



AFRICAN CENTERS OF EXCELLENCE IN BIOINFORMATICS

KAMPALA, UGANDA

Molecular Biology: a tour on virtual cloning and sequence editing

Today's Instructor



Dr. Mariam Quiñones,
Ph.D. in Molecular
Microbiology

Ongoing Computational
Biology projects:

- 16S microbiome/WGS
- Parasite strain evolution

- Bioinformatics and Computational Biosciences Branch (BCBB), NIAID
- National Institutes of Health, Bethesda, MD USA.
- Contact our team via email:
 - Email: bioinformatics@niaid.nih.gov
 - Instructor: mariam.quinones@nih.gov



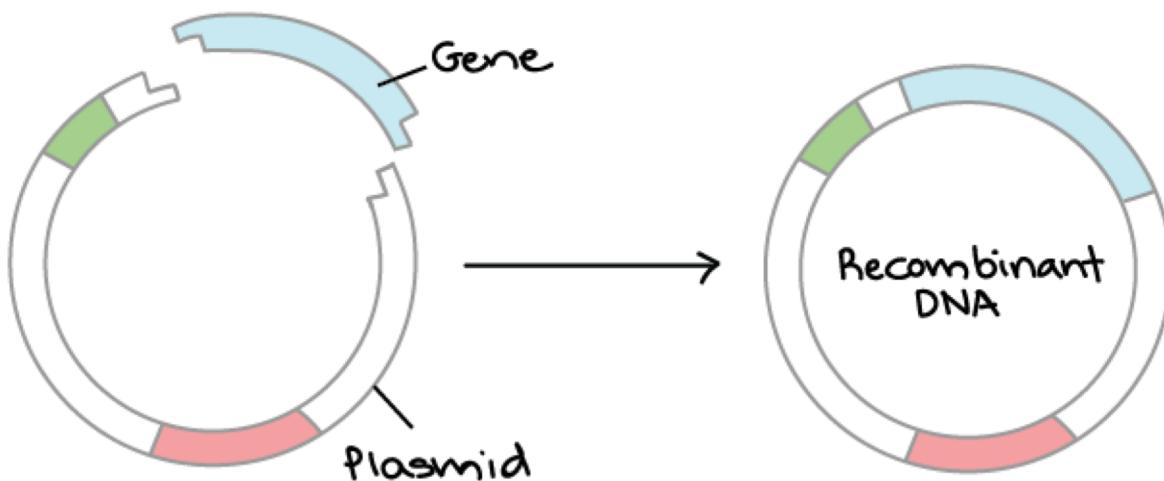
Agenda

- Introduction to cloning using plasmid vectors
- Exercise on cloning:
 - Topo TA cloning
 - Expression vector pET28b(+) – novagen
- Exercise on viewing Sanger sequencing files

DNA Cloning

What is it?

- Making multiple copies of a particular DNA fragment such as a gene
- The DNA is usually cloned by inserting a fragment into a vector that usually a bacteria can replicate



Figures taken from Khan Academy <<https://www.khanacademy.org>>



Major Types of Cloning Vectors

- **Plasmid.** Circular extrachromosomal DNA that autonomously replicates inside the bacterial cell. They can be low or high copy number.
- **Phage.** Linear DNA molecules derived from bacteriophage lambda. Can be replaced with foreign DNA without disrupting its life cycle.
- **Phagemids.** It replicates as a plasmid and gets packaged in the form of single-stranded DNA in viral particles. Incorporates F1 origin of replication but it is used in combination with M13 phage.
- **Cosmids.** Another circular extrachromosomal DNA molecule that combines features of plasmids and phage (lambda). They can carry up to 45 kb of insert
- **Bacterial Artificial Chromosomes.** Based on bacterial mini-F plasmids. Can carry DNA insert up to 350kb.
- **Yeast Artificial Chromosomes.** This is an artificial chromosome that contains telomeres with origins of replication, a yeast centromere (part of a chromosome that links sister chromatids or a dyad), and a selectable marker for identification in yeast cells. Used for cloning into eukaryotic cells.
- **Human Artificial Chromosome.** This type of vector is potentially useful for gene delivery into human cells, and a tool for expression studies and determining human chromosome function. It can carry very large DNA fragment.
- **Retroviral Vectors.** Retroviral vectors are used for introduction of novel or manipulated genes into the animal or human cells. The viral RNA is converted into DNA with the help of reverse transcriptase and henceforth, efficiently integrated into the host cell.



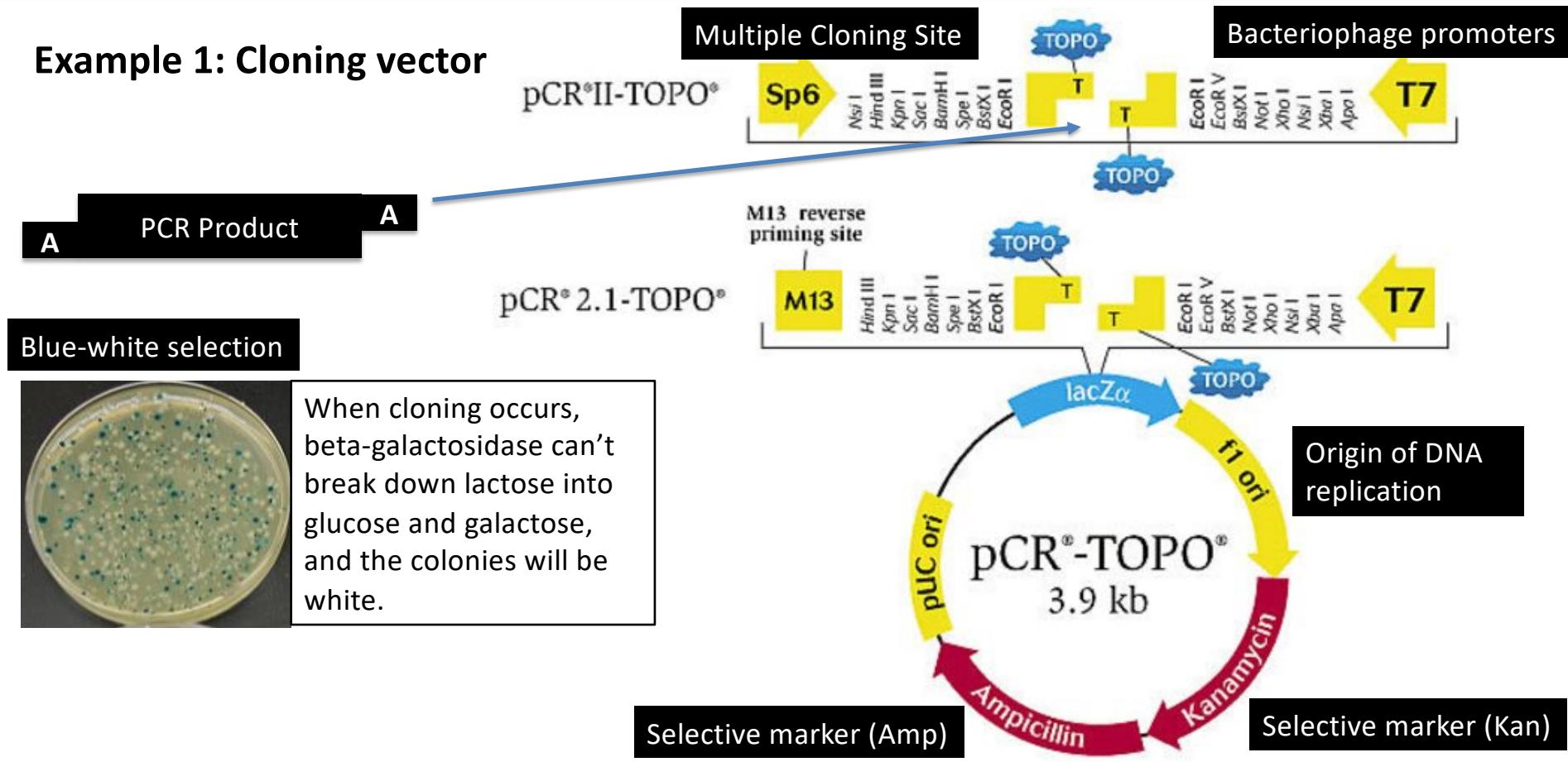
Popular Cloning Techniques

- **Gateway** – transfer DNA between plasmids
- **Gibson Assembly** – allows multiple fragments to be inserted
- **GeneArt** – assembles fragments and vectors in vitro
- **In-Fusion** – based on annealing of complementary ends of an insert and linearized vector
- **TA** – no need for restriction enzymes. Uses Taq.
- **TOPO*** - could be a TA cloning or Zero BLUNT. contain a lethal ccdB gene that is interrupted in cloning
- **Restriction enzyme*** - insert and vector are digested with compatible enzymes

Important features of a plasmid vector

<https://www.thermofisher.com>

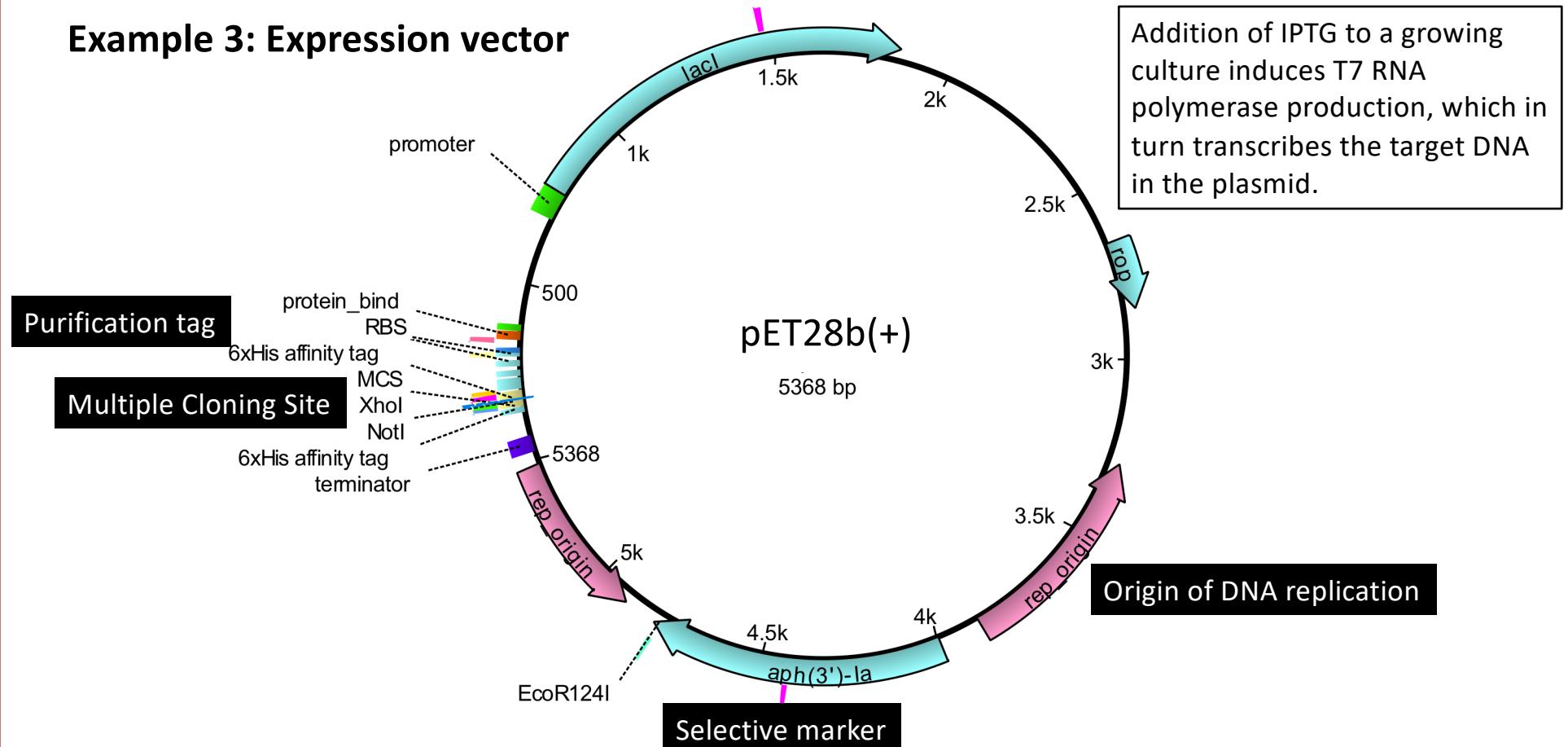
Example 1: Cloning vector



Important features of a plasmid vector

<https://www.novoprolabs.com/vector/V11004>

Example 3: Expression vector



Review: What are restriction enzymes?

- Molecular scissors that cut double stranded DNA molecules at specific points.
- Found naturally in a wide variety of prokaryotes – defense against phages
- An important tool for manipulating DNA.
- Named for bacterial genus, species, strain, and type

Example: EcoR1

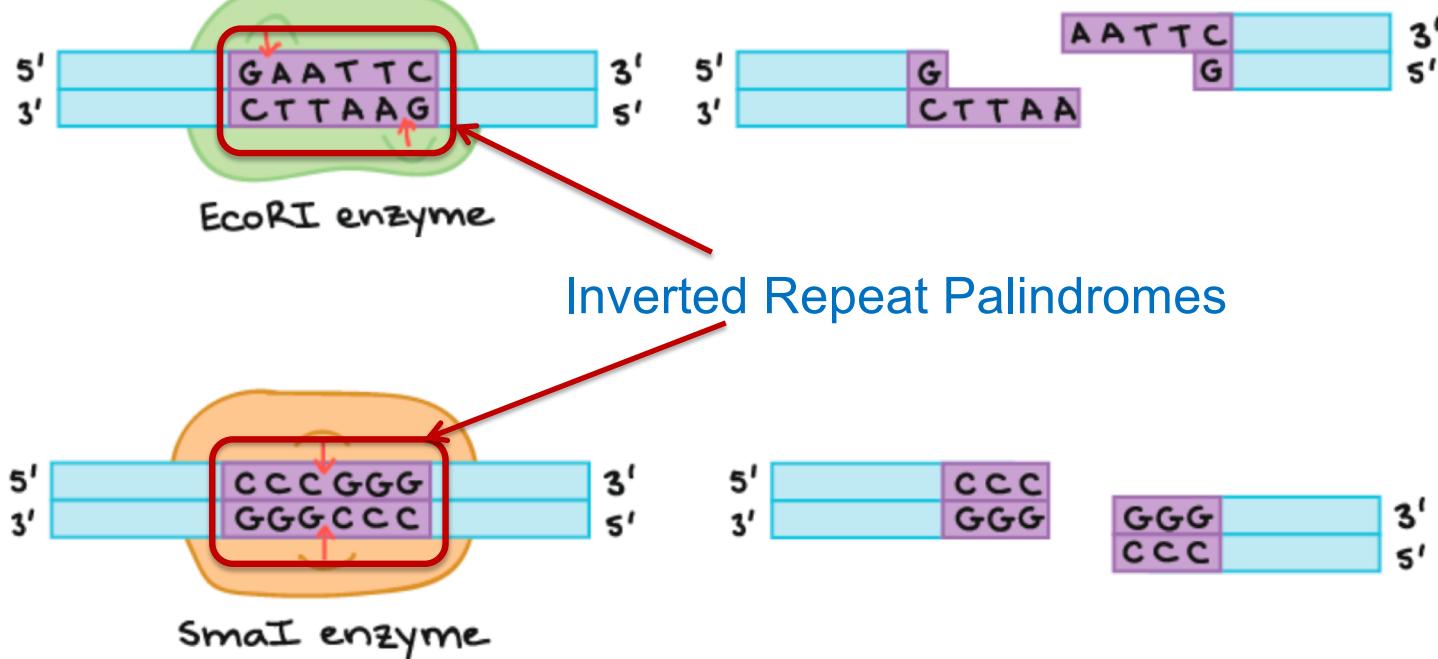
Genus: Escherichia

Species: coli

Strain: R

Order discovered: I (roman numeral 1)

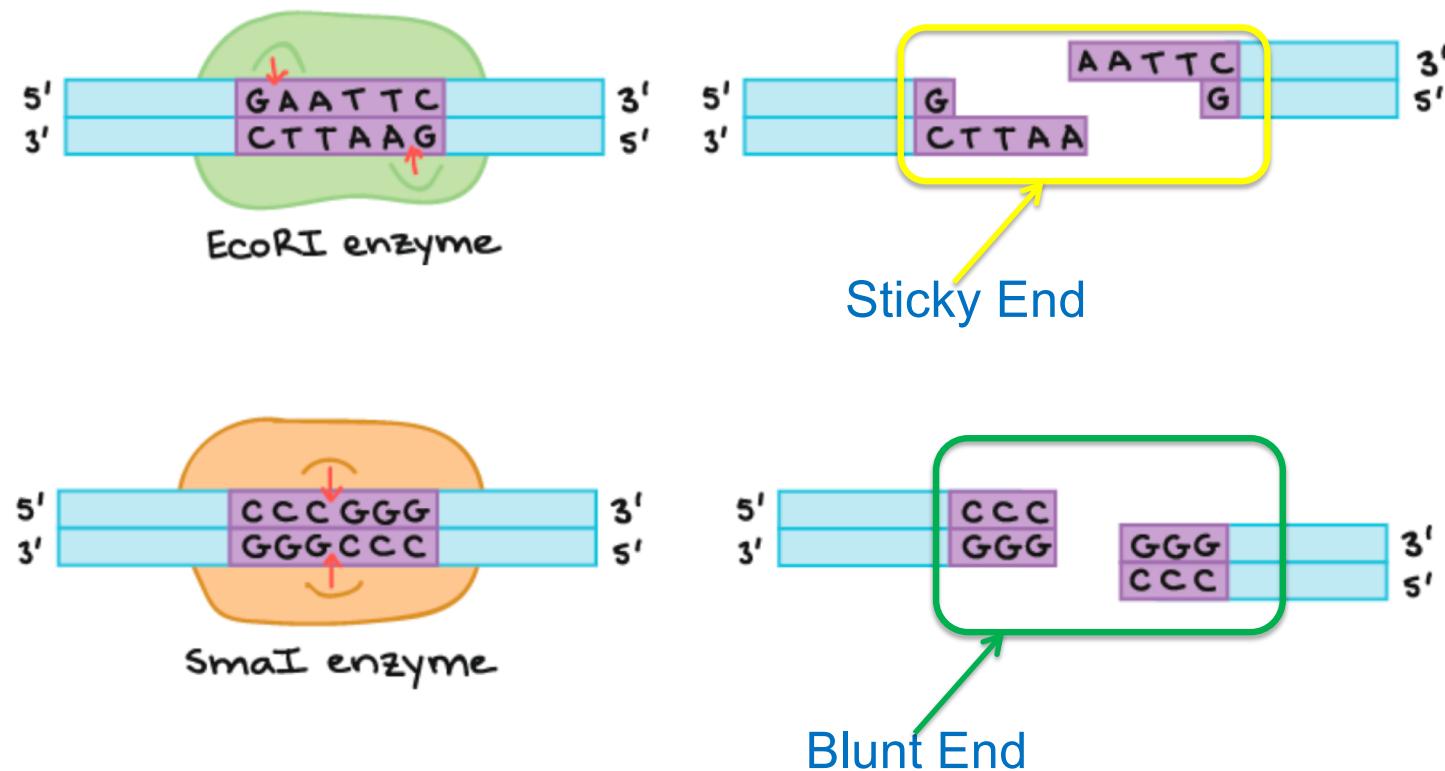
Recognition Sequences



figures from Khan Academy

Originally presented by P. Subramanian

Blunt and Sticky Ends

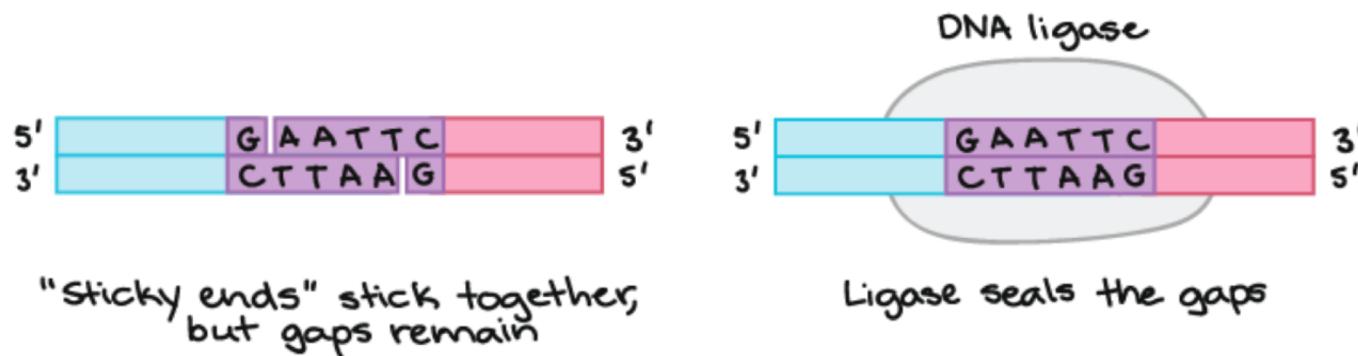


figures from Khan Academy

Originally presented by P. Subramanian

Sticky Ends

- Make it easier for enzyme **DNA ligase** to put the ends back together
- Useful for recombinant DNA and cloning



figures from Khan Academy

Originally presented by P. Subramanian



Exercises

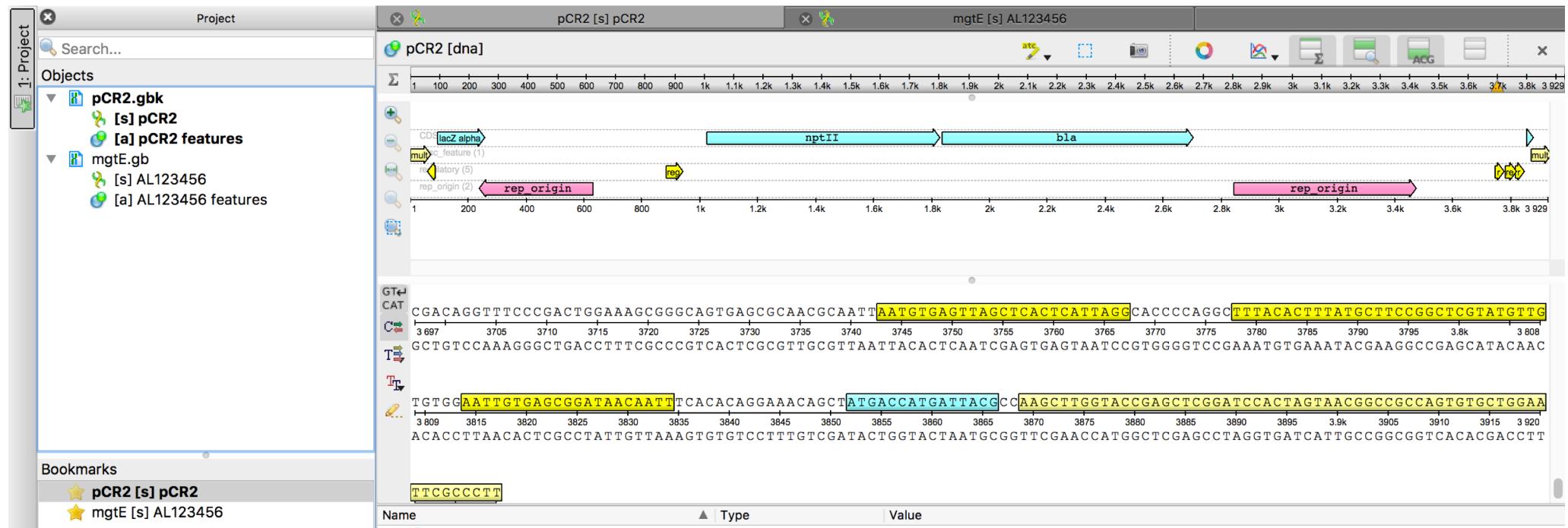
1. Topo cloning of a gene from *M. tuberculosis*.
2. Restriction enzyme cloning of gene EgB8/2 from *Echinococcus granulosus*. This will allow expression of the antigen gene for purification.
3. Aligning and examining Sanger sequences

Reference: <https://ugene.net/wiki/display/UUOUM15/Molecular+Cloning+in+silico>

Exercise 1: Topo cloning of a gene
from *M. tuberculosis*

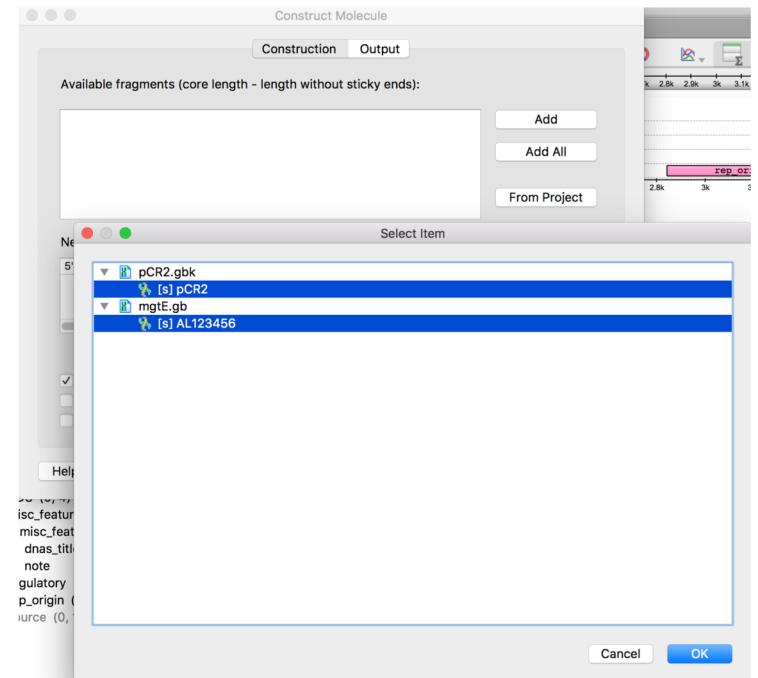
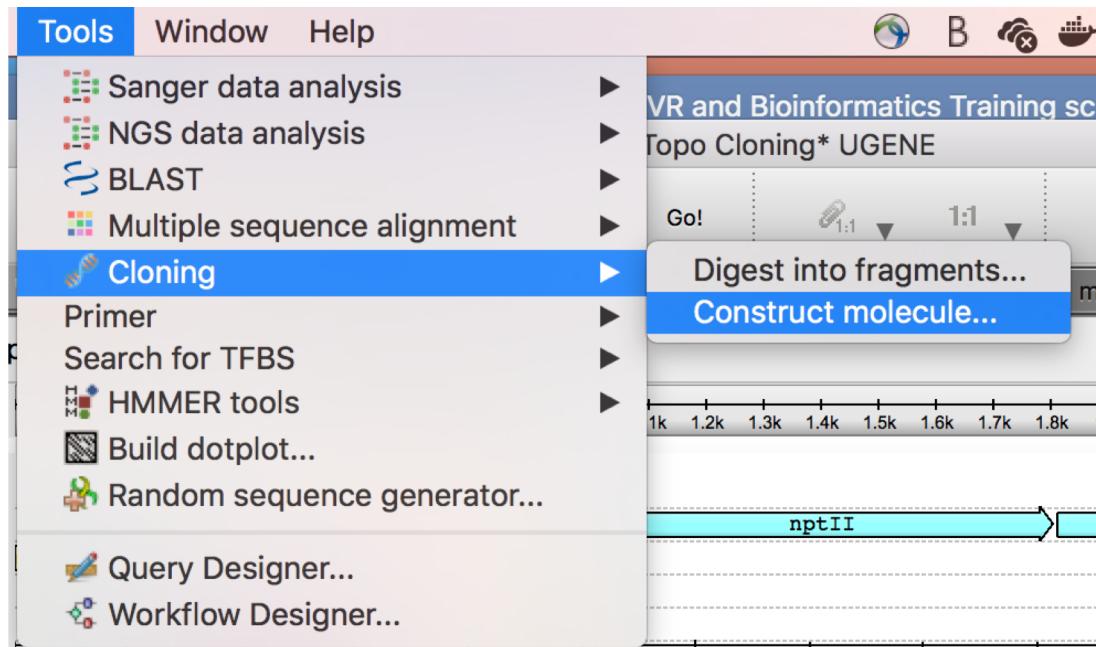
Step 1: Open project Topo.uprj

- Double click on the file with the DNA symbol to activate it
- View gene mgtE (*M. tuberculosis* magnesium transporter) and Topo pCR 2.1-TOPO vector



Step 2: Select option to Construct molecule

Select option “From Project” and select linear molecules



Step 3: Select option to Construct molecule

The image displays two windows from a DNA assembly software:

Create DNA Fragment window:

- Region: Whole sequence
- Start: 1
- End: 3929
- Include Left Overhang: checked
- Include Right Overhang: checked
- Direct (radio button selected)

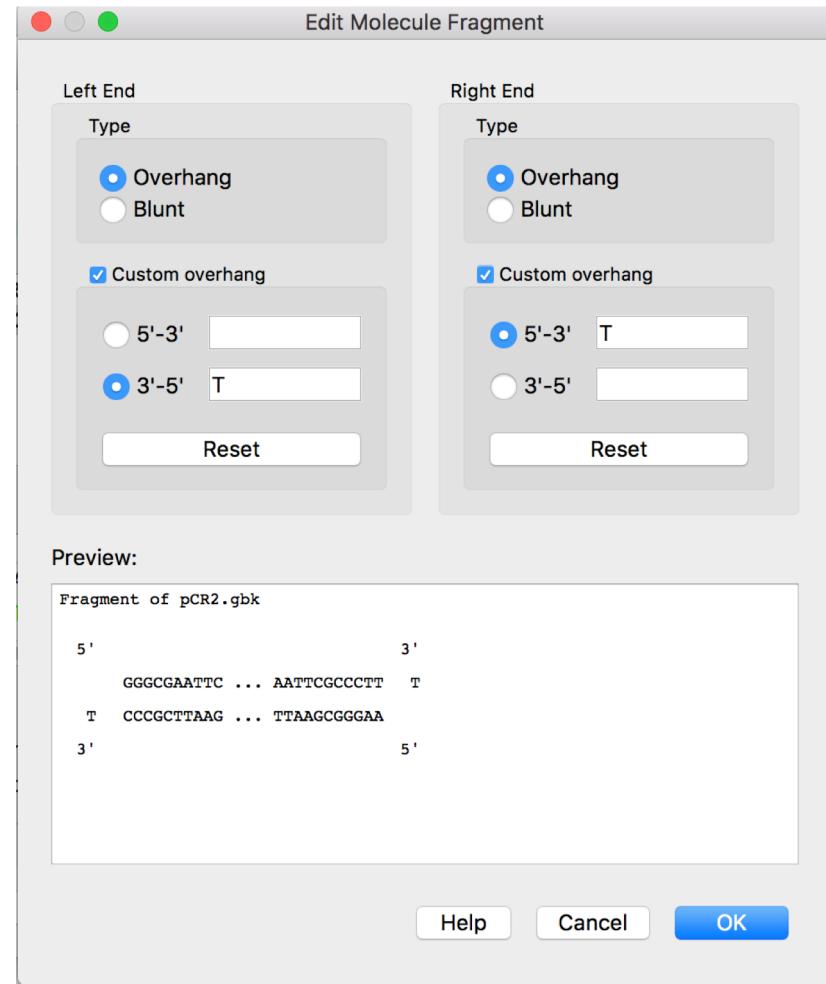
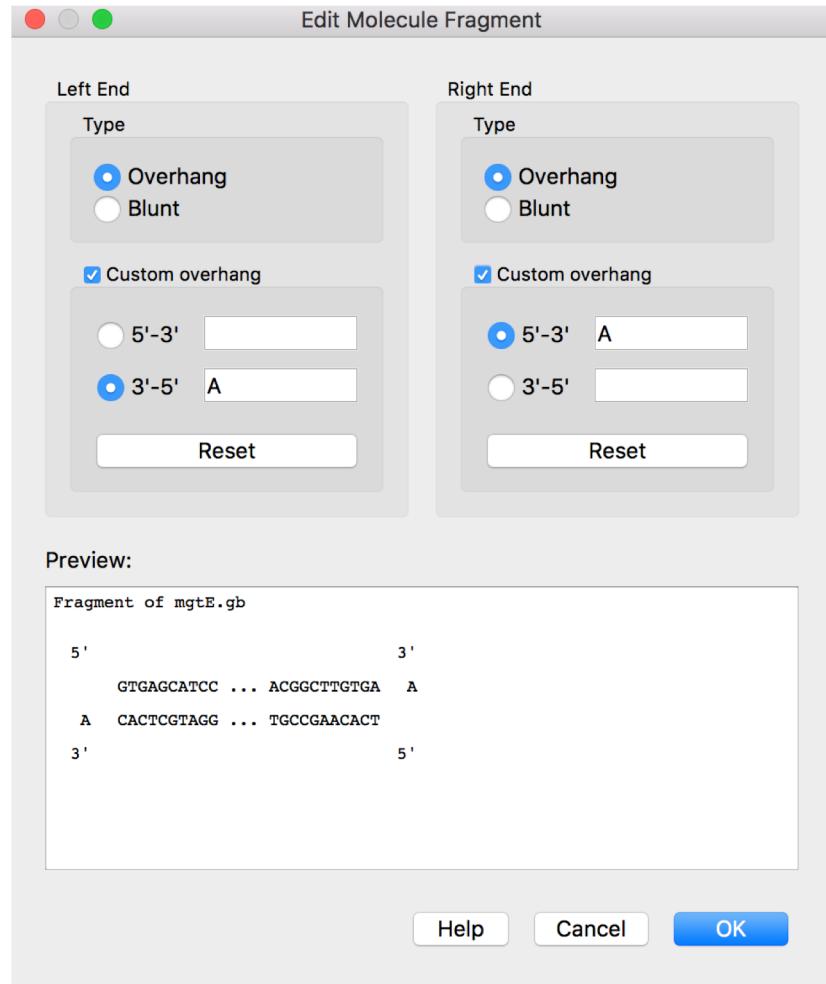
Construct Molecule window:

- Available fragments:
 - pCR2 (pCR2.gbk) Fragment (1-3929)
 - AL123456 (mgtE.gb) Fragment (1-1383)
- Add All (button circled in blue)
- From Project
- New molecule contents:

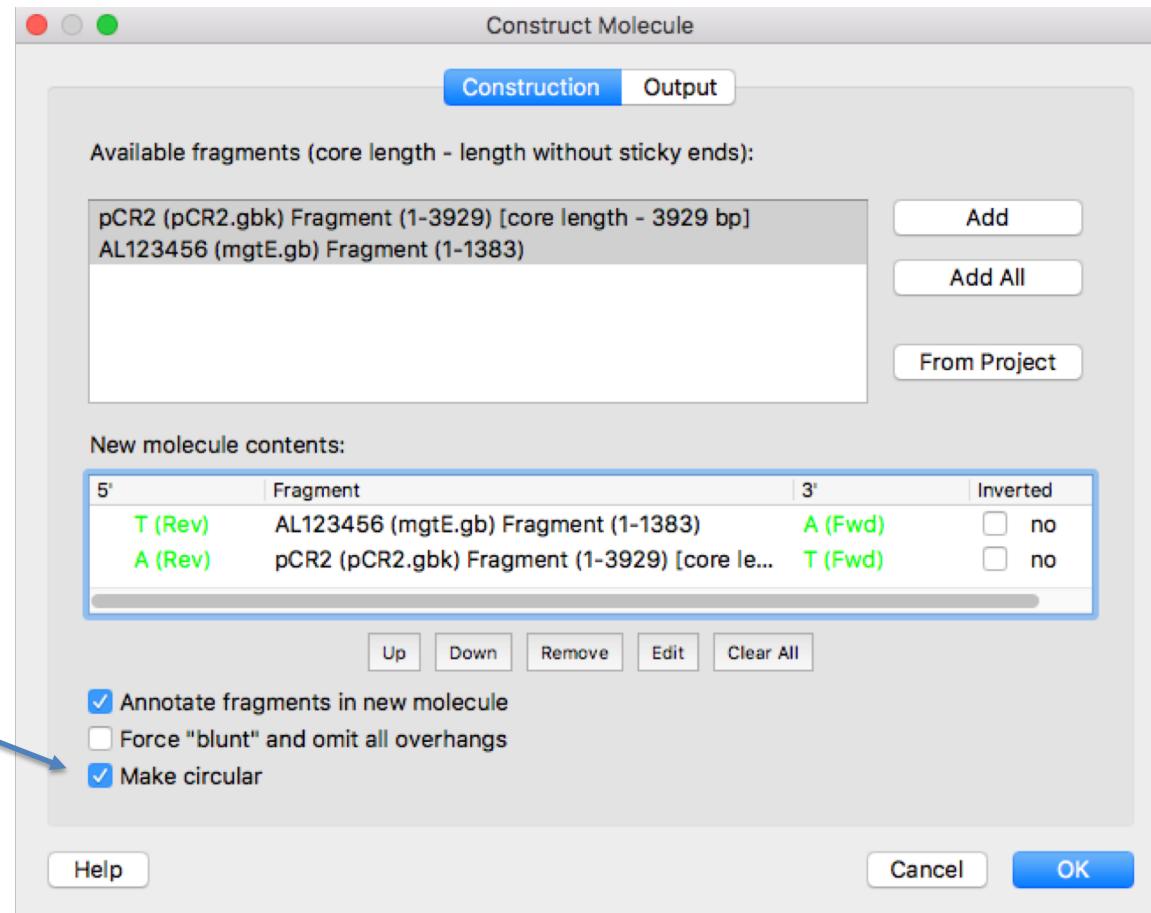
5'	Fragment	3'	Inverted
Blunt	pCR2 (pCR2.gbk) Fragment (1-3929)	Blunt	<input type="checkbox"/> no
Blunt	AL123456 (mgtE.gb) Fragment (1-1383)	Blunt	<input type="checkbox"/> no
- Up, Down, Remove, Edit (button circled in red), Clear All
- Checkboxes:
 - Annotate fragments in new molecule (checked)
 - Force "blunt" and omit all overhangs
 - Make circular
- Help, Cancel, OK

Add molecules, then select to **Edit** (see next slide)

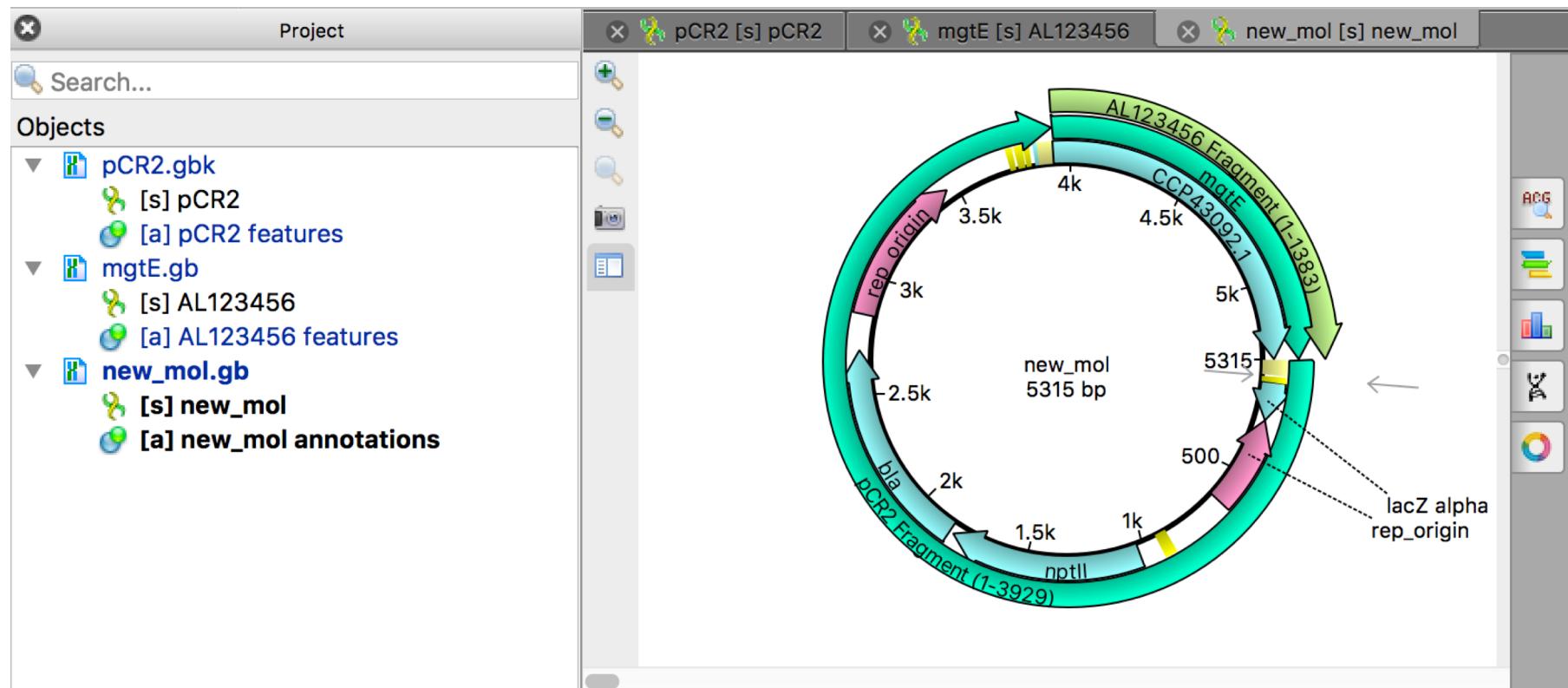
Step 4: Create T overhangs in Vector and A overhangs on insert.



Step 4: Create T overhangs in Vector and A overhangs on insert.



A new molecule is created with the mgtE insert



Exercise 2: Restriction enzyme
cloning of gene EgB8/2 from
Echinococcus granulosus

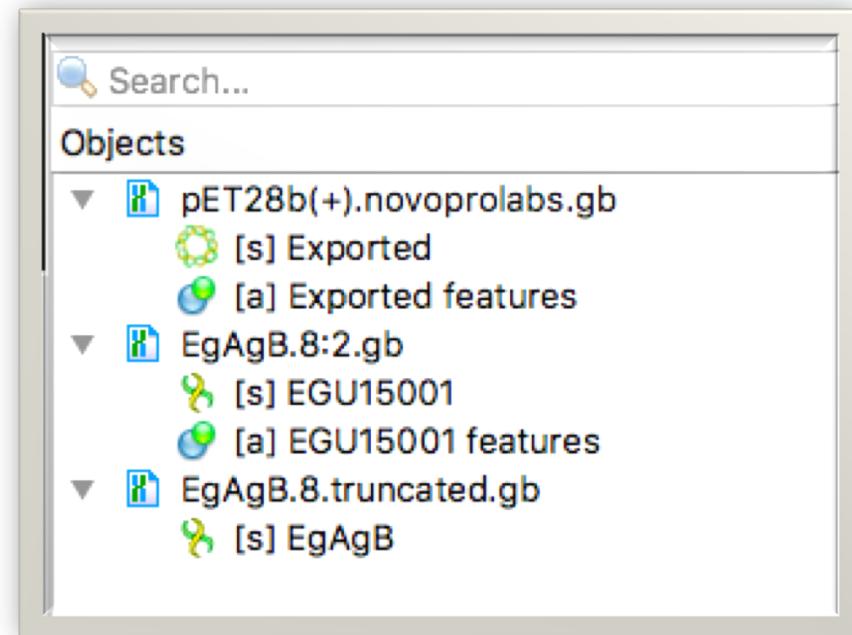
[https://www.ncbi.nlm.nih.gov/pmc
/articles/PMC6396996/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6396996/)

Step 1: Prepare the fragments to be combined

Upload project: projectRE.uprj

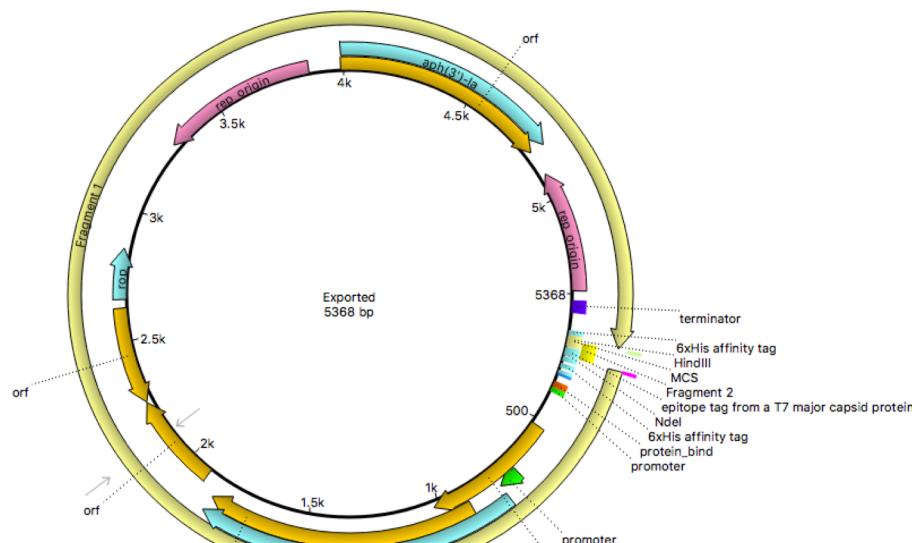
Files:

- 1- pET28b is the vector.
- 2- EgAgB.8 is an antigen used to detect hydatidosis
- 3- EgAgB.8.truncated is a subset of the gene that is truncated to remove the signal peptide. It was also edited to contain NdeI in the 5' and HindIII in the 3' end.

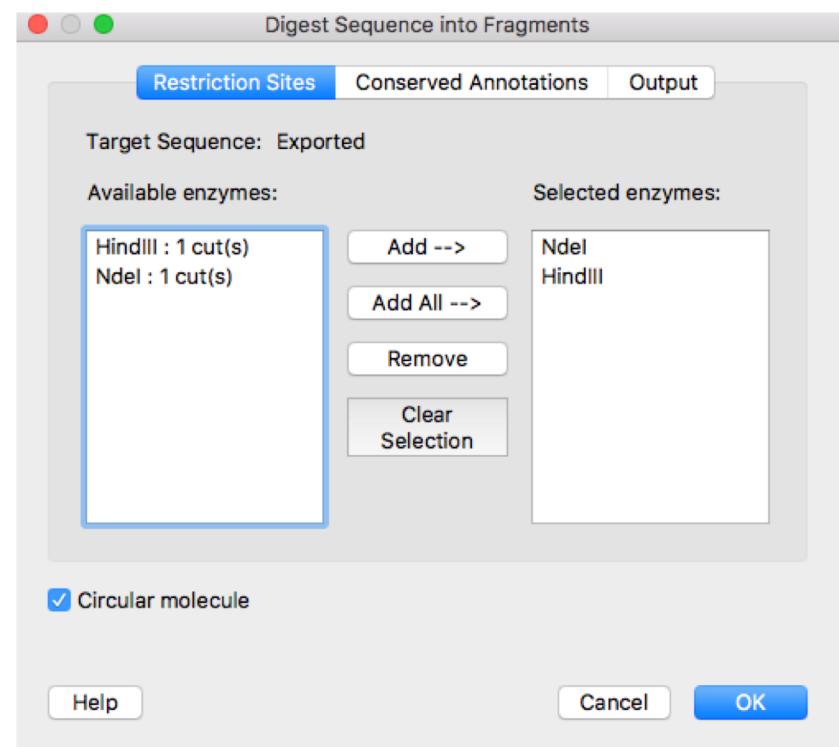


Step 2: Digest the vector and fragments

1- Select the circular molecule by double click, then use menu **Tools** → **Cloning** → **Digest into fragments**. Click OK and two fragments will appear

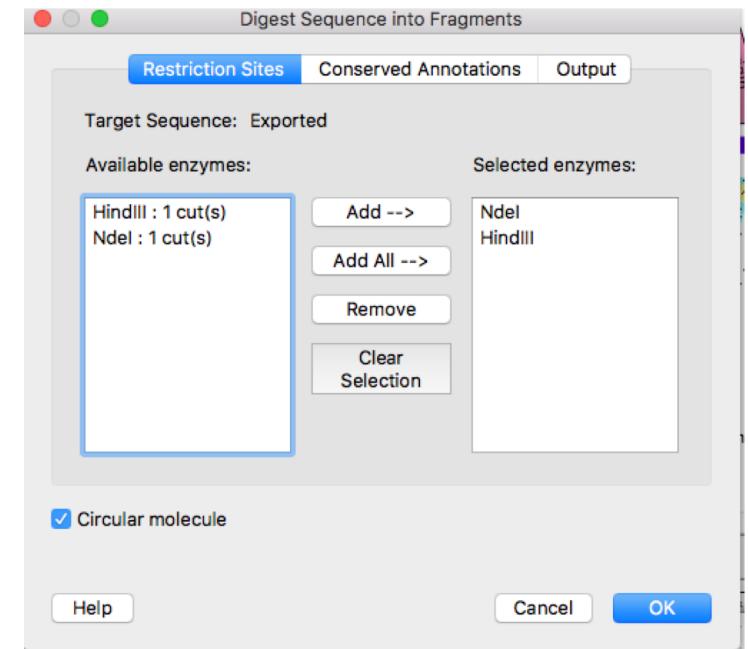


Expected output after trimming by NdeI and HindIII

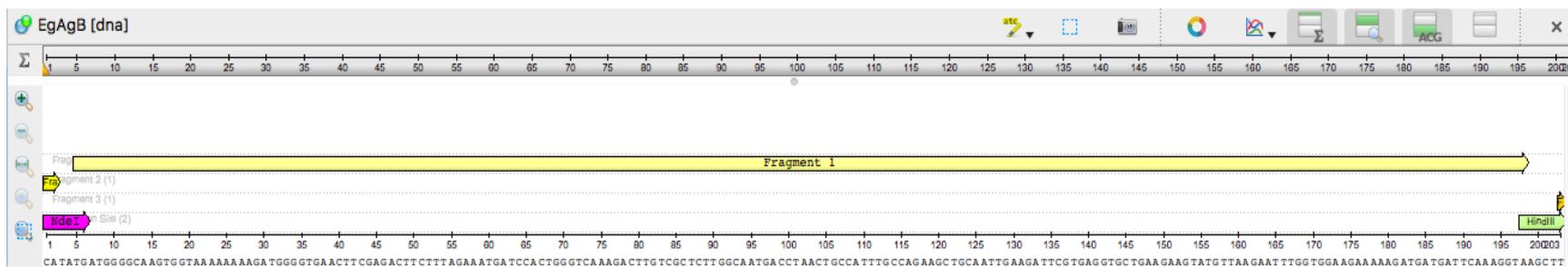


Step 3: Digest the insert to make sticky ends compatible with vector

- 1- Select the EgAgB.8.truncated linear molecule by double click, then use menu **Tools → Cloning → Digest into fragments.** The output should look as the figure below.

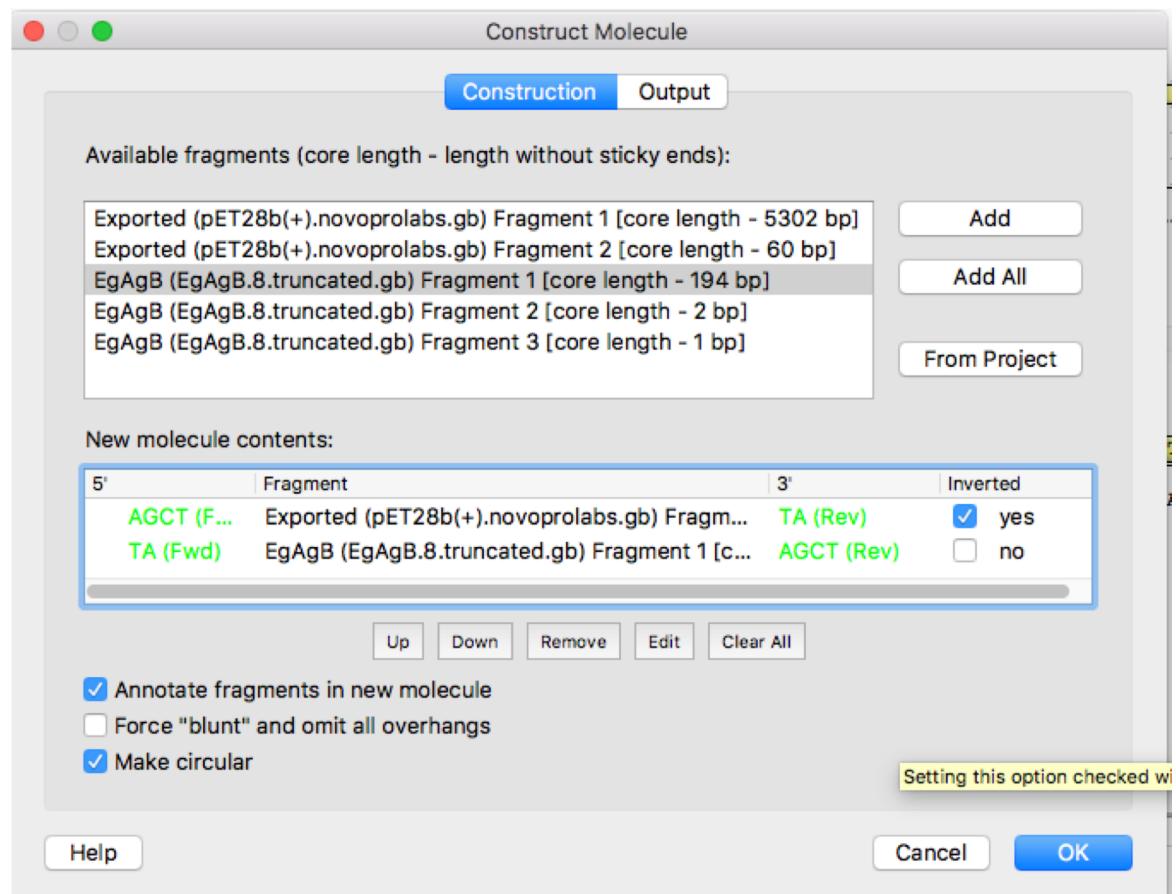


Expected output after trimming by NdeI and HindIII



Step 4: Clone

- 1- Select option **Tools → Cloning → Construct Molecule.**
- 2- Select the large fragments (plasmid and insert) and Add. Proceed to invert the vector and to select “Make circular”.



Step 4: Examine output

1- Output should be in frame with the 6xHis affinity tag



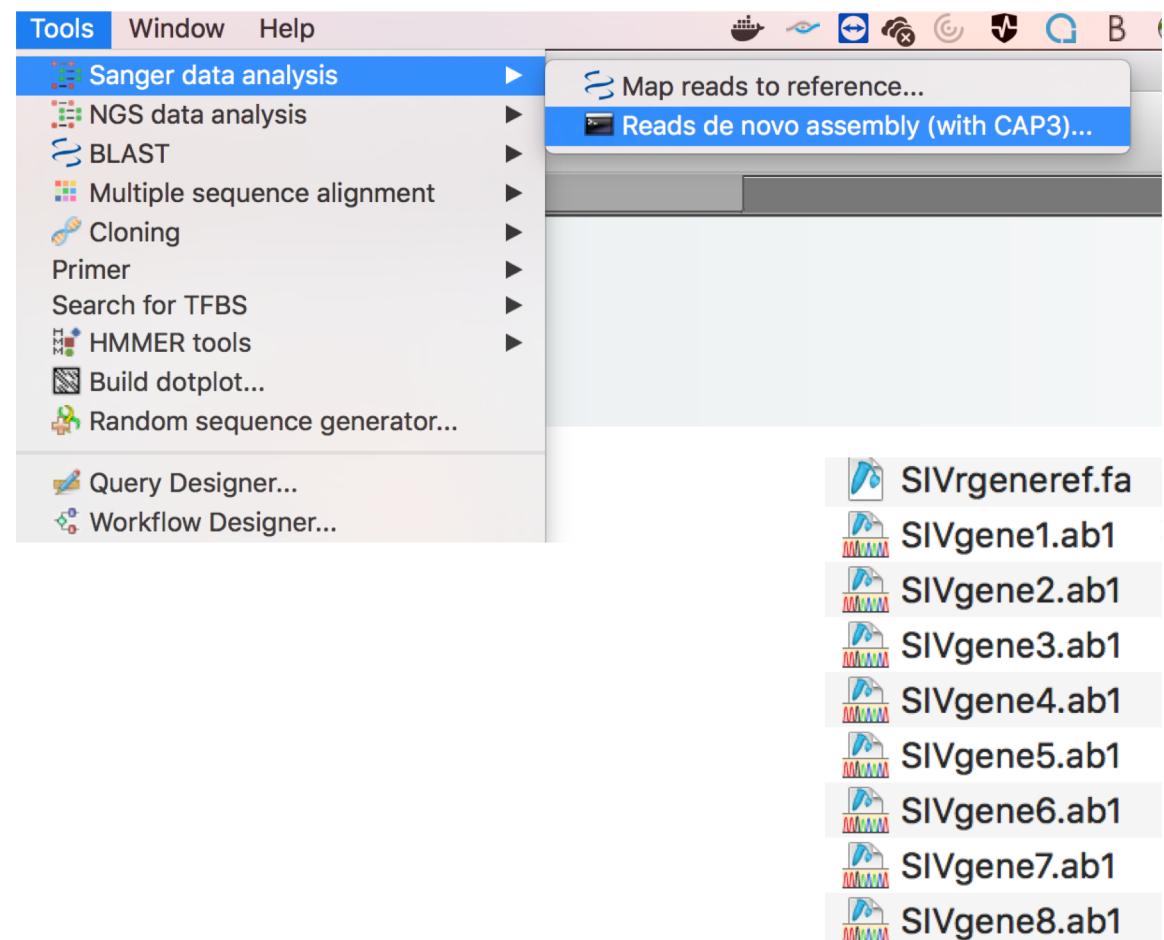
Exercise 3: Aligning and examining Sanger sequences

Note: This demo dataset is not related to the files in Exercises 1 and 2 but it will show the user how to view and align Sanger files (.ab1) using UGENE

Step 1: Select Tools -->
Sanger data analysis →
Reads de novo assembly

Step 2: Navigate to the folder with 8 ".abi" files and one fasta file (reference)

Step 3: Select files, click Open, then click "Run"



Step 4: Decide if to view the Alignment in the Editor or the Browser

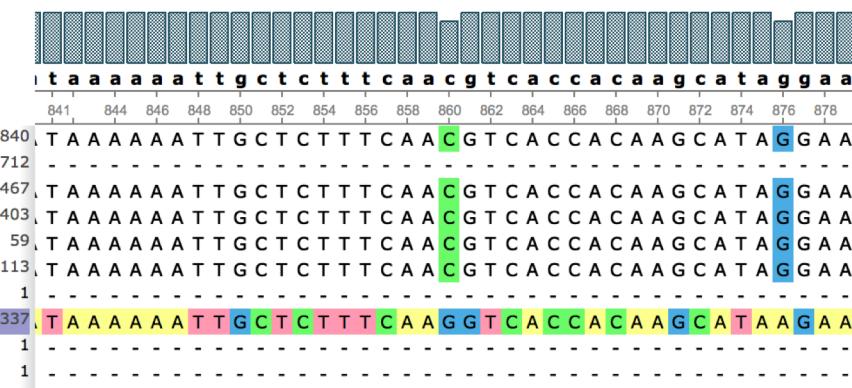
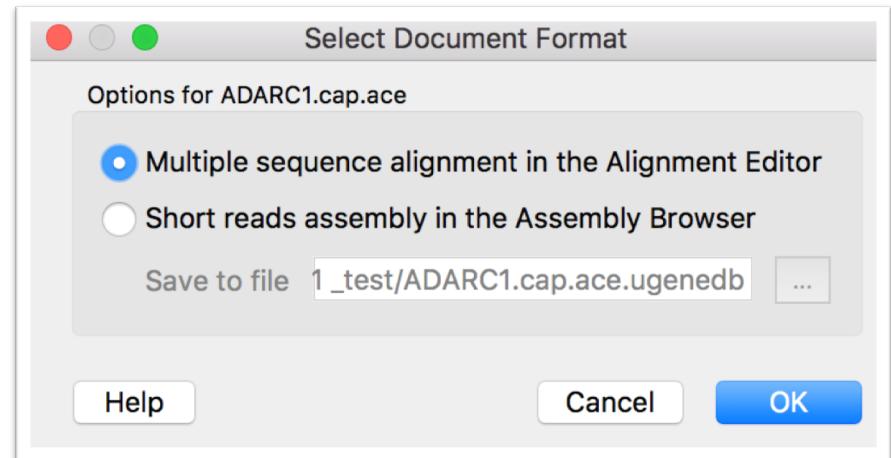
Step 5: Color the contig to find ambiguities related to the reference file. This can be done by selecting the reference, then picking the option “Set this sequence as reference”

002_14664
HIVSFAAA_1382_bases_5822 checksum
002_14668(rev-compl)
002_14666

Go to position...
Add
Copy/Paste
Colors
Highlighting
Edit
Align
Tree
Statistics
View
Export
Advanced

Set this sequence as reference

Consensus mode...





Thank you!