Cloning plan document

# Purpose

This plan described the cloning plan for the 2025 Illinois Tech iGEM project. This will be assembled by GG golden gate methos in the context of the BTK toolkit (bee tool kit?) REF

The aim is to construct a vectors that deliver a payload genes for pesticide derdadation. Three vectors with 3 such payload gene are planned:

1. OPD - name
2. MPD
3. XXX

The first series will be vectors designed to insert via homologous recombination into the S alvi genome

In addition we will develop a second set of vectors (to deliver the same 3 genes, but in a vector that is maintained stably in S alvi ( i..e not insert in the genome\_

[Purpose 1](#_Toc200669379)

[Original plan 2](#_Toc200669380)

[Professor Barrick’s plan 3](#_Toc200669381)

[Final gene version 1 4](#_Toc200669382)

[Things we need to order 6](#_Toc200669383)

[Final plan 8](#_Toc200669384)

[Scheme of overhangs 8](#_Toc200669385)

[development 8](#_Toc200669386)

[final to order 9](#_Toc200669387)

[verify 10](#_Toc200669388)

[Alt plan 13](#_Toc200669389)

# Original plan

Our original plan is to construct the full plasmid using the homology arms pBTK1050 and pBTK1051 from S. alvi, along with the backbone pYTK095. We aim to insert a synthesized DNA fragment containing the promoter, RBS, gene, and terminator into this complete vector. To achieve this, we need a total of four parts: Homology Arm 1, the synthesized DNA, Homology Arm 2, and the backbone.

Among these, we already possess three parts—pBTK1050, pBTK1051, and the backbone pYTK095—and are currently in the process of cloning them.

Here is a simplified summary of the cloning parts:

[pYTK095]

Backbone we are using – overhang(ccct) – N(random sequence) – BsaI rec site – dropout sequence – BsaI rec site - N(random sequence – overhang(taca) – Backbone we are using

[H1]

Dropout sequence - BsaI rec site - N(random sequence) – overhang(ccct) – H1 arm sequence – overhang(aacg) – N(random sequence) - BsaI rec site – Dropout sequence

[promotor/RBS – opd - terminator]

Random sequence for restriction enzyme recognition – overhang(aacg) – Promotor – original overhang(tatg) – opd – original overhang(ggat) – terminator – overhang(gctg) - Random sequence for restriction enzyme recognition

[H2]

Dropout sequence - BsaI rec site - N(random sequence) – overhang(gctg) – H2 arm sequence – overhang(taca) – N(random sequence) - BsaI rec site – Dropout sequence

\*One thing we are not fully confident about is the addition of approximately 25 extra bases upstream of the synthesized DNA to generate sticky ends with BsaI. We are unsure whether this length is sufficient to ensure proper recognition by the restriction enzyme.\*

\* Remove original overhang?

# Professor Barrick’s plan

[plasmid insertion – without suicidal vector]

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ID** | **Type** | **Origin** | **5'-overhang** |  | **BSA1 3'-overhang** |  |
| **pBTK1066** | super dropout | RSF1010 | 5 |  | 1 | CCTC |
| **pBTK1079** | super dropout | RSF1010 | 5 |  | 1 |  |
| **pBTK107** | promoter\_RBS | ColE1 | 2 |  | 2 |  |
| **pBTK301** | spacer-terminator | ColE1 | 6 |  | 7 |  |
| **pBTK205** | CDS | ColE1 | 3 |  | 3 |  |

Professor Barrick recommended first transforming the plasmid into \*E. coli\* to verify its functionality and then inserting it into S. alvi to confirm the plasmid’s overall effectiveness. For this purpose, he suggested using two backbones containing the RSF1010 origin, which functions across a variety of hosts. The promoter and terminator we plan to use are the ones we already have—107 and 301. Additionally, it seems that he recommended pBTK205, which carries GFP optim-1, for use as a positive control in comparison with the experimental group.

[Homologous recombination – with suicidal vector]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ID** | **Type** | **Origin** | **5'-overhang** | **3'-overhang** |
| **pBTK1090** | Homology arm (S. alvi wkB2) | ColE1 | 1 | 1 |
| **pBTK1091** | Homology arm (S. alvi wkB2) | ColE1 | 5 | 5 |
| **pBTK1056** | marker | ColE1 | 4 | 4 |
| **pBTK1129** | transcriptional unit | ColE1 | 2 | 3 |
| **pYTK095** | high copy plasmid backbone | ColE1 | 6 | 8 |

Next, for homologous recombination, Professor Barrick recommended the following plasmids. Among them, pBTK1090 and pBTK1091 contain a different set of homology arms, but he mentioned that they serve a similar purpose to the ones we currently have—pBTK1050 and pBTK1051—with the only difference being the specific parts used. He recommended pBTK1090 in particular because a PhD student in his lab, who will be advising us, has frequently used 1090 in their work. This would make it easier to explain and troubleshoot the process.

# Final gene version 1

>gene-GG-pBB25-OPHeptiope-Ter-GG

\*All sequences contain restriction enzyme site

what RE? BsaAI ... GGTCTC(N1)/(N5)-

show that sequence (some color, underline .. )

>pBB25\_RBS

Gg5’

ggtctcaaacgctttggcagtttattcttgacatgtagtgagggggctggtataatcacatagtactgttatacagaaacagaggagatattacatatgtgagacc

>opdA

ATGGAGAATGAAATGACCACCGAGCCGGGCGATCGAGCCTCGCAGTTTGATGTGGCTATTGTCGGCTGCGGCCCCGTCGGCGCTCTGCTGGCTAATCTACTAAAGCAATACGGGCATAAAGTCGCGGTTCTCGAGCGCGAGCCCGATATTTTCTACGCGCCGCGAGGGATGGGATTTGATGATGAATCGACGCGGATCATGCAATCCGCGGGCATCCTCGATCGCCTGAAAGCGGAAGGGCATATCTATCAGGCCGACCTCGAATTGATTGACCGAAGTGGCAAACGCCTTGGTGGATTCGATCGGCGTAGCGTTGGAGAAGATCTCCTGTCCGGCCTTCATGGCCATCGCCATCTGACCTTGTTCCACCAGCCGAGCTTGGAGGCGACGCTGCGGGAGGAGTTTGCGACTGGCGAAAACGCGGCAACCGCTTATTTCAATCACGAGGTCACTGGCATCACCGACCAAGGCGATCGGGTCGAATTGAACTGCAAGGATCGCGCCACGGACGAGGAACATTCCCTCATCGCCAAATATGTCGTGGGGTGTGACGGTGCGAGGAGCACCGTGCGCAAAACGATGAACGTTCCGAGGATAGACCTGAAATATACTCAAAAGTATCTGGTCGTCGACGCCATCGTCGATGACCCTGTCTATTTCAGGACGATGATACCGCAAGGCGGATATATTCTGCTCGATGGCAAGGAAGCTGGTGTCCTGGCCAAGGGACTGCACGGGCATGTTCGTTTTGATTTTCTGCAGCATTCCGAGACAATCGGGCAGGAGCTGAAGACTGATGAGGACTACCAAAAGGCTGCGAGAGACCTGATCAGGTCGCGAGGTTTCGATCCGGAGAACTTCCGCGTGATCCGAAGCGTCTCATACACGTTTCACGCTGGGATGCCGAGTAAATGGCGGGTCGGACGGCTGATGGTGGCAGGAGACGCAGCGCACCTTACGCCACCGTGGTCGGGGCAAGGCCTGAACATGGGTGTGCGCGACGCAGCCAATCTTTCTTTTAAATTGAATCTGGCGCTTCGAGGCAAGAGCTCCGATCGGATCCTGGACACATATGATGAAGAGCGCAGGCCGCAGAGTCTCGAGACGATCCAGGCTGCCGTCGACATGGGCATCAGAATGCAGAACACGAGTCCGTTGCAGATCGGACTGCGTAATCTGGCCTATGCTTTGTCGCGGAAAAGCAAGTTTGTTAACCGCCTCCTATTCAAAAACTGGATCAGAAAGCCTTCCTACAAAAGCGGCCTTTTGGGGCTGCAGCATCGCTTGTCGGGTGGACCTATGTTTCAGCCGTGGGTGGAGACAGCTGAGGGCAAGCGCGTGCGGATGGACGATCTGATCGGCCTGAACTTCGCGCTCATTTCCACGGACAGTCCAACTGGTCCCGAGGTGAGGCAATTCGTGAGCGAACTTGGGGGCGTCGTTCTCAAACTCGACTGTGATTTTTTCGACCCCAGCGAGACCGTCTGCAAGTGGTACGACGAGCACCGGATCAATGCAGTGCTGCTCCGCCCCGACCGTGTCATTTACGACGCGGGTCGCGACGGTCAGGCGCTATGCCGATCCCTCCTTGCCGAGCTCAGGAAATAG

>opd

ATGCAAACGAGAAGGGTTGTGCTCAAGTCTGCGGCCGCGAGAACTCTGCTCGGCGGCCTGGCTGGGTGCGCGACGTGGCTGGATCGATCGGCACAGGCGATGCGATCAATACGTGCGCGTCCTATCACAATCTCTGAAGCGGGTTTCACACTGACTCACGAGGACATCTCGGCAGCTCGGCAGGATTCTTGCGTGCTTGGCCAGAGTTCTTCGGTAGCGCAAAGCTCTAGCGGAAAAGGCTGTGAGAGGATTGCGCGCCAGAGCGGCTGGCGTGCGAACGATTGTCGATGTGTCGACTTTCGATATCGGTCGCGACGTCAGTTTATTGGCCGAGGTTTCGCGGGCTGCCGACGTTCATATCTGGCGGCGACCGGCTTGTGGTTCGACCCGCCACTTTCGATGCGATTGAGGTATGTAGAGGAACTCACACTAGTTCTTCCTGCGGTGAGATTCAATATGGCATCGAAGTACACCGGAATTAGGGCGGGCATTATCAAGGTCGCGACCACAGGCAAGGCGACCCCCTTTCAGGAGTTAGTGTTAAAGGCGGCCGCCCGGGCCAGCTTGGCCACCGGTGTTCCGGTAACCACTCACACGGCAGCAAGTCAGCGCGATGGTGAGCGAGGCAGGCCGCCATTTTTGAGTCCGAAGCTTGAGCCCTCACGGGTTTGTATTGGTCACAGCGATGATACTGACGATTTGAGCTATCTCACCGCCCTGCTGCGCGGATACCTCATCGGTCTAGACCACATCCCGCACAGTGCGATTGGTCTAGAAGATAATGCGAGTGCATCACCGCTCCTGGGCATCCGTTCGTGGCAAACACGGGCTCTCTTGATCAAGGCGCTCATCGACCAAGGCTACATGAAACAAATCCTCGTTTCGAATGACTGGCTGTTCGGGTTTTCGAGCTATGTCACCAACATCATGGACGTGATGGATCGCGTGAACCCCGACGGGATGGCCTTCATTCACTGA

>mpd

ATGCCCCTGAAGAACCGCTTGCTGGCCCGCCTGTCCTGTGTTGCGGCCGTGGTGGCCGCCACGGCCGCCG

TTGCACCGTTGACGCTGGTGTCCACCGCCCACGCCGCCGCACCGCAGGTGCGCACCTCGGCCCCCGGCTA

CTACCGGATGCTGCTGGGCGACTTCGAAATCACCGCGCTGTCGGACGGCACGGTGGCGCTGCCGGTCGAC

AAGCGGCTGAACCAGCCGGCCCCGAAGACGCAGAGCGCGCTGGCCAAGTCCTTCCAGAAAGCGCCGCTCG

AAACCTCGGTCACCGGTTACCTCGTCAACACCGGCTCCAAGCTGGTGCTGGTGGACACCGGCGCGGCCGG

CCTGTTCGGCCCCACCCTGGGCCGGCTGGCGGCCAACCTCAAGGCCGCAGGCTATCAGCCCGAGCAGGTC

GACGAGATCTACATCACCCACATGCACCCCGACCACGTGGGCGGCTTGATGGTGGGTGAGCAACTGGCGT

TCCCGAACGCGGTGGTGCGTGCGGACCAGAAAGAAGCCGATTTCTGGCTCAGCCAGACCAACCTCGACAA

GGCCCCGGACGACGAGAGCAAAGGCTTCTTCAAAGGCGCCATGGCCTCGCTGAACCCCTATGTGAAGGCC

GGCAAGTTCAAGCCTTTCTCGGGGAACACCGACCTGGTGCCCGGCATCAAAGCGCTGGCCAGCCACGGCC

ACACCCCGGGCCACACCACCTACGTGGTCGAAAGCCAGGGGCAAAAGCTCGCCCTGCTCGGCGACCTGAT

ACTCGTCGCCGCGGTGCAGTTCGACGACCCCAGCGTCACGAACCAGCTCGACATCGACGGCAAGTCCGCT

GCGGTGGAGCGCAAGAAGGCCTTCGCGGATGCCGCCAAGGGCGGCTACCTGATCGCGGCGAGCCACCTGC

CGTTCCCCGGCATCGGCCACATCCGCGCCGAAGGCAAGGGCTACCGTTTCGTGCCGGTGAACTACTCGGT

CGTCAACCCCAAGTGACTG

>FLAG\_tag

GATTATAAAGATGATGATGATAAA

>terminator

Ggtctcatacaccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcggtgaacgctctctactagagtcacactggctcaccttcgggtgggcctttctgcgtttataccgatgagacc

Terminators are usually hairpins, that how they work . This one is too:

<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>

A close-up of a red and black dot

AI-generated content may be incorrect.

## Things we need to order

This sequences contains 5 basepairs in both end and epitope tag

>opd\_gene\_set

We still need GG sequences GG5 and GG3

>Module\_GG5\_BB25\_OPH\_flag\_term\_GG3

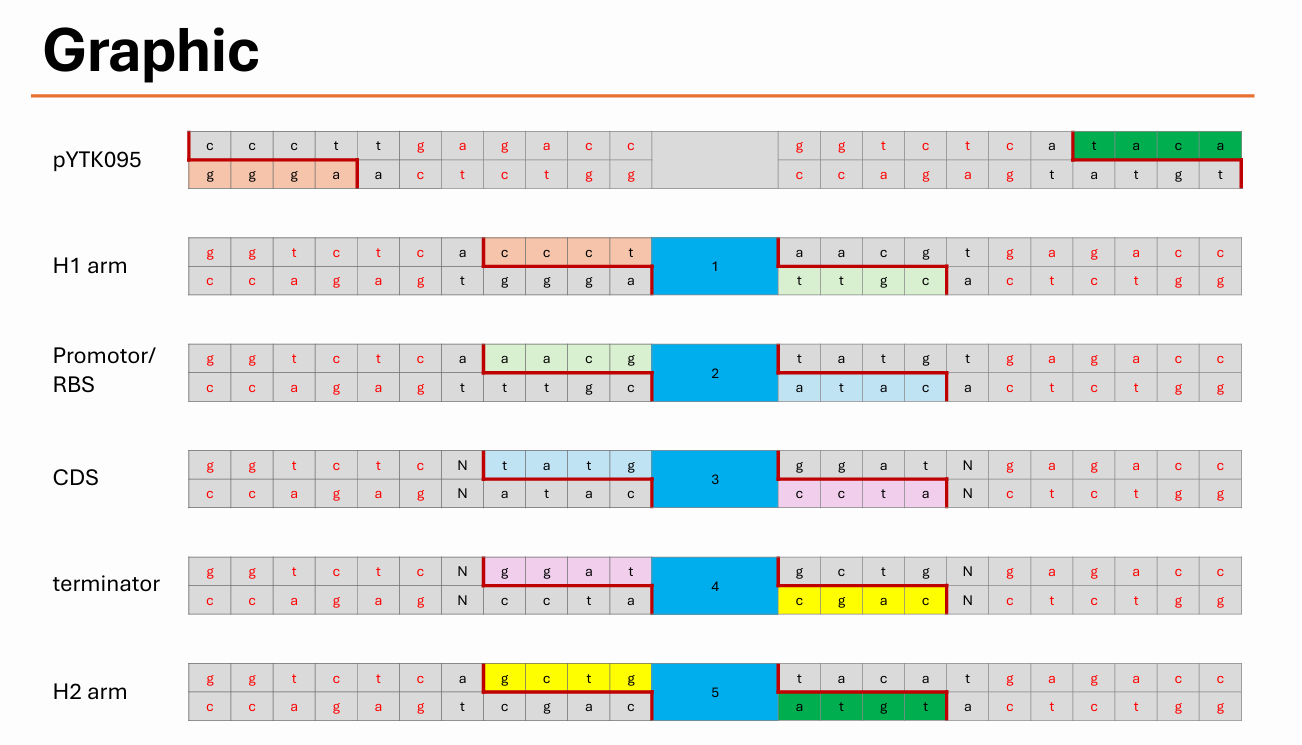
NNGGTCTCN3333ggtctcaaacgctttggcagtttattcttgacatgtagtgagggggctggtataatcacatagtactgttatacagaaacagaggagatattacatatgtgagaccATGGAGAATGAAATGACCACCGAGCCGGGCGATCGAGCCTCGCAGTTTGATGTGGCTATTGTCGGCTGCGGCCCCGTCGGCGCTCTGCTGGCTAATCTACTAAAGCAATACGGGCATAAAGTCGCGGTTCTCGAGCGCGAGCCCGATATTTTCTACGCGCCGCGAGGGATGGGATTTGATGATGAATCGACGCGGATCATGCAATCCGCGGGCATCCTCGATCGCCTGAAAGCGGAAGGGCATATCTATCAGGCCGACCTCGAATTGATTGACCGAAGTGGCAAACGCCTTGGTGGATTCGATCGGCGTAGCGTTGGAGAAGATCTCCTGTCCGGCCTTCATGGCCATCGCCATCTGACCTTGTTCCACCAGCCGAGCTTGGAGGCGACGCTGCGGGAGGAGTTTGCGACTGGCGAAAACGCGGCAACCGCTTATTTCAATCACGAGGTCACTGGCATCACCGACCAAGGCGATCGGGTCGAATTGAACTGCAAGGATCGCGCCACGGACGAGGAACATTCCCTCATCGCCAAATATGTCGTGGGGTGTGACGGTGCGAGGAGCACCGTGCGCAAAACGATGAACGTTCCGAGGATAGACCTGAAATATACTCAAAAGTATCTGGTCGTCGACGCCATCGTCGATGACCCTGTCTATTTCAGGACGATGATACCGCAAGGCGGATATATTCTGCTCGATGGCAAGGAAGCTGGTGTCCTGGCCAAGGGACTGCACGGGCATGTTCGTTTTGATTTTCTGCAGCATTCCGAGACAATCGGGCAGGAGCTGAAGACTGATGAGGACTACCAAAAGGCTGCGAGAGACCTGATCAGGTCGCGAGGTTTCGATCCGGAGAACTTCCGCGTGATCCGAAGCGTCTCATACACGTTTCACGCTGGGATGCCGAGTAAATGGCGGGTCGGACGGCTGATGGTGGCAGGAGACGCAGCGCACCTTACGCCACCGTGGTCGGGGCAAGGCCTGAACATGGGTGTGCGCGACGCAGCCAATCTTTCTTTTAAATTGAATCTGGCGCTTCGAGGCAAGAGCTCCGATCGGATCCTGGACACATATGATGAAGAGCGCAGGCCGCAGAGTCTCGAGACGATCCAGGCTGCCGTCGACATGGGCATCAGAATGCAGAACACGAGTCCGTTGCAGATCGGACTGCGTAATCTGGCCTATGCTTTGTCGCGGAAAAGCAAGTTTGTTAACCGCCTCCTATTCAAAAACTGGATCAGAAAGCCTTCCTACAAAAGCGGCCTTTTGGGGCTGCAGCATCGCTTGTCGGGTGGACCTATGTTTCAGCCGTGGGTGGAGACAGCTGAGGGCAAGCGCGTGCGGATGGACGATCTGATCGGCCTGAACTTCGCGCTCATTTCCACGGACAGTCCAACTGGTCCCGAGGTGAGGCAATTCGTGAGCGAACTTGGGGGCGTCGTTCTCAAACTCGACTGTGATTTTTTCGACCCCAGCGAGACCGTCTGCAAGTGGTACGACGAGCACCGGATCAATGCAGTGCTGCTCCGCCCCGACCGTGTCATTTACGACGCGGGTCGCGACGGTCAGGCGCTATGCCGATCCCTCCTTGCCGAGCTCAGGAAAGATTATAAAGATGATGATGATAAATAGnnnnnnnnnnnnnnnnGgtctcatacaccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcggtgaacgctctctactagagtcacactggctcaccttcgggtgggcctttctgcgtttataccgatgagaccnnnnnGGTCTCn1111

tataccgatgagaccnnnnnn1111nGAGAC

# Final plan

## Scheme of overhangs

1. 1111=====backbone=pYTK095===2222
2. 2222=====homL========3333
3. 3333=====gene========4444
4. 4444=====homR========1111



|  |  |
| --- | --- |
| GG# | sequence |
| 2222 | ccct |
| 3333 | aacg |
| 4444 | gctg |
| 1111 | taca |

## development

>5bp\_spacer\_rec\_site

agttgggtctc

>N\_and\_overhang

aaacg

>promoter\_RBS

ctttggcagtttattcttgacatgtagtgagggggctggtataatcacatagtactgttatacagaaacagaggagatattacatatg

>opd

ATGCAAACGAGAAGGGTTGTGCTCAAGTCTGCGGCCGCGAGAACTCTGCTCGGCGGCCTGGCTGGGTGCG

CGACGTGGCTGGATCGATCGGCACAGGCGATGCGATCAATACGTGCGCGTCCTATCACAATCTCTGAAGC

GGGTTTCACACTGACTCACGAGGACATCTCGGCAGCTCGGCAGGATTCTTGCGTGCTTGGCCAGAGTTCT

TCGGTAGCGCAAAGCTCTAGCGGAAAAGGCTGTGAGAGGATTGCGCGCCAGAGCGGCTGGCGTGCGAACG

ATTGTCGATGTGTCGACTTTCGATATCGGTCGCGACGTCAGTTTATTGGCCGAGGTTTCGCGGGCTGCCG

ACGTTCATATCTGGCGGCGACCGGCTTGTGGTTCGACCCGCCACTTTCGATGCGATTGAGGTATGTAGAG

GAACTCACACTAGTTCTTCCTGCGGTGAGATTCAATATGGCATCGAAGTACACCGGAATTAGGGCGGGCA

TTATCAAGGTCGCGACCACAGGCAAGGCGACCCCCTTTCAGGAGTTAGTGTTAAAGGCGGCCGCCCGGGC

CAGCTTGGCCACCGGTGTTCCGGTAACCACTCACACGGCAGCAAGTCAGCGCGATGGTGAGCGAGGCAGG

CCGCCATTTTTGAGTCCGAAGCTTGAGCCCTCACGGGTTTGTATTGGTCACAGCGATGATACTGACGATT

TGAGCTATCTCACCGCCCTGCTGCGCGGATACCTCATCGGTCTAGACCACATCCCGCACAGTGCGATTGG

TCTAGAAGATAATGCGAGTGCATCACCGCTCCTGGGCATCCGTTCGTGGCAAACACGGGCTCTCTTGATC

AAGGCGCTCATCGACCAAGGCTACATGAAACAAATCCTCGTTTCGAATGACTGGCTGTTCGGGTTTTCGA

GCTATGTCACCAACATCATGGACGTGATGGATCGCGTGAACCCCGACGGGATGGCCTTCATTCAC

>FLAG\_tag

GATTATAAAGATGATGATGATAAA

>stop\_codon

TGA

>20bp\_spacer

GGGTAGGGGGCTTCAATTCG

>terminator

CCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGT

GAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATA

>overhang\_and\_N

gctga

>rec\_site\_and\_5bp\_spacer

gagacctgcat

## final to order

we need to give this a more concise name and make it fasta. I suggest something like “IITiGEM25\_module001” or some such (feel free to develop some other system) and we create a database of these listing the following elements

* Subcomponents – all these elements in color below
* Purpose – this is a golden gate assembly module with the essential element (promoter, gene, epitope tag, terminator) and golden gate flanking sequnces to enable to to be assembles in the BTK toolkit sustem
* creator -sunah lee
* sequence

then design exactly the same thing with the three other genes.

[5bp\_spacer\_BSAi\_site]-[N\_AND\_3333OVERHANG]-[promoter\_RBS]-[opd]-[FLAG\_tag]-[stop\_codon]-[20bp\_spacer]-[terminator]-[4444OVERHANG\_AND\_N]- [BSAi\_site\_and\_5bp\_spacer]

>IITiGEM25\_module001

agttgggtctcAAACGctttggcagtttattcttgacatgtagtgagggggctggtataatcacatagtactgttatacagaaacagaggagatattacatatg**ATG**CAAACGAGAAGGGTTGTGCTCAAGTCTGCGGCCGCGAGAACTCTGCTCGGCGGCCTGGCTGGGTGCGCGACGTGGCTGGATCGATCGGCACAGGCGATGCGATCAATACGTGCGCGTCCTATCACAATCTCTGAAGCGGGTTTCACACTGACTCACGAGGACATCTCGGCAGCTCGGCAGGATTCTTGCGTGCTTGGCCAGAGTTCTTCGGTAGCGCAAAGCTCTAGCGGAAAAGGCTGTGAGAGGATTGCGCGCCAGAGCGGCTGGCGTGCGAACGATTGTCGATGTGTCGACTTTCGATATCGGTCGCGACGTCAGTTTATTGGCCGAGGTTTCGCGGGCTGCCGACGTTCATATCTGGCGGCGACCGGCTTGTGGTTCGACCCGCCACTTTCGATGCGATTGAGGTATGTAGAGGAACTCACACTAGTTCTTCCTGCGGTGAGATTCAATATGGCATCGAAGTACACCGGAATTAGGGCGGGCATTATCAAGGTCGCGACCACAGGCAAGGCGACCCCCTTTCAGGAGTTAGTGTTAAAGGCGGCCGCCCGGGCCAGCTTGGCCACCGGTGTTCCGGTAACCACTCACACGGCAGCAAGTCAGCGCGATGGTGAGCGAGGCAGGCCGCCATTTTTGAGTCCGAAGCTTGAGCCCTCACGGGTTTGTATTGGTCACAGCGATGATACTGACGATTTGAGCTATCTCACCGCCCTGCTGCGCGGATACCTCATCGGTCTAGACCACATCCCGCACAGTGCGATTGGTCTAGAAGATAATGCGAGTGCATCACCGCTCCTGGGCATCCGTTCGTGGCAAACACGGGCTCTCTTGATCAAGGCGCTCATCGACCAAGGCTACATGAAACAAATCCTCGTTTCGAATGACTGGCTGTTCGGGTTTTCGAGCTATGTCACCAACATCATGGACGTGATGGATCGCGTGAACCCCGACGGGATGGCCTTCATTCACGATTATAAAGATGATGATGATAAA**TGA**GGGTAGGGGGCTTCAATTCGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAGCTGAgagacctgcat

## verify

this looks good to me – NM 13 jun 2025

## verify translation and reading frame of epitope tag codon optimization

look for RARE codons CTA CTC AGG

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| S.alvi | | |  |  |  |  |
| **Amino Acid** | **Codon** | **Frequency** |  |  | 33% |  |
| \* | TAA | 0.6385 |  |  | 50% |  |
| \* | TAG | 0.1266 |  |  |  |  |
| \* | TGA | 0.2349 |  |  | ratio | Rare? |
| A | GCA | 0.327 | 4 | 0.25 | 1.308 |  |
| A | GCC | 0.2453 | 4 | 0.25 | 0.9812 |  |
| A | GCG | 0.1037 | 4 | 0.25 | 0.4148 | r |
| A | GCT | 0.3241 | 4 | 0.25 | 1.2964 |  |
| C | TGC | 0.366 | 2 | 0.5 | 0.732 |  |
| C | TGT | 0.634 | 2 | 0.5 | 1.268 |  |
| D | GAC | 0.228 | 2 | 0.5 | 0.456 | r |
| D | GAT | 0.772 | 2 | 0.5 | 1.544 |  |
| E | GAA | 0.7825 | 2 | 0.5 | 1.565 |  |
| E | GAG | 0.2175 | 2 | 0.5 | 0.435 | r |
| F | TTC | 0.2329 | 2 | 0.5 | 0.4658 | r |
| F | TTT | 0.7671 | 2 | 0.5 | 1.5342 |  |
| G | GGA | 0.1894 | 4 | 0.25 | 0.7576 |  |
| G | GGC | 0.2604 | 4 | 0.25 | 1.0416 |  |
| G | GGG | 0.083 | 4 | 0.25 | 0.332 | r |
| G | GGT | 0.4671 | 4 | 0.25 | 1.8684 |  |
| H | CAC | 0.2313 | 2 | 0.5 | 0.4626 | r |
| H | CAT | 0.7687 | 2 | 0.5 | 1.5374 |  |
| I | ATA | 0.1929 | 3 | 0.333333 | 0.5787 |  |
| I | ATC | 0.2177 | 3 | 0.333333 | 0.6531 |  |
| I | ATT | 0.5894 | 3 | 0.333333 | 1.7682 |  |
| K | AAA | 0.8189 | 2 | 0.5 | 1.6378 |  |
| K | AAG | 0.1811 | 2 | 0.5 | 0.3622 | r |
| L | CTA | 0.0502 | 6 | 0.166667 | 0.3012 | R |
| L | CTC | 0.0447 | 6 | 0.166667 | 0.2682 | R |
| L | **CTG** | 0.3951 | 6 | 0.166667 | 2.3706 |  |
| L | CTT | 0.1119 | 6 | 0.166667 | 0.6714 |  |
| L | TTA | 0.2572 | 6 | 0.166667 | 1.5432 |  |
| L | TTG | 0.1409 | 6 | 0.166667 | 0.8454 |  |
| M | ATG | 1 | 1 | 1 | 1 |  |
| N | AAC | 0.2513 | 2 | 0.5 | 0.5026 |  |
| N | AAT | 0.7487 | 2 | 0.5 | 1.4974 |  |
| P | CCA | 0.1971 | 4 | 0.25 | 0.7884 |  |
| P | CCC | 0.1163 | 4 | 0.25 | 0.4652 | r |
| P | CCG | 0.411 | 4 | 0.25 | 1.644 |  |
| P | CCT | 0.2757 | 4 | 0.25 | 1.1028 |  |
| Q | CAA | 0.3421 | 2 | 0.5 | 0.6842 |  |
| Q | CAG | 0.6579 | 2 | 0.5 | 1.3158 |  |
| R | **AGA** | 0.11 | 6 | 0.166667 | 0.66 |  |
| R | AGG | 0.0326 | 6 | 0.166667 | 0.1956 | R |
| R | CGA | 0.0697 | 6 | 0.166667 | 0.4182 | r |
| R | CGC | 0.2665 | 6 | 0.166667 | 1.599 |  |
| R | CGG | 0.1391 | 6 | 0.166667 | 0.8346 |  |
| R | CGT | 0.3821 | 6 | 0.166667 | 2.2926 |  |
| S | AGC | 0.2007 | 6 | 0.166667 | 1.2042 |  |
| S | AGT | 0.2642 | 6 | 0.166667 | 1.5852 |  |
| S | TCA | 0.184 | 6 | 0.166667 | 1.104 |  |
| S | TCC | 0.1022 | 6 | 0.166667 | 0.6132 |  |
| S | TCG | 0.0563 | 6 | 0.166667 | 0.3378 | r |
| S | TCT | 0.1927 | 6 | 0.166667 | 1.1562 |  |
| T | ACA | 0.2847 | 6 | 0.166667 | 1.7082 |  |
| T | ACC | 0.2868 | 4 | 0.25 | 1.1472 |  |
| T | ACG | 0.0959 | 4 | 0.25 | 0.3836 | r |
| T | ACT | 0.3326 | 4 | 0.25 | 1.3304 |  |
| V | GTA | 0.3785 | 4 | 0.25 | 1.514 |  |
| V | GTC | 0.1086 | 4 | 0.25 | 0.4344 | r |
| V | GTG | 0.1943 | 4 | 0.25 | 0.7772 |  |
| V | GTT | 0.3187 | 4 | 0.25 | 1.2748 |  |
| W | TGG | 1 | 6 | 0.166667 | 6 |  |
| Y | TAC | 0.2358 | 2 | 0.5 | 0.4716 | r |
| Y | TAT | 0.7642 | 2 | 0.5 | 1.5284 |  |

<Serial Cloner V2.5> -- <Sat, Jun 14, 2025 10:16 AM>

Restriction map of MOD001 opd 13jun2025 #1.xdna

Showing all restriction sites [using RELibrary as a Restriction Enzyme Library]

###

>BsaI(\*) >GG3 >BB025 promoter & RBS

| | |

ag ttg ggt ctc AAA CGc ttt ggc agt tta ttc ttg aca tgt agt gag ggg gct ggt ata atc aca tag tac tgt tat aca gaa aca gag g < 90

10 20 30 40 50 60 70 80

ag ata tta cat atg ATG CAA ACG AGA **AGG** GTT GTG **CTC** AAG TCT GCG GCC GCG AGA ACT CTG **CTC** GGC GGC CTG GCT GGG TGC GCG ACG T < 180

M Q T R R V V L K S A A A R T L L G G L A G C A T W

100 110 120 130 140 150 160 170

GG CTG GAT CGA TCG GCA CAG GCG ATG CGA TCA ATA CGT GCG CGT CCT ATC ACA ATC TCT GAA GCG GGT TTC ACA CTG ACT CAC GAG GAC A < 270

L D R S A Q A M R S I R A R P I T I S E A G F T L T H E D I

190 200 210 220 230 240 250 260

TC TCG GCA GCT CGG CAG GAT TCT TGC GTG CTT GGC CAG AGT TCT TCG GTA GCG CAA AGC TCT AGC GGA AAA GGC TGT GAG **AGG** ATT GCG C < 360

S A A R Q D S C V L G Q S S S V A Q S S S G K G C E R I A R

280 290 300 310 320 330 340 350

GC CAG AGC GGC TGG CGT GCG AAC GAT TGT CGA TGT GTC GAC TTT CGA TAT CGG TCG CGA CGT CAG TTT ATT GGC CGA GGT TTC GCG GGC T < 450

Q S G W R A N D C R C V D F R Y R S R R Q F I G R G F A G C

370 380 390 400 410 420 430 440

GC CGA CGT TCA TAT CTG GCG GCG ACC GGC TTG TGG TTC GAC CCG CCA CTT TCG ATG CGA TTG **AGG** TAT GTA GAG GAA CTC ACA **CTA** GTT C < 540

R R S Y L A A T G L W F D P P L S M R L R Y V E E L T L V L

460 470 480 490 500 510 520 530

>OPD

|

TT CCT GCG GTG AGA TTC AAT ATG GCA TCG AAG TAC ACC GGA ATT **AGG** GCG GGC ATT ATC AAG GTC GCG ACC ACA GGC AAG GCG ACC CCC T < 630

P A V R F N M A S K Y T G I R A G I I K V A T T G K A T P F

550 560 570 580 590 600 610 620

TT CAG GAG TTA GTG TTA AAG GCG GCC GCC CGG GCC AGC TTG GCC ACC GGT GTT CCG GTA ACC ACT CAC ACG GCA GCA AGT CAG CGC GAT G < 720

Q E L V L K A A A R A S L A T G V P V T T H T A A S Q R D G

640 650 660 670 680 690 700 710

GT GAG CGA GGC **AGG** CCG CCA TTT TTG AGT CCG AAG CTT GAG CCC TCA CGG GTT TGT ATT GGT CAC AGC GAT GAT ACT GAC GAT TTG AGC T < 810

E R G R P P F L S P K L E P S R V C I G H S D D T D D L S Y

730 740 750 760 770 780 790 800

AT CTC ACC GCC CTG CTG CGC GGA TAC CTC ATC GGT **CTA** GAC CAC ATC CCG CAC AGT GCG ATT GGT **CTA** GAA GAT AAT GCG AGT GCA TCA C < 900

L T A L L R G Y L I G L D H I P H S A I G L E D N A S A S P

820 830 840 850 860 870 880 890

CG **CTC** CTG GGC ATC CGT TCG TGG CAA ACA CGG GCT **CTC** TTG ATC AAG GCG **CTC** ATC GAC CAA GGC TAC ATG AAA CAA ATC CTC GTT TCG A < 990

L L G I R S W Q T R A L L I K A L I D Q G Y M K Q I L V S N

910 920 930 940 950 960 970 980

AT GAC TGG CTG TTC GGG TTT TCG AGC TAT GTC ACC AAC ATC ATG GAC GTG ATG GAT CGC GTG AAC CCC GAC GGG ATG GCC TTC ATT CAC G < 1080

D W L F G F S S Y V T N I M D V M D R V N P D G M A F I H D

1000 1010 1020 1030 1040 1050 1060 1070

>FLAG

|

AT TAT AAA GAT GAT GAT GAT AAA TGA GGG TAG GGG GCT TCA ATT CGC CAG GCA TCA AAT AAA ACG AAA GGC TCA GTC GAA AGA CTG GGC C < 1170

Y K D D D D K \*

1090 1100 1110 1120 1130 1140 1150 1160

>term >GG4

| |

TT TCG TTT TAT CTG TTG TTT GTC GGT GAA CGC TCT CTA CTA GAG TCA CAC TGG CTC ACC TTC GGG TGG GCC TTT CTG CGT TTA TAG CTG A < 1260

1180 1190 1200 1210 1220 1230 1240 1250

<BsaI(\*)

|

ga gac ctg cat < 1271

1270

Features :

GG3 : [13 : 16]

GG4 : [1256 : 1259]

BB025 promoter & RBS : [17 : 104]

FLAG : [1080 : 1103]

OPD : [105 : 1079]

term : [1127 : 1255]

4 piece GG

è

1111=====backbone====2222=====homL========3333=====gene========4444=====hom

# Alt plan – no gg yes gibson

Does vector exist in the proper backbone with Hom arms already in it, and some other gene?

No: GG does not work this way, but this could work well with gibson

=====backbone=========homL========3333=====othergene========4444=====hom

Bsai è

1. 4444=====homR=========backbone=========homL========3333
2. 3333=====ourgene========4444

A 2-step GG

=====backbone=========homL========3333=====ourgene========4444=====hom