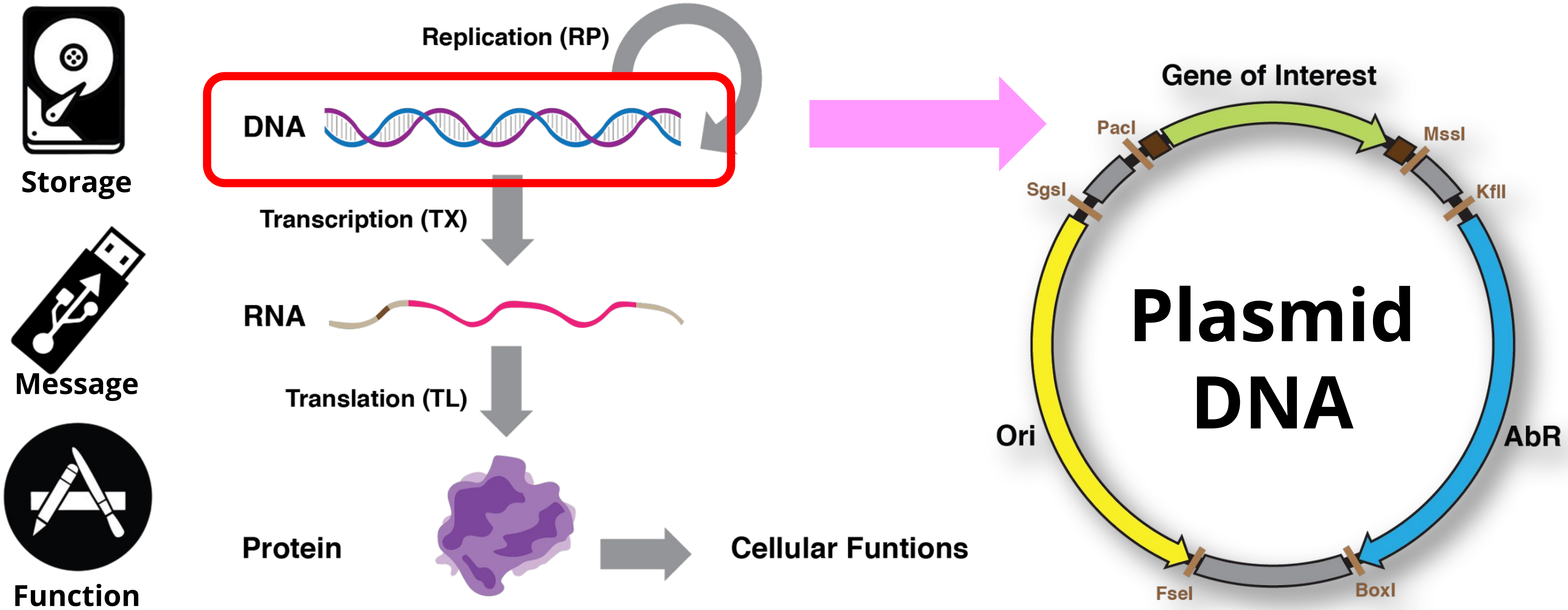
DNA Languages: Reading and Editing

Cholpisit (**Ice**) Kiattisewee
HTGAA Bootcamp
January 23rd, 2026

Overview of the session

- Why do we need specific molecular biology tools? Can't we just use **Notepad, Microsoft Word** or **Google Docs**?
- Show how to use simple **Benchling** Operation
  - Attaching Primers --- find existing primers and pair primers
  - In silico PCR Fragments
- Navigate Simple Assembly Wizard in Benchling
  - Restriction-Ligation (Cut-Connect) of colored proteins
- Discuss some Scenarios while cloning a new plasmid along the way

Most biological information were stored in the form of DNA



Storage

Message

Function

Replication (RP)

DNA

Transcription (TX)

RNA

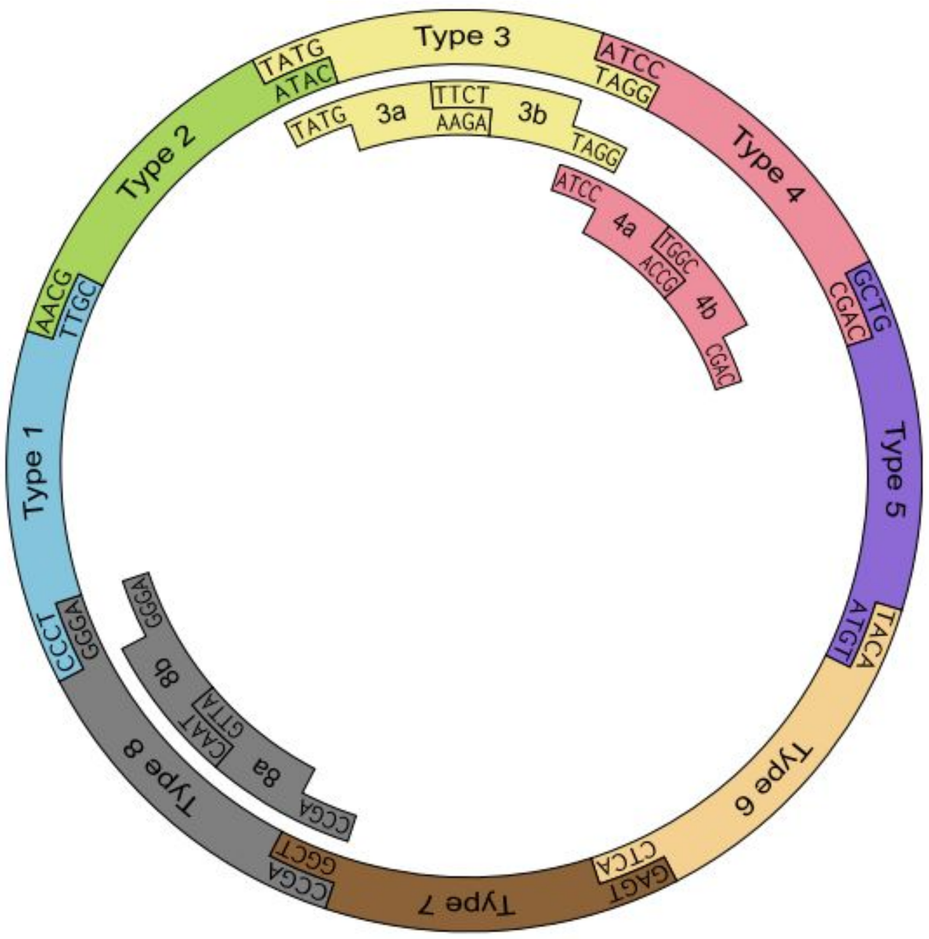
Translation (TL)

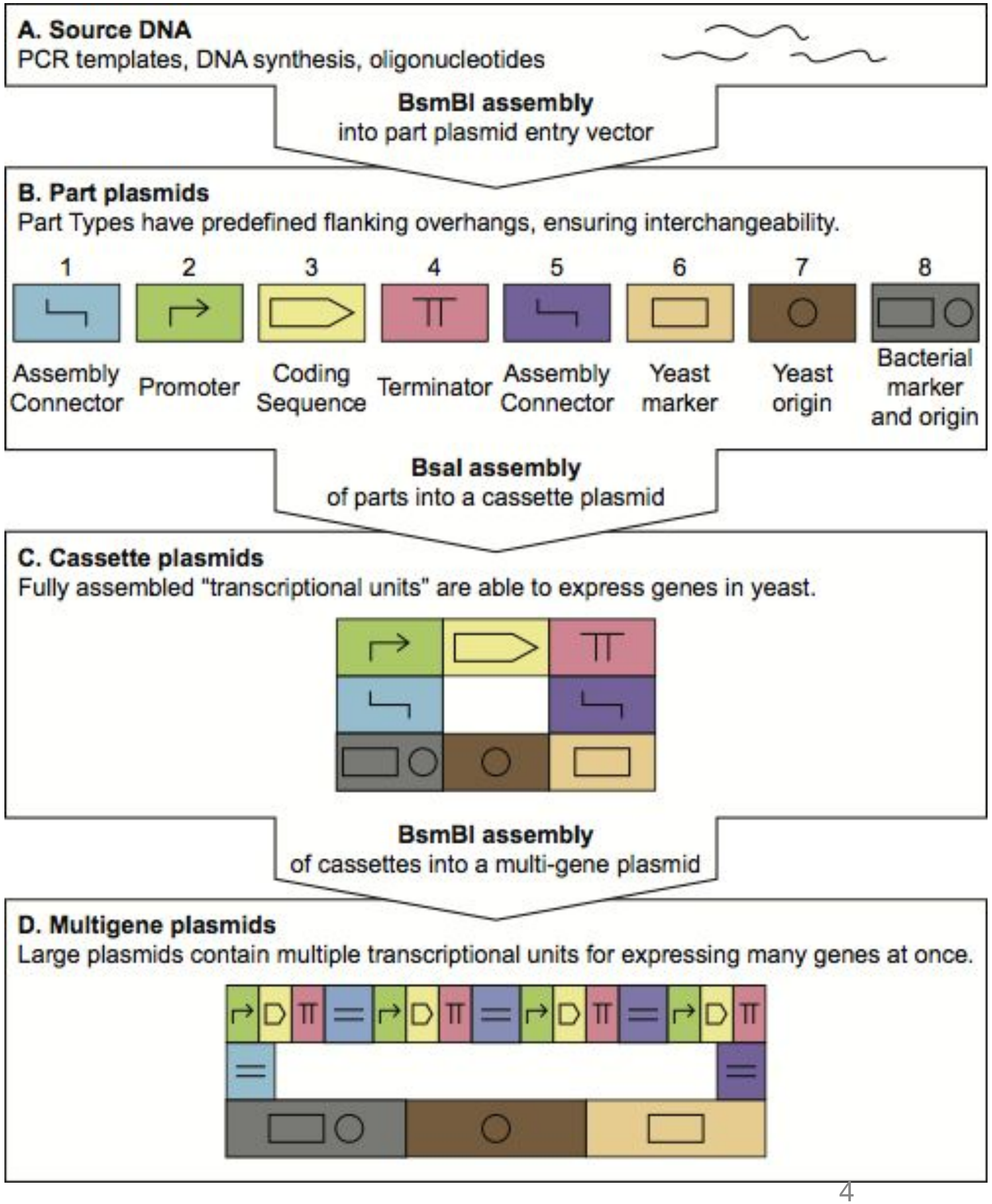
Protein

Cellular Funtions

Plasmid DNA

DNA as LEGO pieces





A. Source DNA
PCR templates, DNA synthesis, oligonucleotides

BsmBI assembly
into part plasmid entry vector

B. Part plasmids
Part Types have predefined flanking overhangs, ensuring interchangeability.

BsaI assembly
of parts into a cassette plasmid

C. Cassette plasmids
Fully assembled "transcriptional units" are able to express genes in yeast.

BsmBI assembly
of cassettes into a multi-gene plasmid

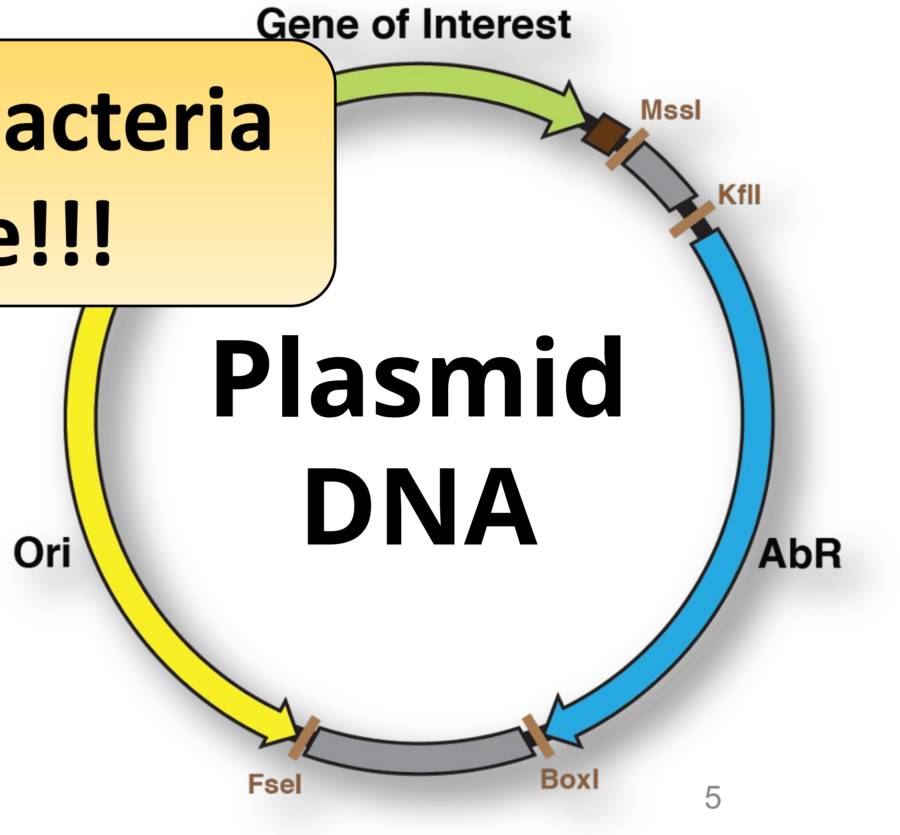
D. Multigene plasmids
Large plasmids contain multiple transcriptional units for expressing many genes at once.

Lee et al. *ACS Synth. Biol.* 2015, 4, 9, 975–986

General Plasmid Architectures

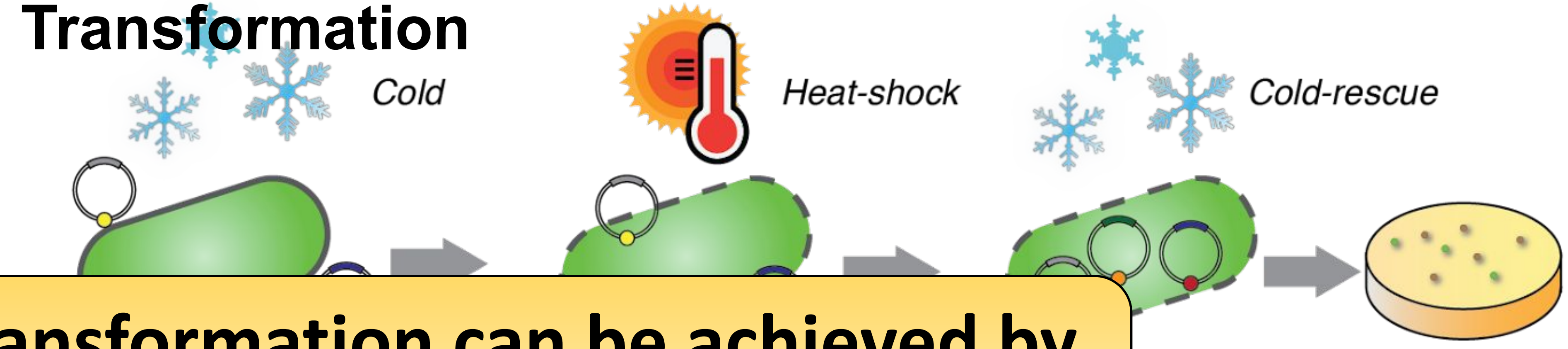
- At their most basic level, plasmids are small circular pieces of DNA that replicate independently from the host's chromosomal DNA
- General eleme
  - Origin of repli
    - And its replic
  - Homologous region (integrative)
  - Selectable marker
  - Gene of interests
    - Multiple Cloning Sites (MCS)

Putting plasmid into bacteria is relatively simple!!!



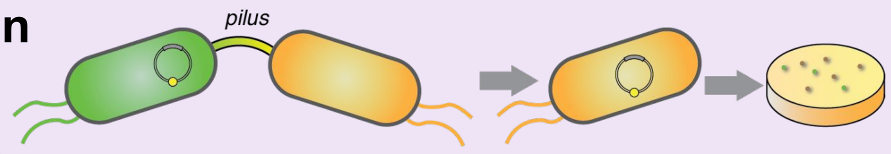
Putting DNA into bacteria (or yeast)

Chemical Transformation

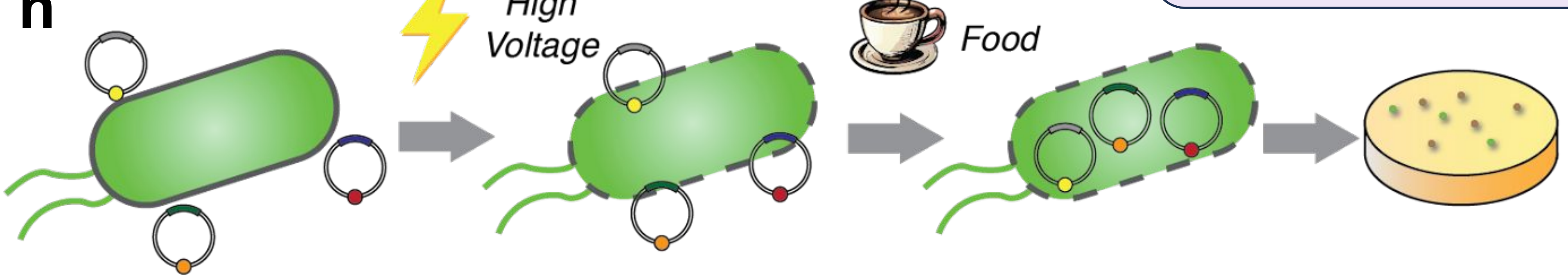


Transformation can be achieved by heat-shock or electric-shock

Conjugation

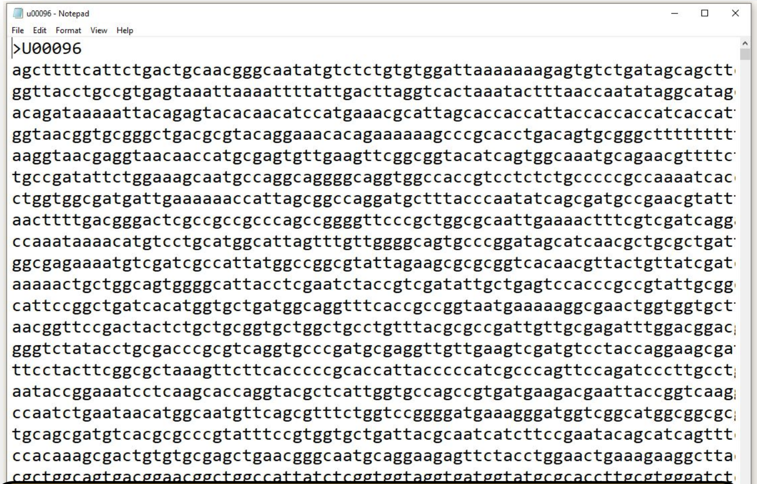


n



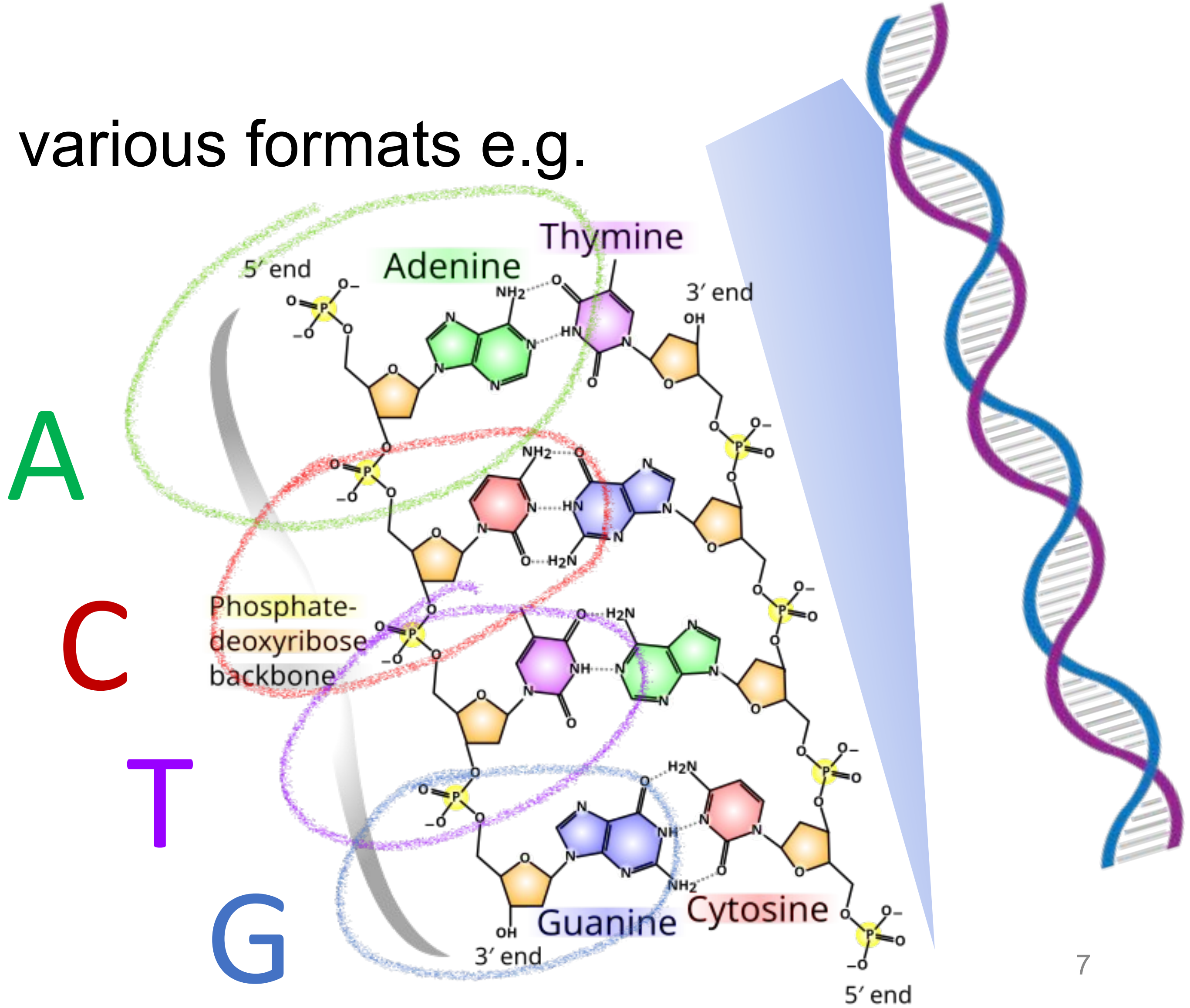
Visualizing DNA Features

- DNA data can be stored in various formats e.g.
    - .fasta
    - .genbank or .gb or .gff
    - .dna (Snapgene)



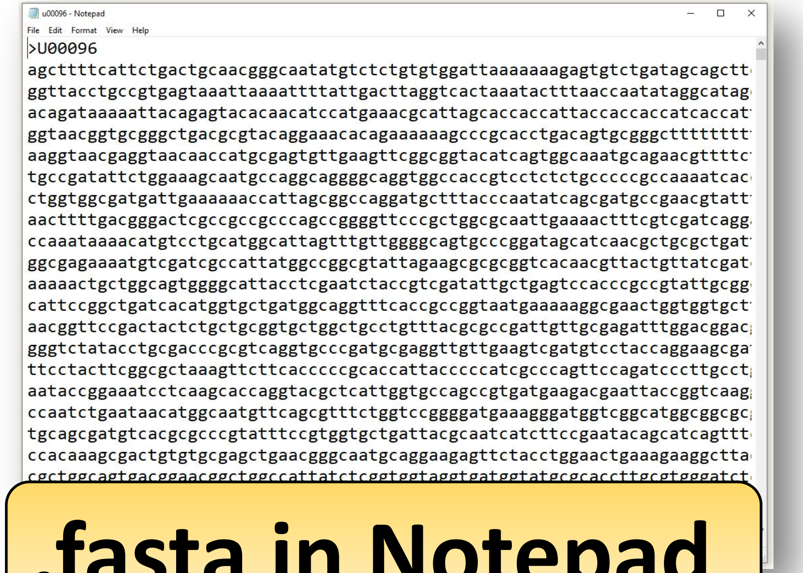
```
>U00096
agcttttcattctgactgcaacgggcaatatgtctctgtgtggattaaaaaaagagtgtctgatagcagctt
ggttacctgccgtgagtaaattaaaattttattgacttaggtcactaaatactttaaccaatataggcatag
acagataaaaattacagagtacacaacatccatgaaacgcattagcaccaccattaccaccaccatcaccat
ggtaacggtgcgggctgacgcgtacaggaaacacagaaaaaagcccgcacctgacagtgcgggcttttttttt
aaggtaacgaggtaacaaccatgcgagtgttgaagttcggcggtacatcagtggcaaatgcagaacgttttc
tgccgatattctggaaagcaatgccaggcaggggcaggtggccaccgtcctctctgcccccgccaaaatcac
ctggtggcgatgattgaaaaaaccattagcggccaggatgctttacccaatatcagcgatgccgaacgtatt
aacttttgacgggactcgccgccgcccagccggggttcccgctggcgcaattgaaaactttcgtcgatcagg
ccaaataaaacatgtcctgcatggcattagtttgttggggcagtgcccggatagcatcaacgctgcgctgat
ggcgagaaaatgtcgatcgccattatggccggcgtattagaagcgcgcggtcacaacgttactgttatcgat
aaaaactgctggcagtggggcattacctcgaatctaccgtcgatattgctgagtccacccgccgtattgcgg
cattccggctgatcacatggtgctgatggcaggtttcaccgccggtaatgaaaaaggcgaactggtggtgct
aacggttccgactactctgctgcggtgctggctgcctgtttacgcgccgattgttgcgagatttggacggac
gggtctatacctgcgacccgcgtcaggtgcccgatgcgaggttgttgaagtcgatgtcctaccaggaagcga
ttcctacttcggcgctaaagttcttcacccccgcaccattaccccatcgcccagttccagatcccttgcct
aataccggaaatcctcaagcaccaggtacgctcattggtgccagccgtgatgaagacgaattaccggtcaag
ccaatctgaataacatggcaatgttcagcgtttctggtccggggatgaaagggatggtcggcatggcggcgc
tgcagcgatgtcacgcgcccgtatttccgtggtgctgattacgcaatcatcttccgaatacagcatcagttt
ccacaaagcgactgtgtgcgagctgaacgggcaatgcaggaagagttctacctggaactgaaagaaggctta
cgctggcagtgacggaacggctggccattatctcggtggtaggtgatggtatgcgcaccttgcgtgggatct
```

.fasta in Notepad



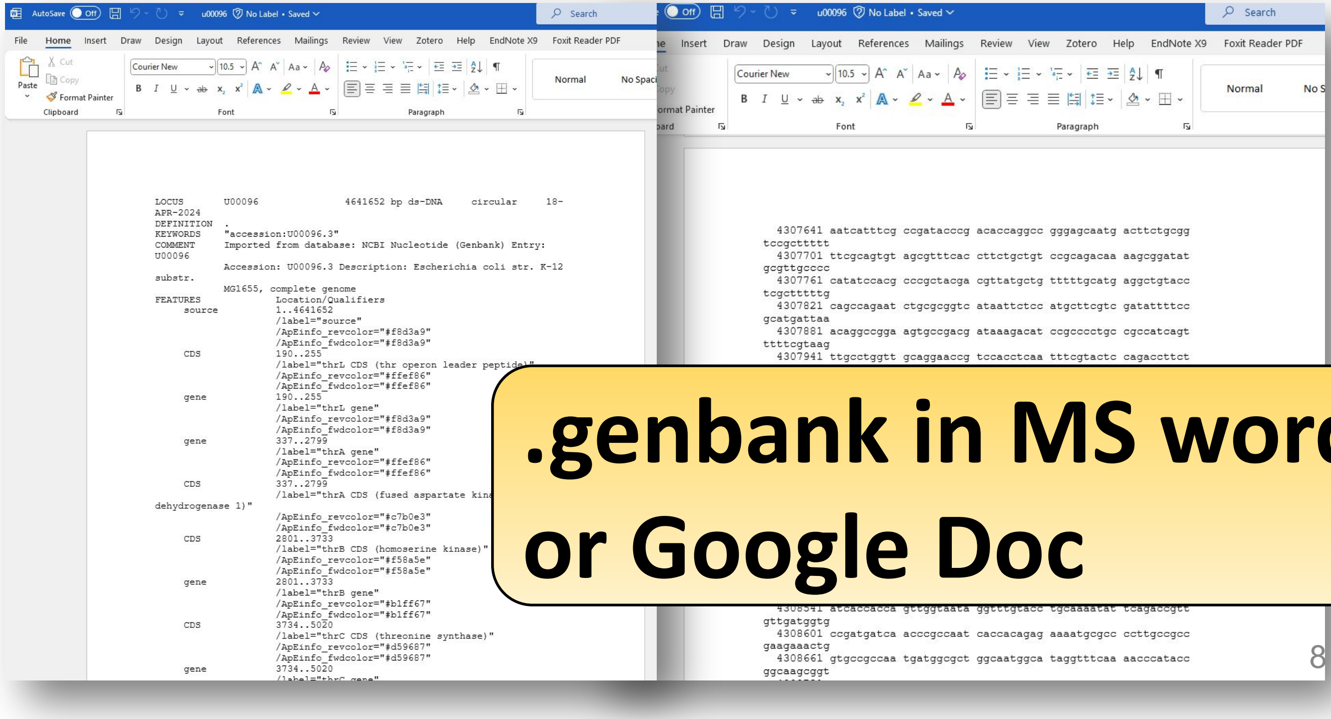
Visualizing DNA Features

- DNA data can be stored in various formats e.g.
  - .fasta
  - .genbank or .gb or .gff
  - .dna (Snapgene)



```
>U00096
agcttttcattctgactgcaacgggcaatatgtctctgtgtggattaaaaaaagagtgtctgatagcagctt
ggttacctgccgtgagtaaattaaaattttattgacttaggtcactaaatactttaaccaatataggcatag
acagataaaaattacagagtacacaacatccatgaaacgcattagcaccaccattaccaccaccatcaccat
ggtaacggtgcgggctgacgcgtacaggaaacacagaaaaaagcccgcacctgacagtgcgggcttttttt
aaggtaacgaggtaacaaccatgcgagtgttgaagttcggcggtacatcagtggcaaatgcagaacgtttc
tgccgatattctggaaagcaatgccaggcaggggcaggtggccaccgtcctctctgcccccgccaaaatcac
ctggtggcgatgattgaaaaaaccattagcggccaggatgctttacccaatatcagcgatgccgaacgtatt
aacttttgacgggactcgccgccgcccagccggggttcccgctggcgcaattgaaaactttcgtcgatcagg
ccaaataaaacatgtcctgcatggcattagtttgttggggcagtgcccggatagcatcaacgctgcgctgat
ggcgagaaaatgtcgatcgccattatggccggcgtattagaagcgcgcggtcacaacgttactgttatcgat
aaaaactgctggcagtggggcattacctcgaatctaccgtcgatattgctgagtccacccgccgtattgcgg
cattccggctgatcacatggtgctgatggcaggtttcaccgccggtaatgaaaaaggcgaactggtggtgct
aacggttccgactactctgctgcggtgctggctgcctgtttacgcgccgattgttgcgagatttggacggac
gggtctatacctgcgacccgcgtcaggtgcccgatgcgaggttgttgaagtcgatgtcctaccaggaagcga
ttcctacttcggcgctaaagttcttcacccccgcaccattaccccatcgcccagttccagatcccttgcct
aataccggaaatcctcaagcaccaggtacgctcattggtgccagccgtgatgaagacgaattaccggtcaag
ccaatctgaataacatggcaatgttcagcgtttctggtccggggatgaaagggatggtcggcatggcggcgc
tgcagcgatgtcacgcgcccgtatttccgtggtgctgattacgcaatcatcttccgaatacagcatcagttt
ccacaaagcgactgtgtgcgagctgaacgggcaatgcaggaagagttctacctggaactgaaagaaggctta
cgctggcagtgacggaacggctggccattatctcggtggtaggtgatggtatgcgcaccttgcgtgggatct
```

.fasta in Notepad


AutoSave Off
u00096 No Label • Saved
Search
File Home Insert Draw Design Layout References Mailings Review View Zotero Help EndNote X9 Foxit Reader PDF
Paste Cut Copy Format Painter
Clipboard
Courier New 10.5
Font
Paragraph
Normal No Spacing

```
LOCUS       U00096               4641652 bp ds-DNA     circular     18-
APR-2024
DEFINITION  .
KEYWORDS    "accession:U00096.3"
COMMENT     Imported from database: NCBI Nucleotide (Genbank) Entry:
U00096
            Accession: U00096.3 Description: Escherichia coli str. K-12
substr.
            MG1655, complete genome
FEATURES             Location/Qualifiers
     source          1..4641652
                     /label="source"
                     /ApEinfo_revcolor="#f8d3a9"
                     /ApEinfo_fwdcolor="#f8d3a9"
     CDS             190..255
                     /label="thrL CDS (thr operon leader peptide)"
                     /ApEinfo_revcolor="#ffef86"
                     /ApEinfo_fwdcolor="#ffef86"
     gene            190..255
                     /label="thrL gene"
                     /ApEinfo_revcolor="#f8d3a9"
                     /ApEinfo_fwdcolor="#f8d3a9"
     gene            337..2799
                     /label="thrA gene"
                     /ApEinfo_revcolor="#ffef86"
                     /ApEinfo_fwdcolor="#ffef86"
     CDS             337..2799
                     /label="thrA CDS (fused aspartate kina
dehydrogenase 1)"
                     /ApEinfo_revcolor="#c7b0e3"
                     /ApEinfo_fwdcolor="#c7b0e3"
     CDS             2801..3733
                     /label="thrB CDS (homoserine kinase)"
                     /ApEinfo_revcolor="#f58a5e"
                     /ApEinfo_fwdcolor="#f58a5e"
     gene            2801..3733
                     /label="thrB gene"
                     /ApEinfo_revcolor="#b1ff67"
                     /ApEinfo_fwdcolor="#b1ff67"
     CDS             3734..5020
                     /label="thrC CDS (threonine synthase)"
                     /ApEinfo_revcolor="#d59687"
                     /ApEinfo_fwdcolor="#d59687"
     gene            3734..5020
                     /label="thrC gene"
```

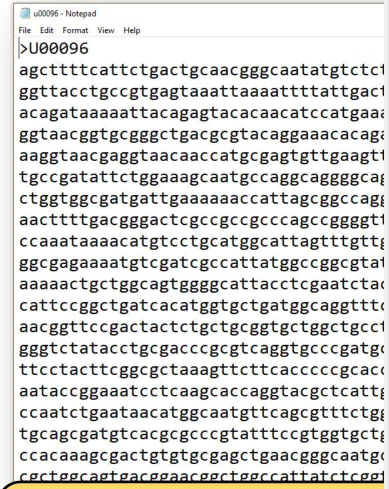
```
  4307641 aatcatttcg ccgataccg acaccaggcc gggagcaatg acttctgcgg
tccgcttttt
  4307701 ttcgcagtgt agcgtttcac cttctgctgt ccgcagacaa aagcggatat
gcgttgcccc
  4307761 catatccacg cccgctacga cgttatgctg tttttgcatg aggctgtacc
tcgctttttg
  4307821 cagccagaat ctgcgcggtc ataattctcc atgcttcgtc gatattttcc
gcatgattaa
  4307881 acaggccgga agtgccgacg ataaagacat ccgcccctgc cgccatcagt
ttttcgtaag
  4307941 ttgcctggtt gcaggaaccg tccacctcaa tttcgtactc cagaccttct
```

```
  4308541 atcaccacca gttggtaata ggtttgtacc tgcaaaatat tcagaccgtt
gttgatggtg
  4308601 ccgatgatca acccgccaat caccacagag aaaatgcgcc ccttgccgcc
gaagaaactg
  4308661 gtgccgccaa tgatggcgct ggcaatggca taggtttcaa aacccatacc
ggcaagcggt
```

.genbank in MS word or Google Doc

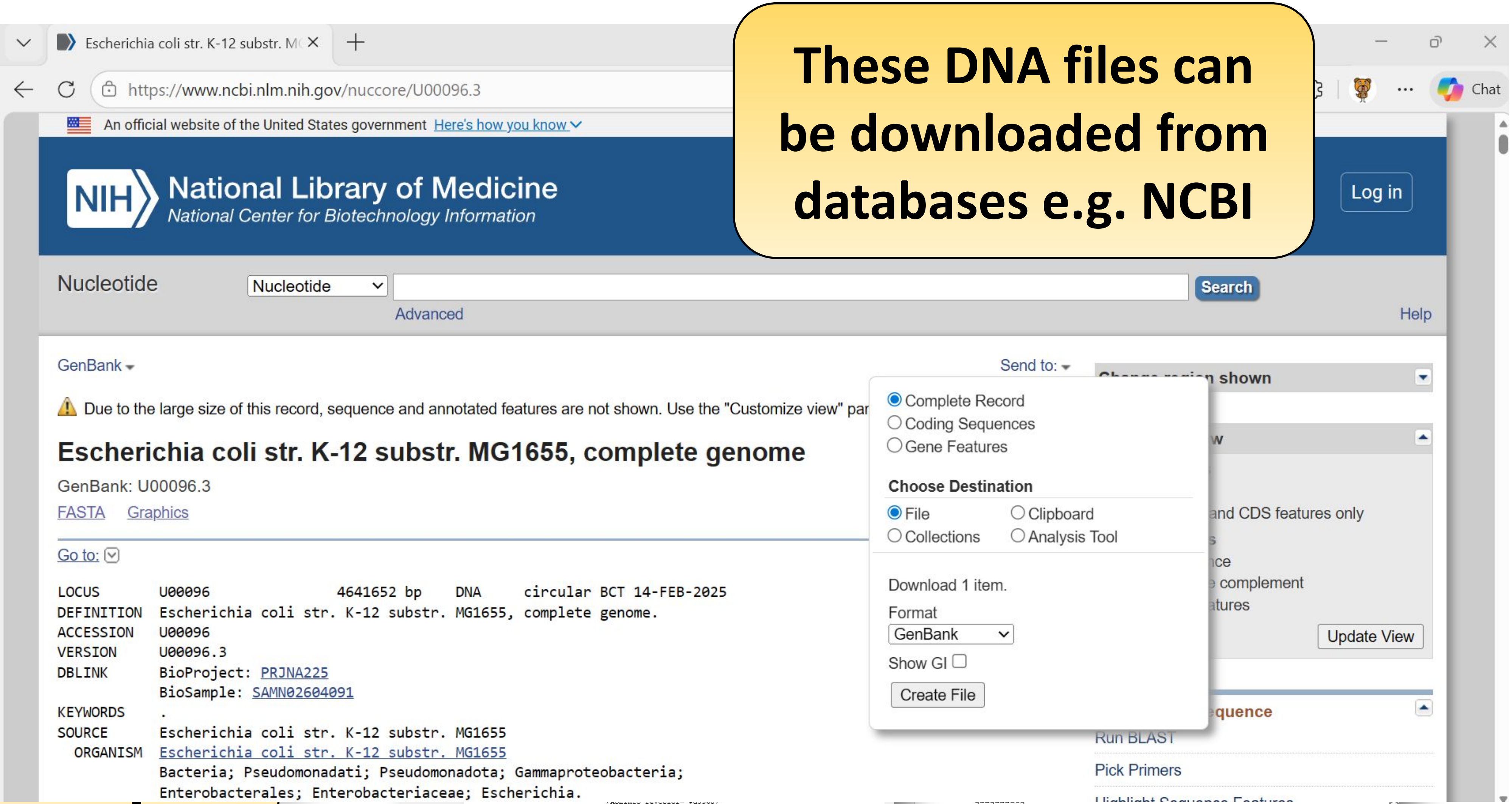
Visualizing DNA Features

- DNA dat
  - ▯ .fasta
  - ▯ .genba
  - ▯ .dna (S



```
>U00096
agcttttcattctgactgcaacgggcaatatgtctc
ggttacctgccgtgagtaaattaaaattttattgac
acagataaaaattacagagtacacaacatccatgaa
ggtaacggtgcgggctgacgcgtacaggaaacacag
aaggtaacgaggtaacaaccatgcgagtgttgaagt
tgccgatattctggaaagcaatgccaggcaggggca
ctggtggcgatgattgaaaaaaccattagcggccag
aacttttgacgggactcgccgccgcccagccggggt
ccaaataaaacatgtcctgcatggcattagtttgtt
ggcgagaaaatgtcgatcgccattatggccggcgta
aaaaactgctggcagtggggcattacctcgaatcta
cattccggctgatcacatggtgctgatggcaggttt
aacggttccgactactctgctgcggtgctggctgcc
gggtctatacctgcgacccgcgtcaggtgcccgatg
ttcctacttcggcgctaaagttcttcacccccgcac
aataccggaaatcctcaagcaccaggtacgctcatt
ccaatctgaataacatggcaatgttcagcgtttctg
tgcagcgatgtcacgcgcccgtatttccgtggtgct
ccacaaagcgactgtgtgcgagctgaacgggcaatg
cgctggcagtgacggaacggctggccattatctcgg
```

.fasta in N


Escherichia coli str. K-12 substr. M
https://www.ncbi.nlm.nih.gov/nuccore/U00096.3
Chat
An official website of the United States government Here's how you know
NIH National Library of Medicine
National Center for Biotechnology Information
Log in
Nucleotide
Nucleotide
Search
Advanced
Help
GenBank
Send to:
Change region shown
w
Update View
equence
Run BLAST
Pick Primers

⚠ Due to the large size of this record, sequence and annotated features are not shown. Use the "Customize view" par

Escherichia coli str. K-12 substr. MG1655, complete genome

GenBank: U00096.3

FASTA Graphics

Go to:

```
LOCUS       U00096             4641652 bp    DNA     circular BCT 14-FEB-2025
DEFINITION  Escherichia coli str. K-12 substr. MG1655, complete genome.
ACCESSION   U00096
VERSION     U00096.3
DBLINK      BioProject: PRJNA225
            BioSample: SAMN02604091
KEYWORDS    .
SOURCE      Escherichia coli str. K-12 substr. MG1655
  ORGANISM  Escherichia coli str. K-12 substr. MG1655
            Bacteria; Pseudomonadati; Pseudomonadota; Gammaproteobacteria;
            Enterobacterales; Enterobacteriaceae; Escherichia.
```

● Complete Record
○ Coding Sequences
○ Gene Features

Choose Destination

● File ○ Clipboard
○ Collections ○ Analysis Tool

Download 1 item.

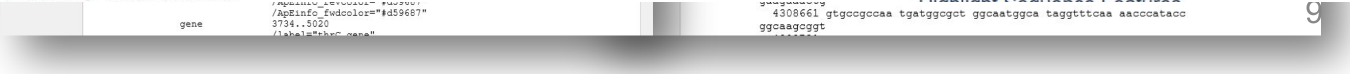
Format: GenBank

Show GI □

Create File

and CDS features only
s
nce
complement
atures

These DNA files can be downloaded from databases e.g. NCBI

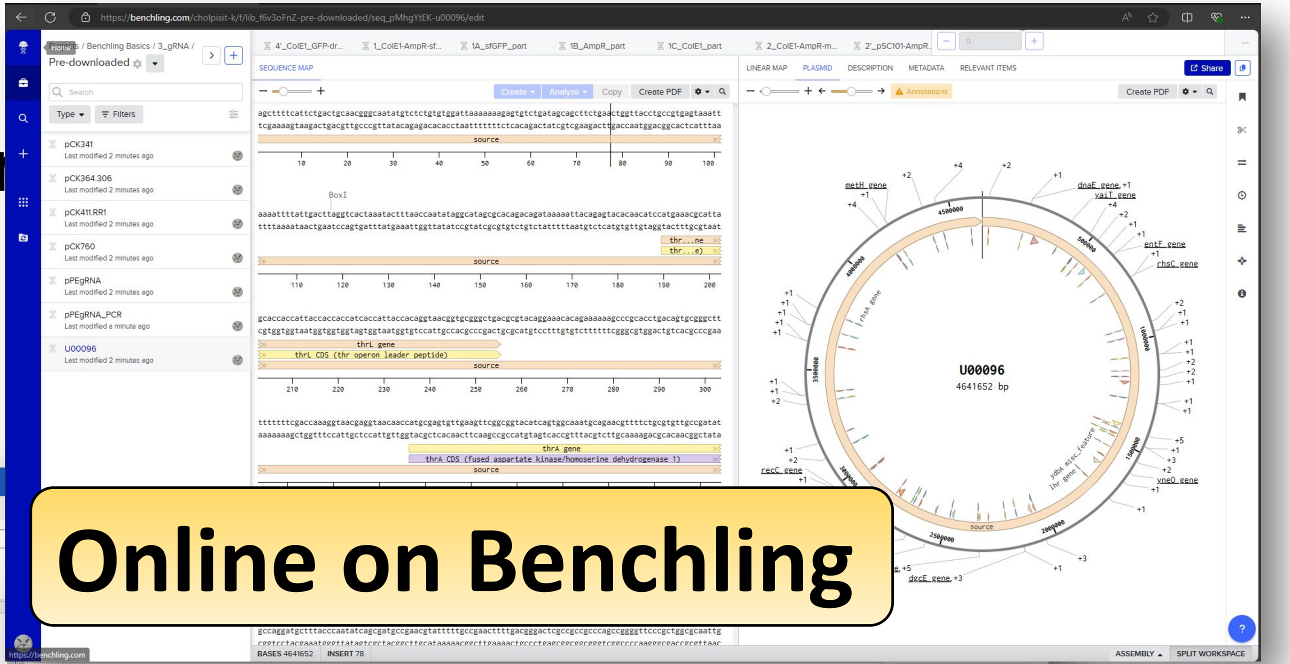


```
/ApEinfo_fwdcolor="#d59687"
gene            3734..5020
```

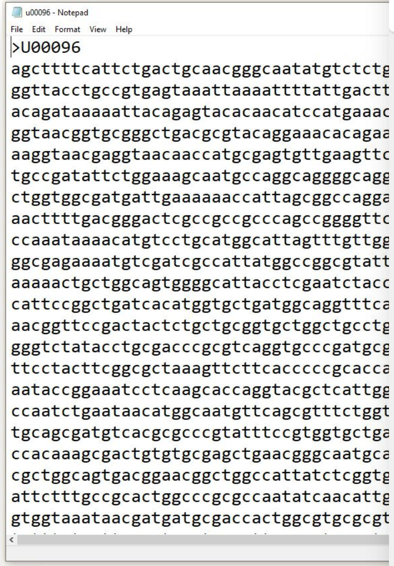
```
4308661 gtgccgccaa tgatggcgct ggcaatggca taggtttcaa aacccatacc
ggcaagcggt
```

Visualizing DNA Features

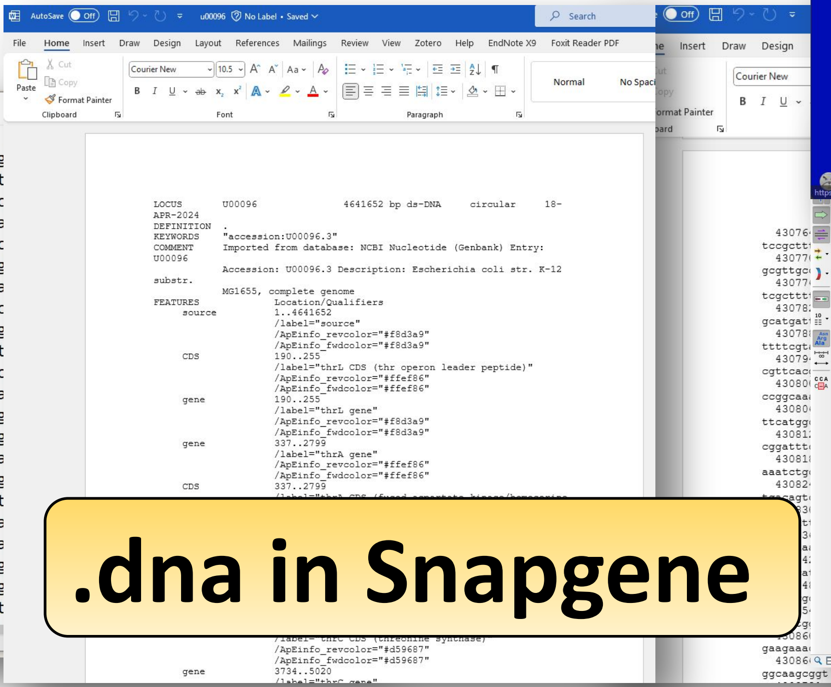
- DNA data can be stored in var
  - .fasta
  - .genbank or .gb or .gff
  - .dna (Snapgene)



Online on Benchling

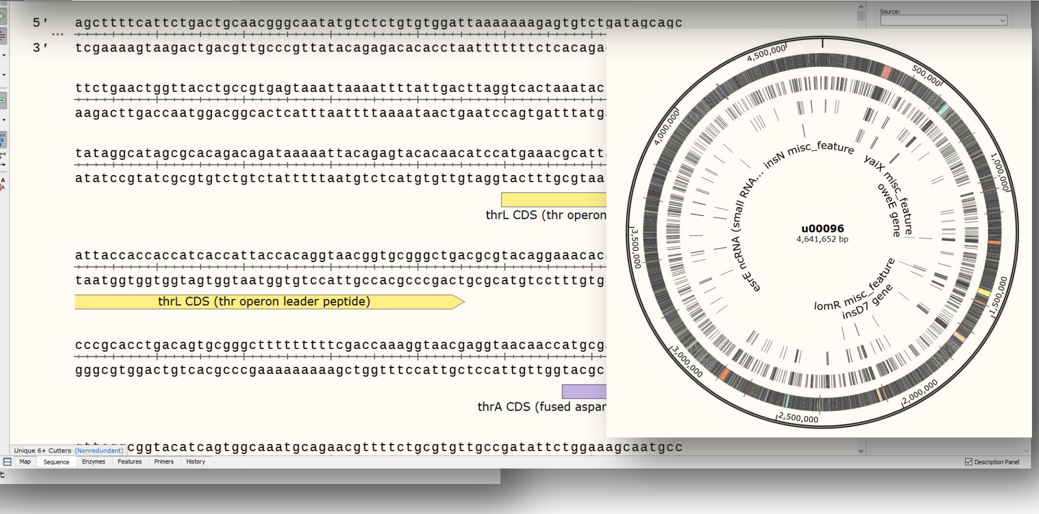


```
>U00096
agcttttcattctgactgcaacgggcaatatgtctctg
ggttacctgccgtgagtaaattaaaattttattgactt
acagataaaaattacagagtacacaacatccatgaaac
ggtaacggtgcgggctgacgcgtacaggaaacacagaa
aaggtaacgaggtaacaaccatgcgagtgttgaagttc
tgccgatattctggaaagcaatgccaggcaggggcagg
ctggtggcgatgattgaaaaaaccattagcggccagga
aacttttgacgggactcgccgccgcccagccggggttc
ccaaataaaacatgtcctgcatggcattagtttgttgg
ggcgagaaaatgtcgatcgccattatggccggcgtatt
aaaaactgctggcagtggggcattacctcgaatctacc
cattccggctgatcacatggtgctgatggcaggtttca
aacggttccgactactctgctgcggtgctggctgcctg
gggtctatacctgcgacccgcgtcaggtgcccgatgcg
ttcctacttcggcgctaaagttcttcaccccgcacca
aataccggaaatcctcaagcaccaggtacgctcattgg
ccaatctgaataacatggcaatgttcagcgtttctggt
tgcagcgatgtcacgcgcccgtatttccgtggtgctga
ccacaaagcgactgtgtgcgagctgaacgggcaatgca
cgctggcagtgacggaacggctggccattatctcggtg
attctttgccgcactggcccgcgccaatatcaacattg
gtggtaaataacgatgatgcgaccactggcgtgcgcgt
```



```
LOCUS       U00096              4641652 bp ds-DNA    circular    18-
APR-2024
DEFINITION  .
KEYWORDS    "accession:U00096.3"
COMMENT     Imported from database: NCBI Nucleotide (Genbank) Entry:
U00096
            Accession: U00096.3 Description: Escherichia coli str. K-12
substr.
            MG1655, complete genome
FEATURES             Location/Qualifiers
     source          1..4641652
                     /label="source"
                     /ApEinfo_revcolor="#f8d3a9"
                     /ApEinfo_fwdcolor="#f8d3a9"
     CDS             190..255
                     /label="thrL CDS (thr operon leader peptide)"
                     /ApEinfo_revcolor="#ffef86"
                     /ApEinfo_fwdcolor="#ffef86"
     gene            190..255
                     /label="thrL gene"
                     /ApEinfo_revcolor="#f8d3a9"
                     /ApEinfo_fwdcolor="#f8d3a9"
     gene            337..2799
                     /label="thrA gene"
                     /ApEinfo_revcolor="#ffef86"
                     /ApEinfo_fwdcolor="#ffef86"
     CDS             337..2799
                     /label="thrC CDS (threonine synthase)"
                     /ApEinfo_revcolor="#d59687"
                     /ApEinfo_fwdcolor="#d59687"
     gene            3734..5020
                     /label="thrC gene"
```

.dna in Snapgene

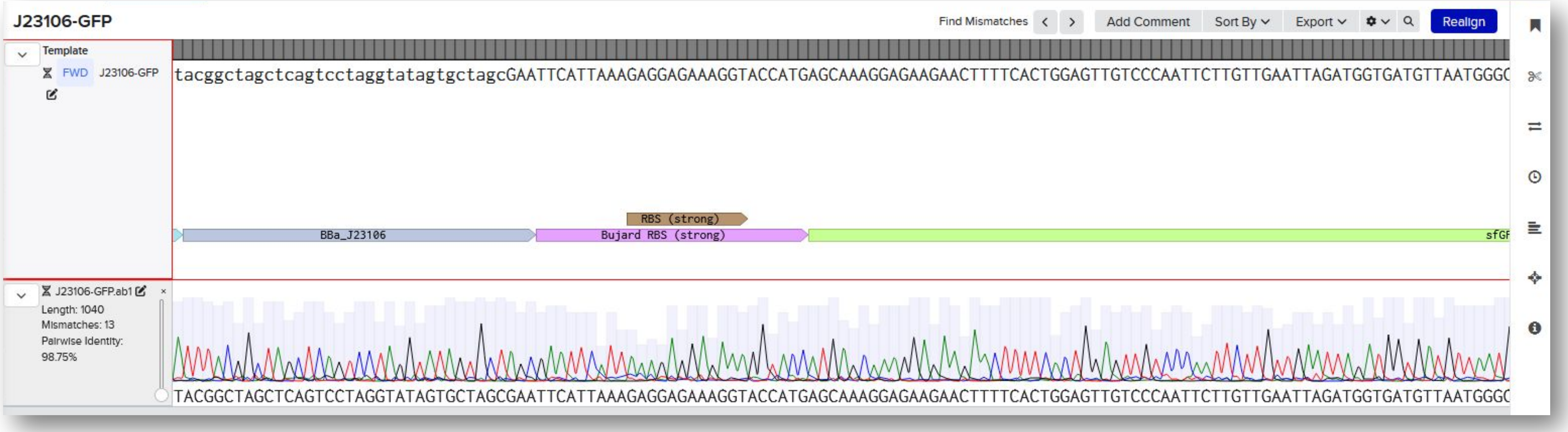


```
5'  agcttttcattctgactgcaacgggcaatatgtctctgtgtggattaaaaaaagagtgtctgatagcagc
3'  tcgaaaagtaagactgacgttgcccgttatacagagacacacctaatttttttctcacaga

    ttctgaactggttacctgccgtgagtaaattaaaattttattgacttaggtcactaaatac
    aagacttgaccaatggacggcactcatttaattttaaaataactgaatccagtgatttatg

    tataggcatagcgcacagacagataaaaattacagagtacacaacatccatgaaacgcatt
    atatccgtatcgcgtgtctgtctatttttaatgtctcatgtgttgtaggtactttgcgtaa
                                                  thrL CDS (thr operon

    attaccaccaccatcaccattaccacaggtaacggtgcgggctgacgcgtacaggaaacac
    taatggtggtggtagtggtaatggtgtccattgccacgcccgactgcgcatgtcctttgtg
       thrL CDS (thr operon leader peptide)

    cccgcacctgacagtgcggctttttttttcgaccaaaggtaacgaggtaacaaccatgcg
    gggcgtggactgtcacgcccgaaaaaaaaagctggtttccattgctccattgttggtacgc
                                                  thrA CDS (fused aspar

    cggtacatcagtggcaaatgcagaacgttttctgcgtgttgccgatattctggaaagcaatgcc
```

Validation of new DNA: Sequencing

DNA sequencing could be performed by different technologies including

1. Sanger sequencing (cheap),
2. Next-gen sequencing (massively parallel)
3. 3rd-gen sequencing (long read)



J23106-GFP

Template

FWD J23106-GFP

tacggctagctcagtcctaggtatagtgctagcGAATTCATTAAAGAGGAGAAAGGTACCATGAGCAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGC

J23106-GFP.ab1

Length: 1040
Mismatches: 13
Pairwise Identity: 98.75%

TACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCGAATTCATTAAAGAGGAGAAAGGTACCATGAGCAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGC

Each nucleotide yield different signal (color). Then, they can be aligned to that of template or expected sequence for validation

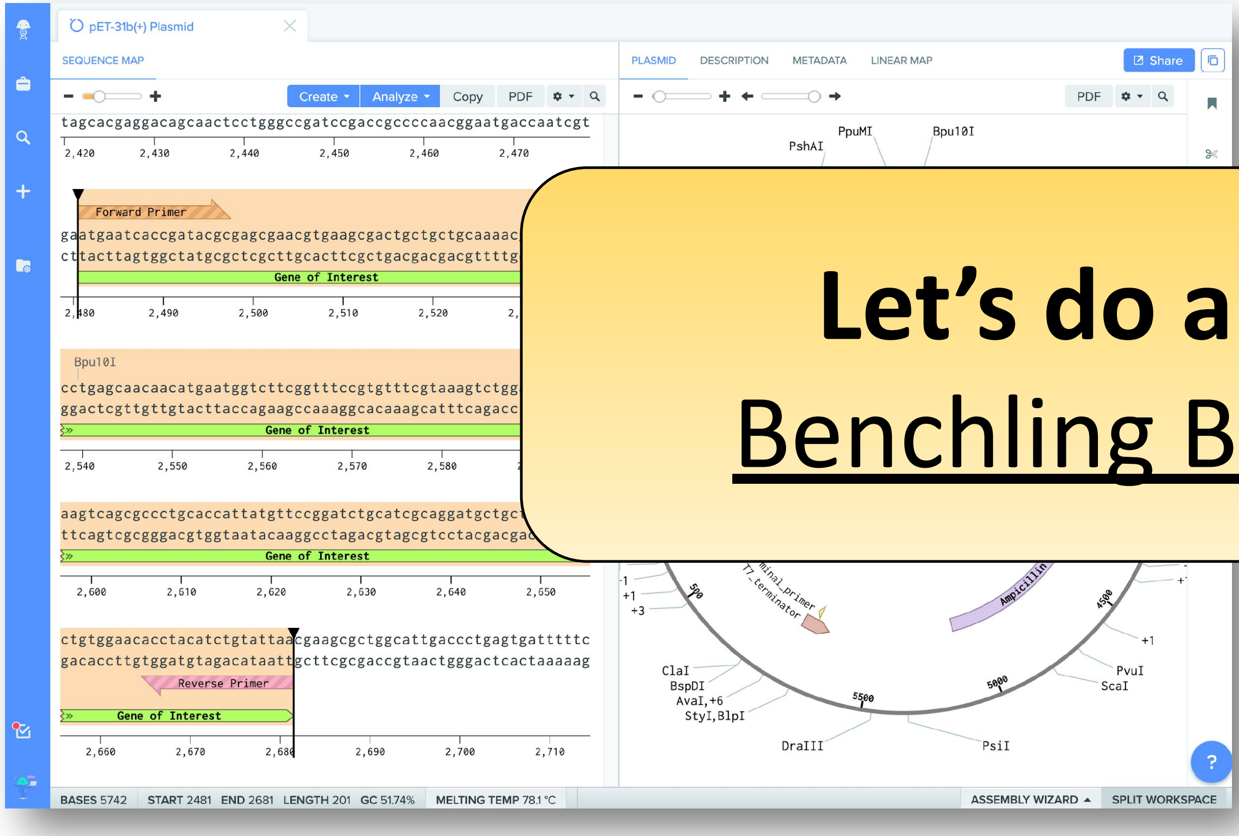
In silico analysis of DNA sequence

Benchling®: Free for academic users

A more powerful version available for **Enterprise user**

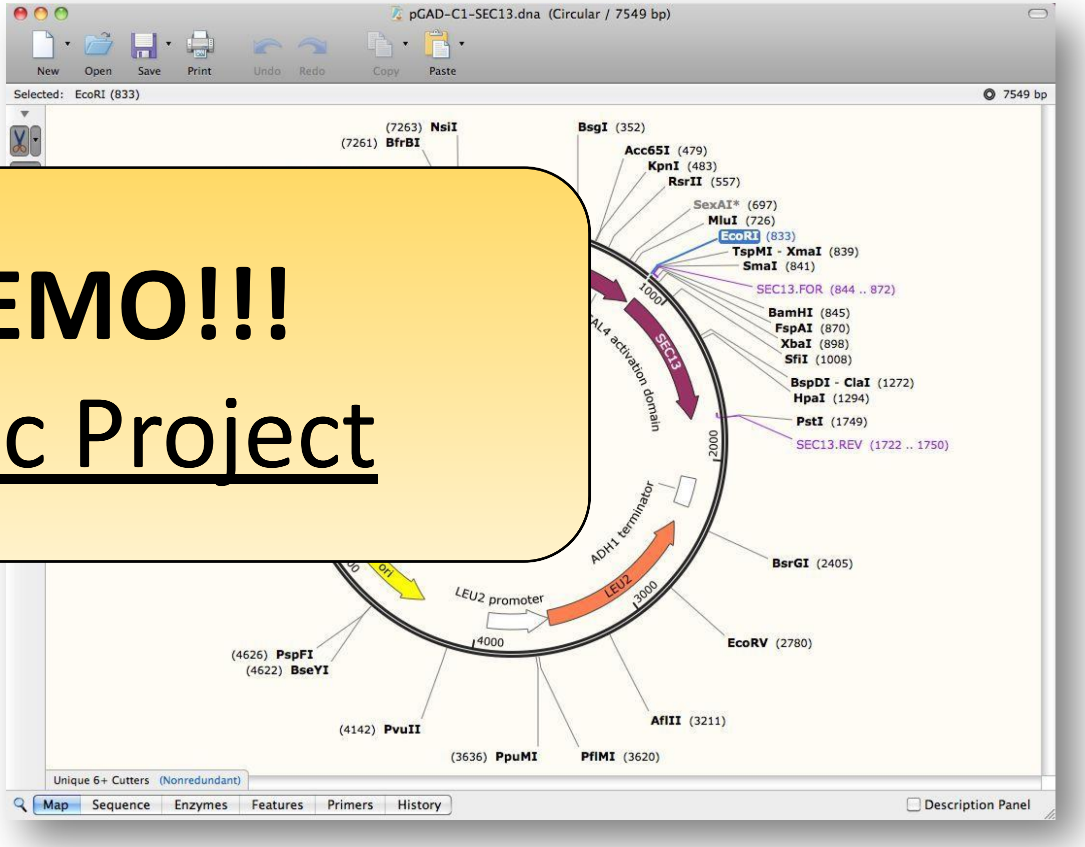
SnapGene Viewer: Free

A more powerful version available as **SnapGene®**



```
tagcacgaggacagcaactcctgggccgatccgaccgccccaacggaatgaccaatcgt
 2,420     2,430     2,440     2,450     2,460     2,470

gaatgaatcaccgatacgcgagcgaacgtgaagcgactgctgctgcaaaac
cttacttagtggctatgcgctcgcttgcacttcgctgacgacgacgttttg
 2,480     2,490     2,500     2,510     2,520

cctgagcaacaacatgaatggtcttcggtttccgtgtttcgtaaagtctgg
ggactcgttgttgtacttaccagaagccaaaggcacaaagcatttcagacc
 2,540     2,550     2,560     2,570     2,580

aagtcagcgccctgcaccattatgttccggatctgcatcgcaggatgctgc
ttcagtcgcgggacgtggtaatacaaggcctagacgtagcgtcctacgacga
 2,600     2,610     2,620     2,630     2,640     2,650

ctgtggaacacctacatctgtattaacgaagcgctggcattgaccctgagtgattttc
gacaccttgtggatgtagacataattgcttcgcgaccgtaactgggactcactaaaaag
 2,660     2,670     2,680     2,690     2,700     2,710
```

BASES 5742 START 2481 END 2681 LENGTH 201 GC 51.74% MELTING TEMP 78.1 °C



pGAD-C1-SEC13.dna (Circular / 7549 bp)

Selected: EcoRI (833) 7549 bp

(7263) NsiI
(7261) BfrBI
BsgI (352)
Acc65I (479)
KpnI (483)
RsrII (557)
SexAI* (697)
MluI (726)
EcoRI (833)
TspMI - XmaI (839)
SmaI (841)
SEC13.FOR (844 .. 872)
BamHI (845)
FspAI (870)
XbaI (898)
SfiI (1008)
BspDI - ClaI (1272)
HpaI (1294)
PstI (1749)
SEC13.REV (1722 .. 1750)
BsrGI (2405)
EcoRV (2780)
AflII (3211)
PflMI (3620)
(3636) PpuMI
(4142) PvuII
(4626) PspFI
(4622) BseYI

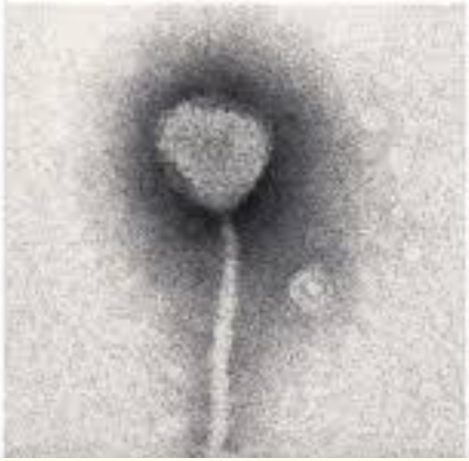
Let's do a DEMO!!!

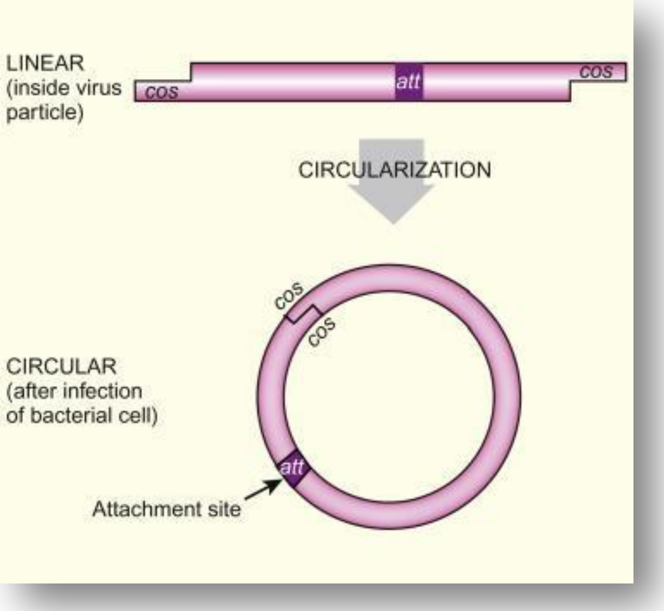
Benchling Basic Project

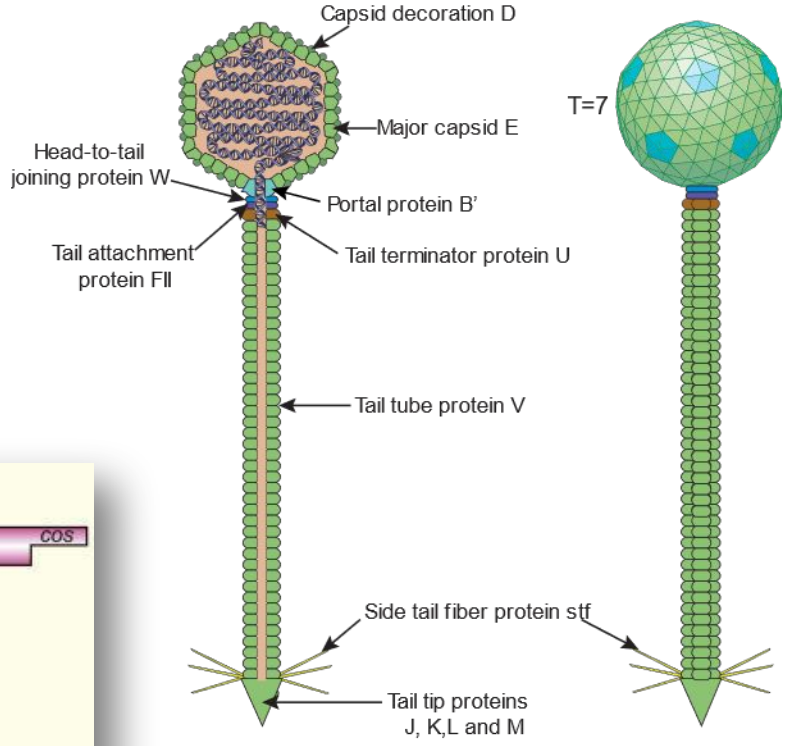
https://benchling.com/

https://www.snapgene.com/snapgene-viewer

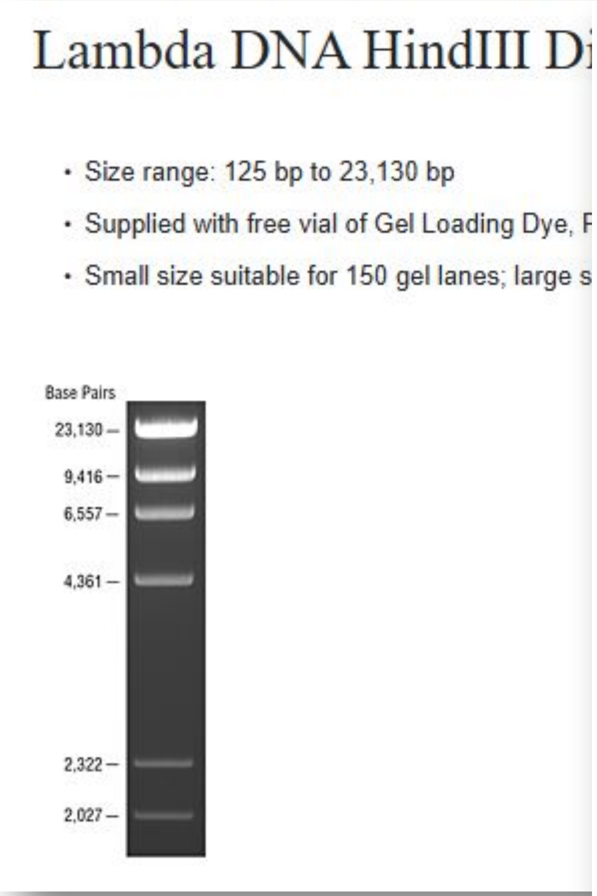
Let’s do DNA Separation: *in silico* gel





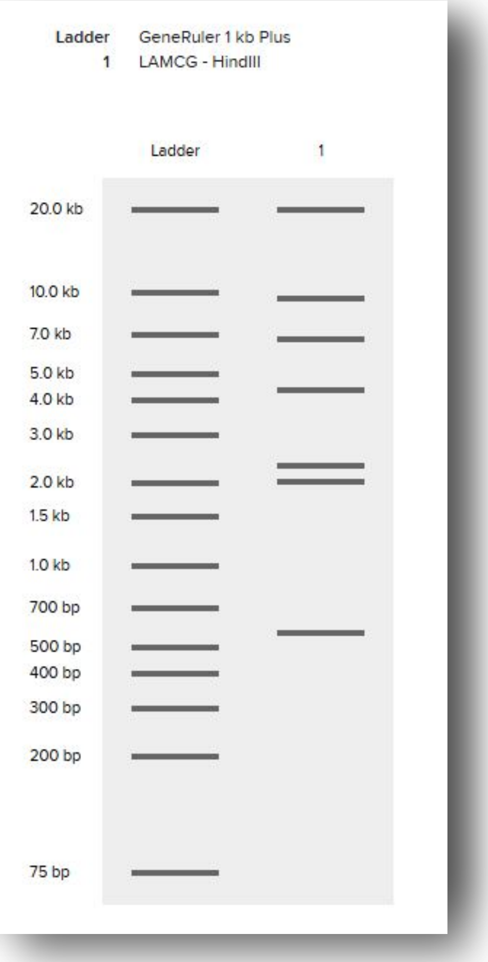


Phage Lambda Genome ~45kb



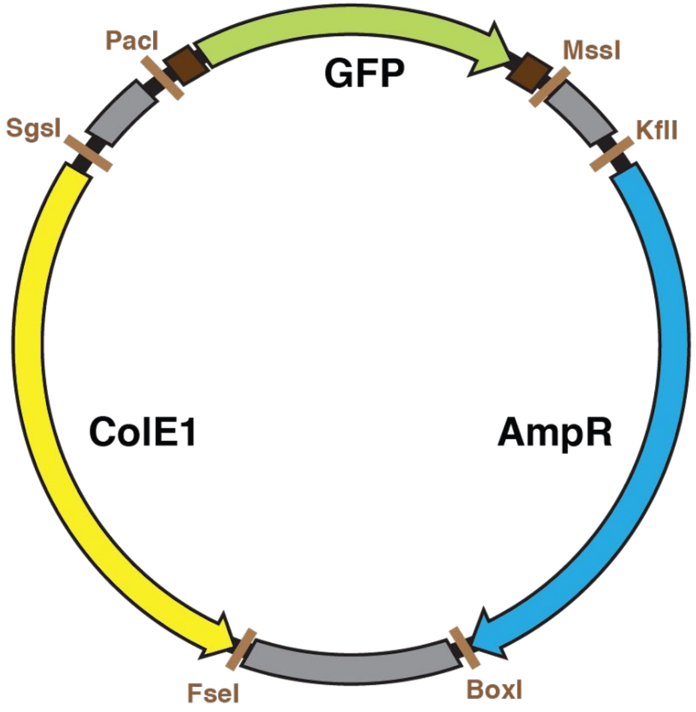
Lambda DNA HindIII Di

- Size range: 125 bp to 23,130 bp
- Supplied with free vial of Gel Loading Dye, P
- Small size suitable for 150 gel lanes; large s



Ladder GeneRuler 1 kb Plus
1 LAMCG - HindIII

Assembly of 3 Fragments into a plasmid – ColE1-AmpR-GFP

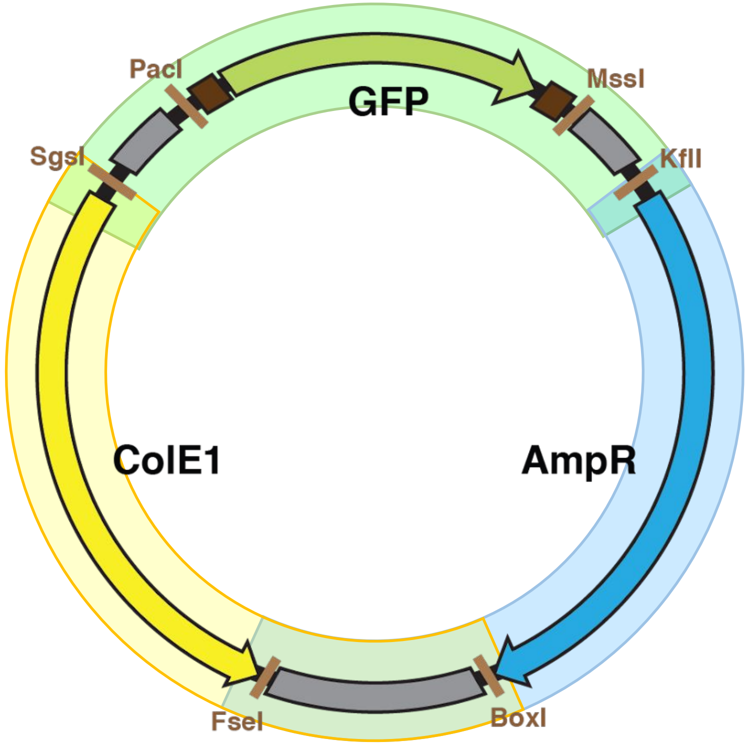


Make different DNA fragments by PCR

DNA fragments could also be made by
1) restriction digest reaction or
2) ordered for chemical synthesis

Links to DNA sequences:
Desired sequence: ColE1-AmpR-sfGFP plasmid
1) sfGFP_part, 2) AmpR_part, 3) ColE1_part

Assembly of 3 Fragments into a plasmid – ColE1-AmpR-GFP

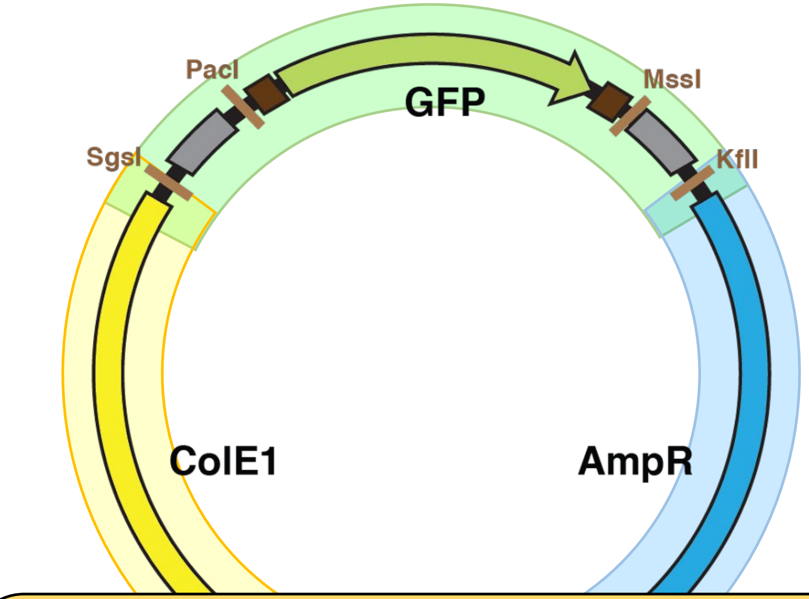


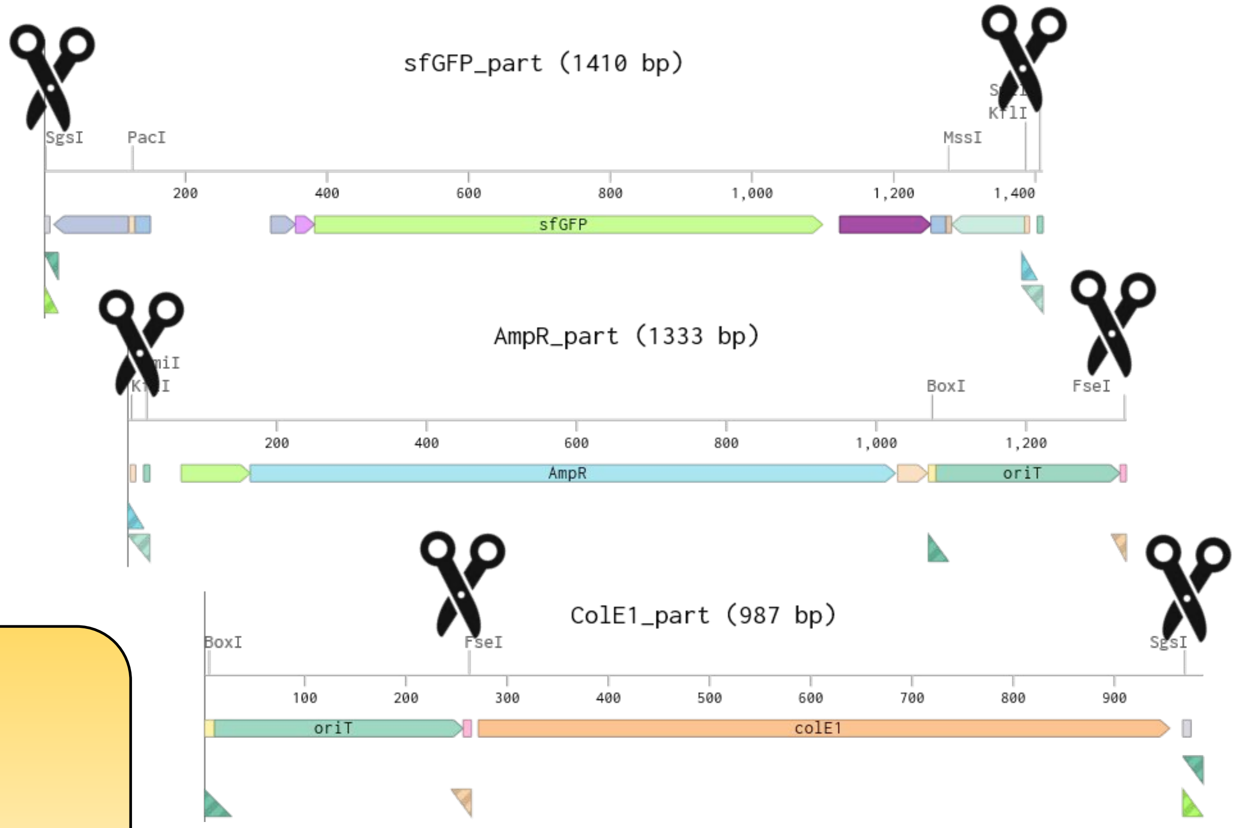
Make different DNA fragments by PCR

DNA fragments could also be made by
1) restriction digest reaction or
2) ordered for chemical synthesis

Links to DNA sequences:
Desired sequence: ColE1-AmpR-sfGFP plasmid
1) sfGFP_part, 2) AmpR_part, 3) ColE1_part

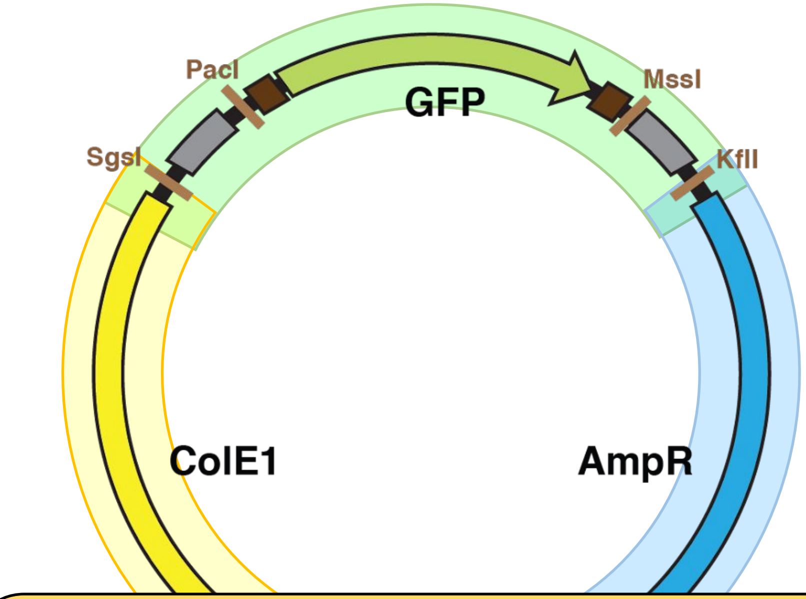
Assembly of 3 Fragments into a plasmid – ColE1-AmpR-GFP

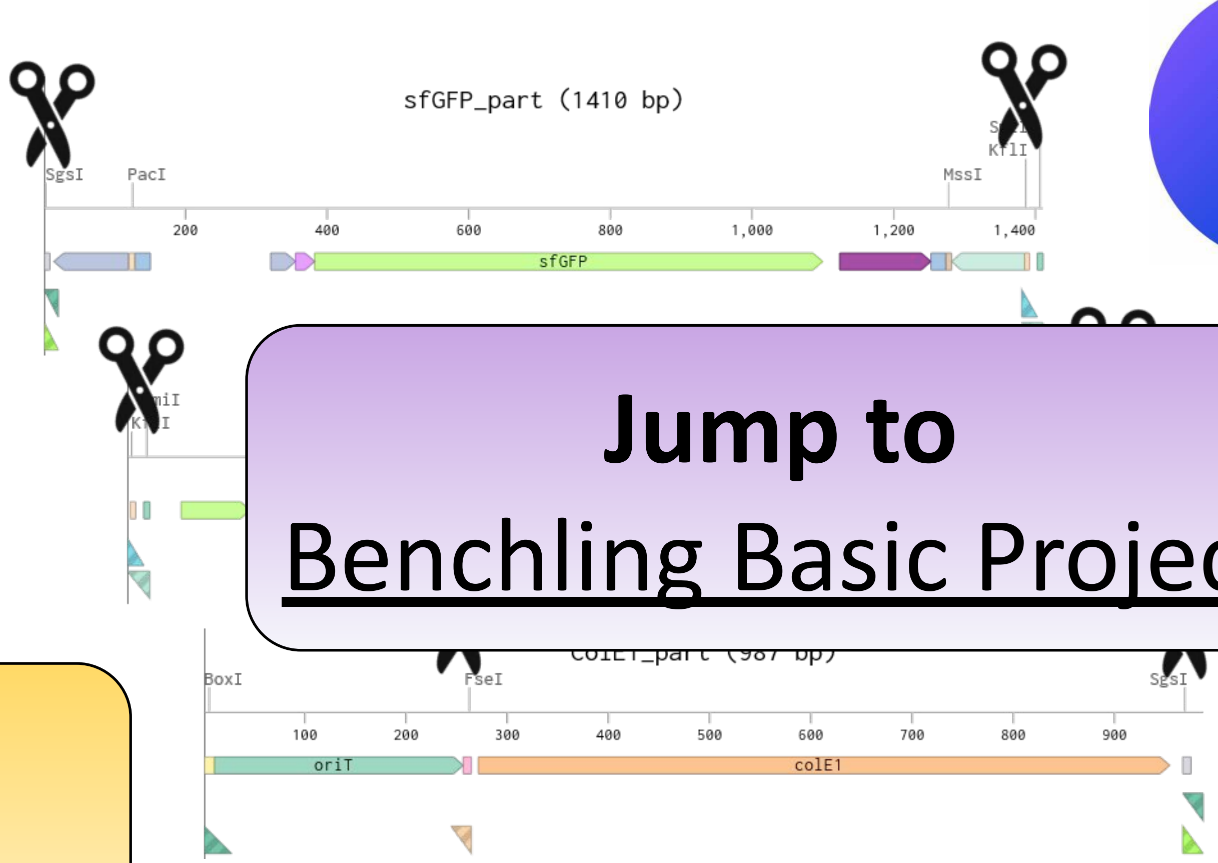




Make appropriate sticky-ends for ligation

Assembly of 3 Fragments into a plasmid – ColE1-AmpR-GFP



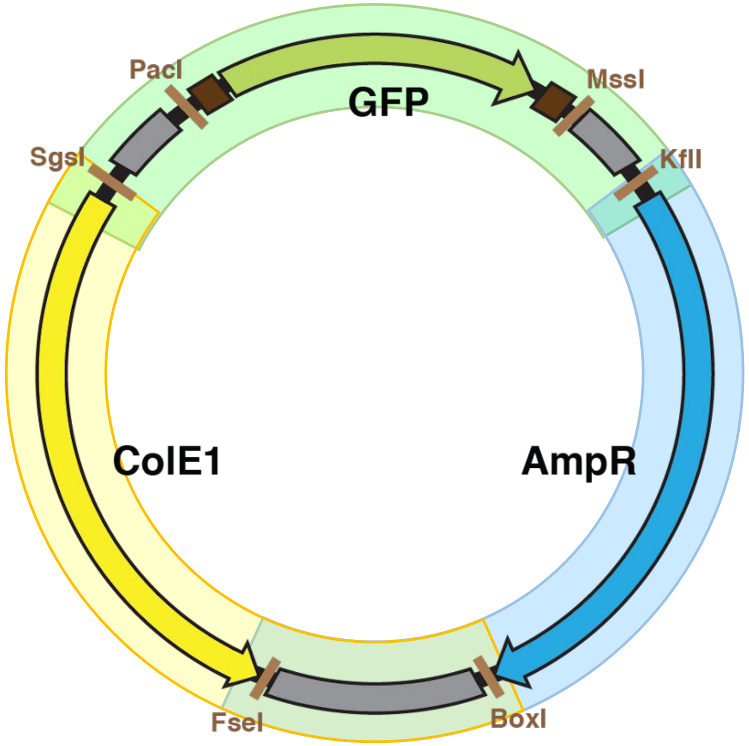


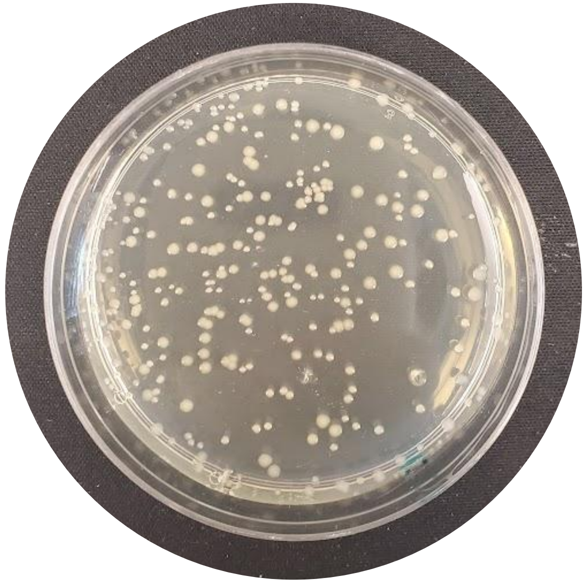
Jump to

Benchling Basic Project

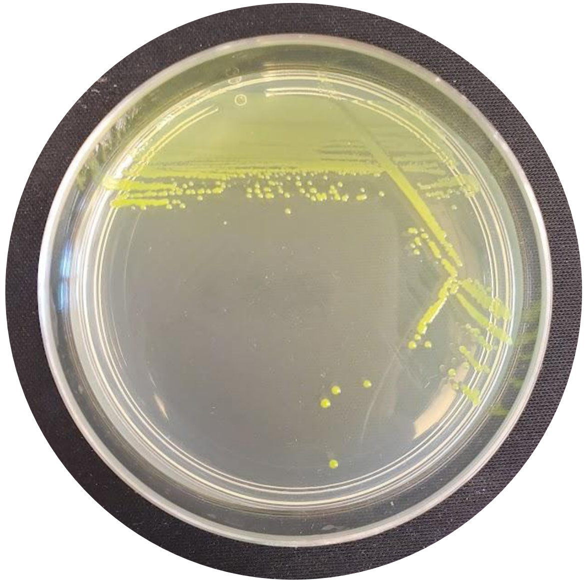
Make appropriate sticky-ends for ligation

Assembly of 3 Fragments into a plasmid – ColE1-AmpR-GFP



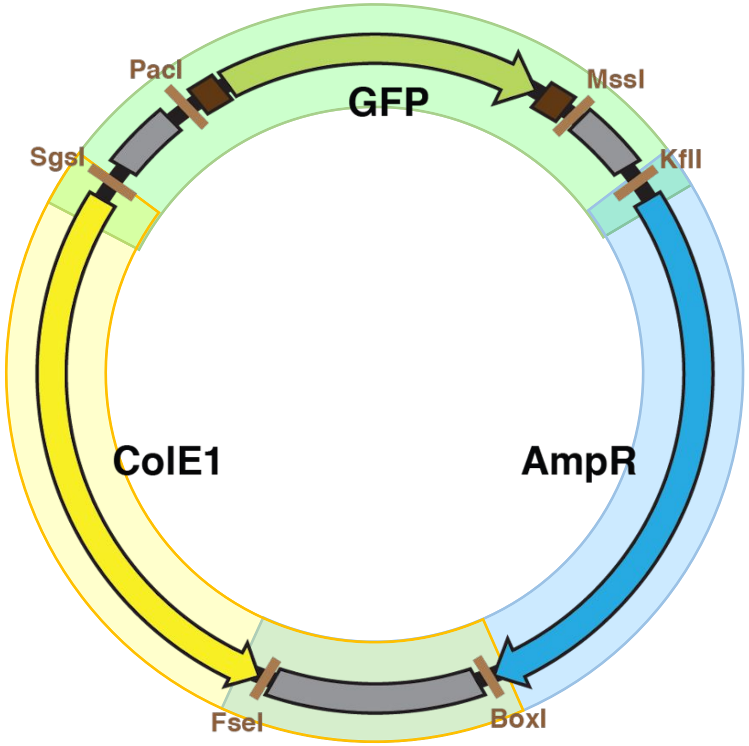


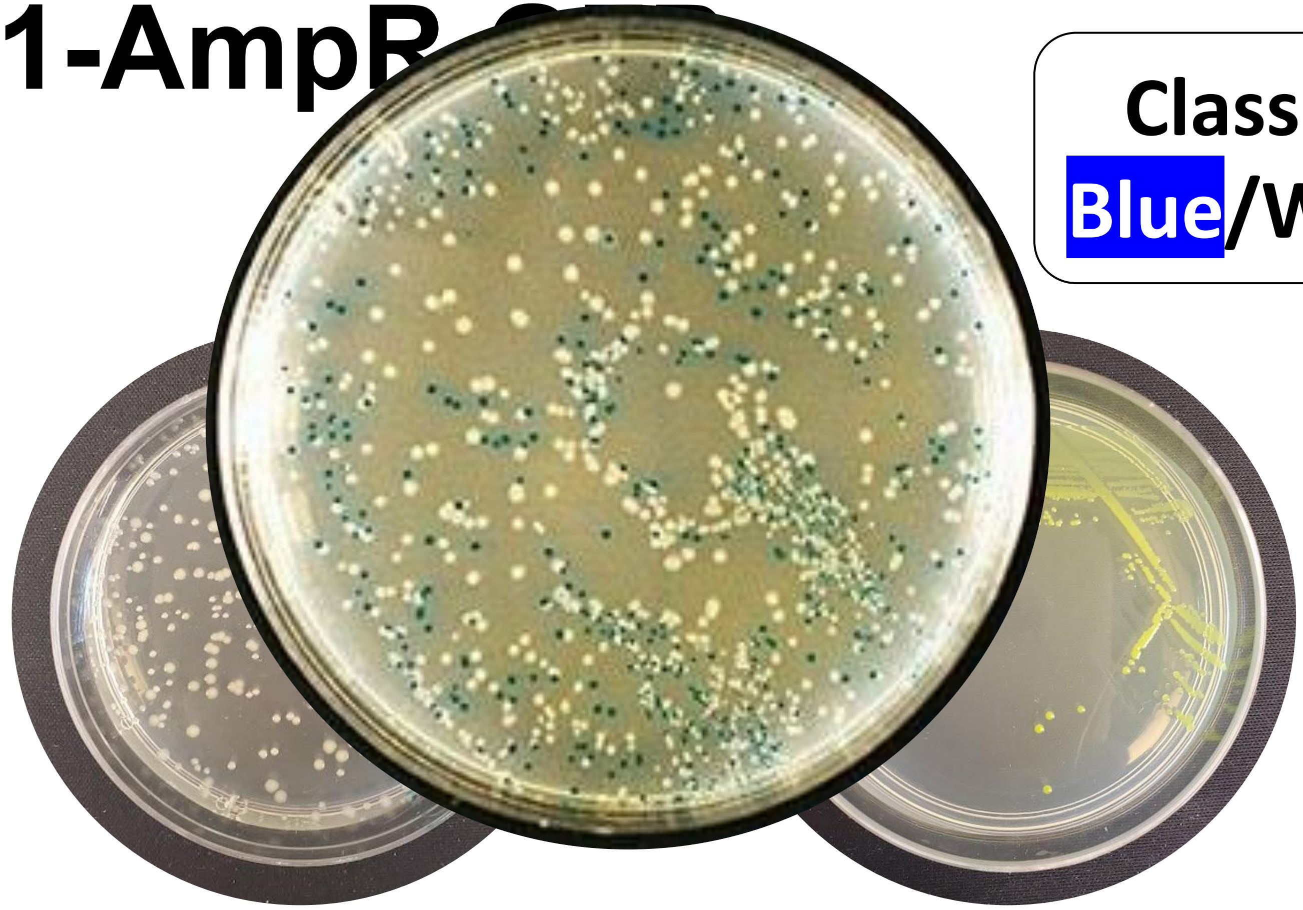
White



Green

Assembly of 3 Fragments into a plasmid – ColE1-AmpR



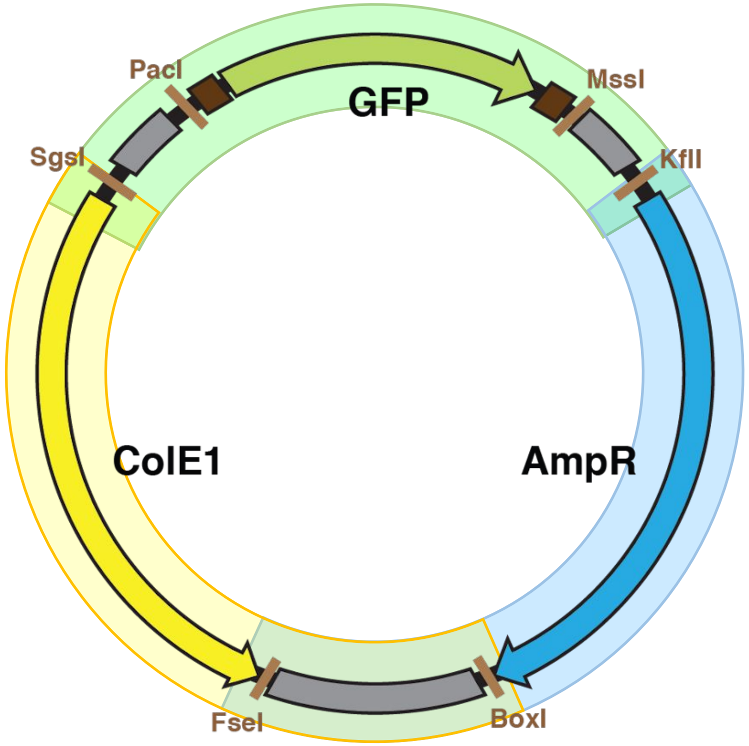


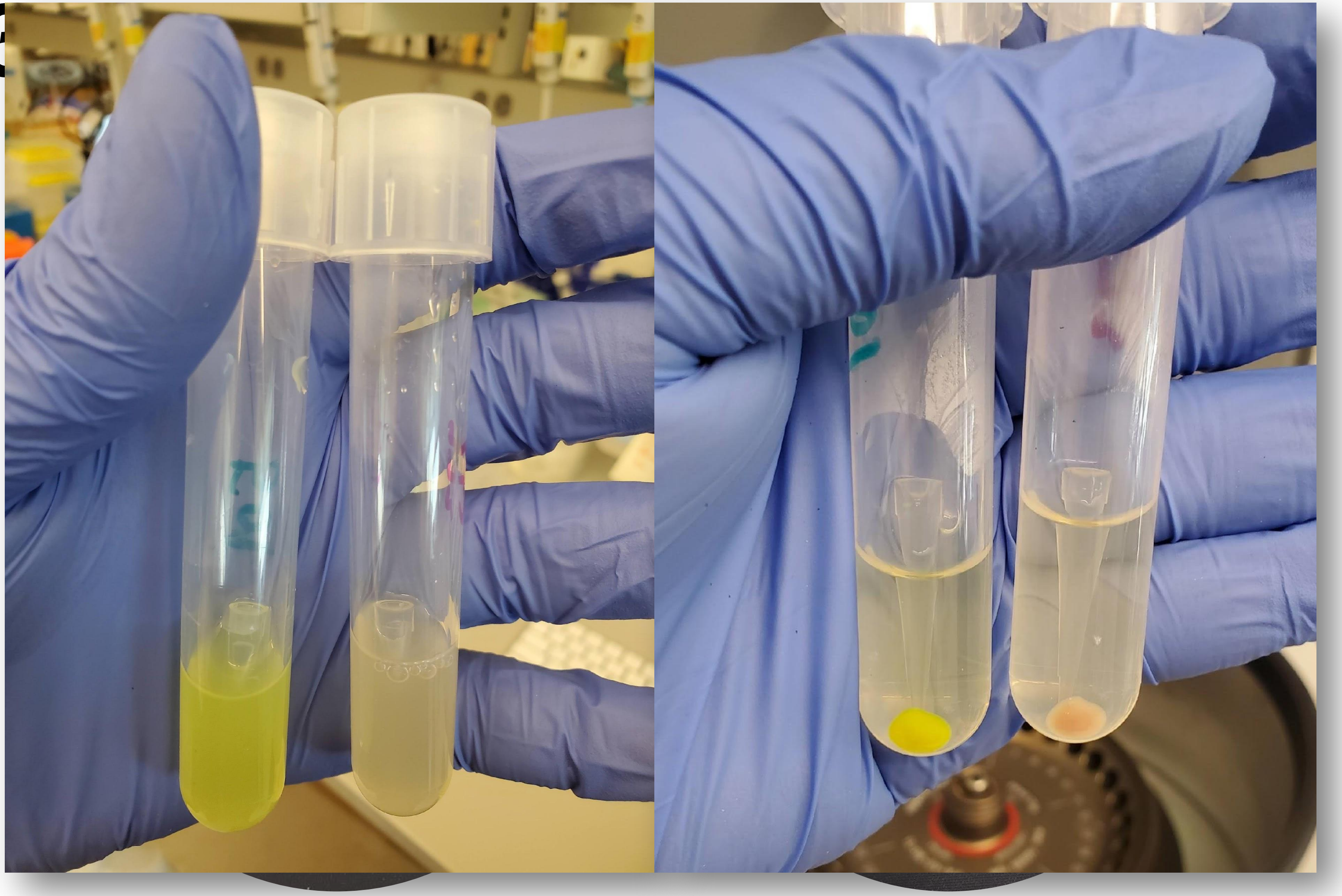
Classical Blue/White

White

Green

Assembly of 3
plasmid – Col





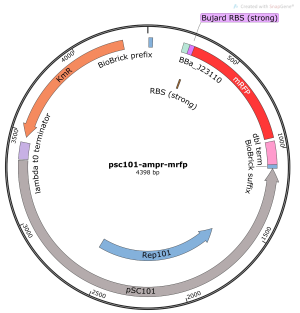
White

Green

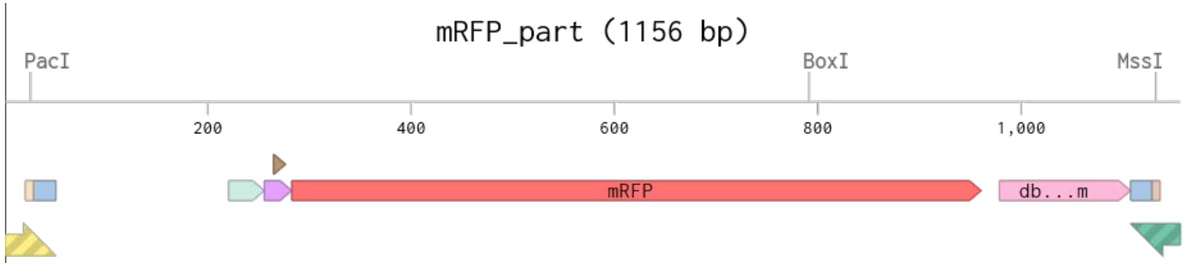
Converting GFP into RFP

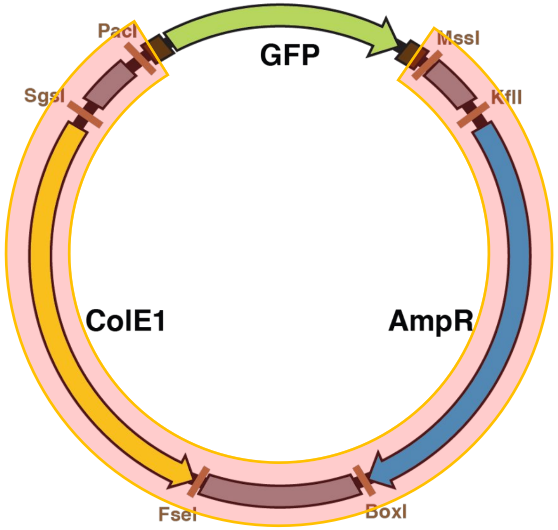
Or DNA Synthesis

Links to DNA sequences:
Starting ColE1-AmpR-sfGFP plasmid
□Digested ColE1-AmpR
mRFP template
□PCR mRFP insert

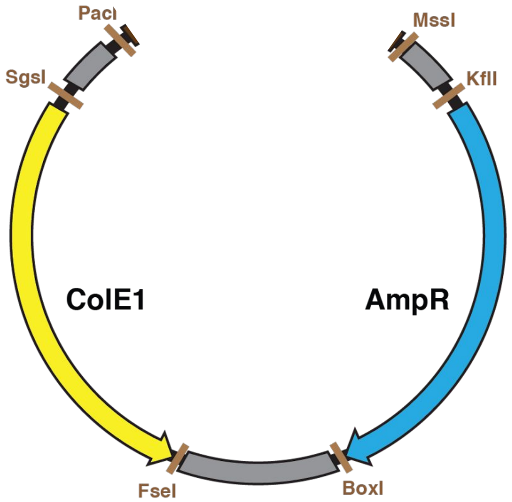


PCR

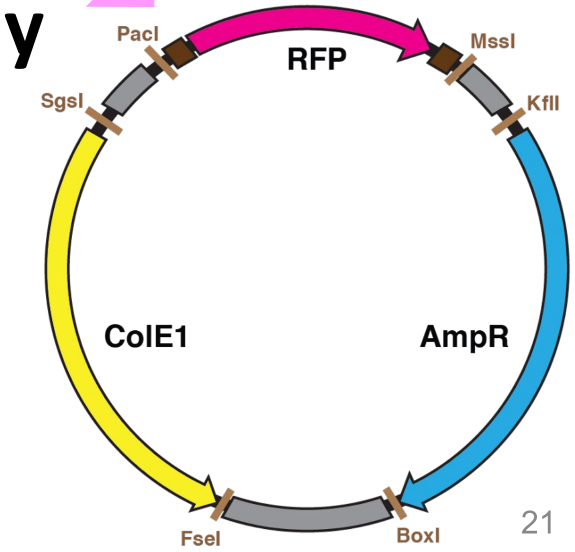




Digest



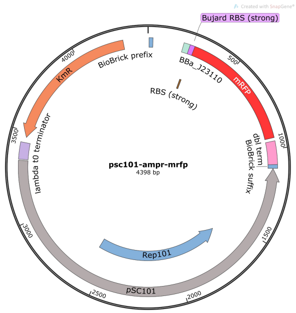
Assembly



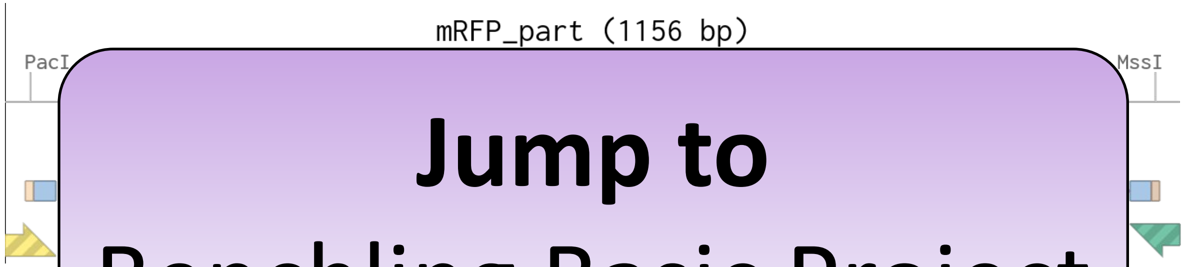
Converting GFP into RFP

Links to DNA sequences:
Starting ColE1-AmpR-sfGFP plasmid
□Digested ColE1-AmpR
mRFP template
□PCR mRFP insert

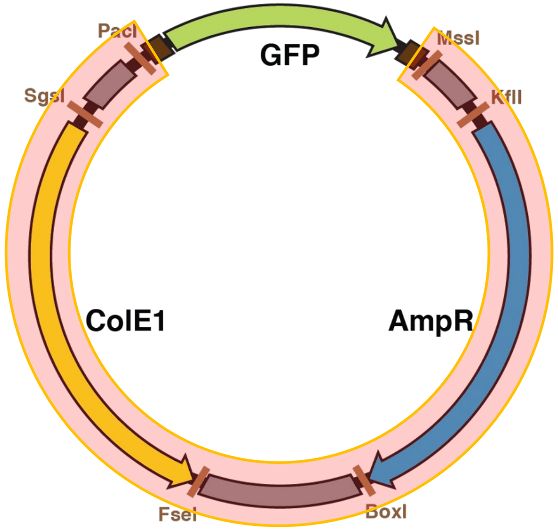
Or DNA Synthesis



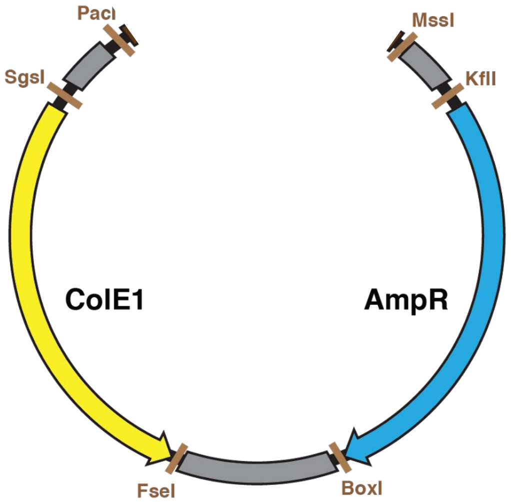
PCR



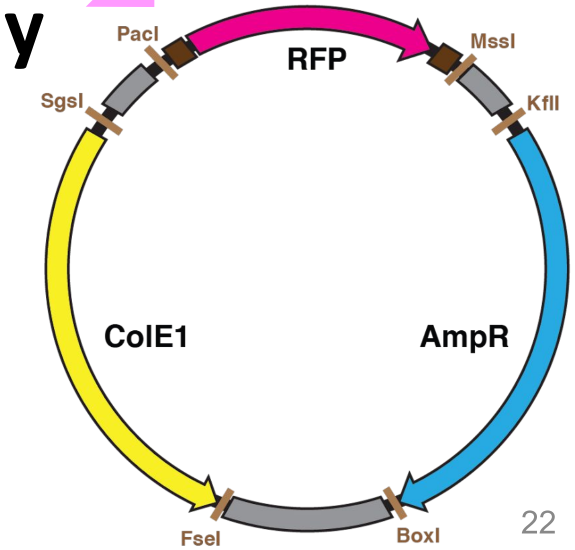
Jump to
Benchling Basic Project



Digest



Assembly

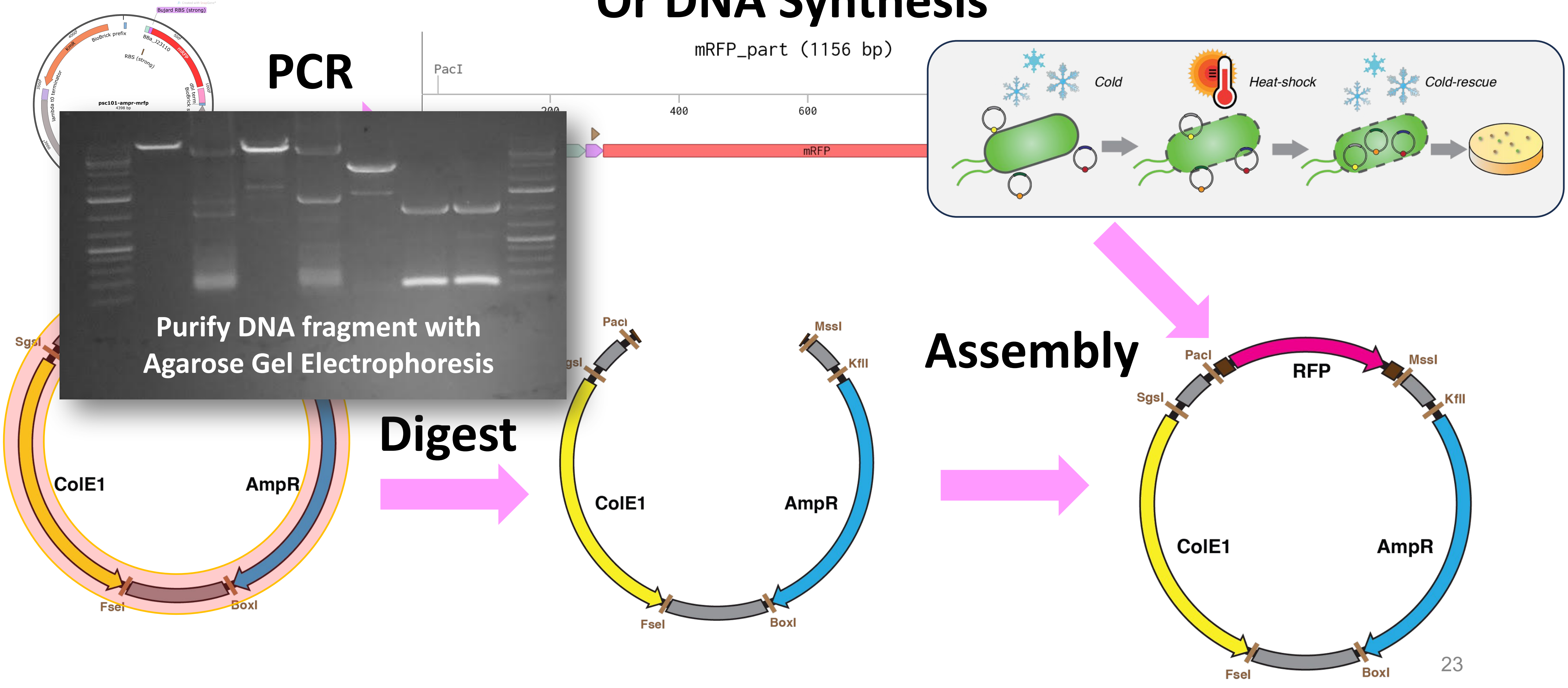


Converting GFP into RFP

Or DNA Synthesis

Links to DNA sequences:

Starting ColE1-AmpR-sfGFP plasmid
□Digested ColE1-AmpR
mRFP template
□PCR mRFP insert



PCR

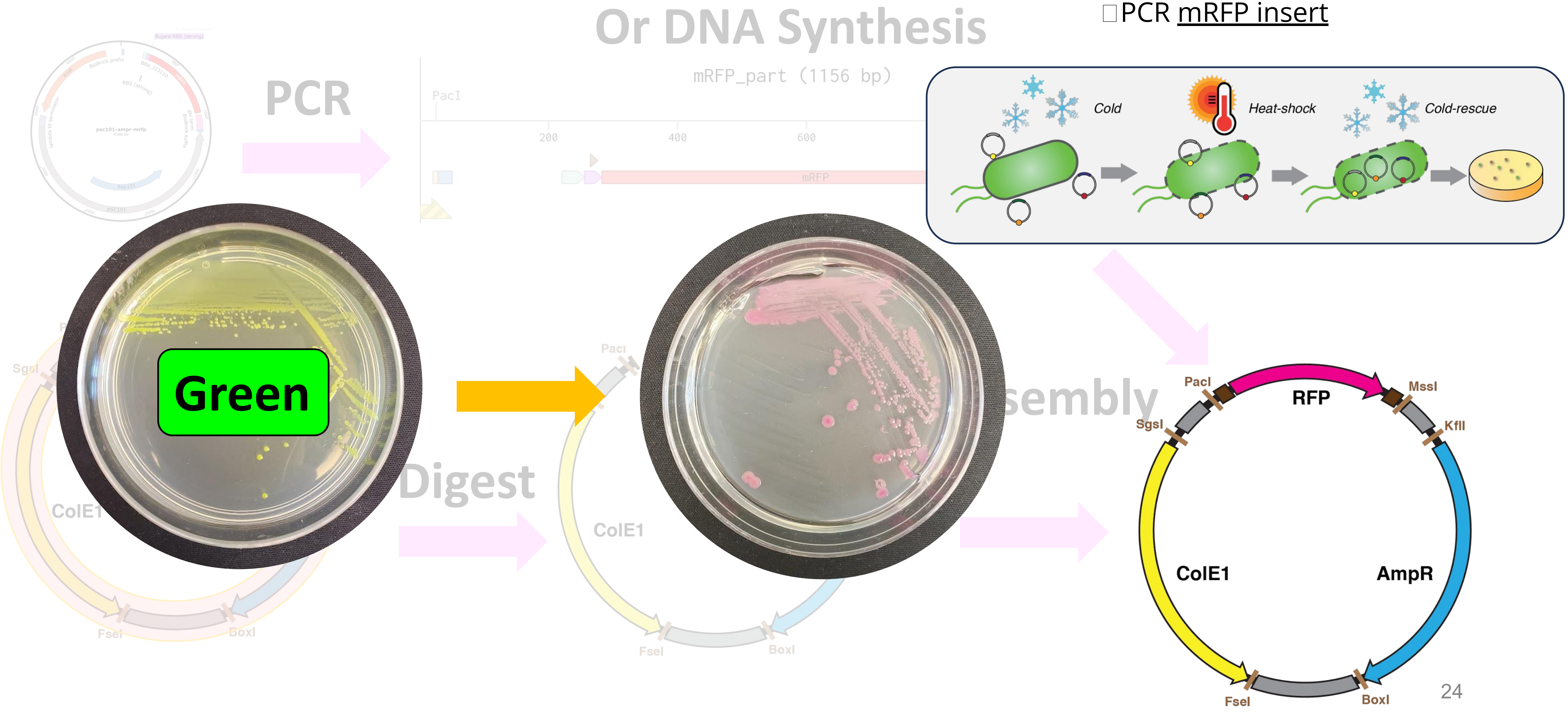
Purify DNA fragment with Agarose Gel Electrophoresis

Digest

Assembly

Converting GFP into RFP

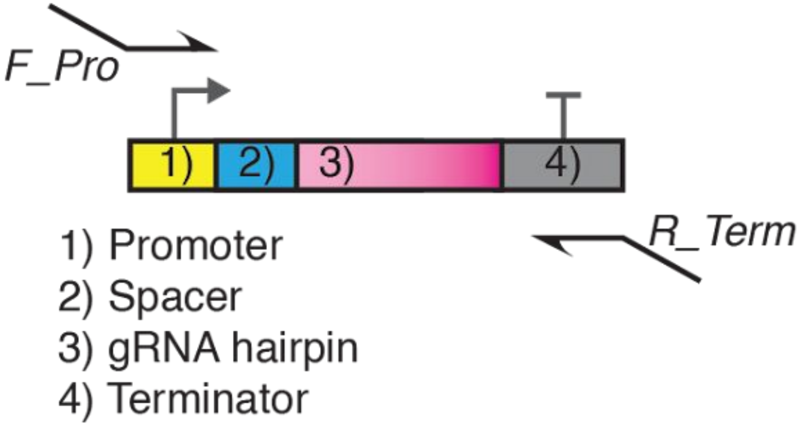
Links to DNA sequences:
Starting ColE1-AmpR-sfGFP plasmid
□Digested ColE1-AmpR
mRFP template
□PCR mRFP insert



Skipped during the talk

See 2025 recording if interested in this demo

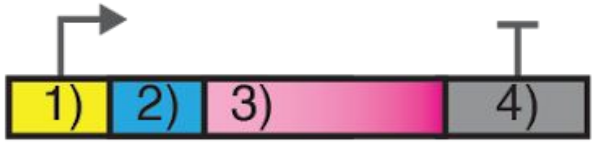
sgRNA expression cassette in bacteria



Spacer
sgRNA
Hairpin

CRISPRi - pCK411.RR1

Ones can also install cut-site between 1) and 2) so that spacer can be changed by PCR with only one oligonucleotide order

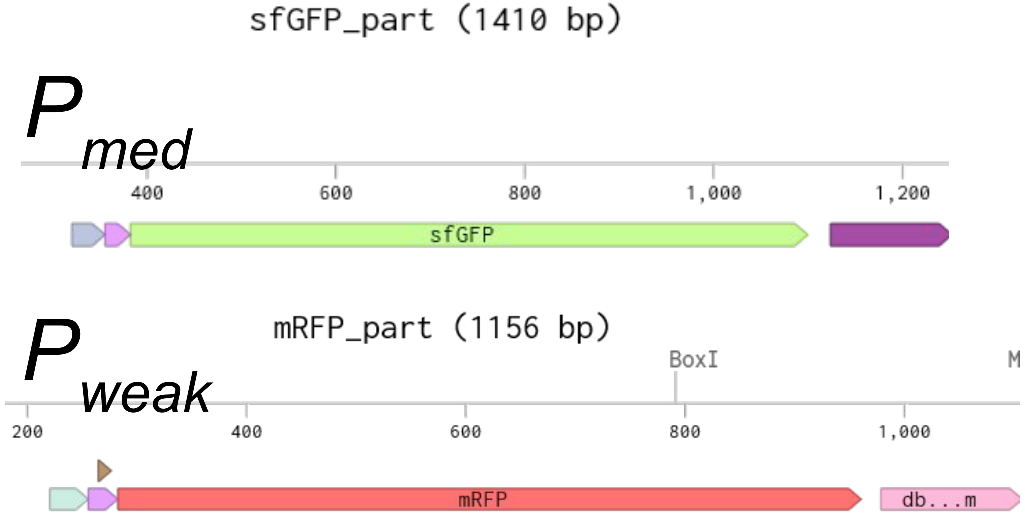


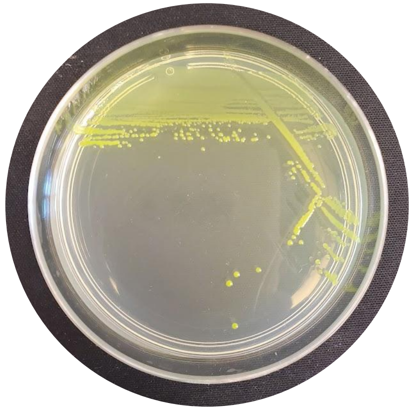
Oligo order is ~$10-20 per 60 bases
20 nt is a regular spacer size
40 nt can be attributed to priming and overhang site

Skipped during the talk

sgRNAs can be used to program phenotypes

Reporter genes Addgene: pCK760





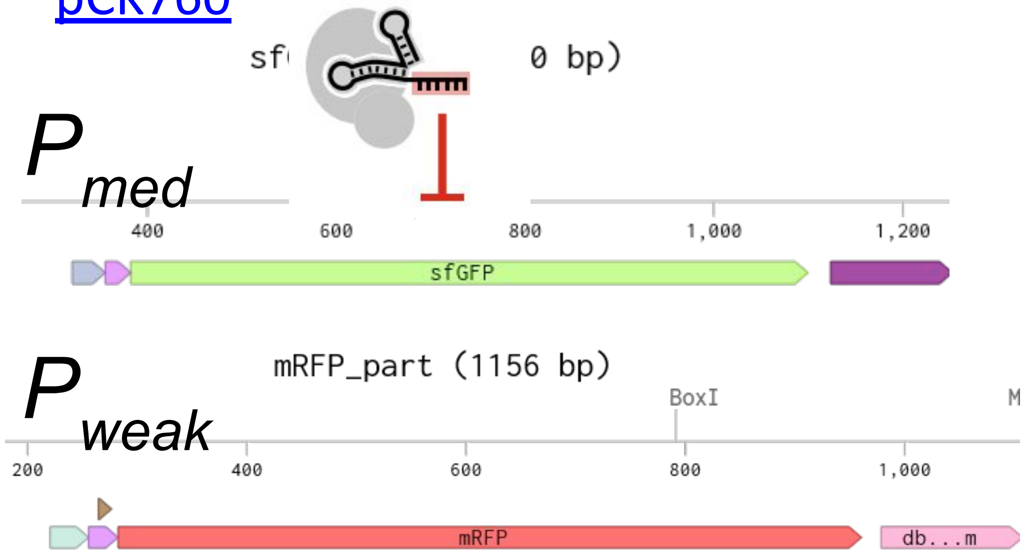
Green

Skipped during the talk

sgRNAs can be used to program phenotypes

CRISPR interference

Reporter genes [Addgene: pCK760]



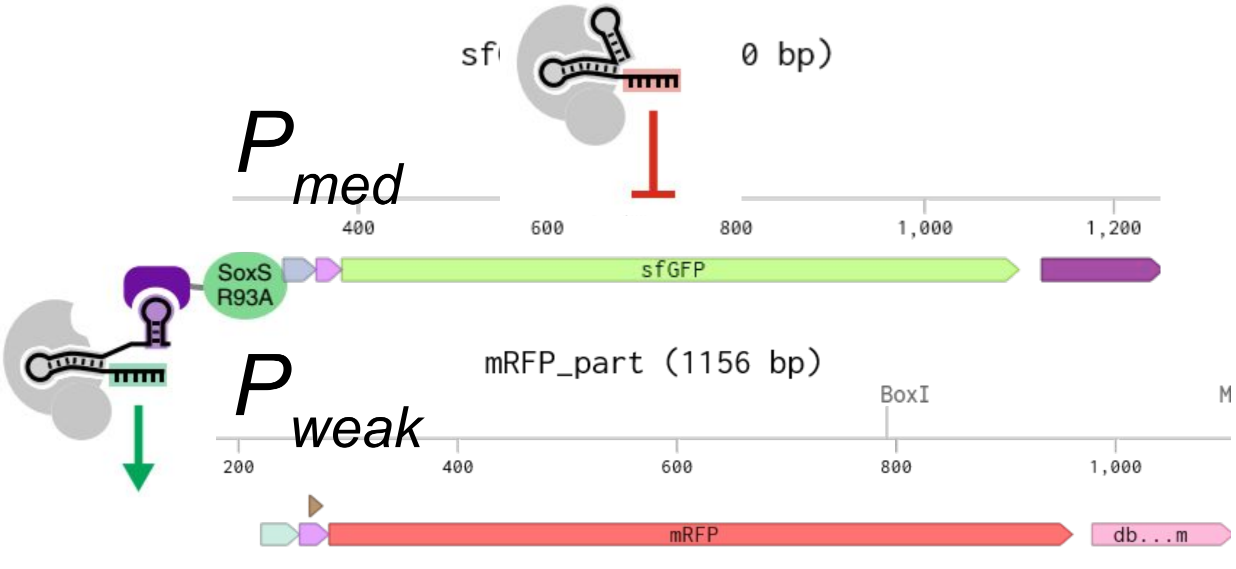
Colorless

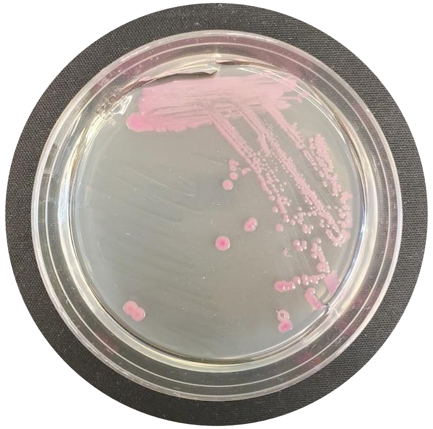
Skipped during the talk

sgRNAs can be used to program phenotypes

CRISPR interference and activation

Reporter genes Addgene: pCK760

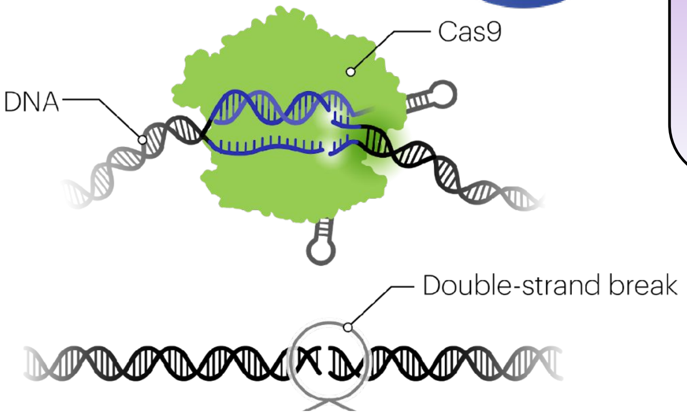




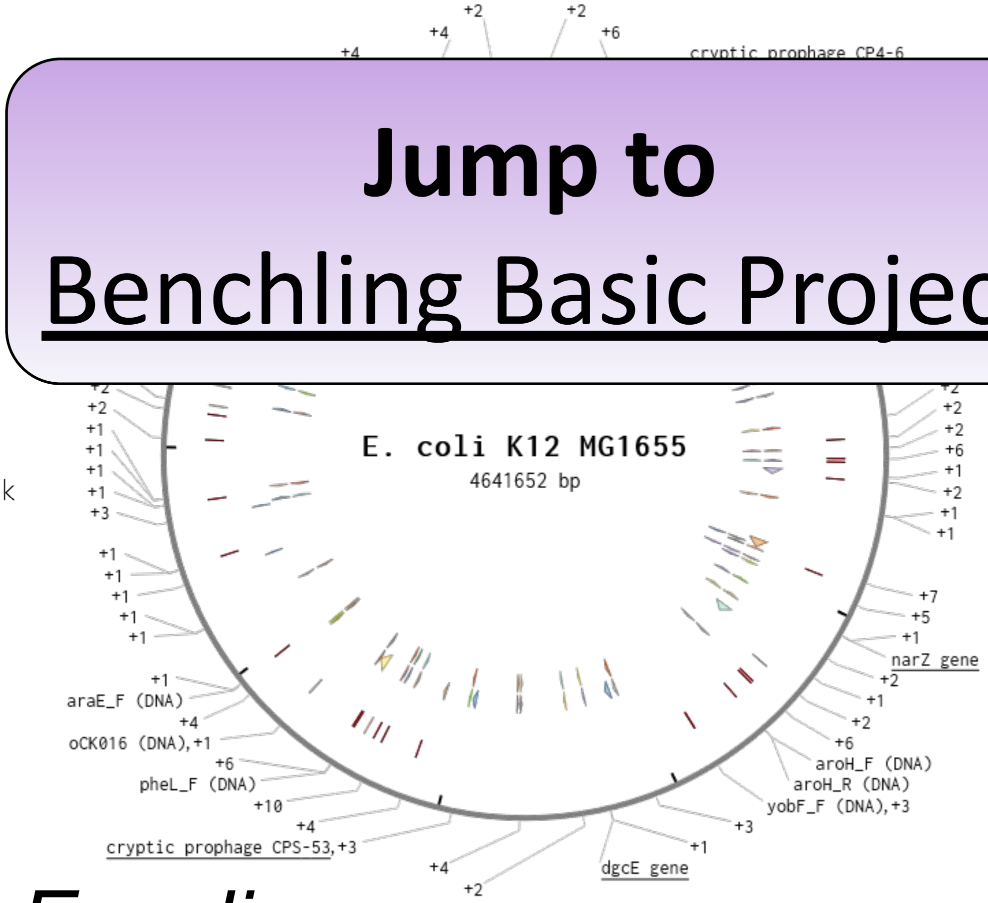
Red

Skipped during the talk

sgRNAs can be used to program phenotypes



Try knocking out ampC gene ---ampicillin resistance



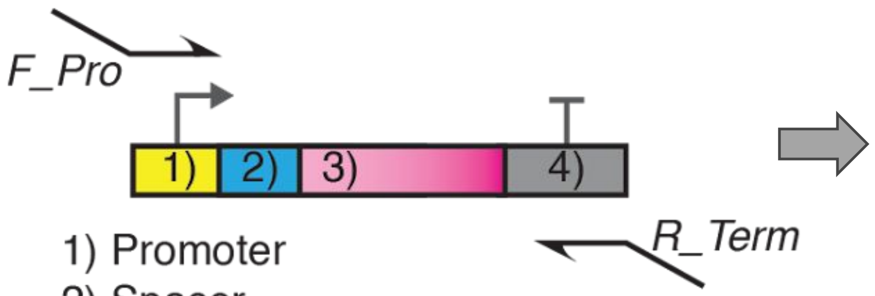
Jump to

Benchling Basic Project

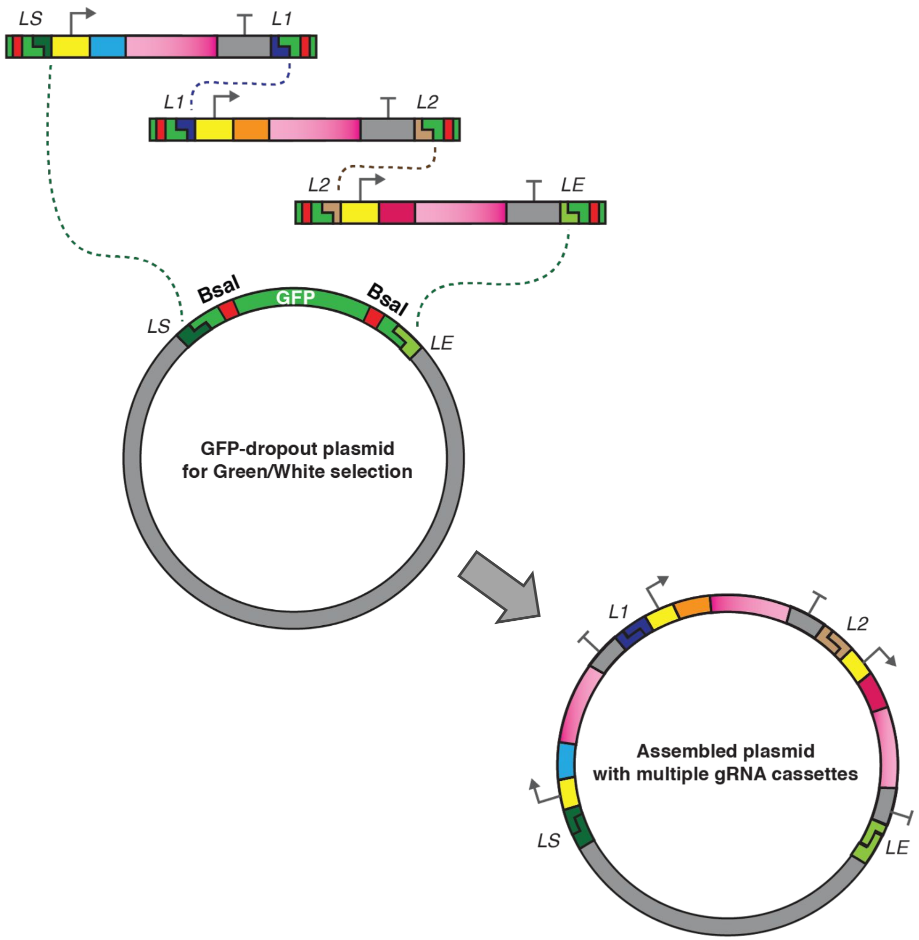
E. coli genome: GenBank: U00096.3

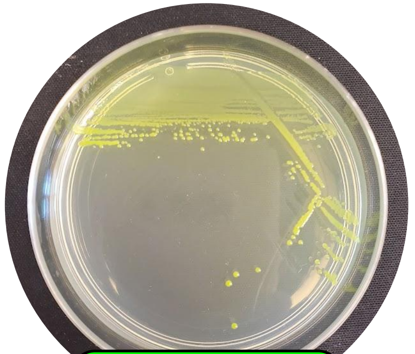
Skipped during the talk

Multiple gRNA cassettes by Golden-Gate



1) Promoter
2) Spacer
3) gRNA hairpin
4) Terminator





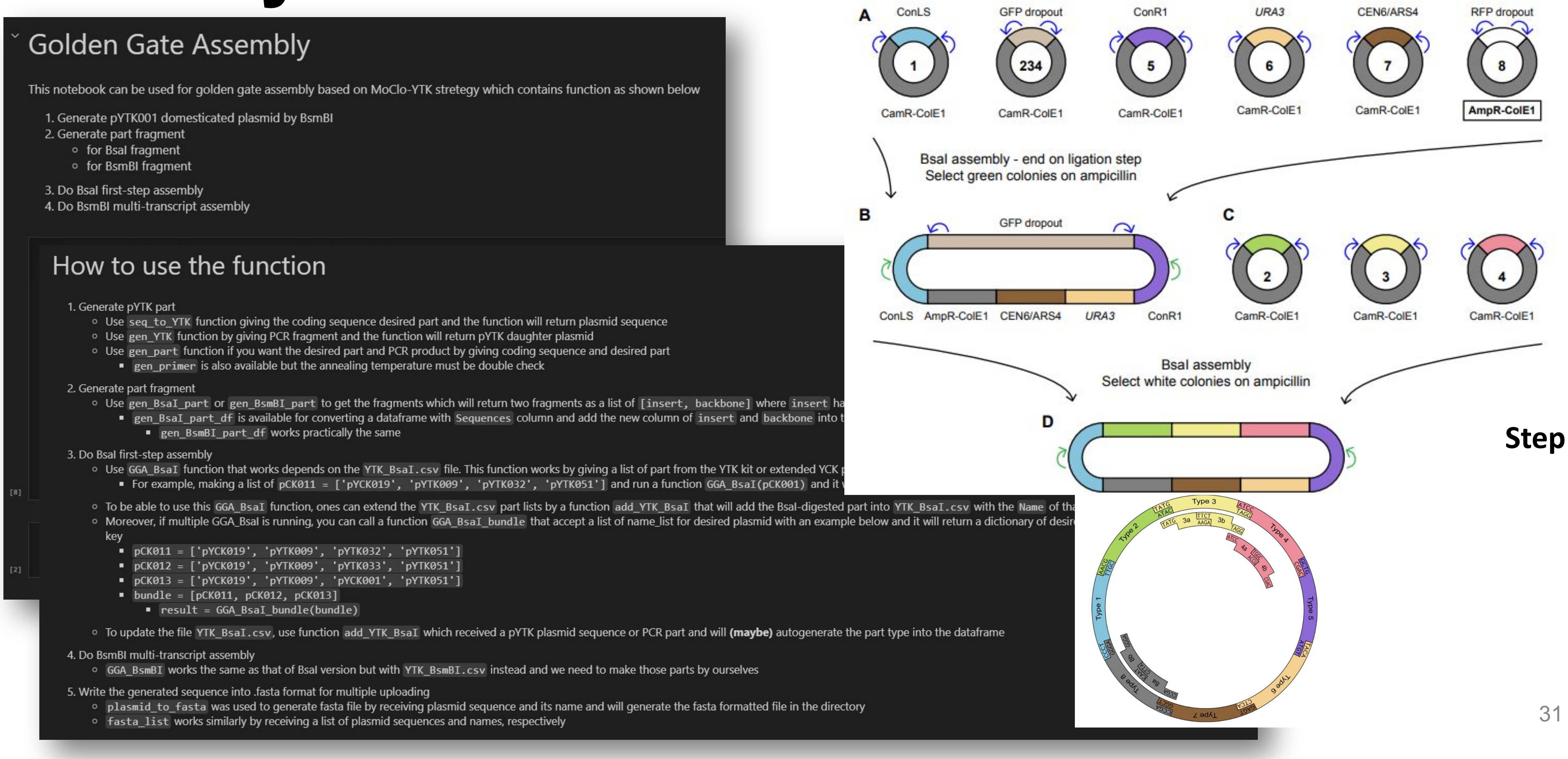
Green

Colorless

3 gRNAs in one cloning reaction

Skipped during the talk

ate Assembly with BioPython



Golden Gate Assembly

This notebook can be used for golden gate assembly based on MoClo-YTK stretegy which contains function as shown below

1. Generate pYTK001 domesticated plasmid by BsmBI
2. Generate part fragment
   - for BsaI fragment
   - for BsmBI fragment
3. Do BsaI first-step assembly
4. Do BsmBI multi-transcript assembly

How to use the function

1. Generate pYTK part
   - Use `seq_to_YTK` function giving the coding sequence desired part and the function will return plasmid sequence
   - Use `gen_YTK` function by giving PCR fragment and the function will return pYTK daughter plasmid
   - Use `gen_part` function if you want the desired part and PCR product by giving coding sequence and desired part
     - `gen_primer` is also available but the annealing temperature must be double check
2. Generate part fragment
   - Use `gen_BsaI_part` or `gen_BsmBI_part` to get the fragments which will return two fragments as a list of `[insert, backbone]` where `insert` ha
     - `gen_BsaI_part_df` is available for converting a dataframe with `Sequences` column and add the new column of `insert` and `backbone` into t
       - `gen_BsmBI_part_df` works practically the same
3. Do BsaI first-step assembly
   - Use `GGA_BsaI` function that works depends on the `YTK_BsaI.csv` file. This function works by giving a list of part from the YTK kit or extended YCK p
     - For example, making a list of `pCK011 = ['pYCK019', 'pYTK009', 'pYTK032', 'pYTK051']` and run a function `GGA_BsaI(pCK001)` and it w
   - To be able to use this `GGA_BsaI` function, ones can extend the `YTK_BsaI.csv` part lists by a function `add_YTK_BsaI` that will add the BsaI-digested part into `YTK_BsaI.csv` with the `Name` of tha
   - Moreover, if multiple GGA_BsaI is running, you can call a function `GGA_BsaI_bundle` that accept a list of name_list for desired plasmid with an example below and it will return a dictionary of desir key
     - `pCK011 = ['pYCK019', 'pYTK009', 'pYTK032', 'pYTK051']`
     - `pCK012 = ['pYCK019', 'pYTK009', 'pYTK033', 'pYTK051']`
     - `pCK013 = ['pYCK019', 'pYTK009', 'pYCK001', 'pYTK051']`
     - `bundle = [pCK011, pCK012, pCK013]`
       - `result = GGA_BsaI_bundle(bundle)`
   - To update the file `YTK_BsaI.csv`, use function `add_YTK_BsaI` which received a pYTK plasmid sequence or PCR part and will **(maybe)** autogenerate the part type into the dataframe
4. Do BsmBI multi-transcript assembly
   - `GGA_BsmBI` works the same as that of BsaI version but with `YTK_BsmBI.csv` instead and we need to make those parts by ourselves
5. Write the generated sequence into .fasta format for multiple uploading
   - `plasmid_to_fasta` was used to generate fasta file by receiving plasmid sequence and its name and will generate the fasta formatted file in the directory
   - `fasta_list` works similarly by receiving a list of plasmid sequences and names, respectively

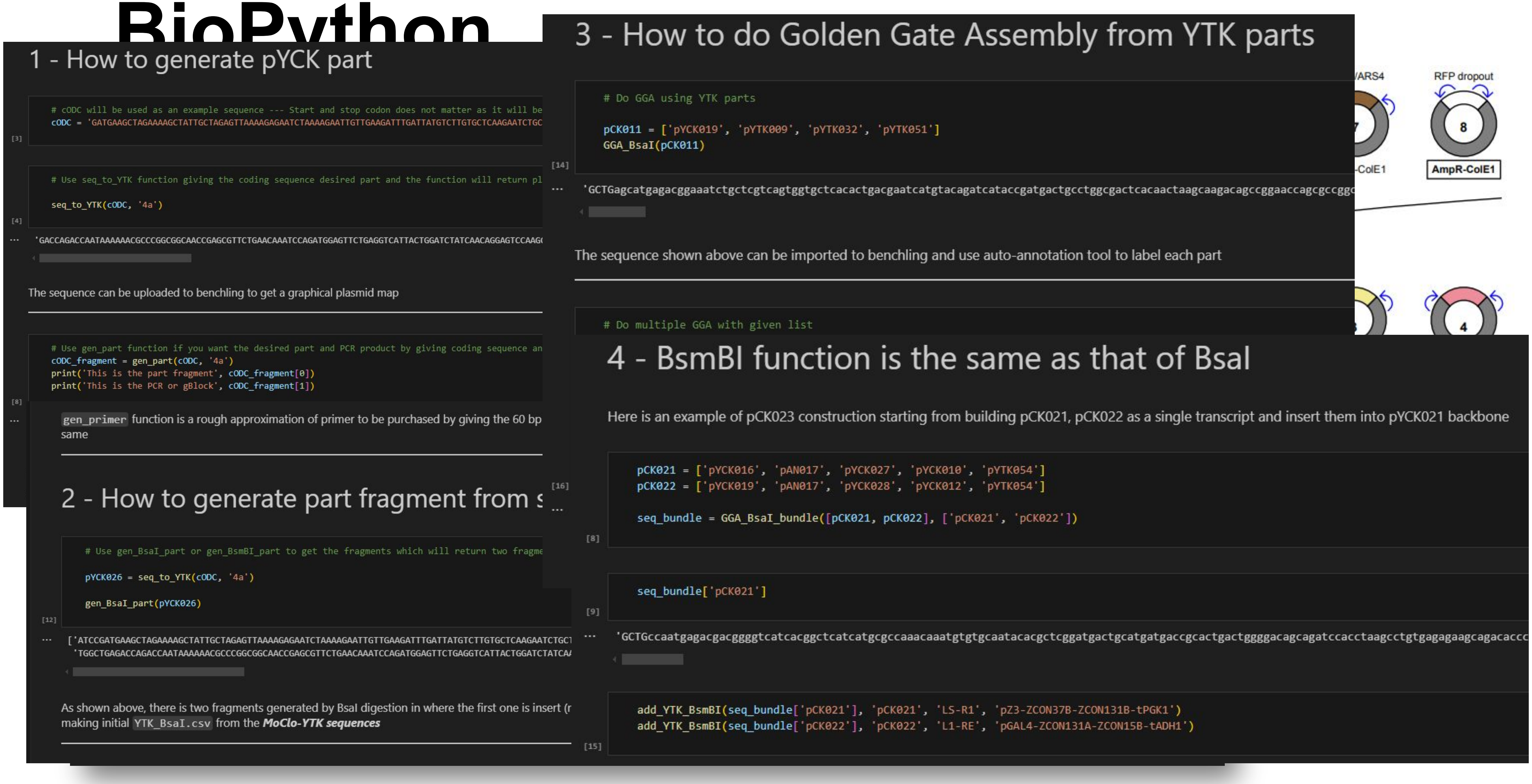
Step 1&2

Step 3

Step 4

Skipped during the talk

ate Assembly with BioPython



1 - How to generate pYCK part

```
# cODC will be used as an example sequence --- Start and stop codon does not matter as it will be
cODC = 'GATGAAGCTAGAAAAGCTATTGCTAGAGTTAAAAGAGAATCTAAAAGAATTGTTGAAGATTTGATTATGTCTTGTGCTCAAGAATCTGC
```
[3]

```
# Use seq_to_YTK function giving the coding sequence desired part and the function will return pl

seq_to_YTK(cODC, '4a')
```
[4]

```
'GACCAGACCAATAAAAAACGCCCGGCGGCAACCGAGCGTTCTGAACAAATCCAGATGGAGTTCTGAGGTCATTACTGGATCTATCAACAGGAGTCCAAG
```

The sequence can be uploaded to benchling to get a graphical plasmid map

```
# Use gen_part function if you want the desired part and PCR product by giving coding sequence an
cODC_fragment = gen_part(cODC, '4a')
print('This is the part fragment', cODC_fragment[0])
print('This is the PCR or gBlock', cODC_fragment[1])
```
[8]

`gen_primer` function is a rough approximation of primer to be purchased by giving the 60 bp same

2 - How to generate part fragment from s

```
# Use gen_BsaI_part or gen_BsmBI_part to get the fragments which will return two fragme

pYCK026 = seq_to_YTK(cODC, '4a')

gen_BsaI_part(pYCK026)
```
[12]

```
['ATCCGATGAAGCTAGAAAAGCTATTGCTAGAGTTAAAAGAGAATCTAAAAGAATTGTTGAAGATTTGATTATGTCTTGTGCTCAAGAATCTGCT
 'TGGCTGAGACCAGACCAATAAAAAACGCCCGGCGGCAACCGAGCGTTCTGAACAAATCCAGATGGAGTTCTGAGGTCATTACTGGATCTATCAA
```

As shown above, there is two fragments generated by BsaI digestion in where the first one is insert (r making initial `YTK_BsaI.csv` from the ***MoClo-YTK sequences***

3 - How to do Golden Gate Assembly from YTK parts

```
# Do GGA using YTK parts

pCK011 = ['pYCK019', 'pYTK009', 'pYTK032', 'pYTK051']
GGA_BsaI(pCK011)
```
[14]

```
'GCTGagcatgagacggaaatctgctcgtcagtggtgctcacactgacgaatcatgtacagatcataccgatgactgcctggcgactcacaactaagcaagacagccggaaccagcgccggc
```

The sequence shown above can be imported to benchling and use auto-annotation tool to label each part

```
# Do multiple GGA with given list
```
[16]

4 - BsmBI function is the same as that of BsaI

Here is an example of pCK023 construction starting from building pCK021, pCK022 as a single transcript and insert them into pYCK021 backbone

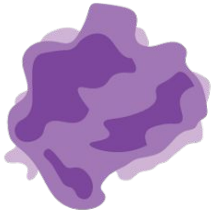
```
pCK021 = ['pYCK016', 'pAN017', 'pYCK027', 'pYCK010', 'pYTK054']
pCK022 = ['pYCK019', 'pAN017', 'pYCK028', 'pYCK012', 'pYTK054']

seq_bundle = GGA_BsaI_bundle([pCK021, pCK022], ['pCK021', 'pCK022'])
```
[8]

```
seq_bundle['pCK021']
```
[9]

```
'GCTGccaatgagacgacggggtcatcacggctcatcatgcgccaaacaaatgtgtgcaatacacgctcggatgactgcatgatgaccgcactgactggggacagcagatccacctaagcctgtgagagaagcagacaccc
```

```
add_YTK_BsmBI(seq_bundle['pCK021'], 'pCK021', 'LS-R1', 'pZ3-ZCON37B-ZCON131B-tPGK1')
add_YTK_BsmBI(seq_bundle['pCK022'], 'pCK022', 'L1-RE', 'pGAL4-ZCON131A-ZCON15B-tADH1')
```
[15]

Key Takeaways



- Use Molecular Biology tools like Benchling or Snapgene for DNA/RNA/Protein works: Read, Compare, Edit, or Design

- Benchling is a simple *in silico* molecular biology tools and is free for academic users

- *In silico* operations can save a lot of experimental time, especially when the procedure didn’t go as planned

Resources

- https://benchling.com/
- https://www.snapgene.com/snapgene-viewer
- https://www.geneious.com/guides/molecular-cloning-methods
- ApE, A Plasmid Editor: A Freely Available DNA Manipulation and Visualization Program
- https://www.addgene.org/mol-bio-reference/cloning/
- https://biopython.org/
- http://plannotate.barricklab.org/
    - For plasmid annotation