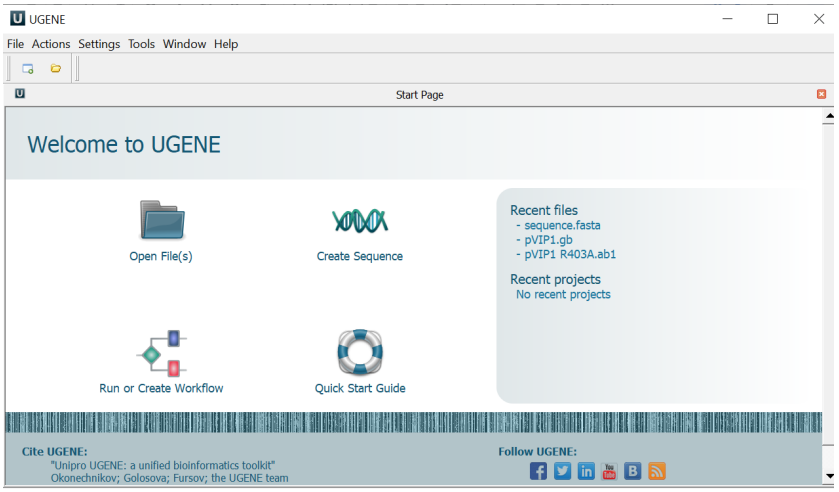
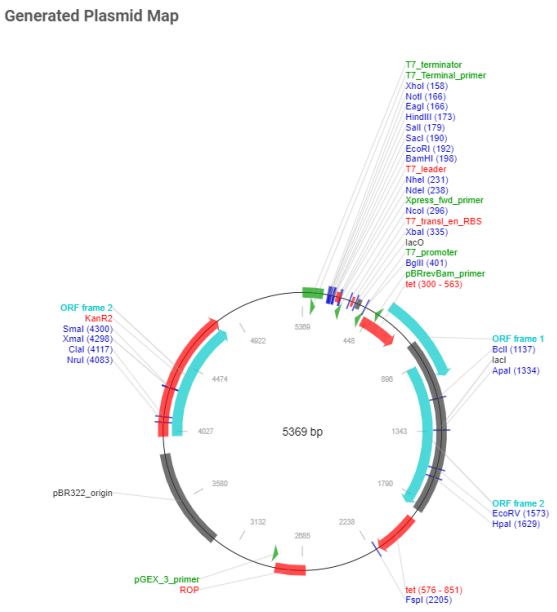
UGENE

Sometimes you have proteins with unusual folding, which need special modifications.



When you buy a commercial plasmid, e.g. pET-28a(+)- you also get a manual with all necessary information.

This is a plasmid map, where restriction endonuclease sites are marked. The sites provide bacteria natural protection from viruses. The more restriction endonuclease bacteria has, the bigger is the probability that bacteria can protect itself from viruses. But with this enzyme the probability of harming bacteria's DNA is also very big, especially for Type II sites consisting only of 4 letters like BamHL. To avoid this, bacteria should have 10 restrictases with sites, which are not equal to sites in the genome that you work with.The text is usually palindromic like TATA (reading from 3`-5` to 5`-3` is equal). Restriction endonuclease can form either blunt or sticky endings.



Let's talk about plasmid map:

1) Plasmid should be transcriptionally active - there should be DNA-dependent RNA landing sites.

2) (Ori.) The plasmid should have sites for landing DNA-dependent DNA polymerase for plasmid division during cell division. There are different types of plasmids. One bacterial cell can contain dozens of high copy plasmids or just a couple of low copy ones. Sometimes you can find quite big plasmids with more than 20 thousand pairs of nucleotides. But if the plasmid is too big - e.g. 40 thousand pairs

of nucleotides - the cell won't replicate it and you'll get a cell population without plasmid cells. The cells from such a population divide more effectively and eventually supplant the plasmid. The analysis of latest articles show that 40 thousand pairs of nucleotides is the maximum number to work with.

3) To divide cells with plasmid and cells without them you need a special marker. For example resistance genes. Usually two types of resistance genes are used with two antibody constructions to resist antibiotics. With their help we get the gene expression product in the antibiotics environment. Protein constructions need linkers like a linear water-soluble part of the protein (GGGS)n. Serin contains OH-group, GGGS is the hydrophilic part. Linker is like a thread between target protein and coloring protein. The activity of target protein is not disturbed due to its binding with the protein-color marker by such linker.

You need a marker that can help with dividing cells with plasmid from cells without one.

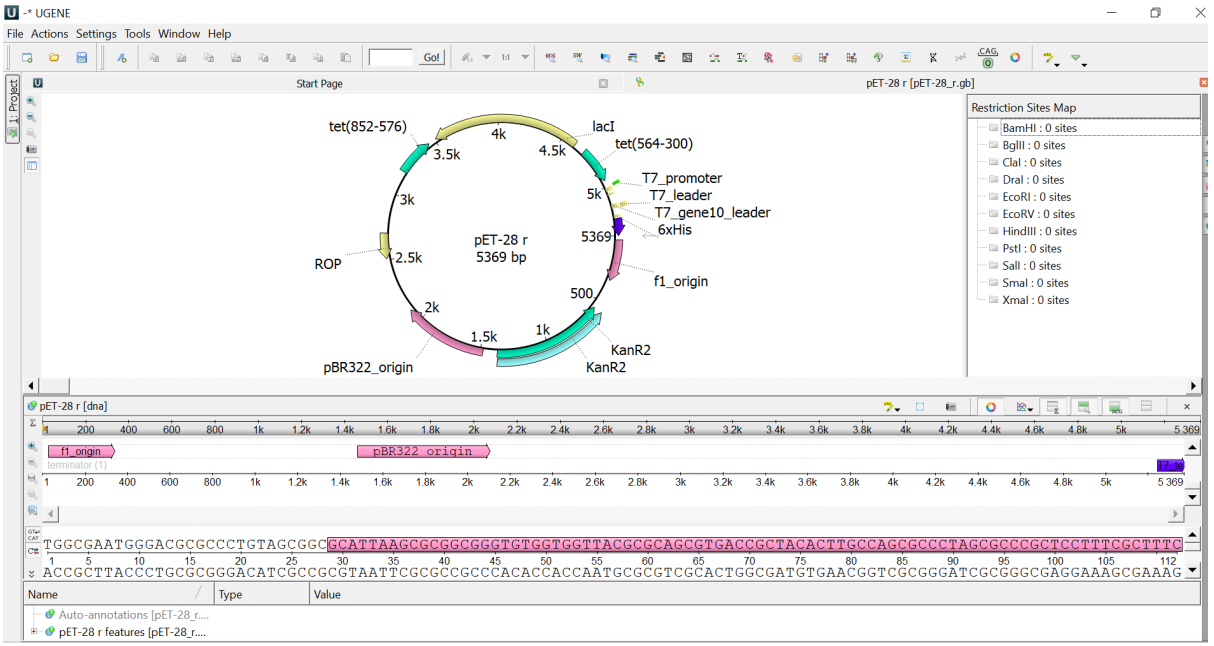
After cutting the piece of DNA by restrictase genetic engineers do ligation with ligases. Phage T4 ligase is most commonly used.

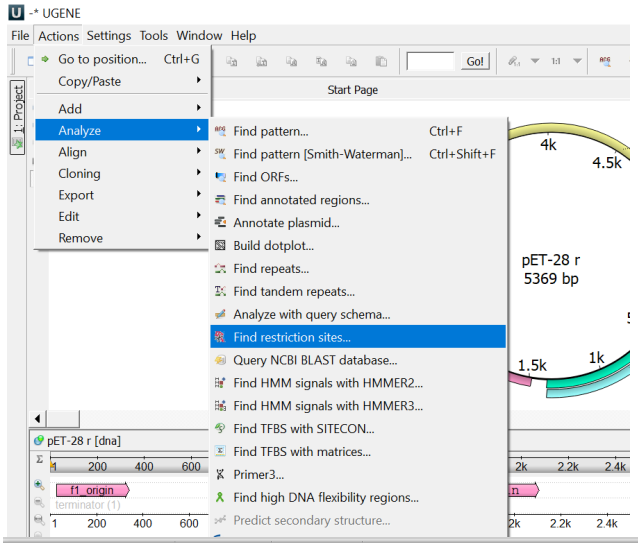
If you get blunt endings after the restriction, the ligation effectiveness will be quite low, because the probability of formation of a triple DNA complex, DNA and ligase is also low. But this happens rarely. Insertion can be performed either in AB direction or BA. Not only will the ligation be low-effective, but you also get a 50% chance of insertion in the wrong direction.

After protein expression you should clean it with histidine tail, that will grab nickel in columns. For example in 300ml of water or 10ml of the environment bacterias flow and reproduce, synthesizing proteins. The question is: how can you isolate the protein? You should pour the lising liquide in the environment with the bacteria that will contain a large spector of proteins.

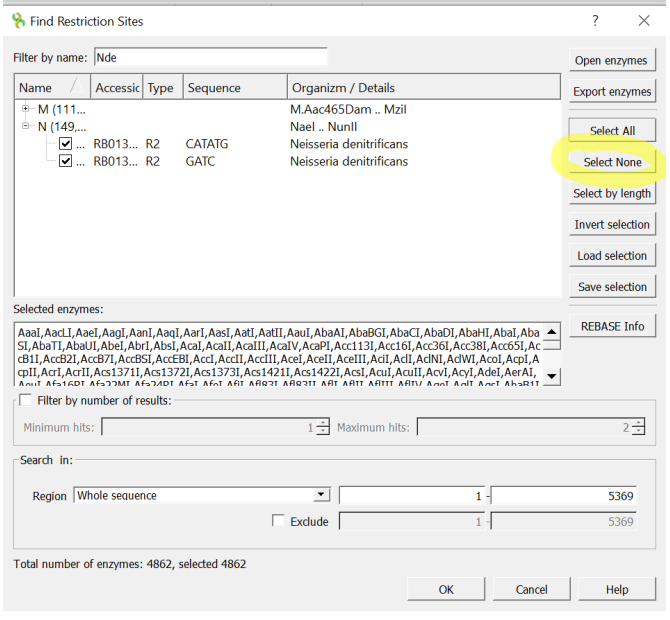
Task 1

Open the plasmid pet-28:





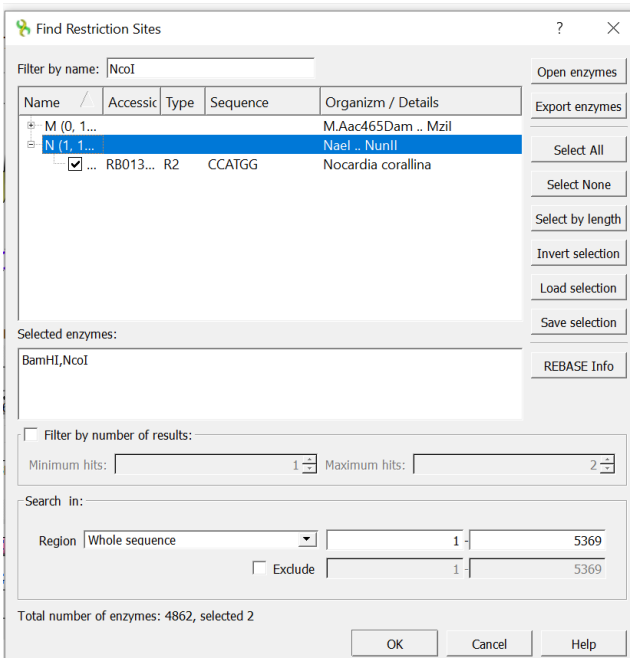
Here you can find the sites of restriction:



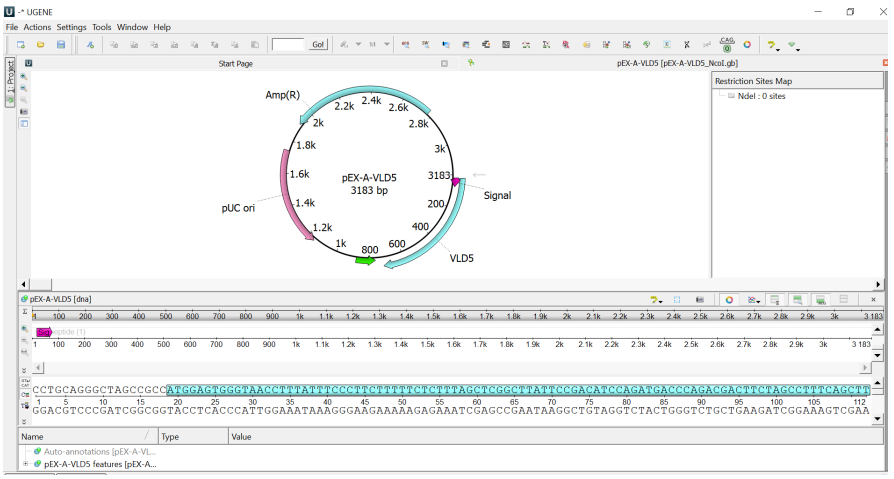
In UGENE all restrictases are listed in alphabitical order

Initially all restrictases are marked as chosen, you should press “Select none” button

We get an empty window without any restrictases and now we can choose the ones that we need. For example NcoI and BamHI

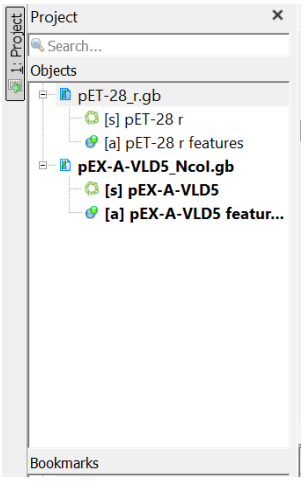


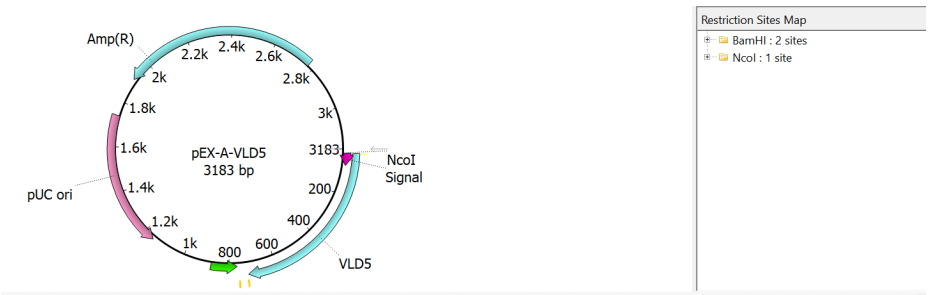
We select a gene



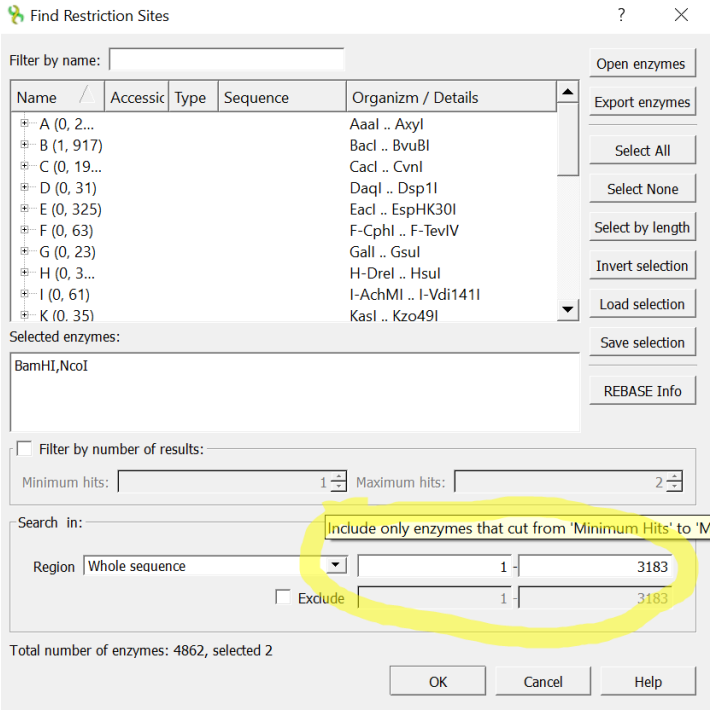
Restriction sites are chosen, so we can open the file with our plasmid. For example pEX-A-VLD5. It's important to open this file in the same project. To do this you just need to push "open" button in the upper-left side of the screen.

On the left side you should see this window:

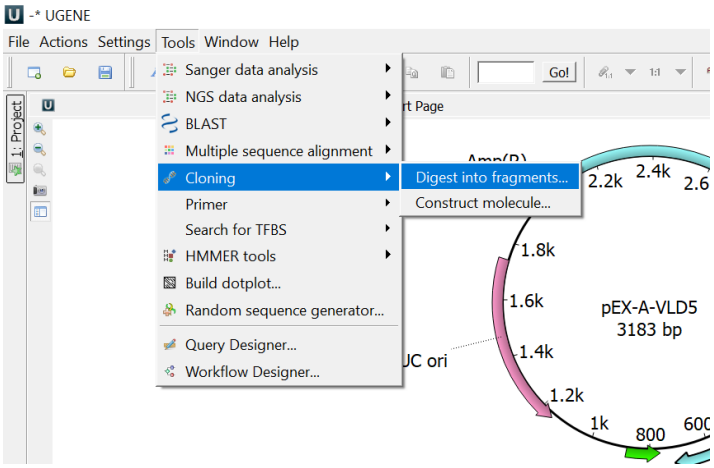




We choose the same restriction sites (BamHI и NcoRI).

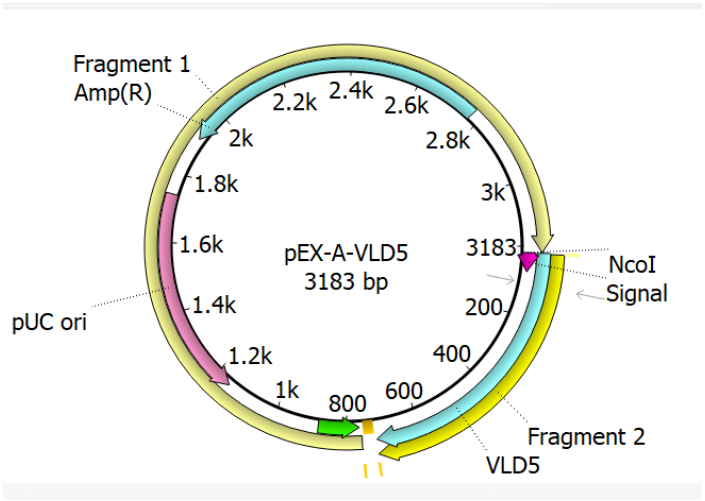


When you search for restriction endonuclease be sure to select right boundaries: from 1 to remaining nucleotide. Be sure that you've chosen the whole length.



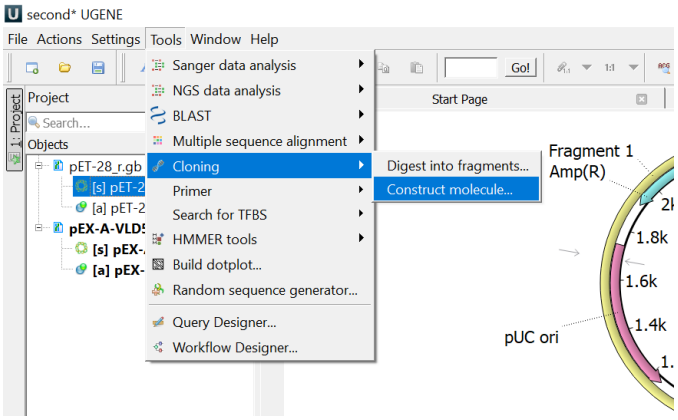
We need to select VLD5 from this plasmid. VLD5 ends with BamHI and begins with Ncol.

In the lab we'll add BamHI and Ncole restrictases, a buffer for their work and water and then we'll carry out hydrolysis. Preliminarily, we want to do the same thing in silico in UGENE.



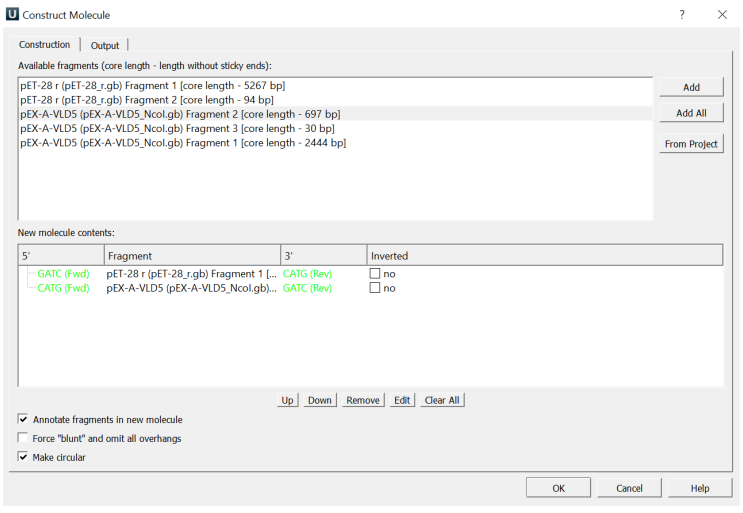
We conducted restriction.

We should look if there's any stop codon in the sequence where we don't need it: they are marked by a star sign.

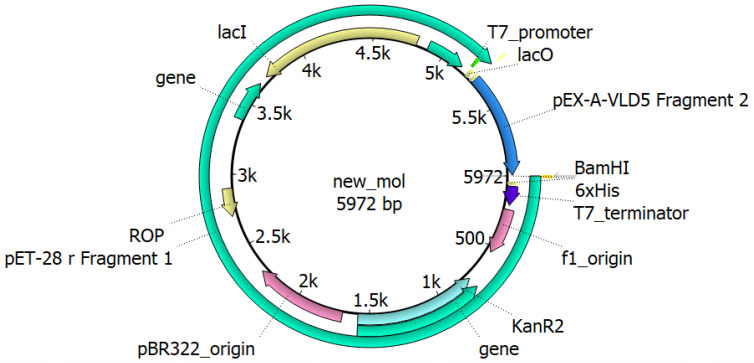


Ligation

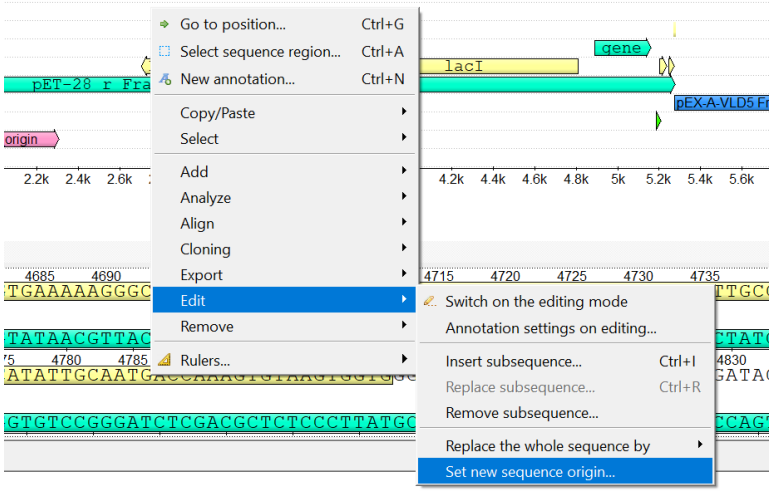
Ligation starts like this: we choose construct molecule



You choose required parts. It is necessary to tick "make circular" or you'll get linear sequence instead of a plasmid.

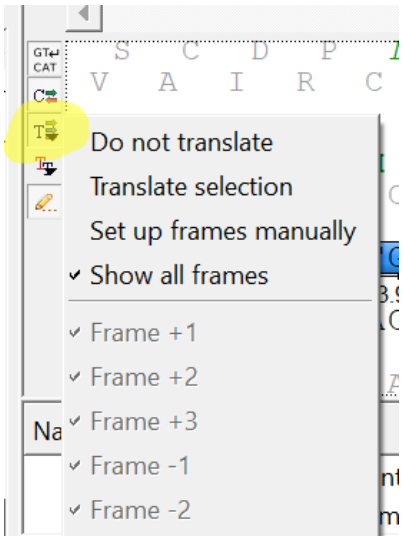


As a result you get new_mol c with cloned gene.

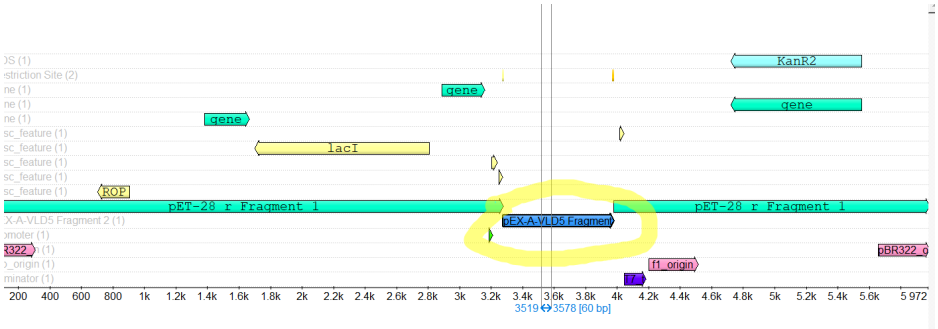


Analysis

To analyze the resulting plasmid with the cloned gene, we should replace the beginning and select another nucleotide as the first one. It will be convenient for visualizing DNA sequence. The first nucleotide may match with the beginning of the VLD5 gene.

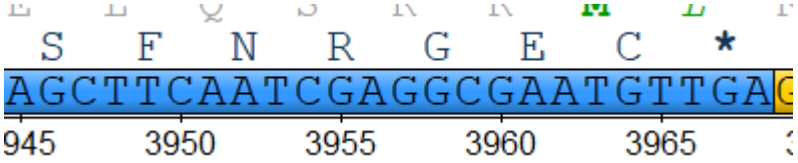


Turn on the display of all aminoacids. You can do it by pressing “show all frames”.



And now we can look on the sequence of our plasmid.

Blue sequence is our insertion, VLD5 gene.



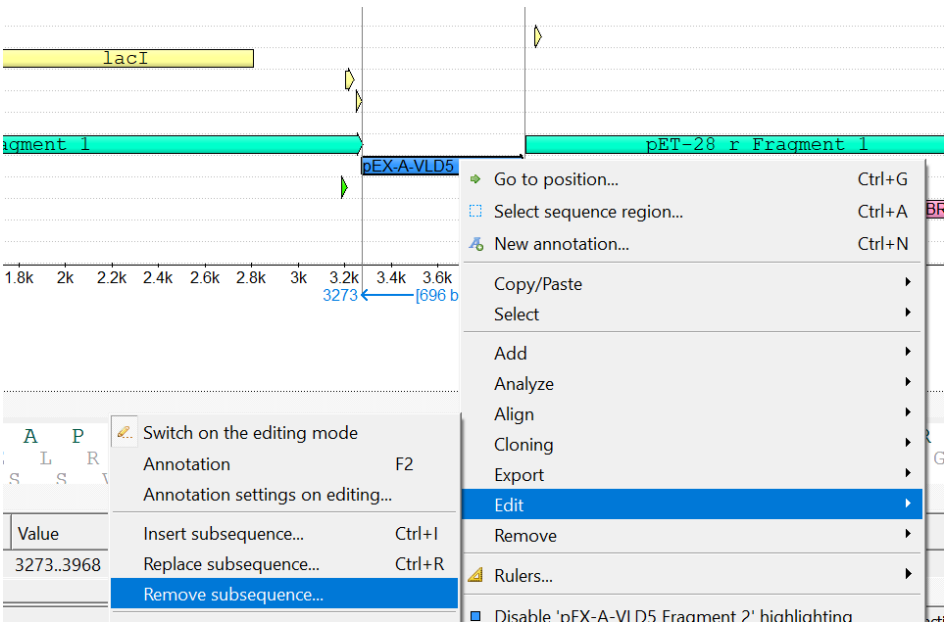
Check if there are any stop codons, there shouldn't be any in this part.

But as we see, there is a stop codon. And we won't get the protein with histidine tail. The gene sequence will stop before the codon and histidine tail will end in another reading frame. Our protein won't get a histidine tail and we won't be able to isolate the protein on the purification stage. We can do the following things:

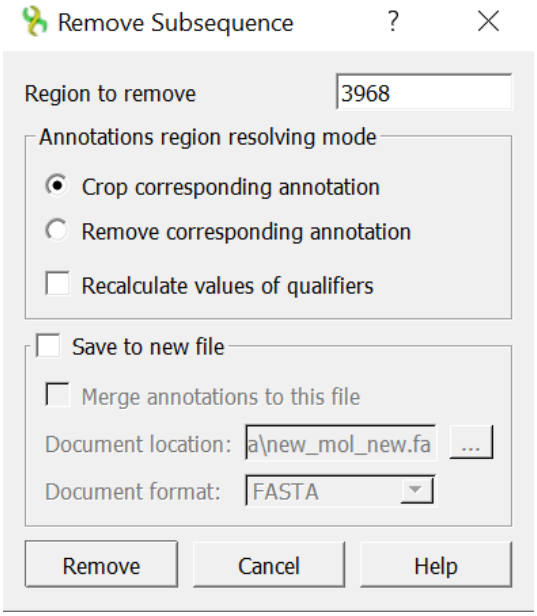
- place His tags in reading frame together with VLD5
- remove stop codon between VLD5 and His-tag

We have the following decision: we spoil the stop codon by replacing one letter. This way we'll get rid of the stop codon and displace the reading frame as we want.

To do this we need to :

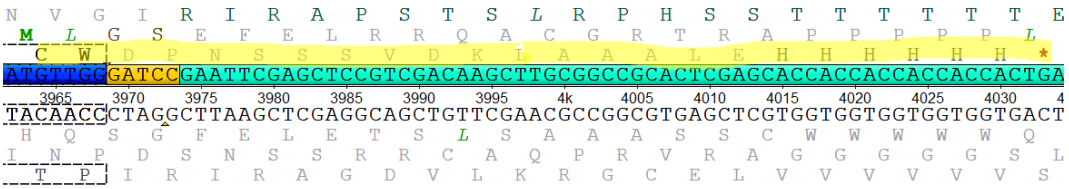


Click on our VLD5 gene and press Remove subsequence



3968 - is the position of the one of the three nucleotides of the stop codon. We remove it, the reading frame moves one step forward and the stop codon disappears.

Now let's look on what we've got: now Histagd is in the same reading frame (highlighted by yellow color), as VLD5 (highlighted by blue color):



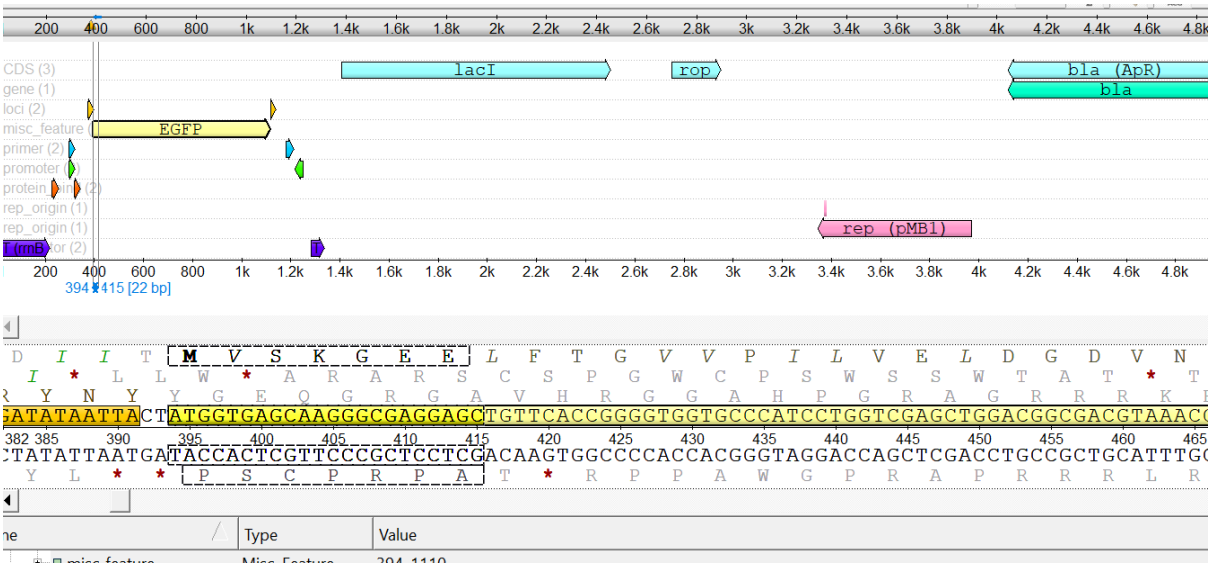
In the real experiment we'll just order the right primer.

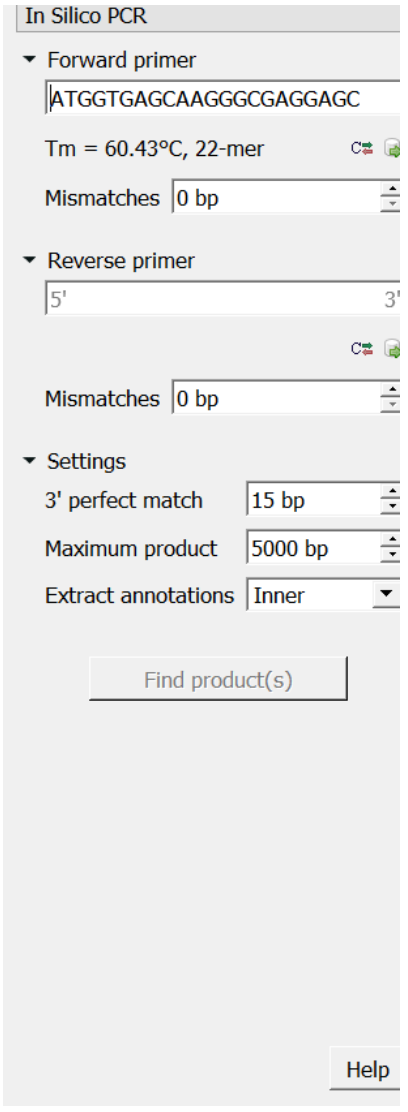
Task 2:

3` OH group is a substrate for DNA-dependent RNA polymerase. If the primer on 3' ending is properly complementary binded with matrix molecules, the polymerase will complete the sequence even if 5' ending is not complementary binded. This is called fusion primer - it has the part that is annealed on a matrix molecule and the tail that we use to form a PCR product with necessary sequences. Fusion primers shouldn't be longer than 50 nucleotides or the probability of unspecific binding with regions of the DNA matrix will be quite big and you won't get the product you need.

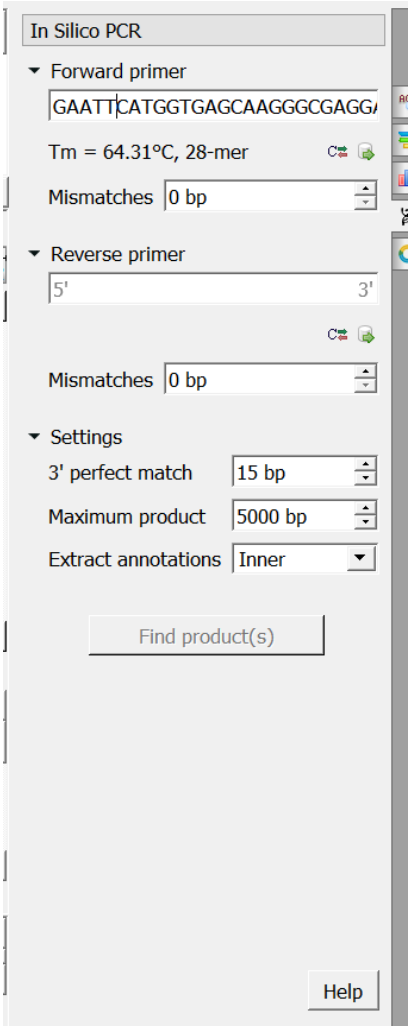
Cloning EGFP:
You want to do the restriction with EcoRI и BstVI.
When you look at restriction sites of the plasmid, you may find out that there's no restriction sites that you need, but you definitely have them in your lab.

Let's make direct primer:

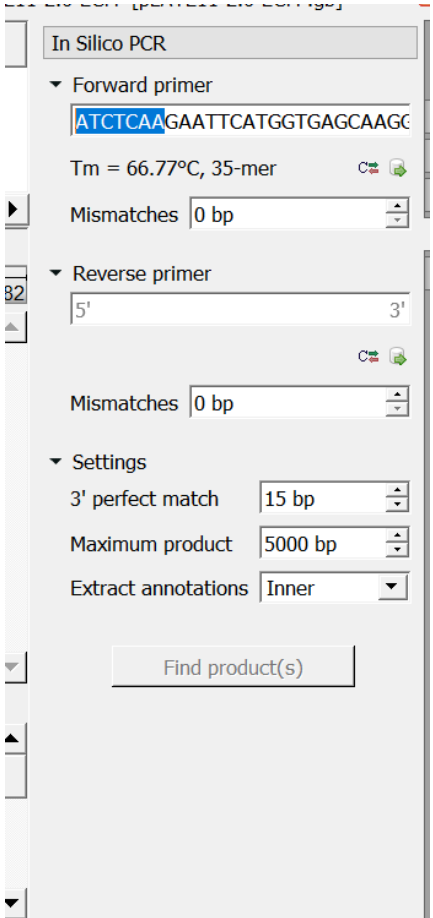




You choose 22 nucleotides from your gene sequence that you want to clone.



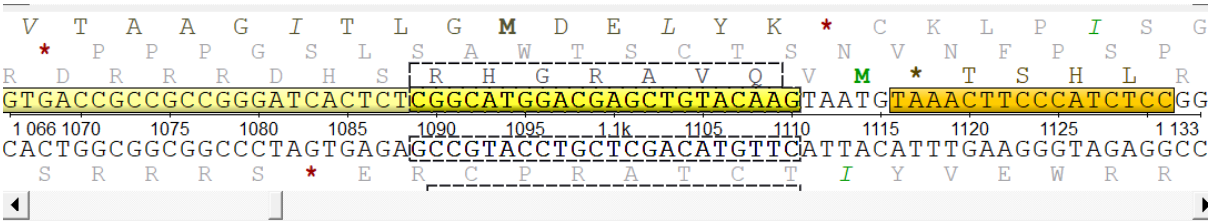
From 5` end we add 6 nucleotides, for example GAATCC - the site of restriction endonuclease for our experience. This site will be used for PCR.

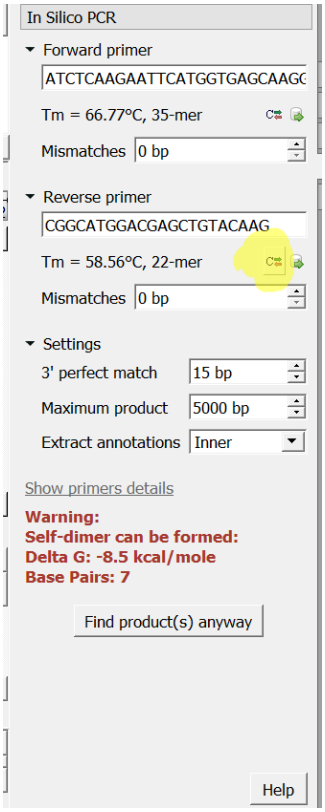


Before guanine you can add 6-8 nucleotides, that are generated randomly, to make restriction site easy to detect for restriction endonuclease.

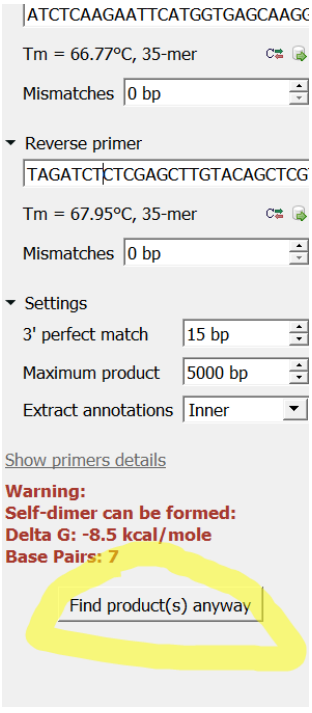
Reverse primer:

Next we look at the end of EGFP gene and create reverse primer that will be annealed on complementary sequence.



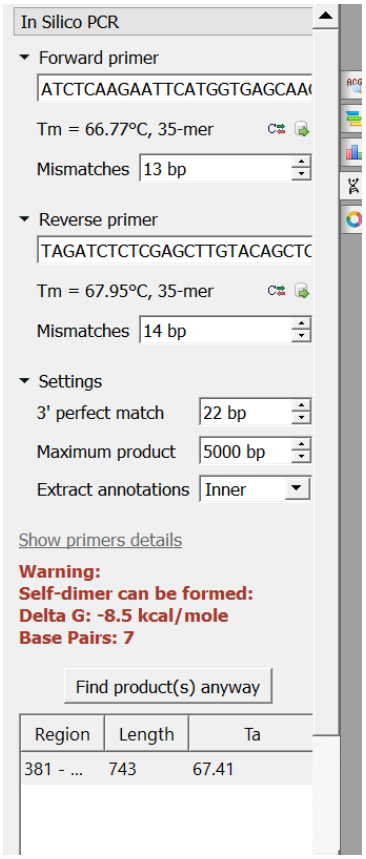


We copy the last 20 nucleotides in the complementary sequence - which is situated below the basic one - to “reverse primer” window and choose the reverse-complement sequence label.



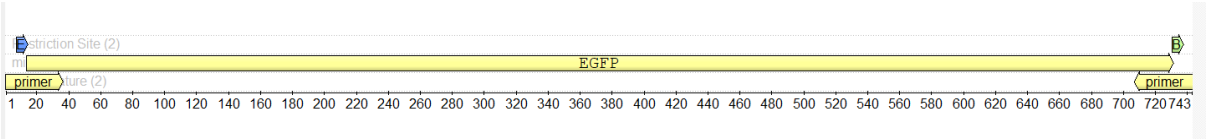
And then we add CTCGAG text to the 5’ end for detection by restrictase. We also add 6-8 random nucleotides to the 5’ ending, for example, TAGATCT

We could choose the button Find products anyway, but…



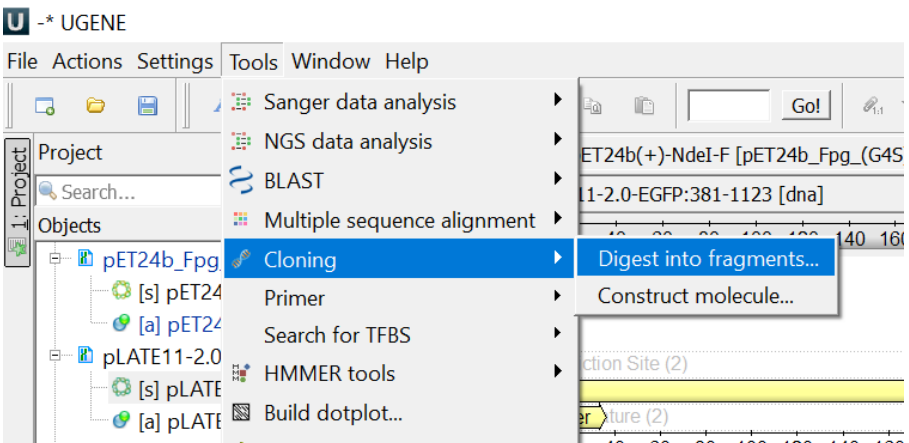
We can't do this if we don't indicate mis match.
We added 13 nucleotides for the direct primer and 14 for the reverse one. You should additionally indicate in settings 3` perfect match 22 bp - this is the primer length we've chosen for the annealing on matrix.

And we got EGFP with primers (if everything is done right):

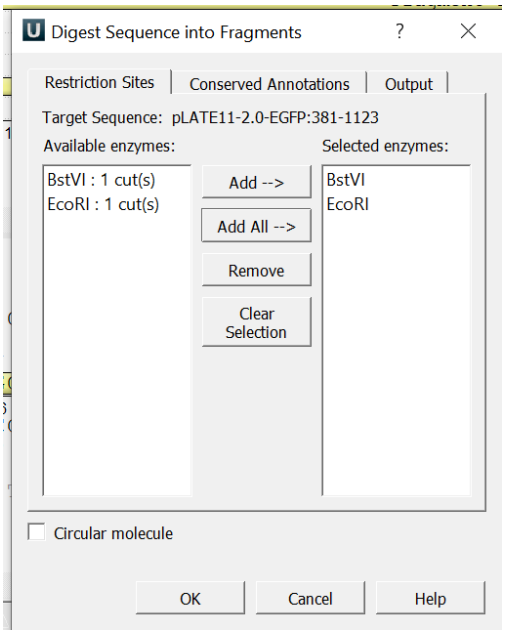


Cloning

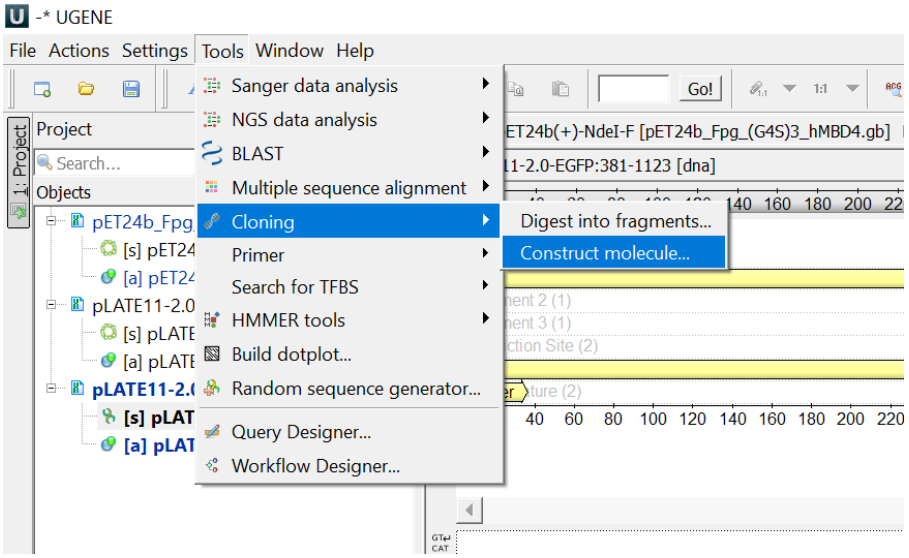
PCR product of EGFP with this primers will contain all the necessary restriction sites. So now we can do the restriction:



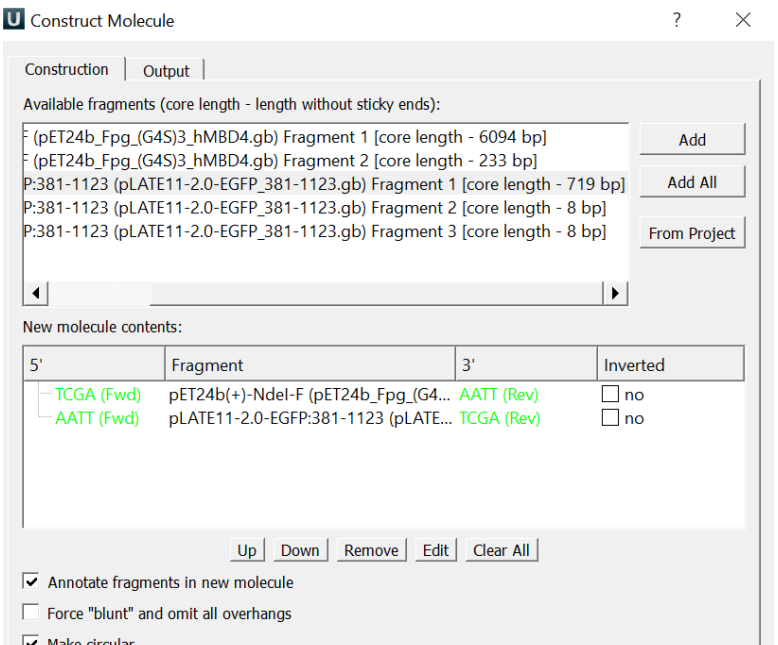
We do restriction with both vector (pET24b) and PCR product EGFP



Restriction with BstVI и EcoRI restrictases

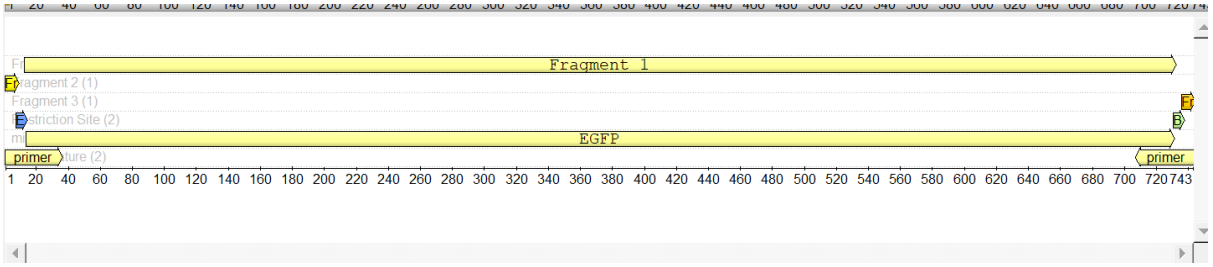


We make a construction with EGFP insertion:

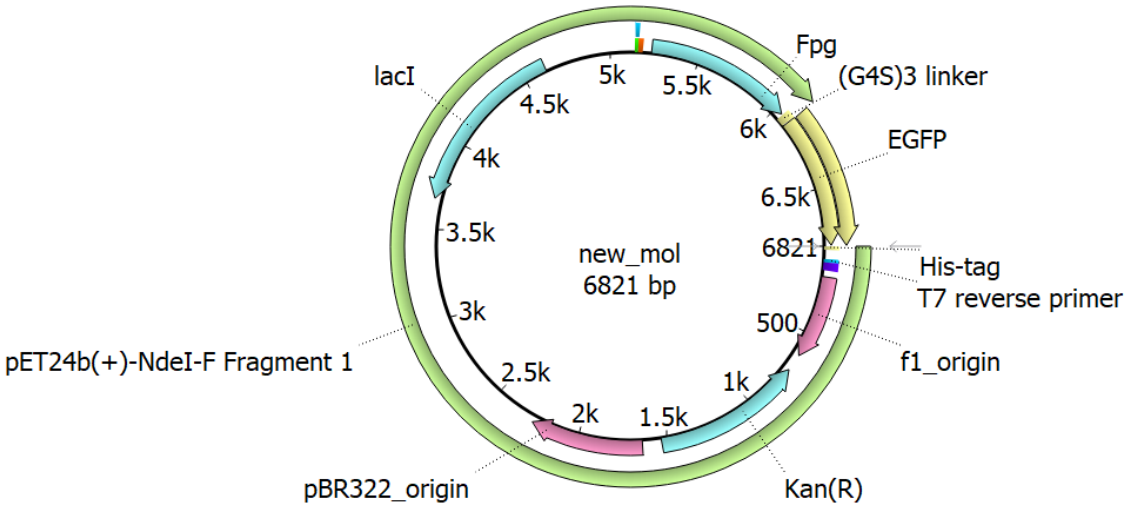


We choose fragments:

All the fragments are signed above the window for choosing fragments. Все фрагменты подписаны чуть выше окна для выбора фрагментов:



The cloning is finished:



Full genome sequencing

MiniSeq: maximum reeds 2 by 150 - for bacterial genome (150 is the length of the reed): 1,5*10^7*150*2 (1,5*10^7 - is the number of the reeds)
MiSeq: maximum reeds are 2 by 300 - this is the most universal platform
NextSeq - for human genome: 2,5*10^7*300*2
HiSeq - up to 20 people for one chip

You can get bacterias for 50 mln. pairs of nucleotides long with 30-fold covering: 5*10^7*30=1.5*10^9 п.н.
Every reed is 500 pairs of nucleotides.
The amount of reeds is 1,5*10^9 / 500 = 3*10^6