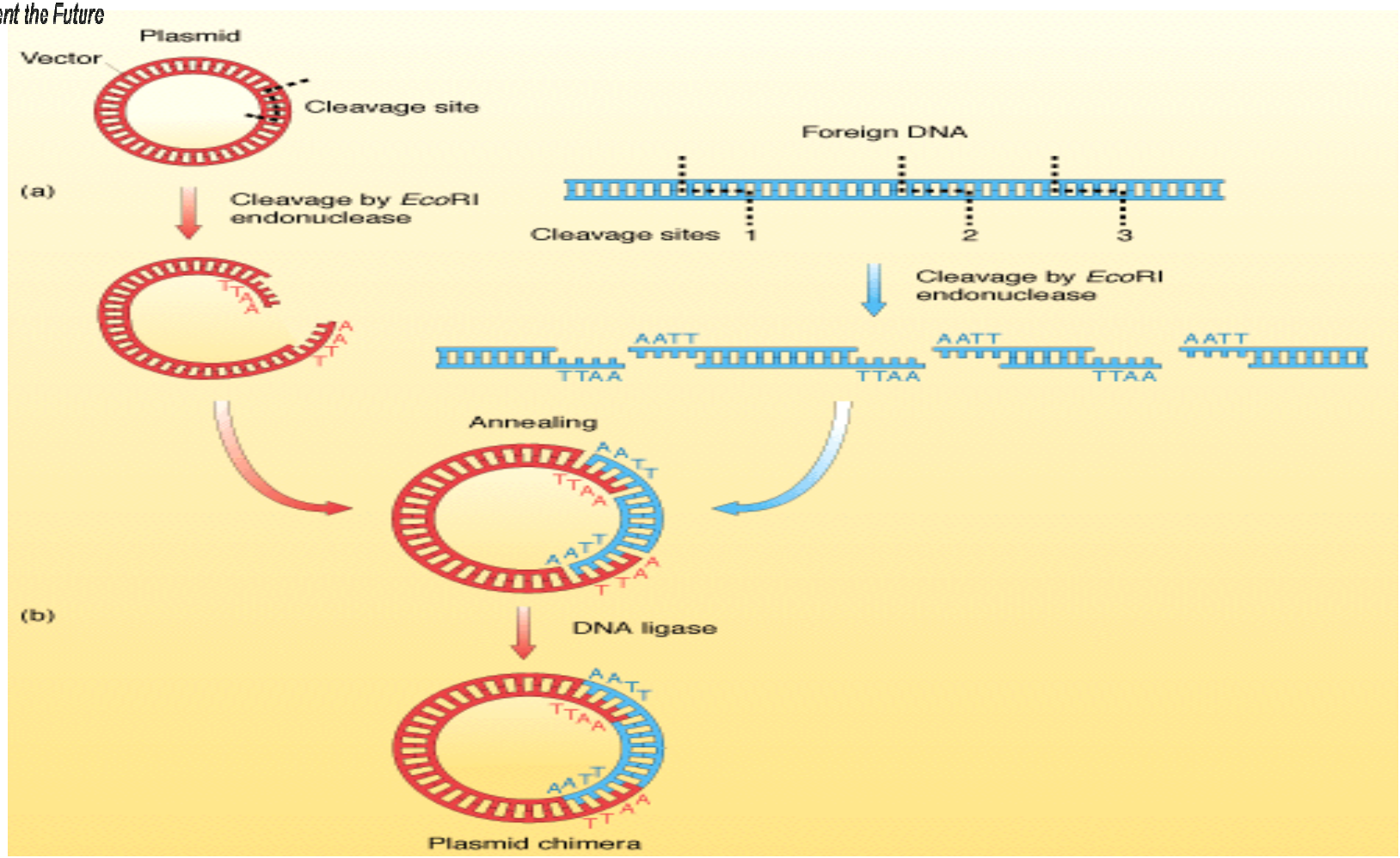


Virginia Tech
Invent the Future

Restriction Map

Noncutters: AarI, AatII, AbsI, AccI, AclI, AflIII, AjuI, AlfI, AloI, AlwNI, ApaI, ApaLI, ApoI, ArsI, AscI, AsuII, AvaI, AvrII, BalI, BamHI, BarI, BbvCI, BcgI, BciVI, BclI, BdaI, BglI, BglII, BplI, Bpu10I, BsaAI, BsaBI, BsaXI, BseMII, BsePI, BseRI, BsgI, BsmI, Bsp1407I, BspHI, BsrBI, BsrDI, BstEII, BstXI, BtgZI, BtrI, BtsI, ClaI, CspCI, DraII, DraIII, DrdI, Eam1105I, EciI, Eco47III, Eco57I, Eco57MI, EcoNI, EcoRI, FseI, GsuI, HaeIV, Hin4I, HindII, HpaI, KpnI, MauBI, MboII, MfeI, MluI, MslI, NaeI, NarI, NcoI, NdeI, NheI, NotI, NruI, NspI, OliI, PacI, PasI, PflMI, PleI, PmaCI, PmeI, PpiI, PpuMI, PshAI, PsiI, PI-PspI, PstI, PvuI, PvuII, RsrII, SacI, SacII, SalI, SanDI, SapI, ScaI, PI-SceI, SexAI, SfiI, SgfI, SgrAI, SgrDI, SmaI, SnaBI, SpeI, SphI, SrfI, Sse8387I, SspI, StuI, SwaI, TaqII, TatI, TfiI, Tsp45I, TspDTI, TspRI, Tth111I, VspI, XbaI, XhoI, XhoII, XmnI

Name	Sequence	Site Length	Overhang	Frequency	Cut Positions
EcoRV	GATATC	6	blunt	1	612
FspAI	RTGCGCAY	8	blunt	1	4
AcyI	GRCGYC	6	five_prime	1	74
AflII	CTTAAG	6	five_prime	1	10
AgeI	ACCGGT	6	five_prime	1	148
AvaII	GGWCC	5	five_prime	1	258
BccI	CCATC	5	five_prime	1	85
BseYI	CCCAGC	6	five_prime	1	131
BspMI	ACCTGC	6	five_prime	1	280
Eco31I	GGTCTC	6	five_prime	1	496
Esp3I	CGTCTC	6	five_prime	1	33
FauI	CCCGC	5	five_prime	1	370


Plasmid
Vector
Cleavage site
(a)
Cleavage by EcoRI endonuclease
Foreign DNA
Cleavage sites
1
2
3
Cleavage by EcoRI endonuclease
AATT
TTAA
Annealing
(b)
DNA ligase
Plasmid chimera


BamHI
G GATCC
CCTAG G
BamHI
G
GATCC
CCTAG
G
Sticky ends

VirginiaTech
Invent the Future

GGATCC ATGCGCACTCTTAAGTCTCTCGTAATCGTCTCGGCTGCGCTGCTGCCGTCCAGCACCGTTACAAGTTCTAAGGATCC
TACGCGTGAGAATTCAGAGAGCATTAGCAGAGCCGACGCGACGACGGCAGGTCGTGGCAATGTTCAAGATTCCTAGG

Put sequence of Non cutting enzymes at both ends of the gene

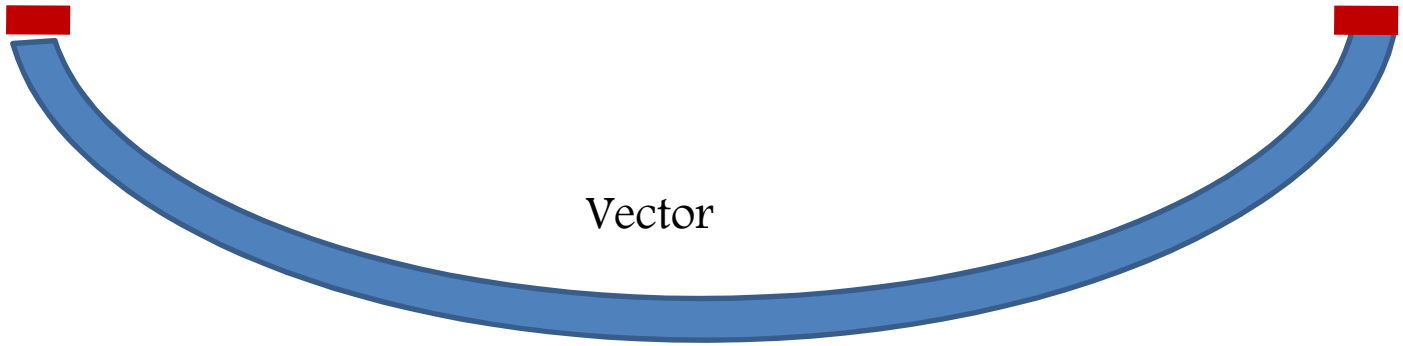
How? PCR and Primers

VirginiaTech
Invent the Future

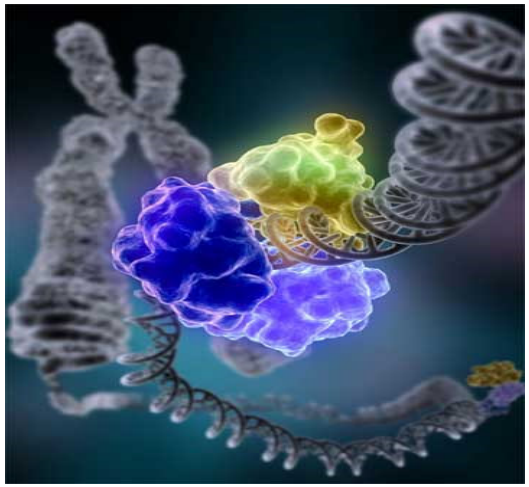
Gene

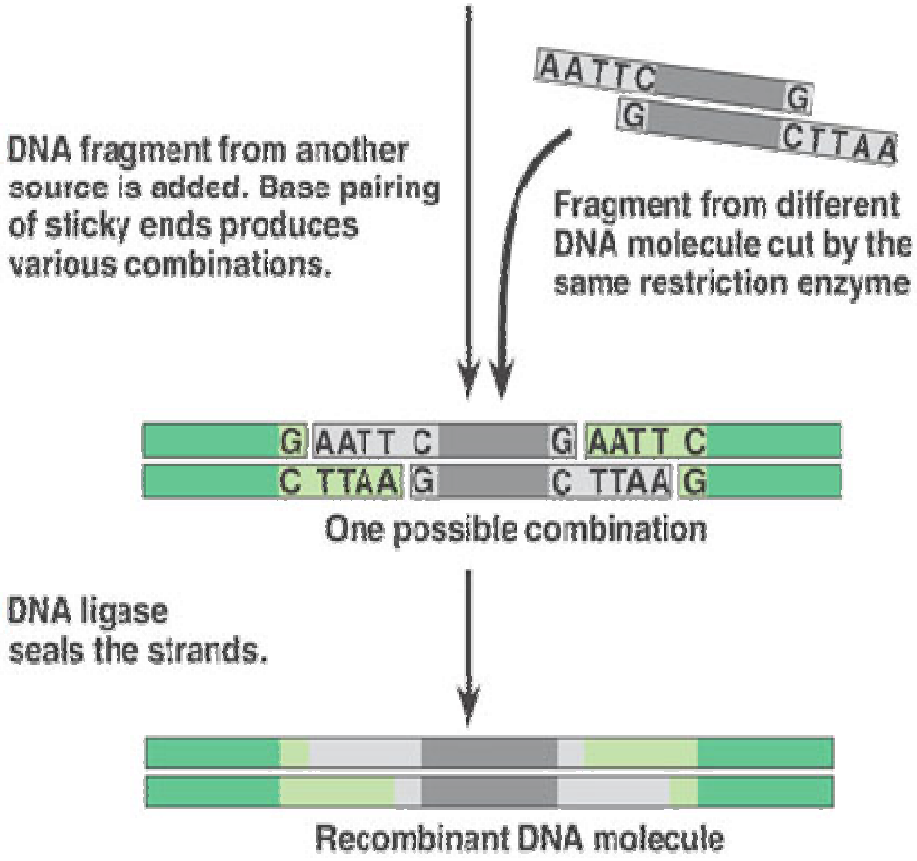
Adding sequence to the ends by PCR

linkers



Pasting DNA





- Complementary ends (sticky ends) H-bond
- Ligase forms phosphodiester bond to seal strands together.

How to chose your enzymes

- Non cutting enzymes
- Enzymes on Vector
- Enzymes that work together (Double digest)
- Price
- Company (I highly recommend Promega)

Don't waste clean thinking on dirty enzymes

Efraim Racker, Cornell University

How use your enzymes

- Ice
- No Vortex
- Set up reaction of 10 µl at least
- 1 µl (10 units) enzyme 1 µl 10xbuffer and 8 µl DNA(you need 1 unit to cut 1 µg of DNA)
- Incubate (usually 37 °C for at least 1 hour)
- Clean your DNA after cutting

Enzyme and buffer

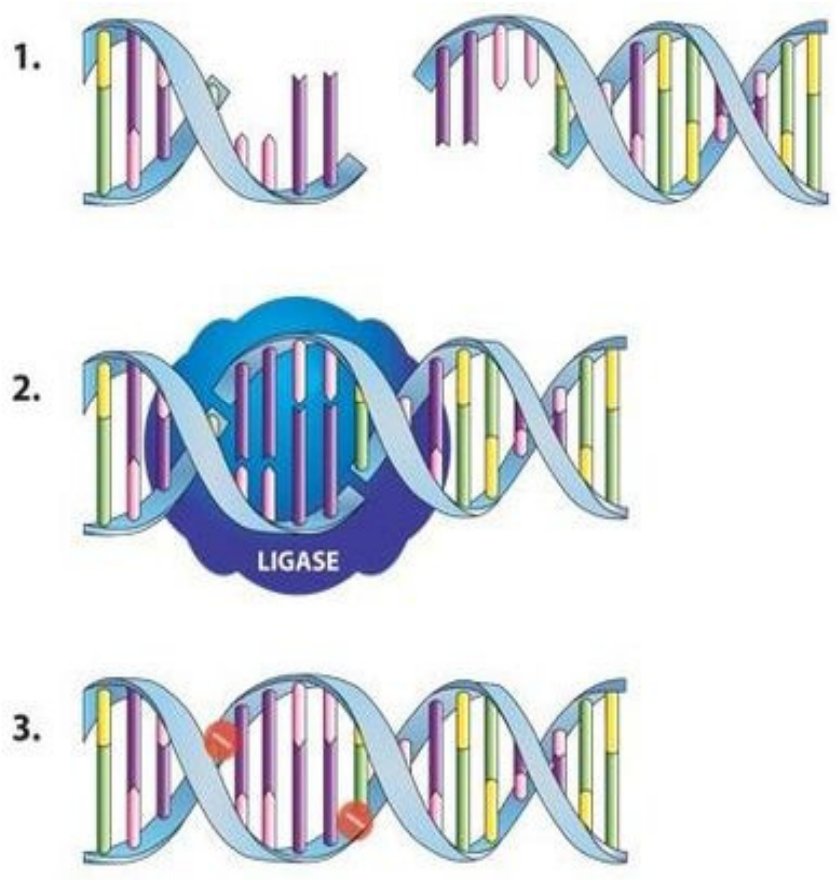
VirginiaTech
Invent the Future

Joining DNA

Ligase enzyme

anneal two pieces of compatible ends of DNA together

Repair Mechanism in biological system

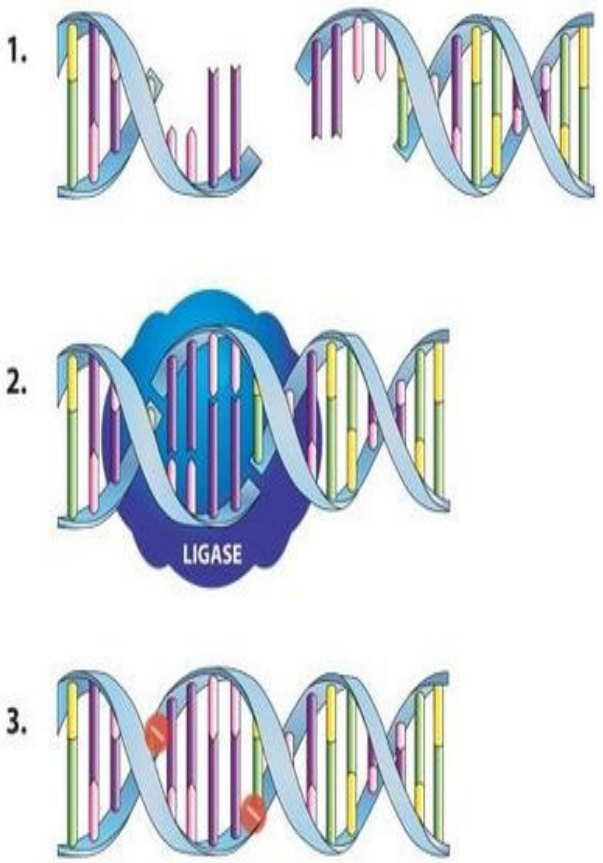


VirginiaTech
Invent the Future

How to join DNA

Get rid of restriction enzymes

1 µl (3 units) enzyme
1 µl 10xbuffer
4 µl gene
4 µl vector (plasmid)



Incubation ligation reaction

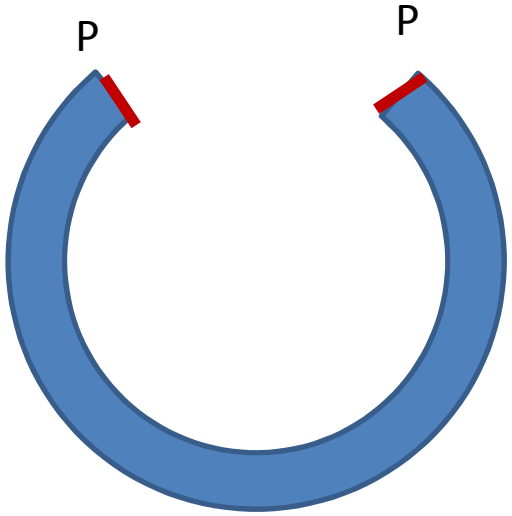
- 4 °C over night or 16 °C over night
- 25 °C 3 hours
- TCL (temperature cycle ligations) Lund et al., 1996

  99 cycles (1 min 30 ° C 1 min 10 ° C) 4 fold increase ligation
- 5 minutes ligation buffer (Polyethylene glycol PEG 6000)

How to check for successful ligation reaction

- Transformation and count number of good colonies

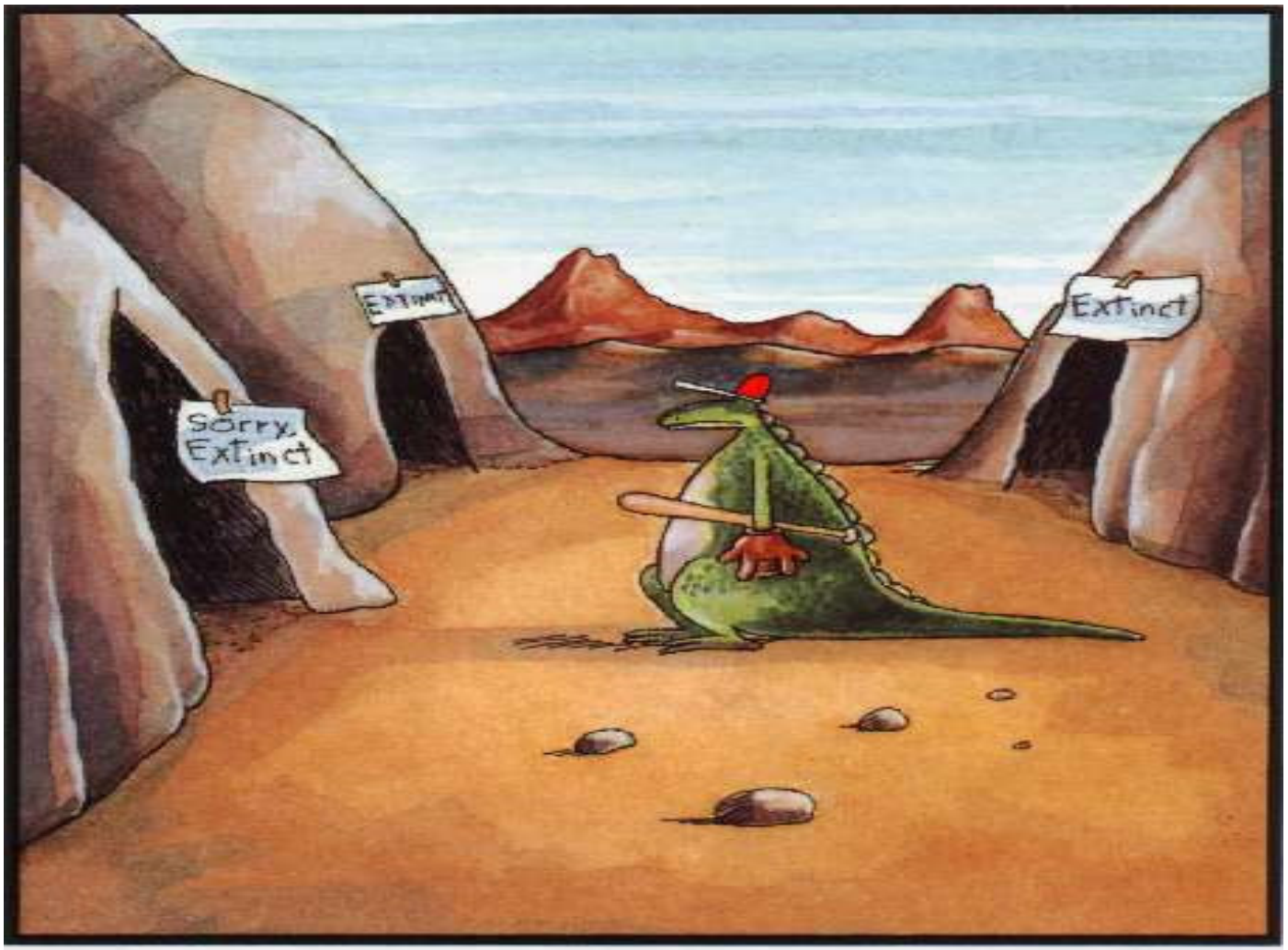
Alkaline phosphatase
dephosphorylation



Removing phosphate group to prevent self ligation of the vector

Add it only to the vector

Questions!!!!



Suddenly, Bobby felt very alone in the world.