Enhancing Platforms at the Interface of Viruses and Directed Evolution

Samantha D. Levinson

A dissertation submitted to the Department of Chemistry

in partial fulfillment of the requirement for the degree of

Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

May 2021

ENHANCING PLATFORMS AT THE INTERFACE OF VIRUSES AND DIRECTED EVOLUTION

by

Samantha D. Levinson

Advisor: Professor Abhishek Chatterjee, Ph.D.

Abstract

Directed evolution is a powerful technique to expand chemical space in biological systems. In particular, this method has been used to develop cellular machinery to enable genetic code expansion (GCE), the incorporation of unnatural amino acids (UAAs) into proteins during the translation process. GCE relies on evolving an aminoacyl tRNA synthetase (aaRS) and tRNA pair from a different domain of life to incorporate a UAA into proteins in their new host, as these evolutionarily distant pairs are less likely to be cross-reactive with host pairs. The aaRS and tRNA must meet a number of conditions to be useful for GCE: the pair must be orthogonal (non-cross-reactive) to the host's native aaRS/tRNA pairs in order to ensure site-specific UAA incorporation; the aaRS must have an active site suited to accept the shape of the UAA; and the tRNA must cooperate with the host ribosome, elongation and release factors, and other translational machinery to efficiently incorporate the UAA into the protein.

Numerous aaRS/tRNA pairs have been evolved to allow incorporation of diverse UAAs in bacteria due to the tractable nature of these organisms for directed evolution experiments. While an aaRS evolved in bacteria to charge a novel UAA can be used in eukaryotes, tRNAs cannot be evolved for GCE in bacteria and then used in eukaryotes

because they will not have evolved in the presence of the correct translational machinery. It is necessary to evolve tRNAs directly in their host cells. Unfortunately for researchers working on GCE in mammalian cells, it is difficult to perform directed evolution on small gene products in these hosts. Transformation efficiency in mammalian cells is poor, and transient transfection yields heterogeneous DNA distribution to target cells, making selection based on performance of individual library members impossible. Viruses are an ideal DNA delivery vector for mammalian cells, as production of recombinant viruses allows control over library member generation, and viruses can be delivered with exquisite copy number control. The Chatterjee lab recently developed a platform, Virus-Assisted Directed Evolution of tRNAs (VADER), using adeno-associated virus (AAV) to evolve tRNAs for GCE directly in mammalian cells.

While VADER is the first directed evolution platform that allows the evolution of small gene products in mammalian cells, its efficiency is limited by its continued reliance on transient transfection to deliver non-library DNA that is necessary for the production of rAAV. To overcome this limitation, baculovirus delivery vectors were developed to boost DNA delivery and AAV capsid production to improve virus production efficiency during selections. VADER allows the evolution of tRNAs to incorporate certain UAAs, but the technique relies on installing a UAA into the AAV capsid, which is sensitive to disruption caused by slight modifications in structure. To expand the scope of VADER to evolve tRNAs for UAAs that cannot be incorporated into the AAV capsid, an alternate selection handle (Assembly Activating Protein, or AAP) was deleted from the genome and provided *in trans* to incorporate 5-hydroxytryptophan (5HTP). Incorporating the UAA into this

flexible protein allows UAA-dependent production of AAV and expands the scope of tRNAs that can be evolved in mammalian cells.

Acknowledgements

I would like to thank my family for supporting me throughout my life, and especially over the last five years. A PhD may seem like an individual endeavor, but it is truly a group effort. Thank you to my mom, Laurie Lauchlan, for telling me at age 12 that I actually had to start doing my homework if I wanted to become a scientist and for working so hard so I could succeed. Thank you to my dad, Harry Levinson, for explaining the natural world to me in technical terms even when I was young and helping me know I could understand anything if I worked hard enough. Special thanks to my sister, Sarah Levinson, for supporting me without fail and showing me that being willing is worth it. Thank you to my Auntie Lynn and Uncle Sao for supporting me through my Master's, and to my Uncle Dave for encouraging me in my studies and in my swimming. Thank you to all of my grandparents and ancestors; I wish you could be here to read this thesis. I love you all.

Thank you to my advisor, Professor Abhishek Chatterjee, for your mentorship and guidance throughout these past five years. I would not be here if not for you.

Many thanks to my labmates. To Sarah Erickson, who blew my mind on my first day in lab by explaining how to tell what volume each pipette could hold, and who became the best labmate I could ask for and an excellent friend. To Rachel Kelemen, Xiaofu Cao, and Yunan Zheng, who taught me how to make viruses. To Rachel Huang and Lily Jewel for their technical assistance with tRNAs. And to Katie Grasso, Shu Wang, Elise Ficaretta, and Conor Loynd for their friendship.

Thank you to my friends. I would not have made it without your support. Thank you to Christine Ondris, who told me that if Britney Spears could make it through 2007, I could make it through 6 months of chemistry. It ended up being a little bit longer than 6

months, and your friendship and support got me through it. Thank you to Rachael and David Gelineau, my second family, for being there for me through all the trials and triumphs. Thank you to Olaf Haraldson for your support and for every single thing made out of stone. Thanks to Alyssa Greene and Andy Crow for friendship, baked goods, and many edits. Thank you to Professor (and friend) Cheryl Hackworth for sharing your love of nature both huge and tiny. Thank you to my coach, Craig Lewin, for being patient with me when I was in lab instead of in the water. Thank you many to the Northern Water Tribe: Fran O'Loughlin, Laurie Craigen, Amy Craigen, Chris Graefe, Karen Nazor, Rena Demeo, Natalie Lang, Stacie McHale, Elaine Howley, Jessica O'Herron, and everyone I've met through open water swimming. You all cheer me on, and I couldn't do it without you.

Thank you to Dr. Cheryl Cavanaugh for reminding me that I am not my grade.

Thank you so much to Daniela Klaz. We may have begun at the end, but your love has gotten me through the wall.

Finally, I would like to thank Professor Glenn Eldridge. This is all your fault.

Table of Contents

List of Figures

Chapter 2 Figures

Chapter 3 Figures

Chapter 4 Figures

Chapter 5 Figures

List of Abbreviations

Standard one-letter and three-letter abbreviations are used for the 20 natural amino acids.

5-FOA	5-fluoroorotic acid
5HTP	5-hydroxy-L-tryptophan
A2	acceptor stem library hit 2
AAP	Assembly Activating Protein
aaRS	aminoacyl-tRNA synthetase
AAV	adeno-associated virus
AAV2	adeno-associated virus 2
AcK	N^{ε}-acetyl-L-lysine
Ac*M*NPV	*Autographa californica* multicapsid nucleopolyhedrovirus
Ad	Adenovirus
ATM	altered translational machinery
AzK	azidolysine
BocK	N^{ε}-Boc-L-Lysine
CAPPIA	cell array protein-protein interaction assay
CMV	cytomegalovirus
CO_2	carbon dioxide
CRISPR	clustered regularly interspaced short palindromic repeats
CSS	consensus splice site
DNA	deoxyribonucleic acid
DBCO	dibenzocyclooctyne
DiazK	diazirine lysine
DMEM	Dulbecco's Modified Eagle Medium
EcWRS	*Escherichia coli* tryptophanyl aminoacyl-tRNA synthetase
EcWtR	*E. coli* tryptophanyl tRNA
EcYtR	*E. coli* tyrosyl tRNA
eGFP	enhanced green fluorescent protein
GPCR	G protein-coupled receptor
GFP	green fluorescent protein
HEK-293T	human embryonic kidney 293 cells with a temperature-sensitive allele of the SV40 T antigen
HIV	human immunodeficiency virus
HSV	herpes simplex virus
HTS25	high-throughput sequencing hit 25
IPTG	Isopropyl β-d-1-thiogalactopyranoside
ITR	inverted terminal repeat
Kb	kilobase
LB	Luria broth
LDL-R	low-density lipoprotein receptor
MbPylRS	*Methanosarcina barkeri* pyrrolysyl aminoacyl-tRNA synthetase
MmPytR	*M. mazei* pyrrolysyl tRNA
mL	milligram

MOI	multiplicity of infection
MTH	metallothionein
NBK	norbornene lysine
NES	nuclear export sequence
ORF	open reading frame
P0	passage 0
P1	passage 1
PBS	phosphate-buffered saline
PCK	photocaged lysine
PEG	polyethylene glycol
PEI	polyethylenimine
PPI	protein-protein interaction
PytR	pyrrolysyl tRNA
rAAV	recombinant adeno-associated virus
RNA	ribonucleic acid
RC2	*RepCap2*
RFU	relative fluorescence units
SEAP	secreted alkaline phosphatase
SCOK	cyclo-octyne lysine
Sf9	*Spodoptera frugiperda* 9 cells
Sf21AE	*S. frugiperda* 21AE cells
snRNP	small nuclear ribonucleoprotein
tRNA	transfer ribonucleic acid
U	unit
UAA	unnatural amino acid
UbiC	Ubiquitin-C
VADER	Virus-Assisted Directed Evolution of tRNAs
VP	viral protein
VSV-G	vesicular stomatitis virus G glycoprotein
WRS	tryptophanyl aminoacyl-tRNA synthetase
WtR	tryptophanyl tRNA
Wt	wild-type
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
YtR	tyrosyl tRNA
µg	microgram
µL	microliter

Chapter 1

Introduction

1.1 Directed evolution

Directed evolution is a method to engineer gene products by a controlled process similar to artificial selection, in which a gene for the biomolecule of interest is modified over subsequent rounds of mutagenesis followed by selection for a desired feature.[1,2] This method of engineering is useful in chemistry because it allows researchers to bypass rational catalyst design[1] to achieve a number of important goals. Through directed evolution, researchers have expanded chemical space in biological systems,[3–5] developed novel methods to create therapeutics,[6] and developed novel functions in biological catalysts.[7–9] Directed evolution is a valuable technique to rapidly develop new chemistries in biological systems.

1.1.1 Methods to evolve gene targets

There are two major approaches to directed evolution: screening and selection. In both methods, the investigator generates a library of variants, usually via random mutagenesis of a particular region of interest in the target gene.[2,10] When using a screening method to search for active variants, or "hits," the result of an assay is directly tied to the target gene's desired function. Screening thereby allows straightforward interpretation of the results of the experiment. However, because a researcher must perform the assay to test each library member individually, the size of the test library is restricted to the number of iterations it is feasible for the researcher to perform. This limitation may prevent investigators from testing and accessing the best possible hits. Unlike in a screening system, a directed evolution experiment using selection requires that the performance of each library member is tied to the survival of an organism or the propagation of a virus. It is necessary for a researcher to develop a selection scheme to provide this link between

propagation and the activity of the library member. While developing such a scheme can be a complex process, this technique allows for a much larger library size because it is not necessary to manually test the performance of each individual library member. Nonfunctional library members will not appear as outputs of each round of selection, and over multiple rounds, the best hits can become enriched in the output, linking desired function with evolutionary fitness.

1.1.2 The state of the field of directed evolution in mammalian cells

Directed evolution schemes exist to engineer target gene products in bacteria, yeast, and phages,[11–16] but a variety of constraints prevents such schemes from being widely used directly in mammalian cells. Transformation efficiency in bacteria allows the creation of large libraries in bacterial cells, where each cell contains only one library member. Mammalian cells exhibit poor transformation efficiency,[17] so transient transfection is often used to introduce foreign DNA into mammalian cells. Transient transfection using such reagents as polyethylenimine results in an extremely heterogeneous distribution of plasmid DNA into target cells, such that some cells receive multiple copies of the plasmid of interest and some cells receive none. Such a plasmid distribution cannot ensure that cells each receive only one library member, making directed evolution impossible because the cell's output phenotype could be the result of the activity of any of a number of library members. Unlike bacteria and yeast, mammalian cells do not amplify and maintain foreign DNA, so retrieving the sequences of successful library members is technically challenging.

Different methods have been developed to circumvent the impracticality of introducing a DNA library by transformation. Wang and Tsien developed a method of directed evolution in B-cells relying on somatic hypermutation.[6] RNA viral vectors such

as Sindbis virus[18] and human immunodeficiency virus (HIV)-1[19] with naturally high genome replication error rates, as well as DNA viruses such as adenovirus with viral polymerases evolved to be more error prone,[20] have been used to deliver the DNA for directed evolution to the host cells. These methods successfully allowed the evolution of a variety of biomolecules of interest, such as a new fluorescent protein,[6] GPCRs,[18] and transcription factors.[18–20] However, the utility of the techniques using somatic hypermutation and the viral vectors previously described are limited to the evolution of large genes. The generation of library members in these techniques rely on error-prone DNA replication in either the genome of the host cell or the viral vectors themselves, such that the mutation frequency required to successfully target small genes would be lethal for the host cell or would prevent replication of the viral vector by damaging other areas of their genomes. In addition, such overall random mutagenesis may not yield improved hits as efficiently as a library mutated in a target area, because random mutagenesis cannot specifically focus on the active site of an enzyme to change its substrate specificity.[13] These limitations have thus far inhibited the development of directed evolution schemes in mammalian cells.

1.2 Unnatural amino acid incorporation in proteins in mammalian cells

1.2.1 How to incorporate unnatural amino acids into proteins

Directed evolution is a useful technique to develop tools that permit the incorporation of unnatural amino acids (UAAs) into proteins via translation. Introducing UAAs site-specifically into proteins allows for expanded chemical space beyond the limited functionalities of the 20 canonical amino acids, so an important goal for research in chemical biology is to incorporate UAAs with high efficiency into proteins of interest.

Unfortunately, while cell-free protein production techniques such as solid-state synthesis allow tight control over the placement of a UAA, these methods are expensive and have low yields. Therefore, in order to efficiently incorporate UAAs into proteins, it is necessary to do so during the translation process. To achieve this goal, researchers must evolve translational machinery in the target organism to optimize UAA incorporation within the cell. Directed evolution of this machinery has allowed researchers to incorporate multiple types of UAAs into proteins in both bacteria and eukaryotic cells, including bio-orthogonal conjugation handles,[21–23] fluorescent probes,[24,25] photo-crosslinkers,[26–28] photocaged groups,[29,30] and residues mimicking post-translational modifications.[31,32]

Introducing a UAA into proteins via translation requires the host cell to have an aminoacyl-tRNA synthetase (aaRS) and tRNA pair that can work with the native translation machinery to incorporate the amino acid at a specific site in a protein. Such aaRS/tRNA pairs must be orthogonal to the native translational machinery: that is, they must not cross-react with any of the existing cellular machinery. The orthogonal aaRS must accept only the UAA, and none of the natural amino acids, whereas native synthetases must not charge the UAA, or else that UAA's incorporation into proteins will not be site-specific. Native synthetases should not be able to charge the orthogonal tRNA with any natural amino acids, nor should the orthogonal aaRS be able to charge a native tRNA with a UAA. The orthogonal tRNA must only incorporate the UAA in response to a blank codon, typically one of the stop codons (often the amber stop codon TAG) or a quadruplet codon.[17] If all of these conditions are met, a cell should be able to incorporate a UAA in response to the appropriate blank codon.

Typically, aaRS/tRNA pairs are taken from a distantly related domain of life and introduced into the host organism to maximize the chance that they are orthogonal in the host species.[17] aaRS/tRNA pairs from archaea and eukaryotes are typically orthogonal in bacteria, and pairs from bacteria are typically orthogonal in eukaryotes, due to the distant evolutionary relationship between these domains of organisms. The pyrrolysyl aaRS/tRNA pair from methanogenic archaea in the family Methanosarcinaceae naturally incorporates pyrrolysine in response to the TAG codon, and this pair is naturally orthogonal in both bacteria and eukaryotes because TAG is a stop codon for species in those domains.[33]

When aaRS/tRNA pairs are incorporated into new hosts, they may not function optimally with the native translational machinery, or they may not incorporate a desired UAA into proteins, so it is necessary to evolve these pairs to optimize their performance and UAA incorporation. Evolution of either the tRNA or the aaRS can be done in bacteria via a two-step selection scheme.[17] During the positive selection step, cells with a blank codon in a gene for antibiotic resistance are transformed with the aaRS/tRNA pair and grown in the presence of antibiotics and the UAA; only cells that receive a functional aaRS/tRNA pair can grow. In the negative selection step, bacteria with a blank codon in a toxic gene are transformed with the aaRS/tRNA pair but are not grown with the UAA. Cells that receive a library member that does not incorporate a natural amino acid in response to the blank codon survive; in this manner, aaRS/tRNA pairs that are active and orthogonal are identified. A similar two-step selection system is used in yeast, where the UAA must be incorporated in a transcriptional activator for necessary metabolic genes during the positive selection, and the negative selection uses a blank codon-containing *URA3* gene and 5-FOA to weed out library members that incorporate natural amino acids

in response to the blank codon.[17] These two-step selection schemes have allowed the successful evolution of a number of aaRS/tRNA pairs for the incorporation of UAAs into proteins at the translational level.

1.2.2 UAA incorporation in mammalian cells

Many more UAAs have been incorporated in bacteria than in eukaryotes because it is easier to perform directed evolution in bacteria than in eukaryotes.[17,34,35] Only a subset of those UAAs that have been incorporated in eukaryotes have been used in mammalian cells. Directed evolution techniques are not available for evolving aaRS/tRNA pairs in mammalian cells due to the above-described technical limitations of performing directed evolution in mammalian cells, so aaRS/tRNA pairs are often evolved in yeast and then used in mammalian cells.[11,13] A recent platform developed by Italia et al.[35] allows the evolution of bacterial synthetases in *Escherichia coli* with Altered Translational Machinery (ATM) to select for new synthetase variants that can be imported into mammalian cells in order to incorporate UAAs which were previously unavailable in these systems. While the ATM platform has the potential to vastly expand the number of synthetases and UAAs that can be used in mammalian cells, it does not solve every problem for UAA incorporation in mammalian cells, as the synthetase is not the only piece of cellular machinery involved in translation.

In order to site-specifically incorporate a UAA, an evolved synthetase need only interact with its tRNA and the UAA, making evolution of synthetases for use in eukaryotes a relatively simple problem to solve using the ATM platform or by performing directed evolution in yeast. However, the tRNA in question must interact with the host ribosome and elongation factors to install its UAA in the protein containing the blank codon. The

translational machinery in mammals with which the tRNA must interact is highly divergent from that of bacteria, and is even divergent between yeast and mammalian cells, which are both eukaryotes. The activity of the orthogonal tRNA can be the limiting component in UAA incorporation in proteins in mammalian cells.[36] It is therefore necessary to evolve the tRNA directly in mammalian cells to improve its performance in incorporating UAAs into proteins in mammalian cells.

1.3 Adeno-associated virus as a tool for DNA delivery & directed evolution

1.3.1 Viruses can be used to transduce mammalian cells

In order to evolve tRNAs in mammalian cells to optimize UAA incorporation, researchers need robust selection schemes that can take place directly in mammalian cells. As discussed above, delivering DNA to mammalian cells is the major challenge in developing successful selection schemes for directed evolution. Transformation efficiency is poor in mammalian cells,[17] and transient transfection provides a highly heterogeneous distribution of DNA to target cells so as to render impossible any selection scheme in which a cell's phenotype is based on the genotype of a single library member that it receives. An additional complication is that, unlike bacteria and viruses, mammalian cells do not maintain plasmids, so retrieving the library members after a round of selection is difficult. Any DNA delivery method for directed evolution in mammalian cells must overcome these obstacles.

One DNA delivery vector that circumvents each of these challenges is the virus. Because viruses have evolved to enter and transduce cells, any virus that can infect mammalian cells can overcome the problem of poor transformation efficiency in these cells. Mammalian viruses have also evolved to amplify and package their genetic material

within the host cell, in this case the mammalian cells that otherwise do not maintain or amplify foreign DNA. This feature allows researchers to retrieve library members after each round of selection and then store them in a relatively stable state; use them to infect new cells for subsequent rounds of selection; or characterize their sequences. Modern molecular cloning techniques make producing recombinant viruses relatively simple, so scientists can add the gene for their biomolecule of interest to the virus genome or replace one of its components with a library member. It is also a straightforward and precise process to titer a virus, so it is possible to have exquisite control over the copy number of virus delivered to cells. These features of viruses allow researchers to deliver at most one library member to host cells during a selection, and also allow them to retrieve and characterize DNA after the selection.

Multiple viruses have been used to deliver DNA to mammalian cells for chemical biology experiments. Adeno-associated virus (AAV) has been widely used as a DNA delivery vector because it is replication-deficient, which gives it a favorably biosafety profile for lab work.[37,38] While it is easy to produce rAAV, its small genome capacity limits the sizes of the genes of the biomolecules of interest it can carry as cargo.[39–42] *Autographa californica* multicapsid nucleopolyhedrovirus (Ac*M*NPV, or baculovirus) is a virus that infects insect cells but has been modified to include the vesicular stomatitis virus G glycoprotein (VSV-G) in its envelope, which allows it to enter any cell that expresses the mammalian low-density lipoprotein receptor (LDL-R).[43–46] Baculovirus has an essentially unlimited capacity for recombination in its double-stranded DNA genome due to its capsid assembly process, which makes it an attractive delivery vector for large genetic cargos.[47] Sindbis virus and HIV-1 are RNA viruses with relatively large genomes that, as previously

discussed, have been used for directed evolution experiments that rely on the error-prone polymerases of these RNA viruses. Adenovirus (Ad) also has a larger cargo capacity, as well as a stable double-stranded DNA genome. HIV, Sindbis virus, and Ad all require additional safety protocols for handling in the lab, as these viruses are all human pathogens.

1.3.2 AAV as a facile tool for mammalian cell transduction

One of the most popular viruses for use in chemical biology experiments is AAV. AAV is a small, non-enveloped virus from the family Parvoviridae, with a single-stranded DNA genome approximately 5 kb in length.[39] The genome consists of three open reading frames (ORFs) flanked by two inverted terminal repeats (ITRs) that act as origins of replication and packaging signals.[48,49]

The ORF *Rep* codes for four non-structural proteins responsible for AAV DNA replication and packaging.[50–53] One of these proteins, Rep78, induces apoptosis by activating caspase-3,[54] so the AAV genome cannot be stably integrated into mammalian cells during rAAV production. *Cap* codes for the three structural proteins of the capsid; VP1, VP2, and VP3 are present in the capsid in a 1:1:10 ratio.[37] The capsid proteins fit tightly together, restricting the size of the cargo in recombinant AAV (rAAV) to strictly 5 kb or smaller and limiting structural modifications to the capsid. The constraints on rAAV cargo size and AAV's genome toxicity result in the need for AAV genes to be delivered separately from the recombinant cargo during rAAV production.

AAV's third ORF codes for the nonstructural protein Assembly Activating Protein (AAP), which is required for capsid assembly, and whose gene sits within *Cap* in a different reading frame with the weak CTG start codon.[40] Its ORF occurring in the same mRNA as the Cap proteins facilitates the timely expression of all proteins required for the capsid

assembly process. AAP contains threonine/serine rich regions between these termini that allow these regions to be quite flexible,[55,56] which may provide opportunity for manipulation.

AAV is a replication-deficient virus and requires a co-infection of a helper virus (typically adenovirus, but sometimes herpesvirus or baculovirus) to replicate within mammalian cells.[38,57–59] The regions of adenovirus that are required for AAV replication are: VA, E2A, and E4.[60] Ad DNA-binding protein, which is coded for in the E2A gene, increases the processivity of AAV DNA replication.[59] Proteins from the E4 ORF and the VA RNA genes promote the second strand synthesis of AAV.[37] Without second strand synthesis, AAV cannot replicate. While AAV was discovered as a contaminant in Ad purifications,[38] adenovirus is not the only helper virus for AAV. Both herpes simplex virus 1 and 2 (HSV-1 and HSV-2) provide complete helper functions for AAV by allowing *Rep* gene expression and recruiting AAV DNA to replication centers.[37,57] Baculovirus can also act as a helper virus for AAV, but the mechanisms by which it does so are unknown.[37]

Because it is replication deficient, AAV is an attractive target for use in clinical research and gene therapies due to its inability to cause harm to researchers and patients.[37] Production of rAAV is also facile, as Rep proteins will package any genes into the capsid as long as they sit within the ITRs and are 5kb or less. AAV is therefore a useful tool to use in transducing mammalian cells.

1.3.3 An AAV-based selection scheme for mammalian cells

Virus-Assisted Directed Evolution of tRNAs (VADER), a new directed evolution technique developed by Kelemen et al.,[61] uses rAAV with wild-type capsids to deliver a library of pyrrolysyl tRNA variants to mammalian cells. In that study, the authors infect

cells at a multiplicity of infection (MOI) of less than one virus per cell so that each cell receives at most one library member. They provide the other genes via transient transfection that AAV needs to replicate and produce a UAA-containing capsid (*AdHelper*, *Rep*, and a TAG-containing *Cap*), as well as the UAA azidolysine (AzK) in the cell media. Only cells that received an active tRNA are able to incorporate an amino acid into response to the TAG codon within *Cap,* and these tRNAs are enriched during this positive selection. The newly packaged capsids of viruses containing an orthogonal tRNA contain the azide functional group, but those containing cross-reactive tRNAs contain only canonical amino acids. The negative selection step in this scheme consists of a bio-orthogonal capture, in which the virus capsids are conjugated to DBCO-biotin and are then captured on a streptavidin resin and released by exposure to light. The hits resulting from this selection can be amplified, sequenced, and further characterized to determine their activity in relation to the wild-type tRNA. In that paper, the authors identify a tRNA, A2, that exhibits more than 300% improved activity over the wild-type pyrrolysyl tRNA via transient transfection.

VADER fills an important gap in the field of directed evolution, in that this technique makes it possible to perform selections directly in mammalian cells in order to optimize UAA-incorporating tRNA function in the intended host. This technique ensures that each cell receives at most one library member, which has been generated in a controlled fashion by molecular cloning. VADER does not rely on somatic hypermutation, which allows VADER to evolve small gene products such as tRNAs. As discussed previously, viruses such as AAV evolved to amplify themselves inside the host cell, so as long as the infected cell receives the other genes necessary for AAV replication *in trans*, the cell will

maintain and replicate the library member enough to retrieve and characterize hits from the selection.

However groundbreaking this technique is, there are still limitations to VADER. While hits can be retrieved from each round of selection using this method, it is necessary to re-amplify, re-clone, and re-package the output virus from each round of selection before proceeding to successive rounds of selection. Each of these steps is labor-intensive and takes considerable time. The negative selection from VADER also relies on capturing the output virus using a bio-orthogonal conjugation handle incorporated into the virus capsid, which consists of tightly-fitted proteins that may not tolerate perturbation by all UAAs of interest.[62–64] Indeed, the AAV2 capsid does not tolerate the insertion of several UAAs of interest to researchers, including 5-hydroxytryptophan (unpublished data). It is necessary to modify the established VADER technique to improve the efficiency of the workflow, and so that it does not rely on UAA insertion into the AAV capsid in order to perform selections on tRNAs for such UAAs that the AAV capsid does not tolerate.

1.4 Expanding the scope of viral vectors in the field of directed evolution

In this thesis, I present work to enhance directed evolution platforms in mammalian cells. I demonstrate that it is possible to re-amplify rAAV using recombinant baculoviruses to streamline the workflow between rounds of selection. This modification to the VADER method allows researchers to proceed more quickly between rounds of selection, making it possible to arrive at an optimized evolutionary solution faster than if re-cloning the selection output was necessary after every round.

In addition to using these baculovirus vectors to streamline the current VADER process, I also demonstrate that it is possible to evolve tRNAs using only the positive

selection scheme from VADER without incorporating a UAA into the AAV capsid. Instead, I incorporate the UAA of interest into the essential protein AAP. The termini of AAP are highly structured and do not tolerate modifications, but the flexible threonine/serine rich region in the middle of the protein tolerates modifications and even large insertions.[56,65] This region of AAP is therefore an ideal location to modify with UAAs that would not be tolerated in the virus capsid, and since AAP is essential to AAV replication,[40] I demonstrate that I can use it as a selection handle to evolve tRNAs. Although a negative selection is not possible with this system because there is no bio-conjugation handle in the AAV capsid here, I show here that the positive selection from VADER is sufficient to evolve a tRNA to incorporate 5-hydroxytrypotphan into proteins in mammalian cells.

Additionally, I present multiple baculovirus vectors that make use of VADER-evolved pyrrolysyl tRNAs to improve the incorporation of diverse UAAs into proteins in mammalian cells. These vectors demonstrate that the tRNAs evolved using the VADER platform suppress TAG stop codons at higher levels than do the wild-type pyrrolysyl tRNA, to a higher degree than previously reported when comparing tRNA performances via transient transfection. Using baculovirus allows the delivery of tRNAs to cells at a controlled level, such that researchers can deliver far fewer copies of the tRNA per cell than is possible using transient transfection. These experiments demonstrate that the evolved tRNA's improvements are especially evident at low copy numbers. I have also incorporated different evolved synthetases into these baculovirus vectors to allow for the incorporation of diverse UAAs, and I have optimized the expression of the protein of

interest by changing the promoter for the aaRS[36] and adding a nuclear export sequence to the terminus of the aaRS to ensure that it localizes to the cytosol.

The use of selections in chemical biology research is not limited to evolving gene products. In this thesis, I demonstrate the utility of AAV-based selections to discover protein binding partners in the mammalian proteome. Using packaged rAAV as the output, I show here the development of a two-hybrid selection platform in which the association of bait and prey proteins is necessary to drive transcription of VP1, an AAV capsid protein that is essential to produce infectious virus.[66] The gene for the prey protein is inserted into the rAAV genome, and the genome will only be repackaged into an AAV capsid if the prey and bait proteins associate, leading to the transcription of the essential VP1 protein. Any virus which is produced will only be infectious if the bait and prey proteins associate, so this system can be used to select for prey-binding partners to any particular bait within the mammalian proteome.

In summary, I demonstrate in this thesis the utility of selection schemes in mammalian cells; I furthermore enhance and expand their applications within current systems. I have developed viral vectors that will be useful to the larger scientific community for stop codon suppression and UAA incorporation. The modifications and enhancements I have made to a protocol for a selection scheme in mammalian cells will make the process more efficient and generalizable. Finally, the two-hybrid system that I have developed based on infectious virus production will allow high-throughput scanning of the human proteome for protein binding partners. These developments expand the scope of directed evolution in mammalian cells.

1.5 References

(1) Arnold, F. H. *Acc. Chem. Res.* **1998**, *31* (3), 125.
(2) Sutherland, J. D. *Curr. Opin. Chem. Biol.* **2000**, *4* (3), 263.
(3) Addy, P. S.; Erickson, S. B.; Italia, J. S.; Chatterjee, A. *J. Am. Chem. Soc.* **2017**, *139* (34), 11670.
(4) Usanov, D. L.; Chan, A. I.; Maianti, J. P.; Liu, D. R. Nat. Chem. **2018**, 10 (7), 704.
(5) Chen, K.; Huang, X.; Kan, S. B. J.; Zhang, R. K.; Arnold, F. H. *Science* **2018**, *360* (6384), 71.
(6) Wang, L.; Tsien, R. Y. *Nat. Protoc.* **2006**, *1* (3), 1346.
(7) Pedersen, Henrick; Hölder, Swen; Sutherlin, Daniel P.; Schwitter, Urs; King, David S.; Schultz, P. G. *Proc. Natl. Acad. Sci.* **1998**, *95* (18), 10523.
(8) Arnold, F. H. *Trends Biotechnol.* **1990**, *8* (C), 244.
(9) Umeno, D.; Arnold, F. H. *J. Bacteriol.* **2004**, *186* (5), 1531.
(10) E. T. Farinas, T. Bulter, F. H. A. *Curr. Opin. Biotechnol.* **2001**, *12* (6), 545.
(11) Chen, P. R.; Groff, D.; Guo, J.; Ou, W.; Cellitti, S.; Geierstanger, B. H.; Schultz, P. G. *Angew. Chemie Int. Ed.* **2009**, *48* (22), 4052.
(12) Bessette, P. H.; Rice, J. J.; Daugherty, P. S. *Protein Eng. Des. Sel.* **2004**, *17* (10), 731.
(13) Liu, W.; Brock, A.; Chen, S.; Chen, S.; Schultz, P. G. *Nat. Methods* **2007**, *4* (3), 239.
(14) Boder, E. T.; Wittrup, K. D. *Nat. Biotechnol.* **1997**, *15* (6), 553.
(15) Esvelt, K. M.; Carlson, J. C.; Liu, D. R. *Nature* **2011**, *472* (7344), 499.
(16) Clackson, T.; Hoogenboom, H. R.; Griffiths, A. D.; Winter, G. *Nature* **1991**, *352* (6336), 624.
(17) Liu, C. C.; Schultz, P. G. *Annu. Rev. Biochem.* **2010**, *79* (1), 413.
(18) English, J. G.; Olsen, R. H. J.; Lansu, K.; Strachan, R. T.; Wacker, D.; Correspondence, B. L. R. *Cell* **2019**, *178*, 748.
(19) Das, A. T.; Zhou, X.; Vink, M.; Klaver, B.; Verhoef, K.; Marzio, G.; Berkhout, B. *J. Biol. Chem.* **2004**, *279* (18), 18776.
(20) Berman, C. M.; Papa, L. J.; Hendel, S. J.; Moore, C. L.; Suen, P. H.; Weickhardt, A. F.; Doan, N. D.; Kumar, C. M.; Uil, T. G.; Butty, V. L.; Hoeben, R. C.; Shoulders, M. D. *J. Am. Chem. Soc.* **2018**, *140* (51), 18093.
(21) Dommerholt, J.; Schmidt, S.; Temming, R.; Hendriks, L. J. A.; Rutjes, F. P. J. T.; van Hest, J. C. M.; Lefeber, D. J.; Friedl, P.; van Delft, F. L. *Angew. Chemie Int. Ed.* **2010**, *49* (49), 9422.
(22) Nguyen, D. P.; Lusic, H.; Neumann, H.; Kapadnis, P. B.; Deiters, A.; Chin, J. W. *J. Am. Chem. Soc.* **2009**, *131* (25), 8720.
(23) Elliott, T. S.; Townsley, F. M.; Bianco, A.; Ernst, R. J.; Sachdeva, A.; Elsässer, S. J.; Davis, L.; Lang, K.; Pisa, R.; Greiss, S.; Lilley, K. S.; Chin, J. W. *Nat. Biotechnol.* **2014**, *32* (5), 465.
(24) Chatterjee, A.; Guo, J.; Lee, H. S.; Schultz, P. G. *J. Am. Chem. Soc.* **2013**, *135* (34), 12540.
(25) Luo, J.; Uprety, R.; Naro, Y.; Chou, C.; Nguyen, D. P.; Chin, J. W.; Deiters, A. *J. Am. Chem. Soc.* **2014**, *136* (44), 15551.
(26) Hino, N.; Okazaki, Y.; Kobayashi, T.; Hayashi, A.; Sakamoto, K.; Yokoyama, S.

Nat. Methods **2005**, *2* (3), 201.
(27) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *J. Am. Chem. Soc.* **2002**, *124* (31), 9026.
(28) Lin, S.; He, D.; Long, T.; Zhang, S.; Meng, R.; Chen, P. R. *J. Am. Chem. Soc.* **2014**, *136* (34), 11860.
(29) Gautier, A.; Nguyen, D. P.; Lusic, H.; An, W.; Deiters, A.; Chin, J. W. *J. Am. Chem. Soc.* **2010**, *132* (12), 4086.
(30) Arbely, E.; Torres-Kolbus, J.; Deiters, A.; Chin, J. W. *J. Am. Chem. Soc.* **2012**, *134* (29), 11912.
(31) Neumann, H.; Peak-Chew, S. Y.; Chin, J. W. *Nat. Chem. Biol.* **2008**, *4* (4), 232.
(32) Kim, C. H.; Kang, M.; Kim, H. J.; Chatterjee, A.; Schultz, P. G. *Angew. Chemie - Int. Ed.* **2012**, *51* (29), 7246.
(33) Krzycki, J. A. *Curr. Opin. Microbiol.* **2005**, *8* (6), 706.
(34) Dumas, A.; Lercher, L.; Spicer, C. D.; Davis, B. G. *Chemical Science*. Royal Society of Chemistry January 1, 2015, pp 50–69.
(35) Italia, J. S.; Addy, P. S.; Wrobel, C. J. J.; Crawford, L. A.; Lajoie, M. J.; Zheng, Y.; Chatterjee, A. *Nat. Chem. Biol.* **2017**, *13* (4), 446.
(36) Zheng, Y.; Lewis, T. L.; Igo, P.; Polleux, F.; Chatterjee, A. *ACS Synth. Biol.* **2017**, *6* (1), 13.
(37) Samulski, R. J.; Muzyczka, N. *Annu. Rev. Virol.* **2014**, *1* (1), 427.
(38) Atchison, R. W.; Casto, B. C.; Hammon, W. M. D. *Science (80-.).* **1965**, *149* (3685), 754.
(39) Srivastava, A.; Lusby, E. W.; Berns, K. I. *J. Virol.* **1983**, *45* (2).
(40) Sonntag, F.; Schmidt, K.; Kleinschmidt, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (22), 10220.
(41) Nakai, H.; Storm, T. A.; Kay, M. A. *Nat. Biotechnol.* **2000**, *18* (5), 527.
(42) Dyka, F. M.; Boye, S. L.; Chiodo, V. A.; Hauswirth, W. W.; Boye, S. E. *Hum. Gene Ther. Methods* **2014**, *25* (2), 166.
(43) Jarvis, D. L. In *Methods in Enzymology*; Academic Press Inc., 2009; Vol. 463, pp 191–222.
(44) Barsoum, J.; Brown, R.; McKee, M.; Boyce, F. M. *Hum. Gene Ther.* **1997**, *8* (17), 2011.
(45) Kaikkonen, M. U.; Räty, J. K.; Airenne, K. J.; Wirth, T.; Heikura, T.; Ylä-Herttuala, S. *Gene Ther.* **2006**, *13* (4), 304.
(46) Kost, T. A.; Condreay, J. P.; Jarvis, D. L. *Nature Biotechnology*. Nature Publishing Group May 5, 2005, pp 567–575.
(47) Kost, T. A.; Condreay, J. P. *Trends in Biotechnology*. Elsevier Current Trends April 1, 2002, pp 173–180.
(48) McLaughlin, S. K.; Collis, P.; Hermonat, P. L.; Muzyczka, N. *J. Virol.* **1988**, *62* (6).
(49) Hauswirth, W. W.; Berns, K. I. *Virology* **1977**, *78* (2), 488.
(50) Pereira, D. J.; McCarty, D. M.; Muzyczka, N. *J. Virol.* **1997**, *71* (2), 1079.
(51) Ni, T. H.; McDonald, W. F.; Zolotukhin, I.; Melendy, T.; Waga, S.; Stillman, B.; Muzyczka, N. *J. Virol.* **1998**, *72* (4), 2777.
(52) Im, D. S.; Muzyczka, N. *Cell* **1990**, *61* (3), 447.
(53) Im, D. S.; Muzyczka, N. *J. Virol.* **1992**, *66* (2).

(54) Schmidt, M.; Afione, S.; Kotin, R. M. *J. Virol.* **2000**, *74* (20), 9441.
(55) Tse, L. V.; Moller-Tank, S.; Meganck, R. M.; Asokan, A. *J. Virol.* **2018**, *92* (14).
(56) Naumer, M.; Sonntag, F.; Schmidt, K.; Nieto, K.; Panke, C.; Davey, N. E.; Popa-Wagner, R.; Kleinschmidt, J. A. *J. Virol.* **2012**, *86* (23), 13038.
(57) Buller, R. M. L.; Janik, J. E.; Sebring, E. D.; Rose, J. A. *J. Virol.* **1981**, *40* (1), 241.
(58) Urabe, M.; Nakakura, T.; Xin, K.-Q.; Obara, Y.; Mizukami, H.; Kume, A.; Kotin, R. M.; Ozawa, K. *J. Virol.* **2006**, *80* (4), 1874.
(59) Ward, P.; Dean, F. B.; O'Donnell, M. E.; Berns, K. I. *J. Virol.* **1998**, *72* (1), 420.
(60) Matsushita, T.; Elliger, S.; Elliger, C.; Podsakoff, G.; Villarreal, L.; Kurtzman, G. J.; Iwaki, Y.; Colosi, P. *Gene Ther.* **1998**, *5* (7), 938.
(61) Kelemen, Rachel E.; Jewel, Delilah; Pasha, Muhammad; Cao, Xiaofu; Chatterjee, A. *Submiss.*
(62) Xie, Q.; Bu, W.; Bhatia, S.; Hare, J.; Somasundaram, T.; Azzi, A.; Chapman, M. S. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (16), 10405.
(63) Agbandje-McKenna, M.; Kleinschmidt, J. *Methods Mol. Biol.* **2011**, *807*, 47.
(64) Kelemen, R. E.; Mukherjee, R.; Cao, X.; Erickson, S. B.; Zheng, Y.; Chatterjee, A. *Angew. Chemie Int. Ed.* **2016**, *55* (36), 10645.
(65) Tse, L. V.; Moller-Tank, S.; Meganck, R. M.; Asokan, A. *J. Virol.* **2018**, *92* (14).
(66) Girod, A.; Wobus, C. E.; Zádori, Z.; Ried, M.; Leike, K.; Tijssen, P.; Kleinschmidt, J. A.; Hallek, M. *J. Gen. Virol.* **2002**, *83* (5), 973.

Chapter 2

Baculovirus vectors to amplify AAV during directed evolution

2.1 Introduction

The principles of directed evolution and the limitations of performing selections directly in mammalian cells are outlined in detail in Chapter 1. This chapter deals with the use of viruses instead of transient transfection to transduce mammalian cells to boost the amplification of library members during and after selections.

2.1.1 Directed evolution in mammalian cells

Directed evolution is a highly important tool in chemical biology that allows researchers to access new chemistries in biological contexts.[1–5] While this technique is highly developed in bacteria, yeast, and phages,[6–12] the controlled evolution of small gene products in mammalian cells was not possible until relatively recently, due to specific restrictions on working with mammalian cells. Because of poor transformation efficiency in mammalian cells,[5] researchers usually introduce foreign DNA into mammalian cells via transient transfection. Unfortunately, this system does not permit the controlled introduction of one library member per cell due to the heterogeneous distribution of DNA uptake during transient transfection. Somatic hypermutation can be used to generate libraries within mammalian cells,[13] but its reliance on random mutagenesis may not yield the kind of targeted libraries that researchers often prefer.[6] This technique is also limited to large gene products because the mutational frequency necessary to generate a library via this method in a small gene product would kill the host cell.

It is relatively simple to transform yeast and bacteria, both of which also maintain and amplify plasmid DNA. Researchers may therefore generate a randomized library via established cloning methods and transform this library into either bacteria or yeast. These cells will have at most one library member per cell, and the researcher may proceed with

selections on this transformed library and retrieve the DNA afterwards for sequence characterization. Mammalian cells, on the other hand, do not maintain foreign DNA as readily, so it is challenging to retrieve and sequence the DNA from any selections undertaken in mammalian cells.

2.1.2 Using viruses to transduce mammalian cells

A solution that overcomes all of these challenges of performing selections in mammalian cells is to transduce the cells with viruses, rather than relying on transformation, transient transfection, or somatic hypermutation. While mammalian cells do not naturally maintain foreign DNA, viruses evolved to replicate themselves within their host cells, so mammalian viruses should be able to amplify their own genetic material once they infect the cell. Retrieving and characterizing the library members becomes simple. It is also possible to control the copy number of viruses delivered to the host cell, as one need only infect cells with a titered virus at a level of fewer than one virus per cell, to ensure that each cell has at most one library member when selection begins. Finally, since recombinant viruses are easy to produce via molecular cloning, it is possible to generate a library of small gene products that would be inaccessible to directed evolution via somatic hypermutation, or even continuous evolution techniques using viral vectors that depend upon error-prone polymerases.[14–16] Viruses are the ideal vector for directed evolution of small gene products in mammalian cells.

Adeno-associated virus (AAV) is an excellent virus for use in mammalian cell-based directed evolution experiments. AAV is a small, non-enveloped virus from the family Parvoviridae with a 5kb single-stranded DNA genome.[17] This genome consists of three open reading frames (ORFs). The ORF *Rep* codes for four nonstructural proteins

expressed from the p5 and p19 promoters that replicate and package the genome into the capsid.[18–21] *Cap* is expressed from the constitutively active p40 promoter and codes for the three capsid proteins VP1, VP2, and VP3.[22,23] Assembly Activating Protein (AAP) is also expressed from the p40 promoter, from a weak CTG start codon in a different reading frame from the capsid proteins.[23,24] All of these genes are flanked by two inverted terminal repeats (ITRs) that act as origins of replication and packaging signals.[17,25,26] AAV is a replication deficient virus that requires helper genes from adenovirus, baculovirus, or herpesvirus to replicate inside mammalian cells.[27–30] These genes can be provided by transient transfection to eliminate the need for coinfection by a helper virus during AAV production.[31]

AAV is of particular interest to researchers for its potential uses for gene therapies due to its low immunogenicity and lack of cytotoxicity due to being replication deficient.[23,32] This replication deficiency and low toxicity to human and animal subjects makes AAV an attractive vector to use in directed evolution applications as well. rAAV is simple to create: any genes up to 5kb can be packaged inside an AAV capsid as its single-stranded DNA genome as long as they sit within the ITRs. Since it is easy and relatively safe to manipulate rAAV, this virus was the vector of choice for Virus-Assisted Directed Evolution of tRNAs (VADER).[33]

2.1.3 VADER: an AAV-based directed evolution platform

VADER is a technique that relies on AAV as a vector to deliver libraries of small gene products to cells.[33] In this work, a library of mutant pyrrolysyl tRNAs (PytR) was packaged along with a fluorescent reporter into wild type AAV capsids. Since it is simple to measure an infective titer for this virus by flow cytometry, it is possible to infect a

population of cells at a multiplicity of infection (MOI) less than 1, such that each cell will be infected by at most one virus. This level of control is necessary to ensure that one cell does not receive 2 or more library members. The authors provided *AdHelper*, helper genes from adenovirus that AAV needs to replicate,[31,34] along with *Rep* and *Cap* with a T454TAG mutation. They also added the unnatural amino acid (UAA) azidolysine (AzK) into the cell media, so that any cell that received a virus containing an active PytR variant would produce AAV with AzK in the capsid. This exposed AzK allowed the authors to conjugate the virus to DBCO-biotin and perform a pulldown on a streptavidin resin to capture only virus that contained AzK in the capsid. The pulldown excluded all viruses that had cross-reactive PytR library members that incorporated canonical amino acids in response to the TAG codon in *Cap*. The active library members could be re-amplified out of the AAV genome and re-cloned into plasmids for sequencing, further characterization, and transient transfection into mammalian cells to produce AAV for subsequent rounds of selection.

Re-cloning the outputs of each round of selection is necessary because after the negative selection's streptavidin pulldown and photo-release, there are so few viral particles remaining that if one were to proceed directly to the next round of selection with those output viruses, there would be too few cells both infected and transfected to undergo another round of selection. Indeed, it is necessary to concentrate the output virus to even proceed with re-cloning.[33] While re-cloning the output virus is desirable for sequencing and testing hits in transient transfection experiments, it is a time-consuming and labor-intensive process that delays further selections. It is therefore necessary to develop a procedure to re-amplify output AAV between rounds of selection that does not depend on re-cloning the library.

While VADER is an important new technique in the field of directed evolution, there are limitations to the protocol. During the positive selection, cells are infected with a library of tRNA-containing rAAV, and then transfected with plasmids containing the other genes necessary to replicate and repackage rAAV. This system limits the amplification of rAAV in multiple ways. Not every cell which is infected by a library member-containing virus will receive the transfected plasmids because of the heterogeneous distribution of DNA during transient transfection.[35–37] Transfection reagents like PEI (Figure 2.1) block the entry of viruses into cells, so it is possible that not every library member makes it into a cell to undergo selection, which creates the additional drawback of possibly missing highly active variants during selection. With more complex selections involving more genetic cargo, it might be necessary to use multiple plasmids to deliver all the necessary DNA, and transfection efficiency drops with each additional plasmid used. If researchers wish to optimize this selection system, it will be necessary to break its dependence on transient transfection.

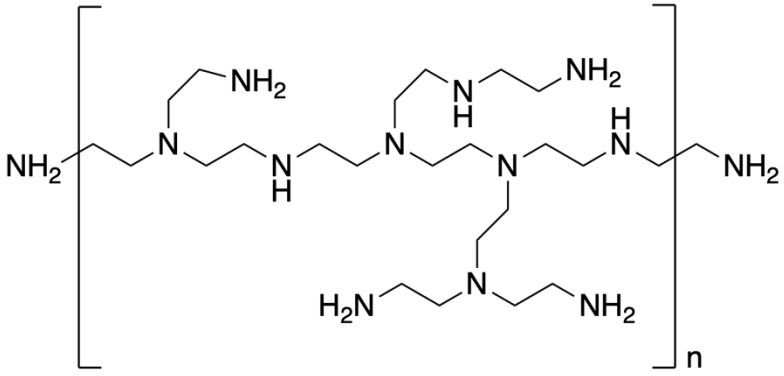


Figure 2.1. Structure of polyethylenimine (PEI)

2.1.4 DNA delivery by baculovirus

While AAV is a small virus with strictly limited cargo capacity, other viruses are not so limited in the amount of recombinant DNA they can package. Baculovirus is one such virus that has a much larger cargo capacity. The *Autographa californica* multicapsid nucleopolyhedrovirus (Ac*M*NPV), the most commonly used baculovirus, has a large double stranded circular genome almost 134kb in length.[38–40] Since this virus infects insects, it will replicate in insect cells but not in mammalian cells, and therefore poses minimal risk to human researchers, a feature which makes this virus useful in laboratory settings.[39] Baculovirus has a large, potentially unlimited cargo capacity of at least 38kb of DNA; this large cargo capacity is likely due to its capsid assembly process, which seems to extend to accommodate extra DNA.[41] This feature makes it an attractive vector to deliver DNA to mammalian cells.

Recombinant baculovirus is relatively simple to produce. Because the virus genome is so large, direct cloning into it is difficult; therefore, shuttle plasmids containing the genes of interest can be transformed into *Escherichia coli* cells which maintain a bacmid containing the rest of the baculovirus genome. The shuttle plasmid undergoes recombination with the bacmid, resulting in a recombinant bacmid that can be isolated and transfected into insect cells.[38,42] Recombinant baculovirus are typically grown in Sf9 or Sf21AE cells derived from *Spodoptera frugiperda* ovarian tissue, which can be conveniently cultured either in suspension or adherent.[39] Even though baculovirus with a wild type envelope can infect mammalian cells, transduction efficiency is low. Barsoum et al. engineered a baculovirus vector to express the vesicular stomatitis virus G glycoprotein (VSV-G) in its envelope, which allows the virus to enter any cell that expresses the low density lipoprotein receptor (LDL-R).[43] This modification greatly improves baculovirus

transduction in mammalian cells by improving viral entry into cells by augmenting gp46-mediated endosomal release.[44,45] Recombinant baculovirus with their large cargo capacities have become useful and facile vectors for research in mammalian cells.

It is possible to produce rAAV on a large scale in Sf9 cells using baculovirus to provide helper genes as well as the AAV genes.[46–48] Sf9 cells can be co-infected with two recombinant baculoviruses: one containing the rAAV genome containing the genes of interest, and the other containing helper genes along with *Rep* and *Cap*.[48] Since Sf9 cells can be grown in suspension to scale up production, this system allows the production of high titer rAAV. This system provides a model for approaching rAAV re-amplification. The output of VADER selections cannot be re-amplified in Sf9 cells, as AAV does not infect insect cells efficiently, but it is possible to use recombinant baculovirus to re-amplify AAV in mammalian cells because of the VSV-G modification that allows baculovirus to infect mammalian cells.

Here I present a method to efficiently re-amplify rAAV in mammalian cells using baculovirus vectors to provide *AdHelper* genes and the AAV genes *Rep* and *Cap*. I demonstrate the utility of co-infection with baculovirus over transient transfection to produce rAAV during the positive selection process in VADER and to re-amplify rAAV in mammalian cells after a round of selection. In addition, I show that the expression of capsid proteins from their endogenous promoters expressed from baculovirus vectors limits the re-amplification of rAAV and provide an engineered baculovirus vector to overcome this limitation.

2.2 Results and Discussion

2.2.1 Determining the effects of addition of PEI on AAV infection

A drawback of the VADER protocols of infecting cells with rAAV containing the tRNA library and then providing the other genes necessary to make AAV via transient transfection is that the presence of PEI or other transfection reagents blocks AAV entry into the cell. This limitation prevents researchers from ensuring that all library members enter cells and are subject to selection, so the best hits from a randomized library could potentially be missed. Transfection with more plasmids is also less efficient; as such, optimizing complex selections using transient transfection would also be more difficult. In order to prevent these problems, dependence on transient transfection during selections must be reduced.

To quantify the effect of PEI on AAV infection efficiency, I infected HEK-293 cells with rAAV-GFP at either 1 or 5 MOI, and added PEI simultaneously, 4 hours later (to simulate VADER's transfection protocol), or not at all (Figures 2.2 and 2.3). Pictures of infected cells expressing the GFP carried as cargo in the rAAV genome show that the addition of PEI inhibits viral transduction of HEK cells (Figure 2.2). Delaying the addition of PEI improves infectivity, but it does not rescue it to the levels achieved when PEI is not added. Using GFP fluorescence as a proxy for infection efficiency, it appears that an MOI of 5, and adding the PEI 4 hours after infection, is necessary to achieve infection results similar to those when using an MOI of 1, without PEI (Figure 2.3). The addition of PEI clearly inhibits transduction of mammalian cells by AAV.

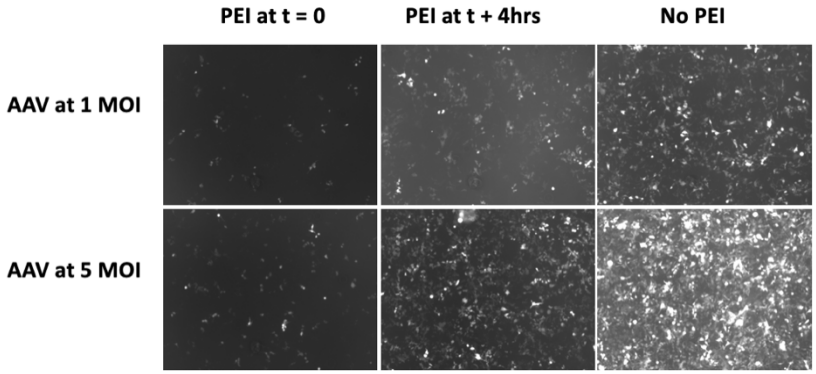


Figure 2.2. The effects of PEI on transduction by rAAV. HEK-293T cells were infected with rAAV-GFP at either 1 or 5 MOI. Cells also received no PEI, PEI at the time of infection (t = 0), or PEI 4 hours post-infection (t + 4hrs).

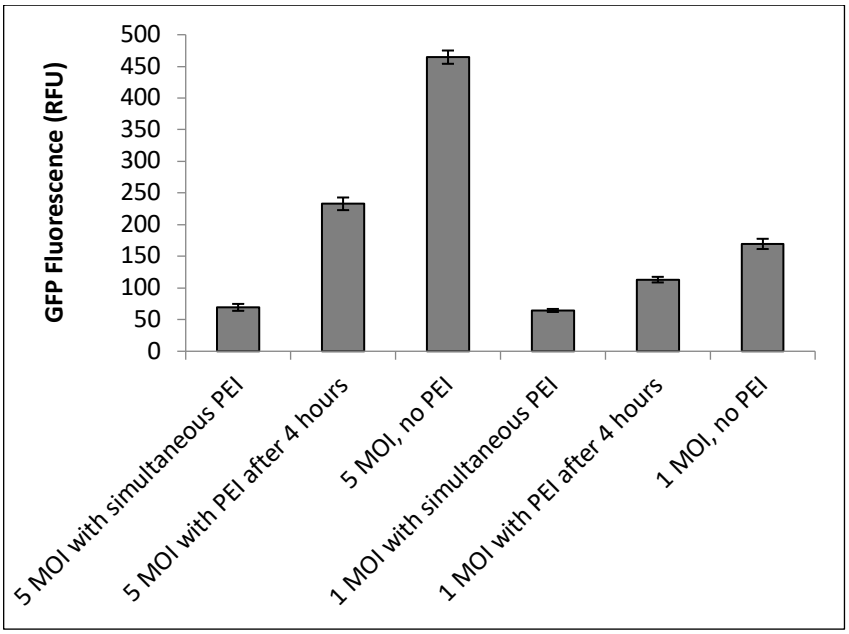


Figure 2.3. The effects of PEI on GFP production in infected cells. Cells infected with rAAV-GFP and treated with PEI were lysed, and GFP fluorescence was measured.

2.2.2 Comparing pAcBac vectors to original rAAV production plasmids

In order to improve the efficiency of rAAV transduction of mammalian cells during directed evolution experiments, it is necessary to provide the other genes (*AdHelper*, *Rep*, and *Cap*) necessary to make rAAV by a means other than transient transfection. Baculovirus vectors present a promising solution to this problem because of their

practically unlimited cargo capacity, their favorable biosafety profile, and the ease with which they can be manipulated to deliver cargo to mammalian cells.

I prepared baculovirus vector shuttle plasmids containing these genes and compared the production of infective rAAV via original infection with rAAV and transient transfection using the original pHelper and pIDT-MbPylRS-RC2wt plasmids versus the shuttle plasmids pAcBac-Helper and pAcBac-MbPylRS-RC2wt. The pAcBac shuttle plasmids contained the same cargo as the pIDT and pHelper plasmids, but in the plasmid backbone that can undergo recombination with a bacmid to produce recombinant baculovirus.

After 72 hours, the cells packaging rAAV show similar numbers of fluorescent cells, indicating that the cells were infected by similar amounts of input AAV, and that the input virus recombinant genome undergoes replication, transcription, and translation at similar rates (Figure 2.4). However, cells infected with AAV harvested from those packaging cells and produced using the pAcBac shuttle plasmids exhibit far less fluorescence than cells infected with the same volume of AAV produced with the original VADER plasmid system (Figure 2.5), and far fewer new cells seemed to be infected with that harvested AAV (Figure 2.4).

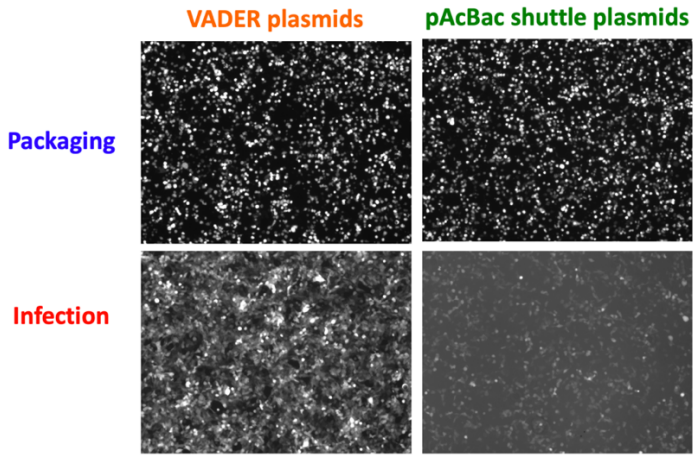


Figure 2.4. Cells packaging rAAV and cells infected with harvested lysate. Cells were imaged 72 hours after being infected with rAAV-GFP and transfected with either the original VADER plasmids or pAcBac shuttle plasmids. Packaging cells were lysed, and new cells were infected with that lysate and imaged after 48 hours.

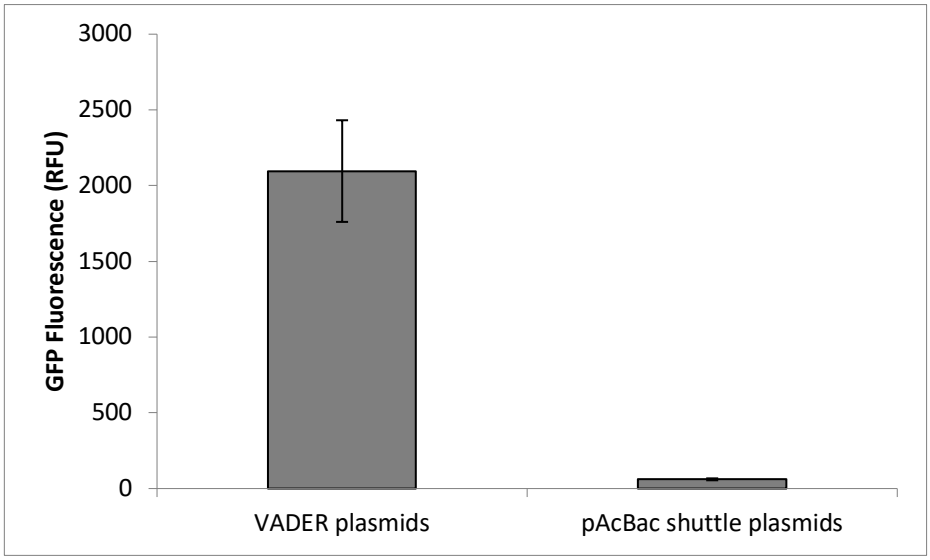


Figure 2.5. GFP fluorescence of cells infected with lysate from cells packaging AAV. New cells were infected with lysate from cells that were infected with rAAV-GFP and transfected with either the original VADER plasmids or with the pAcBac shuttle plasmids. Cells were imaged and lysed after 48 hours, and GFP fluorescence measurements were taken.

In order to determine which plasmid (pAcBac-Helper or pAcBac-MbPylRS-RC2wt) is at fault for reducing the amount of rAAV produced via transient transfection, I infected packaging cells with rAAV-GFP and then transfected with the following pairs of plasmids: pHelper and pIDT-MbPylRS-RC2wt, pAcBac1-Helper and pIDT-MbPylRS-RC2wt, pHelper and pAcBac-MbPylRS-RC2wt, and pAcBac1-Helper and pAcBac-MbPylRS-RC2wt (Figure 2.6). I compared the fluorescence of cells which were infected with lysate from the packaging cells (Figure 2.7). The packaging cells which were transfected with pAcBac-MbPylRS-RC2wt produced much less AAV, regardless of which helper plasmid they received, than cells transfected with pIDT-MbPylRS-RC2wt (based

on the results from cells infected with their lysate). An anti-AAV western blot of lysate from packaging cells shows that much less capsid protein is produced when the AAV genes are provided in the pAcBac backbone (Figure 2.8).

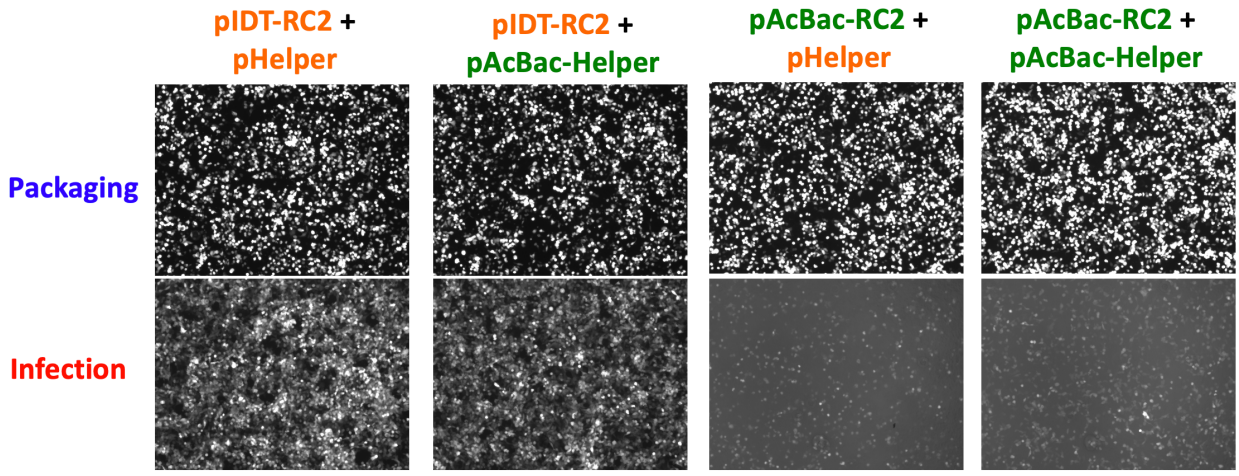


Figure 2.6. Determining the source of limited rAAV production. Cells were infected with rAAV-GFP and transfected with *AdHelper* and *RC2* from either one of the original VADER plasmids or one of the pAcBac shuttle plasmids. After 72 hours, cells were imaged and lysed, and new cells were infected with that lysate. These cells were imaged and lysed after 48 hours, and GFP fluorescence measurements were taken.

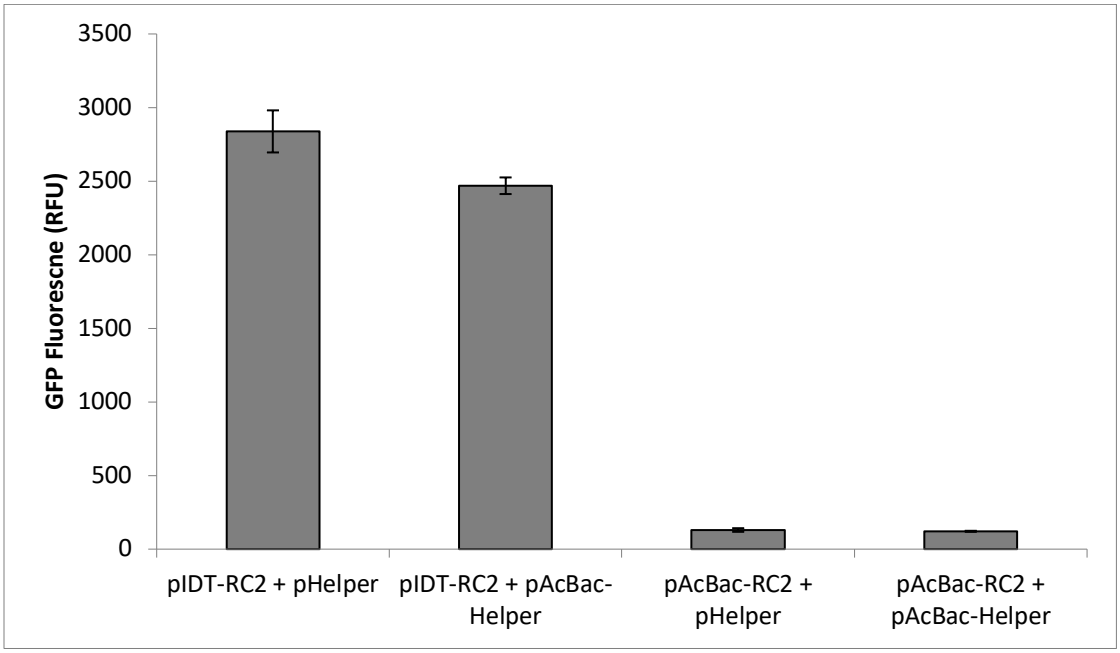


Figure 2.7. GFP fluorescence of cells to test the source of limited rAAV production. Cells were infected with lysate from packaging cells transfected with combinations of the original VADER plasmids and the pAcBac shuttle plasmids to determine the source of reduced rAAV production. After 48 hours, cells were imaged and lysed, and GFP fluorescence was measured.

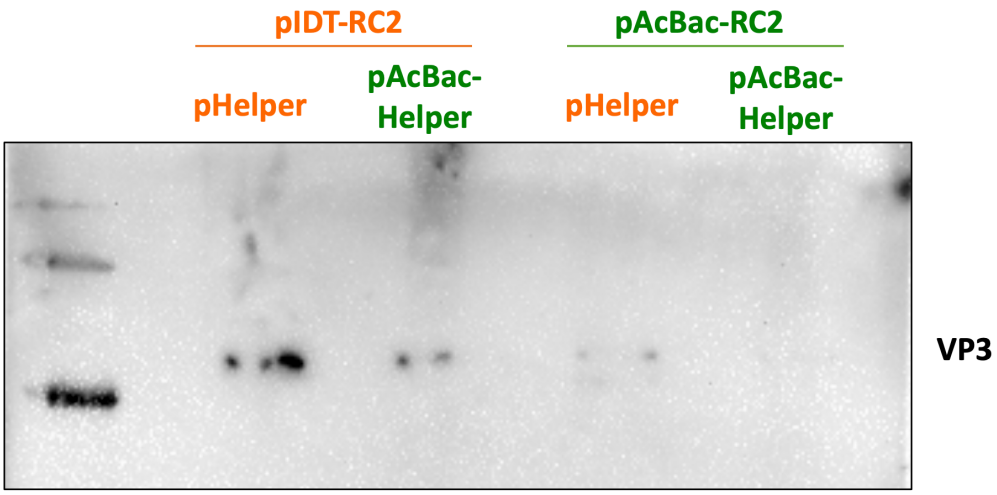


Figure 2.8. Anti-AAV western blot. An anti-VP western blot was performed on lysate from cells packaging rAAV.

To solve this problem of capsid protein limitation, I cloned out the *Cap* genes from AAV2 and placed them under the control of the strong constitutive CMV promoter in a pAcBac shuttle plasmid to maximize production of capsid proteins. I also added the U1 snRNP consensus binding splice site to improve the transcription efficiency of the *Cap* genes. I cloned *Rep* in separately, keeping these genes under the control of their endogenous p5 and p19 promoters to create the final plasmid pAcBac1-Rep-CMV-CapWT-CSS-CMV-MbPylRS (see Appendix for construct map and sequence). I found that overexpressing *Cap* genes form the CMV promoter rescued the production of capsid proteins (Figure 2.9) and also rescued the production of infectious virus (Figure 2.10, Figure 2.11).

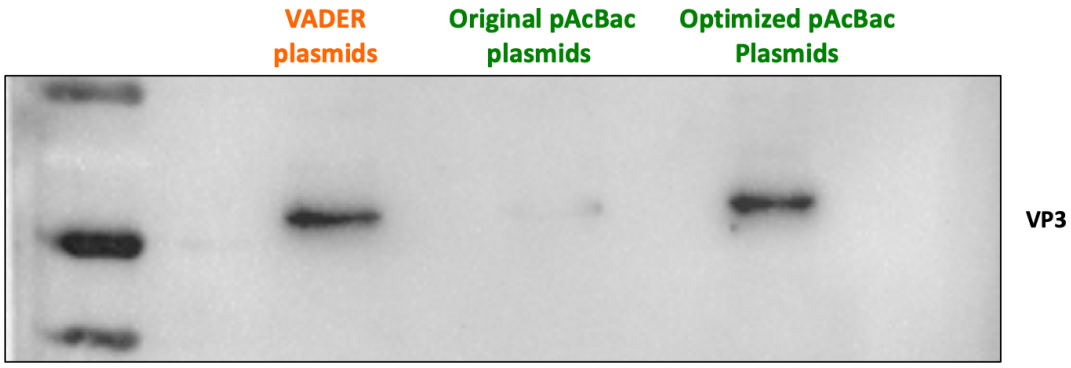


Figure 2.9. Anti-AAV western blot. An anti-AAV western blot was performed on cells packaging rAAV using the VADER plasmids, the original pAcBac shuttle plasmids, or the optimized pAcBac shuttle plasmids.

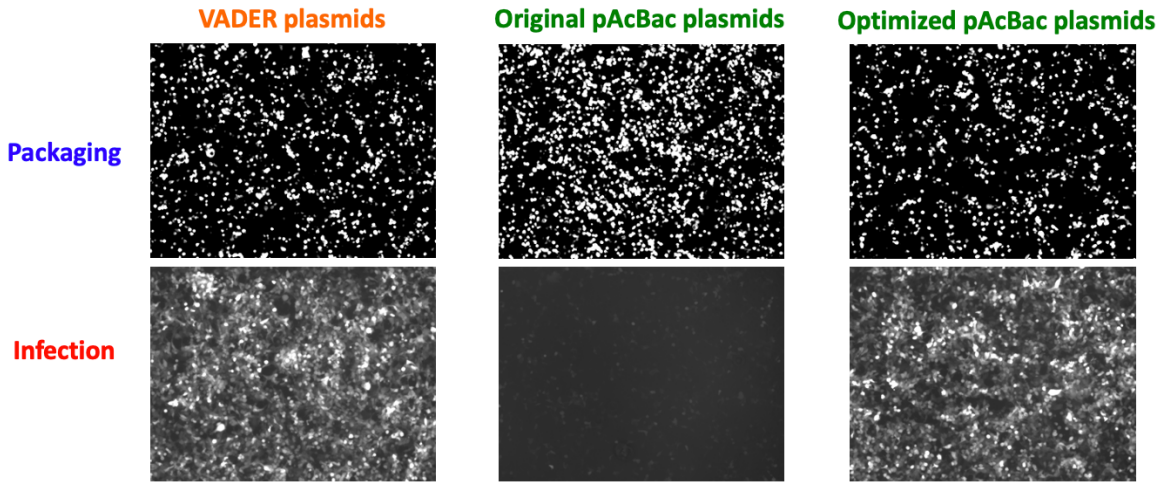


Figure 2.10. Transfecting cells with optimized plasmids. Cells were infected with rAAV and transfected with the VADER plasmids (pHelper and pIDT-RC2-MbPylRS), the original pAcBac shuttle plasmids (pAcBac-Helper and pAcBac1-RC2-MbPylRS), or the optimized pAcBac shuttle plasmids (pAcBac-Helper and pAcBac1-Rep-CMV-CapWT-CSS-CMV-MbPylRS). After 72 hours, the packaging cells were imaged and lysed, and new cells were infected with that lysate. These cells were imaged after 48 hours.

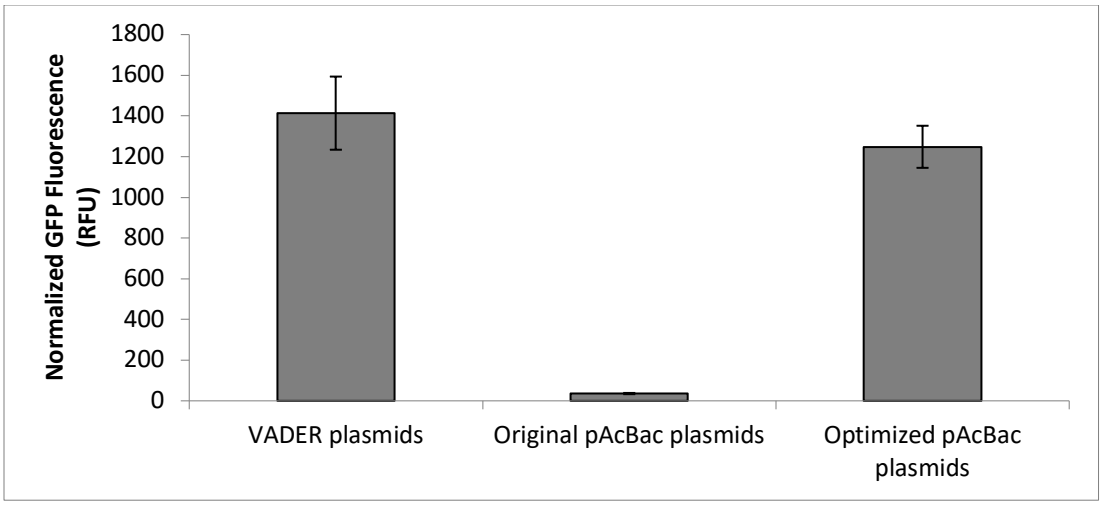


Figure 2.11. GFP fluorescence of cells to test the efficacy of the optimized pAcBac shuttle plasmids. New cells were infected with lysate from cells packaging AAV. After 48 hours, these cells were imaged and lysed, and GFP fluorescence was taken.

I then compared the production of rAAV produced by cells that were infected with rAAV-mCherry and which received *AdHelper*, *Rep*, and *Cap* via the following different methods: transfection with the original VADER plasmid system, transfection with my optimized pAcBac shuttle plasmids, or infection with P1 baculovirus produced from those optimized pAcBac plasmids. Based on images of cells infected with the same volume of lysate from the cells that packaged the rAAV (Figure 2.12), and based on mCherry fluorescence of the cells infected with that lysate (Figure 2.13), it appears that transfection with the optimized pAcBac plasmids produced a similar amount of AAV as transfection with the original VADER plasmids. However, infection with the optimized P1 baculoviruses vastly improved the amount of AAV produced, increasing mCherry fluorescence by 5-fold.

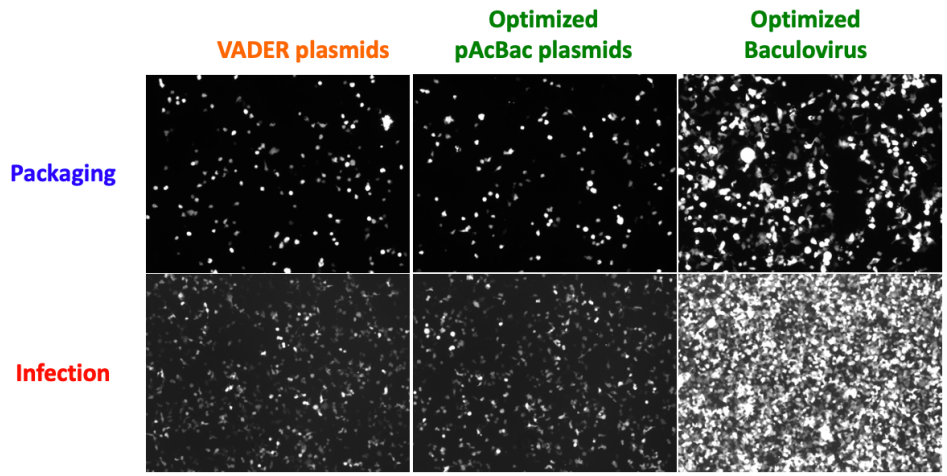


Figure 2.12. Cells packaging rAAV and cells infected with rAAV-containing lysate. Cells were infected with rAAV-mCherry and transfected with the original VADER plasmids, transfected with the optimized pAcBac shuttle plasmids, or infected with baculovirus made from the optimized pAcBac shuttle plasmids. After 72 hours, the packaging cells were imaged and lysed, and new cells were infected with the lysate. These cells were imaged after 48 hours.

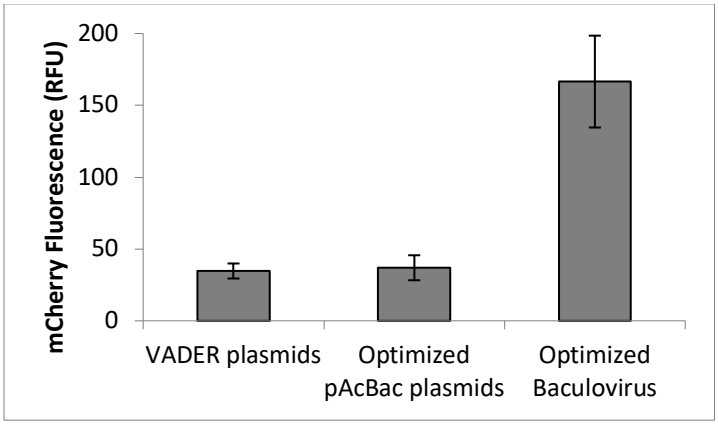


Figure 2.13. mCherry fluorescence of cells infected with rAAV-containing lysate. Cells were infected with rAAV produced by packaging cells to test the effectiveness of the optimized shuttle plasmids and the baculovirus produced from those shuttle plasmids. After 48 hours, these cells were imaged and lysed, and mCherry fluorescence was taken.

2.2.3 Comparing pAcBac vectors to original VADER plasmids for stop codon suppression

In the original VADER protocol, rAAV is packaged into wild type capsids, and it is only during the positive selection itself that viral capsids containing AzK are produced. It is therefore necessary to determine whether the baculovirus infection system will

outperform VADER's original transient transfection system in producing more output rAAV.

I compared the amount of output virus produced in cells infected by rAAV-mCherry-PytR and either infected with P1 baculovirus produced using the plasmids pAcBac1-Helper and pAcBac1-Rep-CMV-Cap454TAG-CSS-CMV-MbPylRS or transfected with pHelper and pIDT-MbPylRS-RC2-454TAG. Instead of AzK, which is an excellent selection handle for VADER but a suboptimal substrate for MbPylRS, I used BocK as the UAA for this experiment, as bio-conjugation was not required and as BocK is a superior substrate for the synthetase (Figure 2.14). Surprisingly, the 454TAG version of the optimized baculovirus did not lead to an increase in the amount of infectious AAV produced (Figure 2.15).

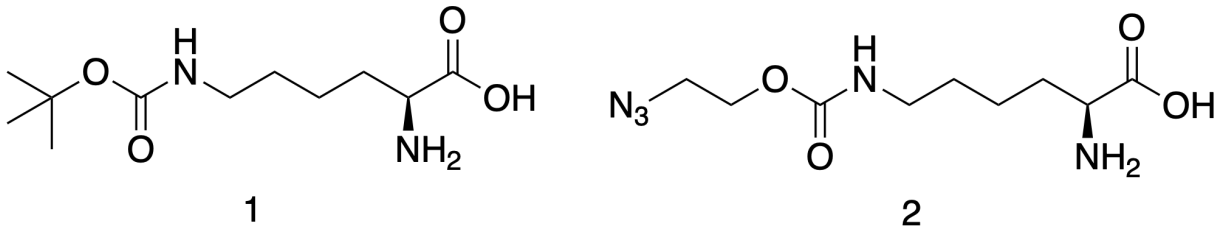


Figure 2.14. Structures of BocK (1) and AzK (2). BocK was used as the unnatural amino acid in all experiments in this chapter.

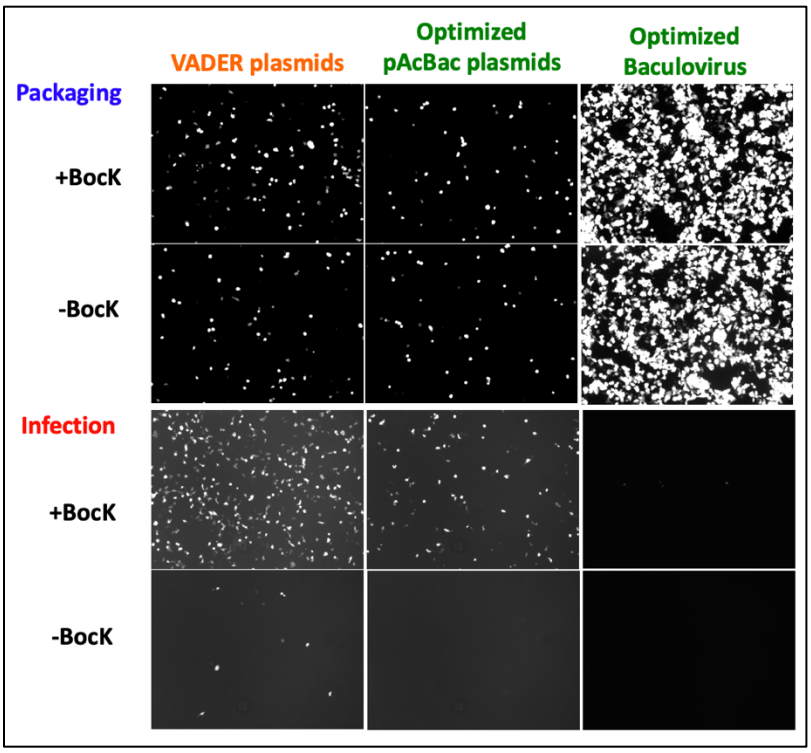


Figure 2.15. Cells packaging rAAV with either wild type or stop codon-containing capsids and cells infected with rAAV-containing lysate. Cells were infected with rAAV and either transfected with plasmids or infected with baculovirus. Top set of images: wild type capsids. Bottom set of images: T454TAG capsids.

An anti-AAV western blot on lysate from cells packaging new AAV after being infected with rAAV-mCherry-PytR and either infected with baculovirus or transfected with either set of plasmids shows that there is still insufficient capsid protein produced when the capsid proteins contain a TAG codon at position 454, even when those genes are being expressed from the highly active CMV promoter (Figure 2.16). It will be necessary to solve the problem of continued capsid protein limitation to move forward with baculovirus-assisted selections like this in the future.

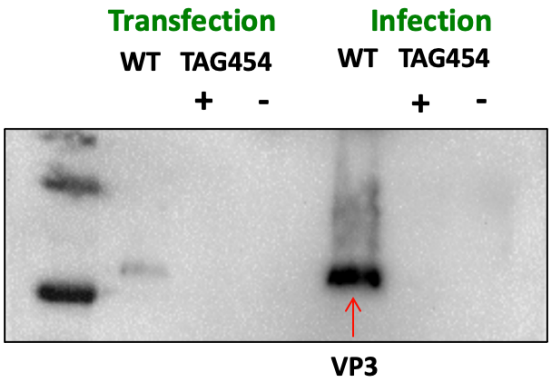


Figure 2.16. Anti-AAV western to determine whether 454TAG capsid proteins are limiting. An anti-AAV western blot was performed on cells infected with rAAV-mCherry and either transfected with the optimized pAcBac shuttle plasmids or infected with the optimized baculovirus.

2.2.4 The source of the improved AAV production

I developed these baculovirus delivery vectors to solve the problem of limited AAV production during the positive selection step in VADER, a problem which has two sources. First, PEI limits AAV entry into the cell. Second, transient transfection delivers DNA to cells in a highly heterogeneous fashion, so that most cells will not receive all the DNA they need to produce more rAAV. Providing the other necessary genes (*AdHelper*, *Rep*, and *Cap*) via baculovirus infection instead of by transient transfection should allow researchers to provide these genes in a much more homogenous manner; it should furthermore also allow improved AAV entry into the packaging cells. More host cells should therefore be involved in packaging AAV.

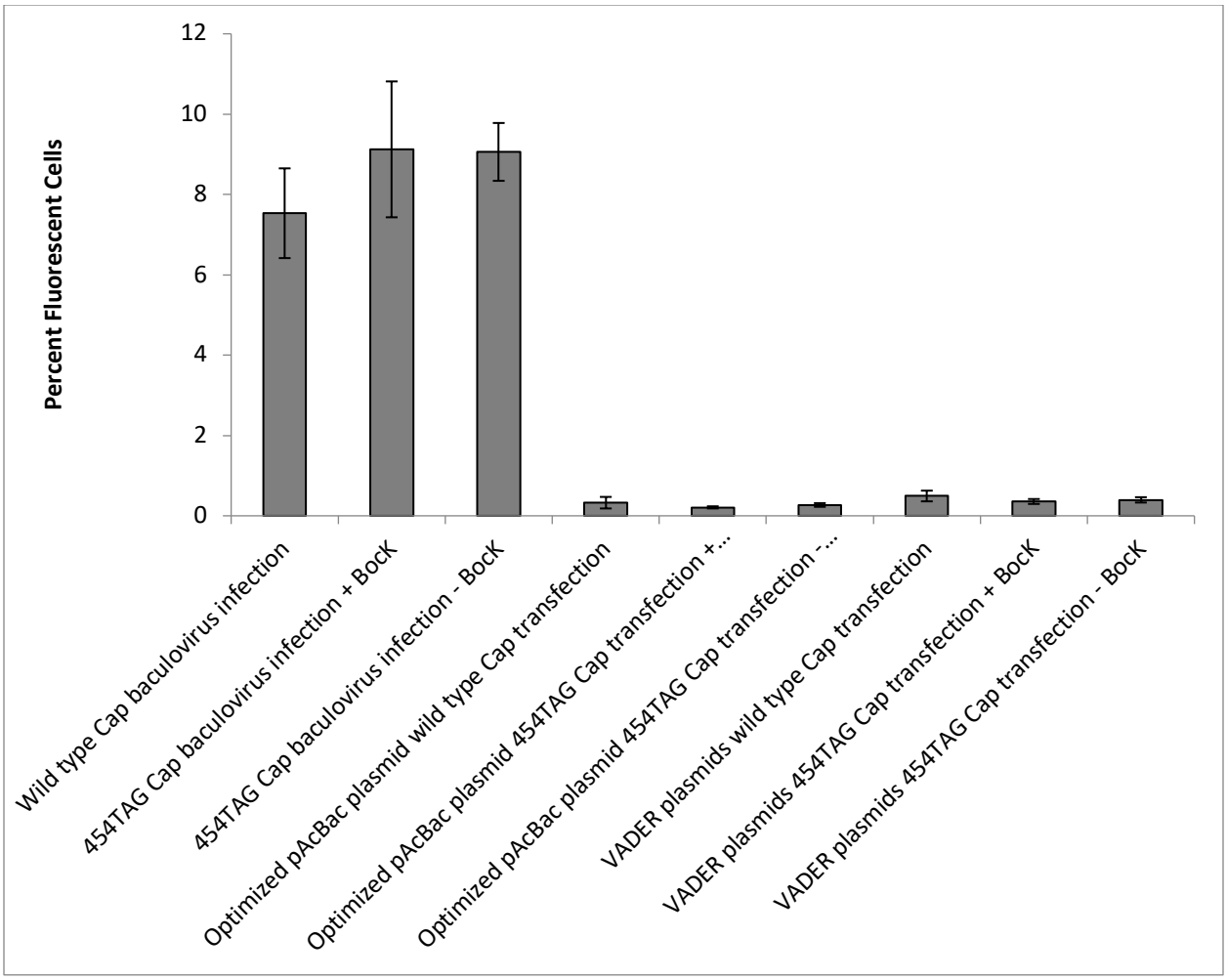


Figure 2.17. Flow cytometry analysis of cells producing rAAV-mCherry. Cells were infected with rAAV-mCherry and either infected with baculovirus or transfected with plasmids to deliver the other genes to make AAV. These cells were harvested and analyzed by FACS to determine the percent fluorescent cells.

To determine whether it is indeed the case that more host cells receive all of the necessary genes to package rAAV, I infected HEK-293T cells with rAAV-mCherry-PytR at 0.01 MOI and provided *AdHelper*, *Rep*, and *Cap* in one of the following manners: infection of P1 baculovirus (pAcBac-Helper and either pAcBac-Rep-CMV-CapWT-CSS-MbPylRS or pAcBac-Rep-CMV-Cap454TAG-CSS-MbPylRS), transfection with baculovirus shuttle plasmids (with either wild type *Cap* or 454TAG *Cap*), or transfection with the original VADER plasmids (pHelper and either pIDT-MbPylRS-RC2wt or pIDT-

MbPylRS-RC2-454TAG). I analyzed these packaging cells by FACS and found that there were at least 15x more fluorescent cells in treatments where the host cells had been infected by baculovirus, as opposed to having been transfected with plasmids and PEI (Figure 2.17). These cells are fluorescent because they received all the necessary genes for amplifying the rAAV genome: the mCherry cargo; *Rep*, which replicates the cargo between the ITRs, making it available for transcription to produce the fluorescent protein; and *AdHelper*, without which Rep proteins are not active. This result indicates that one of the sources of increased rAAV production is more cells receiving all of the necessary genes to amplify rAAV.

2.3 Conclusions

The results of these experiments demonstrate that infection with baculovirus is a much more efficient method to amplify rAAV during and between positive selections than is transient transfection. Not only does far less input AAV need to be added during the positive selection when co-infecting with baculovirus, but the other genes (*AdHelper*, *Rep*, and *Cap*) are provided much more homogenously, so that more cells can be involved in packaging rAAV. This improvement will allow researchers to reamplify rAAV between rounds of selection and to ensure that each potential hit can be accessed during directed evolution. This technique promises to allow access to complex selection systems involving large genetic cargos (such as the CRISPR-Cas system) due to the vast capacity of baculovirus.

2.3.1 Ongoing work and future directions

Because there is not sufficient UAA-containing capsid protein produced using the CMV-Cap baculovirus vectors, it is still necessary to optimize the expression of stop

codon-containing capsid proteins with this baculovirus vector. Future work will also involve using baculovirus to reamplify AAV between rounds of selection to proceed with directed evolution experiments more quickly.

2.4 Acknowledgements

Dr. Rachel Kelemen trained me to make rAAV and gave me access to her plasmid and primer stocks. Dr. Yunan Zheng trained me to make recombinant baculovirus.

2.5 Experimental Procedures

2.5.1. Strains and cell lines

Clonings and plasmid propagations were performed in electrocompetent DH10B *E. coli* cells. Bacmid transposition was performed in chemically competent DH10Bac cells (Invitrogen).

Sf9 cells (Thermo Fisher) were maintained at 27°C in Sf-900 III serum free medium (Gibco), in suspension. HEK-293T (ATCC) cells were maintained at 37°C and 5% CO_2 in DMEM-high glycose (HyClone) supplemented with 10% fetal bovine serum (Corning) and antibiotic-antimycotic (100 U/mL of penicillin, 100 μg/mL of streptomycin, and 0.25 μg/μL of Fungizone).

2.5.2. Plasmids

All plasmid maps and sequences can be found in the Appendix.

The plasmids pAAV-GFP+1xEcYtR, pAAV-mCherry+1xMmPytR, pHelper, pIDT-MbPylRS-RC2wt, pIDT-MbPylRS-RC2-454TAG, pAcBac1-Helper, pAcBac1-MbPylRS-RC2wt, and pAcBac1-MbPylRS-RC2-454TAG were propagated from Dr. Rachel Kelemen's stocks.

To construct the pAcBac1-Rep-CMV-CapWT-CSS-CMV-MbPylRS and pAcBac1-Rep-CMV-Cap454TAG-CSS-CMV-MbPylRS constructs, the consensus splice donor site (CSS) was added to the 5' end of the *Cap* genes from, downstream of the transcription start site, with the primers Consensus-splice-donor-site-sequence-F. This cassette was amplified with the primers NheI-p40-F and Cap-SalI-R, and this construct was cloned into the pIDT backbone (pIDTSMART-MbPylRS) using the NheI and SalI cut sites to produce the plasmid pIDTSMART-CMV-CAP-CSS. The cassette, along with the CMV promoter and Bgh polyA sequence, was cut with the SpeI and AvrII cut sites, and cloned into pAcBac-MbPylRS-RC2, which had been cut with AvrII. *Rep* was then amplified with the primers SbfI-Rep-F and NotI-Rep-R and cloned into both the wild type and the 454TAG constructs using the SbfI and NotI cut sites.

2.5.3. Unnatural amino acids

The N^{ε}-Boc-L-Lysine used in these experiments was purchased from Chem-Impex International (catalog number 00363).

2.5.4. Producing recombinant baculovirus

The pAcBac1 shuttle plasmids were chemically transformed into DH10Bac cells (Invitrogen) and grown for 48 hours on LB agar containing 50μg/mL kanamycin, 7μg/mL gentamicin, 10μg/mL tetracycline, 100μg/mL X-gal, and 40μg/mL IPTG. One white colony per construct was chosen and streaked out on fresh medium. After 48 hours of additional growth, one white colony per construct was cultured overnight in LB with kanamycin, gentamicin, and tetracycline. The bacmids were miniprepped according to Invitrogen's Bac-to-Bac Expression protocol.

To prepare P0 recombinant baculovirus, 4μg of miniprepped bacmid was transfected into adherent Sf9 cells using X-tremeGENE transfection reagent (Sigma-Aldrich). After 96 hours, the media containing P0 virus was harvested and stored at 4ºC away from light. P1 recombinant baculovirus was prepared by infecting shaking flasks of Sf9 cells with P0 virus. After 72 hours, the cultures were spun to pellet out the Sf9 cells, and the supernatant was spun for 60 minutes at 39,800g to pellet the P1 virus. The virus was resuspended in 10% DMSO in PBS and flash-frozen for storage at -80°C. It was titered using the Clontech BacPAK™ Baculovirus Rapid Titer Kit.

2.5.5. Preparing recombinant AAV

To produce rAAV with wild-type capsids, the following plasmids were transfected into HEK-293T cells using PEI (Sigma): pHelper, pIDT-MbPylRS-RC2wt, and either pAAV-Cargo-GFP+1xEcYtR or pAAV-Cargo-mCherry+1xMmPytR. After 72 hours, cells were harvested by scraping and pelleted, saving the culture media. The cell pellets were resuspended in PBS with 300mM NaCl, freeze-thawed twice in a dry ice/ethanol bath and a 37°C water bath to lyse virus-containing cells, and then spun to pellet the cell fragments. The supernatant was then added back to the culture media, and this lysate was incubated at room temperature for 30 minutes with Pierce Universal Nuclease to degrade any unpackaged nucleic acids. After incubation with nuclease, PEG was added to the lysate to 11.5%, and the lysate was precipitated overnight at 4°C. The precipitated lysate was spun to pellet the virus, resuspended in PBS, and flash frozen for storage at -80°C. Infective titer for rAAV containing the fluorescent reporters GFP or mCherry was determined by FACS after infecting HEK-293T cells with rAAV, using sodium butyrate.

To harvest AAV produced on a small scale (in a 12-well plate), cells were lysed with CelLytic™ M (Sigma-Aldrich) with 0.00001% Pierce Universal Nuclease. The resultant lysate was flash-frozen and stored at -80°C.

2.5.6. Determining the effects of addition of PEI on AAV infection

HEK-293T cells were infected with either 1 or 5 MOI of rAAV-GFP. Cells also received PEI at the same time as infection with AAV, 4 hours after infection, or not at all. After 48 hours, cells were imaged and lysed, and GFP fluorescence was measured.

2.5.7. Comparing pAcBac shuttle plasmids to original rAAV production plasmids

HEK-293T cells were infected with 5 MOI rAAV-GFP. After 4 hours, cells were transfected with one of the following pairs of plasmids: pHelper and pIDT-MbPylRS-RC2; pAcBac1-Helper and pAcBac1-MbPylRS-RC2wt; or pAcBac1-Helper and pAcBac1-Rep-CMV-CapWT-CSS-CMV-MbPylRS. The infected and transfected cells were imaged and lysed 72 hours later. Western blots were performed on the lysate with the B1 anti-VP antibody which binds to the C terminus of the capsid proteins. 10% of the lysate was used to infect new HEK-293T cells. 48 hours later, infected cells were imaged and lysed, and GFP fluorescence was measured.

2.5.8. Comparing transfection to infection with baculovirus

HEK-293T cells were infected with 5 MOI rAAV-GFP. Cells that were infected with baculovirus were infected simultaneously with 100 MOI P1 virus prepared from the shuttle plasmids pAcBac1-Helper and either pAcBac1-Rep-CMV-CapWT-CSS-CMV-MbPylRS or pAcBac1-Rep-CMV-Cap454TAG-CSS-CMV-MbPylRS. Cells that were transfected were transfected 4 hours later with one of the following pairs of plasmids: pHelper and either pIDT-MbPylRS-RC2 or pIDT-MbPylRS-RC2-454TAG; pAcBac1-

Helper and either pAcBac1-Rep-CMV-CapWT-CSS-CMV-MbPylRS or pAcBac1-Rep-CMV-Cap454TAG-CSS-CMV-MbPylRS. To test the production of rAAV containing a UAA in the capsid, N_{α}-Boc-L-Lysine (BocK) was added to the cell media to a final concentration of 1.0 mM. 72 hours later, cells were imaged and lysed. 10% of that lysate was used to infect new HEK-293T cells, and an anti-VP western blot was also performed on the lysate. After 48 hours, the infected cells were imaged and lysed, and GFP fluorescence was measured.

2.6 References

(1) Addy, P. S.; Erickson, S. B.; Italia, J. S.; Chatterjee, A. *J. Am. Chem. Soc.* **2017**, *139* (34), 11670.
(2) Usanov, D. L.; Chan, A. I.; Maianti, J. P.; Liu, D. R. *Nat. Chem.* **2018**, *10* (7), 704-714.
(3) Chen, K.; Huang, X.; Kan, S. B. J.; Zhang, R. K.; Arnold, F. H. *Science* **2018**, *360* (6384), 71.
(4) Umeno, D.; Arnold, F. H. *J. Bacteriol.* **2004**, *186* (5), 1531.
(5) Liu, C. C.; Schultz, P. G. *Annu. Rev. Biochem.* **2010**, *79* (1), 413.
(6) Liu, W.; Brock, A.; Chen, S.; Chen, S.; Schultz, P. G. *Nat. Methods* **2007**, *4* (3), 239.
(7) Chen, P. R.; Groff, D.; Guo, J.; Ou, W.; Cellitti, S.; Geierstanger, B. H.; Schultz, P. G. *Angew. Chemie Int. Ed.* **2009**, *48* (22), 4052.
(8) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G. *Science (80-.).* **2003**, *301* (5635), 964.
(9) Esvelt, K. M.; Carlson, J. C.; Liu, D. R. *Nature* **2011**, *472* (7344), 499.
(10) Bessette, P. H.; Rice, J. J.; Daugherty, P. S. *Protein Eng. Des. Sel.* **2004**, *17* (10), 731.
(11) Boder, E. T.; Wittrup, K. D. *Nat. Biotechnol.* **1997**, *15* (6), 553.
(12) Clackson, T.; Hoogenboom, H. R.; Griffiths, A. D.; Winter, G. *Nature* **1991**, *352* (6336), 624.
(13) Wang, L.; Tsien, R. Y. *Nat. Protoc.* **2006**, *1* (3), 1346.
(14) English, J. G.; Olsen, R. H. J.; Lansu, K.; Strachan, R. T.; Wacker, D.; Correspondence, B. L. R. *Cell* **2019**, *178*, 748.
(15) Berman, C. M.; Papa, L. J.; Hendel, S. J.; Moore, C. L.; Suen, P. H.; Weickhardt, A. F.; Doan, N. D.; Kumar, C. M.; Uil, T. G.; Butty, V. L.; Hoeben, R. C.; Shoulders, M. D. *J. Am. Chem. Soc.* **2018**, *140* (51), 18093.
(16) Das, A. T.; Zhou, X.; Vink, M.; Klaver, B.; Verhoef, K.; Marzio, G.; Berkhout, B. *J. Biol. Chem.* **2004**, *279* (18), 18776.
(17) Srivastava, A.; Lusby, E. W.; Berns, K. I. *J. Virol.* **1983**, *45* (2).
(18) Pereira, D. J.; McCarty, D. M.; Muzyczka, N. *J. Virol.* **1997**, *71* (2), 1079.
(19) Ni, T.-H.; McDonald, W. F.; Zolotukhin, I.; Melendy, T.; Waga, S.; Stillman, B.; Muzyczka, N. *J. Virol.* **1998**, *72* (4), 2777.
(20) Im, D. S.; Muzyczka, N. *Cell* **1990**, *61* (3), 447.
(21) Im, D. S.; Muzyczka, N. *J. Virol.* **1992**, *66* (2).
(22) Tratschin, J. D.; West, M. H.; Sandbank, T.; Carter, B. J. *Mol. Cell. Biol.* **1984**, *4* (10), 2072.
(23) Samulski, R. J.; Muzyczka, N. *Annu. Rev. Virol.* **2014**, *1* (1), 427.
(24) Sonntag, F.; Schmidt, K.; Kleinschmidt, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (22), 10220.
(25) McLaughlin, S. K.; Collis, P.; Hermonat, P. L.; Muzyczka, N. *J. Virol.* **1988**, *62* (6).
(26) Hauswirth, W. W.; Berns, K. I. *Virology* **1977**, *78* (2), 488.
(27) Atchison, R. W.; Casto, B. C.; Hammon, W. M. D. *Science (80-.).* **1965**, *149* (3685), 754.

(28) Buller, R. M. L.; Janik, J. E.; Sebring, E. D.; Rose, J. A. *J. Virol.* **1981**, *40* (1), 241.
(29) Urabe, M.; Nakakura, T.; Xin, K.-Q.; Obara, Y.; Mizukami, H.; Kume, A.; Kotin, R. M.; Ozawa, K. *J. Virol.* **2006**, *80* (4), 1874.
(30) Ward, P.; Dean, F. B.; O'Donnell, M. E.; Berns, K. I. *J. Virol.* **1998**, *72* (1), 420.
(31) Matsushita, T.; Elliger, S.; Elliger, C.; Podsakoff, G.; Villarreal, L.; Kurtzman, G. J.; Iwaki, Y.; Colosi, P. *Gene Ther.* **1998**, *5* (7), 938.
(32) Wang, D.; Tai, P. W. L.; Gao, G. *Nature Reviews Drug Discovery*. Nature Publishing Group May 1, 2019, pp 358–378.
(33) Kelemen, Rachel E.; Jewel, Delilah; Pasha, Muhammad; Cao, Xiaofu; Chatterjee, A. *Submiss.*
(34) Xiao, X.; Li, J.; Samulski, R. J. *J. Virol.* **1998**, *72* (3), 2224.
(35) Cohen, R. N.; van der Aa, M. A. E. M.; Macaraeg, N.; Lee, A. P.; Szoka, F. C. *J. Control. Release* **2009**, *135* (2), 166.
(36) Carpentier, E.; Paris, S.; Kamen, A. A.; Durocher, Y. *J. Biotechnol.* **2007**, *128* (2), 268.
(37) Glover, D. J.; Leyton, D. L.; Moseley, G. W.; Jans, D. A. *J. Gene Med.* **2010**, *12* (1), 77.
(38) Jarvis, D. L. In *Methods in Enzymology*; Academic Press Inc., 2009; Vol. 463, pp 191–222.
(39) Kost, T. A.; Condreay, J. P. *Trends in Biotechnology*. Elsevier Current Trends April 1, 2002, pp 173–180.
(40) Ayres, M. D.; Howard, S. C.; Lopez-Ferber, M.; Lopez-Ferber, M.; Possee, R. D. *Virology* **1994**, *202* (2), 586.
(41) Cheshenko, N.; Krougliak, N.; Eisensmith, R. C.; Krougliak, V. A. *Gene Ther.* **2001**, *8* (11), 846.
(42) Kost, T. A.; Condreay, J. P.; Jarvis, D. L. *Nature Biotechnology*. Nature Publishing Group May 5, 2005, pp 567–575.
(43) Barsoum, J.; Brown, R.; McKee, M.; Boyce, F. M. *Hum. Gene Ther.* **1997**, *8* (17), 2011.
(44) Kaikkonen, M. U.; Räty, J. K.; Airenne, K. J.; Wirth, T.; Heikura, T.; Ylä-Herttuala, S. *Gene Ther.* **2006**, *13* (4), 304.
(45) Chatterjee, A.; Xiao, H.; Bollong, M.; Ai, H. W.; Schultz, P. G. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (29), 11803.
(46) Kotin, R. M. *Hum. Mol. Genet.* **2011**, *20* (R1), R2.
(47) Urabe, M.; Ding, C.; Kotin, R. M. *Hum. Gene Ther.* **2002**, *13* (16), 1935.
(48) Smith, R. H.; Levy, J. R.; Kotin, R. M. *Mol. Ther.* **2009**, *17* (11), 1888.

Chapter 3

Development of an alternate selection handle to evolve tRNAs in mammalian cells

3.1 Introduction

The concepts of directed evolution, especially the difficulties of performing selections directly in mammalian cells, have been outlined in detail in Chapter 1. Chapter 2 outlined the use of viruses as DNA delivery vectors to perform selections in mammalian cells. This chapter focuses on the use of an alternative selection handle in a viral protein to allow the evolution of tRNAs to incorporate UAAs that are not tolerated in the AAV capsid.

3.1.1 VADER as a solution for evolving small gene products

The development of directed evolution schemes in mammalian cells has historically been limited by the constraints of delivering DNA to mammalian cells. Several selection schemes have been developed recently for mammalian cells that use viruses to deliver library members to the host cells, but these techniques have relied on error-prone replication to generate variation in library members.[1–3] Somatic hypermuation is a technique that also solves the problem of delivering DNA to cells, as it does not require continued supplemental DNA delivery to cells for the selection process to occur.[4] While these selection schemes work well for large gene products such as GPCRs and transcription factors, they are unsuitable for the evolution of small gene products because the random mutation rate necessary to drive variation in small genes such as tRNAs would be too high to permit the continued function of essential genes, as deleterious missense and nonsense mutations would accumulate in the viral or cellular genome. A new selection scheme is needed to evolve small gene products such as tRNAs.

VADER is a technique which solves the problem of delivering library members to mammalian host cells while not relying on random mutagenesis to generate variation in the library members.[5] Since the selection scheme does not make use of random

mutagenesis, it can be used to perform directed evolution on tRNAs, which could not be accessed via somatic hypermutation or existing available virus-based selection systems. VADER is a powerful technique that has yielded multiple improved tRNAs that install lysine variants in response to TAG codons. These tRNAs improve the efficiency of UAA incorporation in mammalian cells through their enhanced activities and are excellent tools for the scientific community working on stop codon suppression.

3.1.2 The limitations of VADER: the AAV capsid

As it currently stands, the VADER scheme works well for evolving tRNAs to install UAAs that fit into the AAV capsid. However, the AAV capsid does not tolerate the incorporation of all UAAs. The AAV capsid is composed of 60 individual proteins that fit tightly and rigidly together. It is easy to perturb the complex manner in which the capsid proteins fit together, which results in unassembled capsids and no functional virions.[6,7] It is also easy to modify the capsid protein structure in such a way that they can no longer perform the functions necessary for viral infectivity, such as by disrupting VP1's phospholipase activity that is necessary for endosomal escape or encapsulating the single-stranded DNA genome.[8–11] If a UAA does not fit into the capsid correctly, it could destroy the capsid's ability to assemble or the virion's ability to infect a cell. It is therefore not possible to use these UAAs as reactive handles in VADER for the bio-orthogonal pulldown, so researchers cannot perform directed evolution on the tRNAs to install these UAAs into proteins in mammalian cells.

5-hydroxy-L-tryptophan (5HTP, Figure 3.1) is one such UAA that is not tolerated in the AAV capsid (Figure 3.2), even at position 454, which is a highly permissive site for BocK and AzK incorporation.[12] It is likely that the shape of this UAA perturbs the assembly

of the capsid, even though residues in this position should be surface-exposed and not involved in the inter-protein junctions. Installing 5HTP at position 454 does not result in infective virions, so it is not possible to use this UAA in VADER. This situation is unfortunate, because 5HTP is a useful bioconjugation handle.[13–15] The chemical biology community would benefit from the ability to evolve a tryptophanyl tRNA (WtR) to improve the incorporation of 5HTP in proteins in mammalian cells for the purposes of site-specific labeling. It is necessary to modify the VADER technique to access the WtR for the purposes of directed evolution.

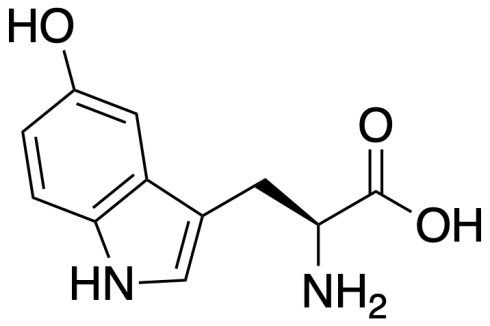


Figure 3.1. Structure of 5-hydroxytryptophan (5-HTP). 5HTP is one of the UAAs that is not tolerated in the AAV capsid.

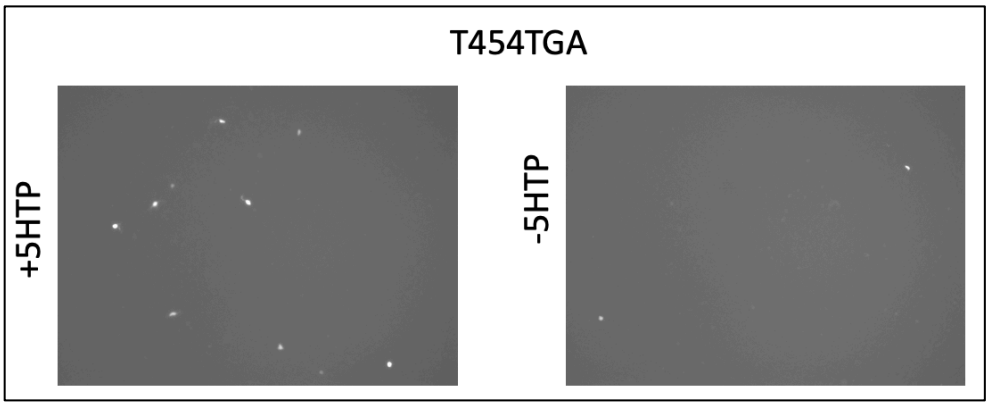


Figure 3.2. Incorporation of 5HTP into the AAV2 capsid. AAV2 does not tolerate 5HTP in its capsid, as evidenced by the lack of production of infectious virus even when provided with the UAA.

Delicate capsid assembly is not the only limitation to successful virion assembly during the VADER process. Capsid proteins are required at an extremely high level in order to produce enough virus to capture during selections. As I showed in Chapter 2 of this dissertation, production of capsid proteins containing UAAs is a limiting factor in being able to produce infectious virus in the VADER process. Being able to install a UAA into a protein selection handle which is required at much lower levels than are the capsid proteins would free researchers from the constraint of needing to produce such high levels of protein using tRNAs which may not yet be as active as native tRNAs. The ideal selection handle for UAAs such as 5HTP would be flexible, so in order to eliminate worries about structural perturbation, and it would be a catalytic protein, so that it would be required at much lower levels.

3.1.3 The ideal selection handle: Assembly Activating Protein

One protein in the AAV proteome that satisfies the requirements of being a structurally flexible catalyst is Assembly Activating Protein (AAP). The gene that codes for this protein is located within the *Cap* ORF, but within an alternate reading frame with a CTG start codon (Figure 3.3).[16] During AAV2 replication, AAP is localized in the host nucleolus and targets newly-synthesized Cap proteins to that location, where capsid assembly takes place.[16] AAP's highly structured N- and C-termini are relatively conserved between different AAV serotypes.[17,18] The termini of AAP likely interact with the capsid proteins and act as scaffolds to promote capsid protein interaction along the axis of 2-fold symmetry.[18]

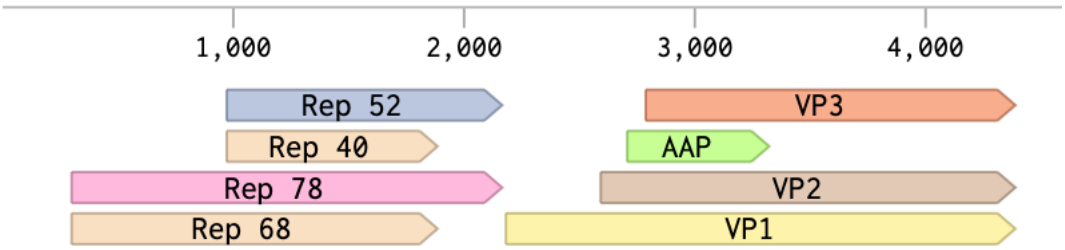


Figure 3.3. Structure of the AAV genome. The *Cap* gene in the AAV genome contains the genes for the three capsid proteins VP1, VP2, and VP3, as well as the gene for AAP in a different reading frame.

While AAP is required for AAV2 capsid assembly, it is likely needed at much lower levels than are the capsid proteins, as AAP acts catalytically to assemble proteins into the viral capsid.[16,19] AAP's CTG start codon is relatively weak, which is another clue indicating that this protein is needed at lower levels. AAP interacts closely with the capsid proteins during the assembly process, but its termini are not as structurally rigid as the AAV capsid itself, and there are several threonine/serine rich regions in the protein that are quite flexible.[18] This protein is therefore a strong candidate for the structurally flexible catalyst that would tolerate the incorporation of 5HTP.

3.1.4 A selection platform based on AAP

As described above, the gene that codes for AAP sits within the ORF for the capsid proteins, but in an alternative reading frame. This genetic organization permits a way to silence the expression of AAP without silencing the expression of the capsid proteins: the AAPstop60 platform.[20,21] A mutation in reside 64 of AAP from TCA to TGA results in an S64TGA mutation in AAP (Figure 3.4). As AAP's gene is in a different reading frame than the genes for the capsid proteins, there is a silent GTC to GTG mutation at residue 239 of the VP proteins. Use of the AAPstop60 system has allowed researchers to silence AAP

without perturbing the expression of the capsid protein and to demonstrate that AAP is required to for the production of infectious AAV2 capsids.[20]

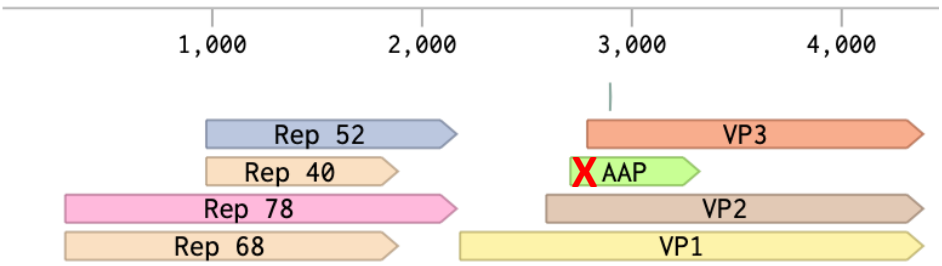


Figure 3.4. The structure of the AAPstop60 genome. A point mutation produces a S64TGA mutation in AAP while resulting in a silent mutation in the VP genes because these genes are in different reading frames within the same ORF.

This system will be useful for developing an AAP-based selection scheme. It is possible to silence the endogenous AAP in the AAV2 genome with this S64TGA mutation and then provide supplemental AAP *in trans* with a blank codon mutation of choice, as well as a tRNA that can suppress that blank codon. In this manner, it is possible to tie the production of infectious AAV2 to the tRNA's ability to incorporate a UAA into AAP during translation.

In order to make use of AAP in selection schemes to evolve tRNAs, it is necessary to find an optimum location to position a stop codon in the gene for AAP, where the UAA will be introduced. The structures of the termini of AAP are highly conserved between AAV serotypes,[18] which indicates that these exact structures are essential for the protein's capsid assembly functions. Since these regions are essential, requiring the incorporation of a UAA in the N-terminus might prove to be a stringent selection, which could benefit the evolution of a highly active tRNA. If these regions are too structurally rigid to permit the incorporation of a UAA such as 5HTP, there are flexible regions downstream of the N-

terminus that are rich in threonine and serine residues, which tolerate large insertions which the AAV capsid does not.[17,18] These T/S-rich regions might be ideal candidate regions to incorporate 5HTP into AAP.

In this chapter, I present a modified selection scheme in which I use AAP as a selection handle to evolve mutant WtRs to incorporate 5HTP in proteins in mammalian cells. I silence the expression of AAP from the AAV2 genome using the AAPstop60 method, and I provide wild-type AAP back *in trans* to demonstrate that it is possible to rescue production of infectious AAV using this supplemental AAP. I optimize the location of a TAG stop codon in AAP and show that UAA-containing AAP can still facilitate the production of infectious AAV2. By changing this TAG stop codon to TGA, I am able to incorporate 5HTP into AAP using the *Escherichia coli* tryptophanyl tRNA (EcWtR) with a UCA anticodon. Finally, I am able to use this platform to subject a library of EcWtR variants to positive selection and identify hits that incorporate 5HTP into proteins in mammalian cells at a highly improved level.

3.2 Results and Discussion

3.2.1 Producing infectious AAV with TAG-containing AAP

To determine whether it is possible to produce infectious AAV when there is a stop codon in AAP, I transfected cells with plasmids containing an AAV2 genome with an S64TGA mutation in AAP that is silent in the *Cap* genes and provided supplemental either wild type or L12TAG AAP *in trans* under the control of the strong CMV promoter. The AAV produced in this way has wild type capsids, but its production depends on UAA incorporation (in this case, BocK) into AAP. I compared these results to those achieved

with the VADER plasmid system, which relies on a T454TAG mutation in the capsid proteins (Figure 3.5).

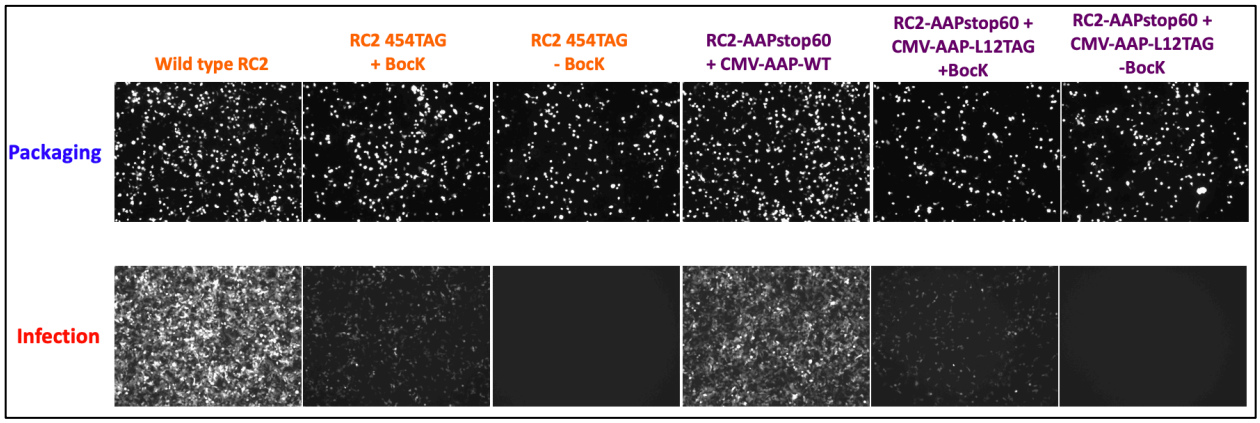


Figure 3.5. Producing AAV dependent on supplemental AAP. AAV was produced using the AAPstop60 system, providing supplemental AAP *in trans*. Production of AAV in this way was comparted to production of AAV using the RC2-454TAG system from VADER. After 72 hours, the packaging cells were imaged and lysed, and new cells were infected with the lysate. These cells were imaged after 48 hours.

The plasmid system producing rAAV using *Cap* genes with the T454TAG mutation produced 7.44% of the infectious virus when supplemented with BocK as did the plasmid system using wild type *Cap* genes, as measured by fluorescence imaging and GFP fluorescence. When wild type AAP was provided *in trans* along with the AAPstop60 system, cells produced 57.5% of the infectious virus as made with wild type AAV2 genes (Figure 3.6). With the L12TAG in AAP, cells only produced 6.39% of the virus as they had with supplemental wild type AAP, which is not an improvement over the *Cap* T454TAG system. Even though AAP is likely used catalytically in capsid assembly, it appears that the low levels of UAA-containing AAP produced using this system were not sufficient to assembly infectious AAV capsids at levels approaching that when wild type AAP is provided.

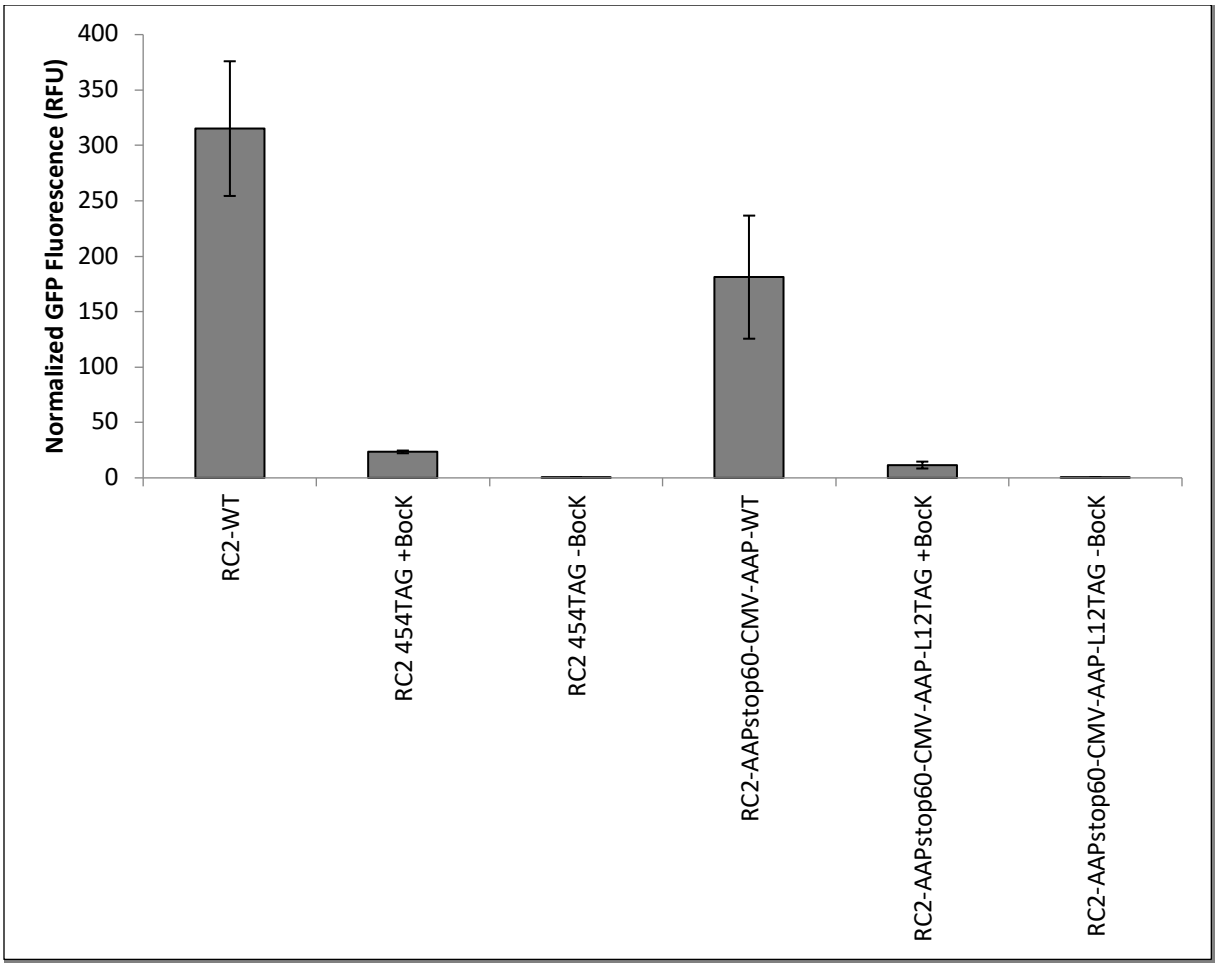


Figure 3.6. GFP fluorescence of cells infected with AAV produced using the AAPstop60 system. Cells were infected with rAAV produced by packaging cells to test the effectiveness of the AAPstop60 plus supplemental AAP to produce rAAV with wild type capsids. After 48 hours, these cells were imaged and lysed, and GFP fluorescence was taken.

3.2.2 Determining the optimum site for a TAG mutation in AAP

While it is possible to produce some infectious AAV by installing a UAA in the N-terminus of AAP, there may be other locations within the protein that are better suited to UAA incorporation. To that end, I created supplemental AAP variants with TAG mutations in the flexible T/S-rich regions in the protein, as these regions are highly tolerant of mutations and insertions,[17] while the termini of the protein are much more conserved

between serotypes and are not mutation-tolerant.[17,18] The mutations tested with supplemental AAP provided *in trans* were: T78TAG, T97TAG, T110TAG, T124TAG, and T177TAG (Figure 3.7).

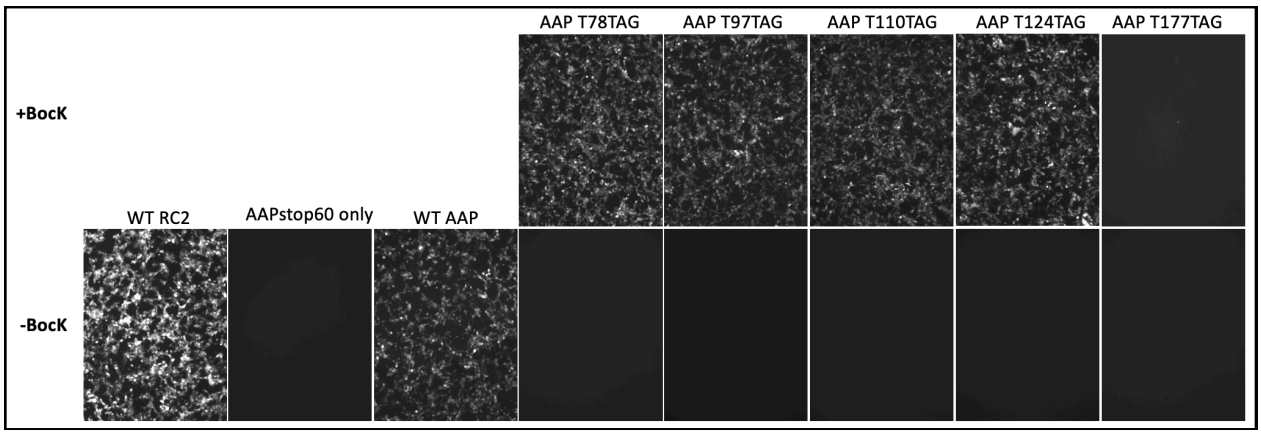


Figure 3.7. Producing AAV using AAP with TAG mutations. AAV was produced using the AAPstop60 system, providing supplemental AAP with TAG stop codons in different locations in the T/S rich regions. These cells were infected with lysate from cells that packaged rAAV and were imaged after 48 hours post-infection.

For the T78TAG, T97TAG, T110TAG, and T124TAG mutations, the amount of infectious AAV produced with the addition of BocK was approximately equal to that produced when wild type AAP was provided, as evidenced by fluorescence images and GFP fluorescence readings (Figure 3.8). AAP with the T177TAG mutation produced negligible amounts of infectious virus. These surprising and encouraging results indicate that it is possible to tie the production of infectious AAV dependent on stop codon suppression in AAP and still achieve reasonably high AAV titers.

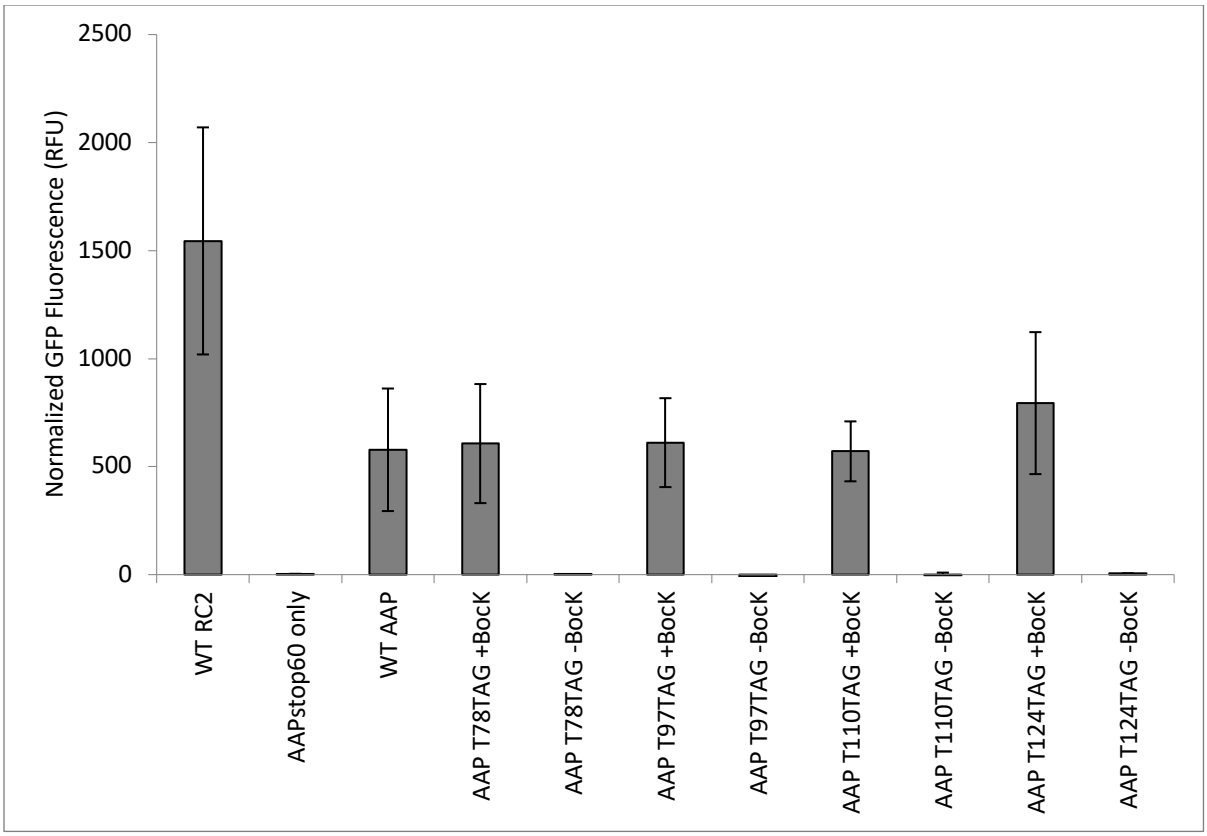


Figure 3.8. GFP fluorescence of cells infected with AAV produced with AAP TAG variants. Cells were infected with rAAV produced by packaging cells to determine the optimum site for a TAG mutation. After 48 hours, these cells were imaged, and GFP fluorescence was taken.

3.2.3 Incorporating 5-hydroxytryptophan into AAP

Having determined that it is possible to make the production of AAV dependent on stop codon suppression in AAP, I moved to test whether it was possible to make the production of AAV dependent on the incorporation of 5-hydroxytrypophan in AAP. The AAV capsid does not tolerate 5HTP incorporation, but because AAP tolerates mutations and insertions in its highly flexible T/S rich regions, it is likely that it would tolerate the mutation of one amino acid to 5HTP at one of these sites: T97TGA, T110TGA, and T124TGA (Figure 3.9).

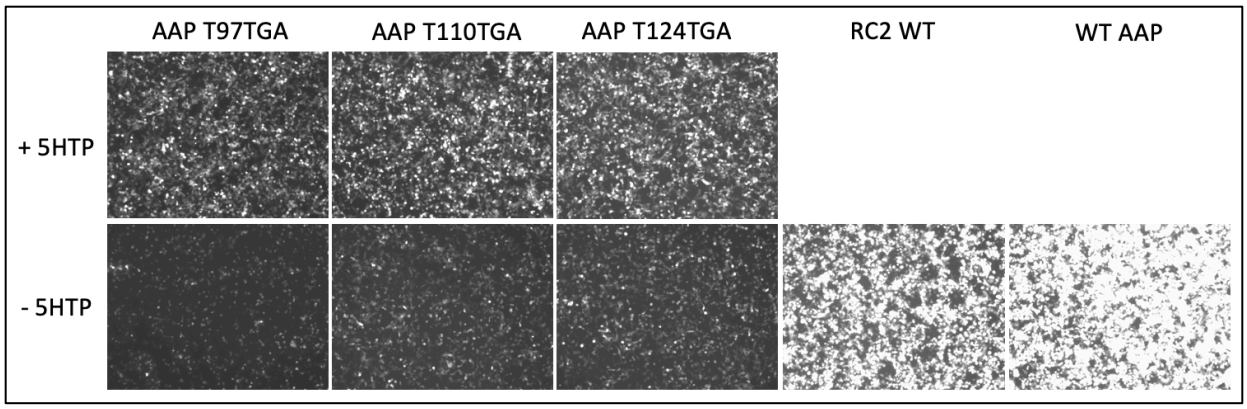


Figure 3.9. Infection of cells with virus produced using AAP with TGA mutations. AAV was produced using the AAPstop60 system, with supplemental AAP TGA mutants provided on the same plasmid for improved DNA delivery. The above cells were infected with lysate from cells that packaged rAAV and were imaged after 48 hours post-infection.

In this experiment, I provided supplemental AAP on the same plasmid as the mutant RC2-AAPstop60 AAV genome. Addition of wild type AAP rescued infective AAV production to wild type levels in this case, while it had not done so when provided on a different plasmid *in trans* in the previous experiments with AAPtag, likely due to the random delivery of DNA which occurs with transient transfection. The stop codon in the supplemental AAP was changed to TGA, as the EcWtR anticodon for stop codon suppression is UCA. Here, infective AAV production reached only approximately 10% of that achieved with wild type AAP, which is reasonable to expect because TGA suppression is not as efficient as TAG suppression. All three TGA mutants tested performed similarly in the presence of 5HTP, with little AAV production in the absence of 5HTP (Figure 3.10). Since all three mutants performed similarly well, I chose to move forward with the T97TGA mutant in selections for more efficient tryptophanyl tRNAs.

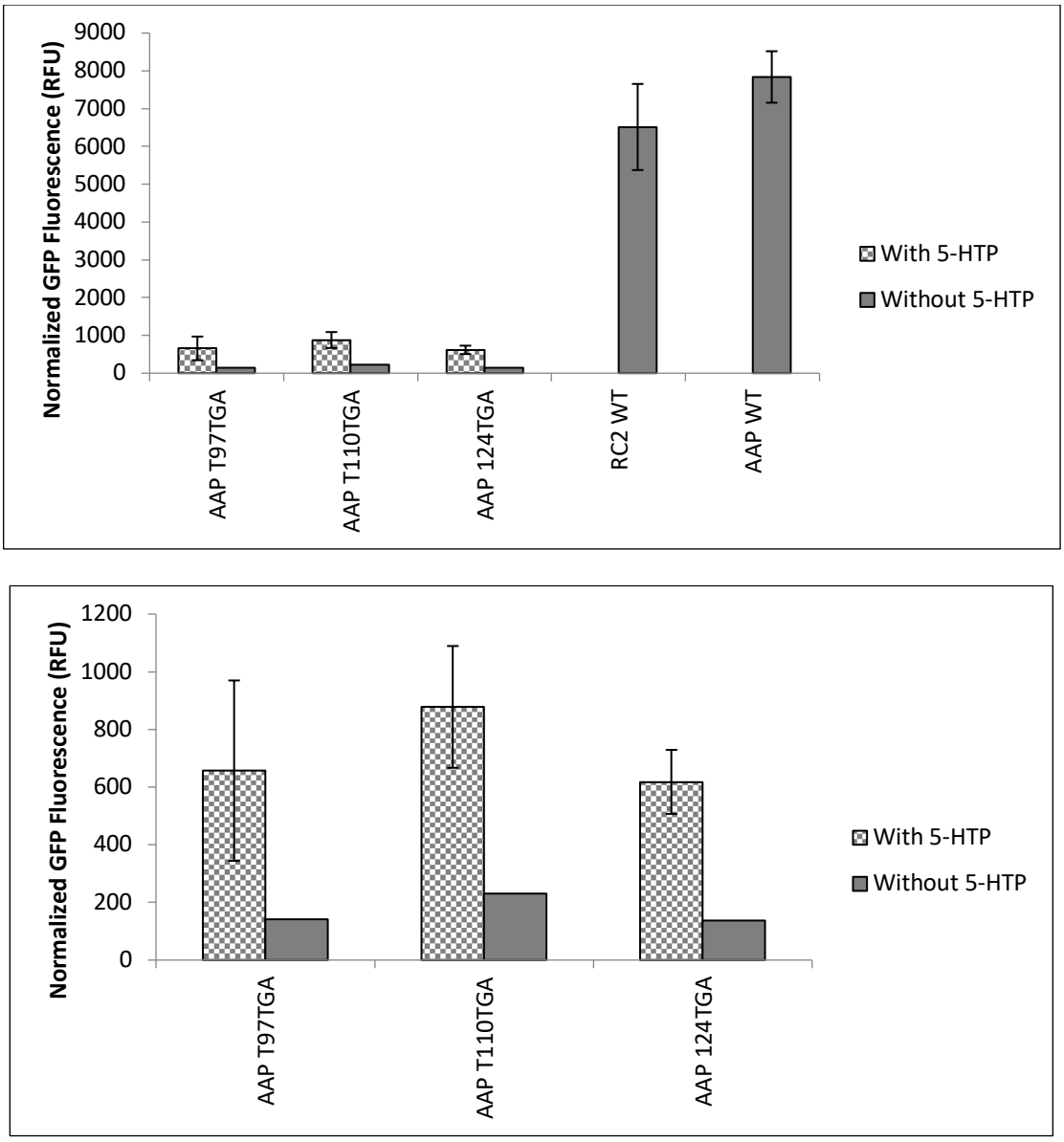


Figure 3.10. GFP fluorescence of cells infected with AAV produced with AAP TGA variants. Cells were infected with rAAV produced by packaging cells to determine whether AAP could tolerate insertion of 5HTP at various sites. After 48 hours, these cells were imaged, and GFP fluorescence was taken.

3.2.4 Mock selections to determine if VADER is possible with AAP as a selection handle

Before proceeding with selections on real libraries to evolve the *E. coli* tryptophanyl tRNA, it was necessary to test the selection parameters with a mock selection. To that end, I infected HEK-293T cells with rAAV containing as cargo GFP and the

MmPytR with a CUA anticodon at an MOI of 5 and rAAV containing as cargo mCherry and the EcWtR with a UCA anticodon at an MOI of 0.05 for a ratio of 100:1 PytR to WtR. I provided 5HTP in the cell media and the other genes necessary to make AAV via transient transfection, as established in the original VADER protocol, with the plasmids pHelper and pIDTSMART-RC2-AAPstop60-CMV-AAP97tga-CMV-WRS. The WRS should not charge the PytR, and the PytR should not incorporate any amino acid at the TGA stop codon in AAP. In the mock selection, the PytR mimics any inactive tRNA library member, while the WtR should be active.

The production of virus in each round of this selection scheme depends on the incorporation of 5HTP into AAP, but not into the capsid of the virus, as the structure of the viral capsid does not permit the disruption caused by the introduction of 5-HTP. Therefore, the viral capsids produced from this selection scheme are wild type capsids, and it is not possible to perform a pulldown to capture viruses that contain only active but not cross-reactive tRNAs. This version of the VADER selection scheme will rely only on positive selections to continuously enrich active tRNAs.

After harvesting the output virus from one round of the mock selection, I infected cells with the output virus as well as the original input virus and analyzed the ratio of green to red cells by flow cytometry. This ratio will indicate the ratio of PytR to WtR, as the genes for fluorescent reporter and tRNA are packaged together in the genome of the rAAV. The ratio of green to red cells from the input virus was 101:1, which demonstrates that the original calculations for infections were accurate. For the output virus, the ratio of green to red cells was 0.823:1, for a 121-fold enrichment of mCherry and WtR-containing rAAV (Figure 3.11). This result indicates that it is possible to achieve at least 100-fold enrichment

of active tRNAs after just one round of positive selection, which will allow researchers to evolve the previously un-evolvable WtR in mammalian cells.

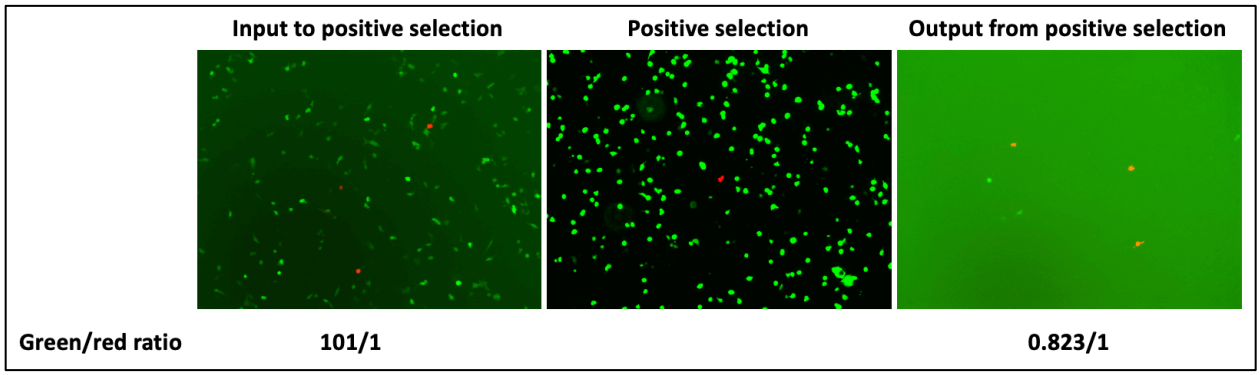


Figure 3.11. Enrichment of WtR+mCherry over PytR+GFP in a mock selection. An inactive tRNA (PytR) was provided to cells at an MOI 100 times that of the active tRNA (WtR). These tRNAs underwent a mock selection in HEK-293T cells. The resulting output virus was used to infect new cells. Cells infected with the input and output virus mixes were analyzed by FACS to determine the ratio of green (PytR) to red (WtR) viruses.

3.3 Conclusion

With this result in hand, it will be possible to evolve libraries of WtR variants to efficiently install 5HTP in proteins in mammalian cells. The selection scheme here will only involve positive selections, but with over 100-fold enrichments of active tRNAs over inactive ones, repeated rounds of positive selection should yield highly active WtR variants. This tool should enable researchers to achieve high yields of 5HTP-containing proteins, which aid the development of novel bioconjugation chemistries.

3.3.1 Ongoing work and future directions

A library of WtR variants with mutations in the acceptor stem has been developed and will be subject to selection via the method developed here. Because there is no negative selection step in this selection scheme, all selections will be repeated without adding 5HTP, so that if a tRNA variant is cross-reactive with a native aaRS and canonical amino acid, it has the opportunity to be charged with that amino acid as well. High-throughput

sequencing will make it possible to distinguish between variants that are active but not cross-reactive and variants that are active and cross-reactive.

AAP can also be used for a variety of future applications for selections in mammalian cells. For example, split halves of AAP can be fused to bait and potential prey proteins in a two-hybrid system. Since AAP can tolerate the insertions of large proteins in its Thr/Ser rich regions,[17] it is likely that it will still perform its functions with the addition of the associated bait and prey proteins in this region. This two-hybrid selection can use the production of AAV as the reporter to screen the mammalian proteome for binding partners.

3.4 Acknowledgments

Dr. Rachel Kelemen and Dr. James Italia gave me access to their plasmid stocks. Dr. Kelemen also trained me to perform selections using rAAV. Rachel Huang provided invaluable assistance in constructing the WtR library that will be used for future selections.

3.5 Experimental Procedures

3.5.1. Strains and cell lines

Clonings and plasmid propagations were performed in electrocompetent DH10B *E. coli* cells.

HEK-293T (ATCC) cells were maintained as described in Chapter 2.

3.5.2. Plasmids

All plasmid maps and sequences can be found in the Appendix.

The plasmids pAAV-GFP+1xMmPytR, pAAV-GFP+1xEcWtRtga, pAAV-mCherry+1xEcWtRtga, pIDTSMART-8xPytR-ITR-GFP, pHelper, pIDTSMART-

MbPylRS-RC2wt, and pIDTSMART-MbPylRS-RC2-454TAG were propagated from Dr. Rachel Kelemen's stocks.

To construct pIDTSMART-MbPylRS-RC2-AAPstop60, I introduced the mutation into the *Cap* gene using the primers AAPstop60-F and AAPstop60-R. I amplified this *Cap* containing the mutant gene for AAP using the terminal primers RC2int-HindIII-F and RC2-SbfI-R and cloned it back into the vector using the HindIII and SbfI cut sites.

I introduced a Kozak sequence to the 5' end of the gene for AAP and amplified the construct using the primers NheI-Kozak-AAP-WT-F and XhoI-AAP-R. To generate pIDTSMART-CMV-AAPwt, I cloned the resulting cassette into a pIDT vector using the NheI and XhoI restriction sites. I used mutagenesis primers to generate the following variants for the pIDTSMART-CMV-AAP plasmid: L12TAG, T78TAG, T97TAG, T110TAG, T124TAG, T177TAG, T97TGA, T110TGA, and T124TGA.

To generate pIDTSMART-MbPylRS-RC2-AAPstop60-CMV-AAP (both wild type and T97TGA variants), I cut the plasmids pIDTSMART-CMV-AAP (either wild type AAP or T97TGA) and pIDTSMART-MbPylRS-RC2-AAPstop60 with NheI and SphI and ligated the AAP construct into the AAPstop60 vector.

The *E. coli* tryptophanyl aaRS was amplified out of pBK-MCS-EcWRS from Dr. James Italia's stock using the primers AvrII-CMV-WRS-F and SpeI-WRS-R. It was then cloned into the vector pIDTSMART-MbPylRS-RC2-AAPstop60-CMV-AAP-97TGA which had been cut with AvrII.

3.5.3. Unnatural amino acids

BocK was purchased from Chem-Impex International (catalog number 00363). 5HTP was purchased from Chem-Impex International (catalog number 00607).

3.5.4 Preparing recombinant AAV

For details on producing recombinant AAV, see Chapter 2 of this thesis.

3.5.5 Producing infectious AAV with TAG-containing AAP

To determine whether it is possible to produce infectious AAV whose packaging depends on stop codon suppression in AAP, HEK-293T cells were transfected with the following sets of plasmids: pHelper and pIDTSMART-MbPylRS-RC2wt; pHelper, pIDTSMART-MbPylRS-RC2-454TAG, and pAAV-mCherry+1xMmPytR; pHelper, pIDTSMART-MbPylRS-RC2-AAPstop60, and pIDTSMART-AAPwt; or pHelper, pIDTSMART-MbPylRS-RC2-AAPstop60, pIDTSMART-AAP-L12TAG, and pAAV-mCherry+1xMmPytR. N^{ε}-Boc-L-Lysine (BocK) was added to the cell media to test the UAA-dependence of AAP production. Cells were lysed with CelLytic™ M (Sigma-Aldrich) with 0.00001% Pierce Universal Nuclease. The resultant lysate was flash-frozen and stored at -80°C. Lysate was thawed and used to infect new HEK-293T cells with sodium butyrate to 1.0 mM. After 48 hours, cells were imaged and lysed, and GFP fluorescence was measured.

3.5.6 Determining the optimum site for a TAG mutation in AAP

To find an improved location for the TAG mutation in AAP, cells were transfected with the following sets of plasmids: pHelper and pIDTSMART-MbPylRS-RC2wt; pHelper and pIDTSMART-MbPylRS-RC2-AAPstop60; pHelper, pIDTSMART-MbPylRS-RC2-AAPstop60, and pIDTSMART-AAPwt; or pHelper, pAAV-Cargo-GFP+1xMmPytR, pIDTSMART-MbPylRS-RC2-AAPstop60, and pIDTSMART-AAPtag with one of the following mutations: T78TAG, T97TAG, T110TAG, T124TAG, or T177TAG. N_{α}-Boc-L-Lysine (BocK) was added to the cell media to test the UAA-dependence of AAP

production. Cells were lysed with CelLytic™ M (Sigma-Aldrich) with 0.00001% Pierce Universal Nuclease. The resultant lysate was flash-frozen and stored at -80°C. Lysate was thawed and used to infect new HEK-293T cells with sodium butyrate to 1.0 mM. After 48 hours, cells were imaged and lysed, and GFP fluorescence was measured.

3.5.7 Incorporating 5-hydroxytryptophan into AAP

Cells were transfected with the following sets of plasmids: pHelper and pIDTSMART-MbPylRS-RC2wt; pHelper, pIDTSMART-MbPylRS-RC2-AAPstop60, and pIDTSMART-AAPwt; pHelper, pAAV-GFP+1xExWtRtga, and pIDTSMART-RC2-AAPstop60-CMV-AAP-CMV-WRS with one of the following mutations in AAP: T97TGA, T110TGA, or T124TGA. 5-hydroxy-L-tryptophan (5HTP) was added to the cell media to test the UAA-dependence of AAP production. Cells were lysed with CelLytic™ M (Sigma-Aldrich) with 0.00001% Pierce Universal Nuclease. The resultant lysate was flash-frozen and stored at -80°C. Lysate was thawed and used to infect new HEK-293T cells with sodium butyrate to 1.0 mM. After 48 hours, cells were imaged and lysed, and GFP fluorescence was measured.

3.5.8 Mock selections to determine if VADER is possible with AAP as a selection handle

HEK-293T cells were simultaneously infected with 5 MOI of rAAV-GFP-PytR and 0.05 MOI of rAAV-mCherry-WtR. These cells were transfected with pHelper and pIDTSMART -RC2-AAPstop60-CMV-AAP-CMV-WRS, and 5HTP was added to the cell media to 1.0 mM. Output virus was harvested, PEG precipitated, and flash frozen for storage at -80°C.

Fresh HEK-293T cells were infected with the input virus (calculated to deliver rAAV-GFP-PytR and rAAV-mCherry-WtR to the selection cells at a ratio of 100:1) or the

virus harvested after selection, with sodium butyrate to 1.0 mM. These infected cells were analyzed using flow cytometry to determine the ratio of green to red cells, which would correspond to the ratio of cells infected by rAAV-GFP-PytR to cells infected by rAAV-mCherry-WtR.

3.6 References

(1) English, J. G.; Olsen, R. H. J.; Lansu, K.; Strachan, R. T.; Wacker, D.; Roth, B. L. *Cell* **2019**, *178*, 748.

(2) Berman, C. M.; Papa, L. J.; Hendel, S. J.; Moore, C. L.; Suen, P. H.; Weickhardt, A. F.; Doan, N. D.; Kumar, C. M.; Uil, T. G.; Butty, V. L.; Hoeben, R. C.; Shoulders, M. D. *J. Am. Chem. Soc.* **2018**, *140* (51), 18093.

(3) Das, A. T.; Zhou, X.; Vink, M.; Klaver, B.; Verhoef, K.; Marzio, G.; Berkhout, B. *J. Biol. Chem.* **2004**, *279* (18), 18776.

(4) Wang, L.; Tsien, R. Y. *Nat. Protoc.* **2006**, *1* (3), 1346.

(5) Kelemen, Rachel E.; Jewel, Delilah; Pasha, Muhammad; Cao, Xiaofu; Chatterjee, A. *Submiss.*

(6) Kern, A.; Schmidt, K.; Leder, C.; Müller, O. J.; Wobus, C. E.; Bettinger, K.; Von der Lieth, C. W.; King, J. A.; Kleinschmidt, J. A. *J. Virol.* **2003**, *77* (20), 11072.

(7) Grieger, J. C.; Snowdy, S.; Samulski, R. J. *J. Virol.* **2006**, *80* (11), 5199.

(8) Bleker, S.; Sonntag, F.; Kleinschmidt, J. A. *J. Virol.* **2005**, *79* (4), 2528.

(9) Wu, P.; Xiao, W.; Conlon, T.; Hughes, J.; Agbandje-McKenna, M.; Ferkol, T.; Flotte, T.; Muzyczka, N. *J. Virol.* **2000**, *74* (18), 8635.

(10) Sonntag, F.; Bleker, S.; Leuchs, B.; Fischer, R.; Kleinschmidt, J. A. *J. Virol.* **2006**, *80* (22), 11040.

(11) Girod, A.; Wobus, C. E.; Zádori, Z.; Ried, M.; Leike, K.; Tijssen, P.; Kleinschmidt, J. A.; Hallek, M. *J. Gen. Virol.* **2002**, *83* (5), 973.

(12) Kelemen, R. E.; Mukherjee, R.; Cao, X.; Erickson, S. B.; Zheng, Y.; Chatterjee, A. *Angew. Chemie Int. Ed.* **2016**, *55* (36), 10645.

(13) Addy, P. S.; Erickson, S. B.; Italia, J. S.; Chatterjee, A. *J. Am. Chem. Soc.* **2017**, *139* (34), 11670.

(14) Sarathi Addy, P.; Italia, J. S.; Chatterjee, A. *ChemBioChem* **2018**, *19* (13), 1375.

(15) Addy, P. S.; Erickson, S. B.; Italia, J. S.; Chatterjee, A. In *Bioconjugation, Methods in Molecular Biology*; Humana Press Inc., 2019; Vol. 2033, pp 239–251.

(16) Sonntag, F.; Schmidt, K.; Kleinschmidt, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (22), 10220.

(17) Tse, L. V.; Moller-Tank, S.; Meganck, R. M.; Asokan, A. *J. Virol.* **2018**, *92* (14).

(18) Naumer, M.; Sonntag, F.; Schmidt, K.; Nieto, K.; Panke, C.; Davey, N. E.; Popa-Wagner, R.; Kleinschmidt, J. A. *J. Virol.* **2012**, *86* (23), 13038.

(19) Samulski, R. J.; Muzyczka, N. *Annu. Rev. Virol.* **2014**, *1* (1), 427.

(20) Maurer, A. C.; Pacouret, S.; Karla, A.; Diaz, C.; Blake, J.; Andres-Mateos, E.; Vandenberghe, L. H. *Cell Reports* **2018**, *23*, 1817.

(21) Maurer, A. C.; Cepeda Diaz, A. K.; Vandenberghe, L. H. *J. Virol.* **2019**, *93* (7), 2013.

Chapter 4

Development of an optimized baculovirus vector for incorporation of lysine analogs

4.1 Introduction

Chapter 2 of this thesis examined improvements to the VADER selection protocol using baculovirus vectors. This chapter will focus on the development of optimized baculovirus vectors to improve the incorporation of a variety of UAAs in mammalian cells, using tRNAs evolved using the VADER selection platform.

4.1.1 Different methods to deliver genetic code expansion machinery to mammalian cells

There are multiple vector types for expressing the machinery needed to install UAAs in proteins in mammalian cells. Each vector has benefits and drawbacks to its use, especially in the context of delivering small gene products such as tRNAs, which may be required at higher copy numbers to improve stop codon suppression.

4.1.1a Transient transfection

Because mammalian cells exhibit poor transformation efficiency,[1] the most commonly used method to deliver DNA to mammalian cells is transient transfection. Transfection is a simple procedure to perform in mammalian cells, as the only materials required are purified plasmids and transfection reagents. Researchers can deliver the orthogonal aaRS/tRNA pair and the gene of interest containing a stop codon, using one or two plasmids, and achieve relatively high transfection efficiencies. However, if the experiment requires more extensive genetic cargo, such as any work involving the CRISPR-Cas system, it may be necessary to use additional plasmids to deliver all of the cargo to the host cells. Since transfection efficiency drops with each additional plasmid used, and since DNA distribution to host cell nuclei during transient transfection is highly heterogeneous,[2–4] using this method may not result in optimum UAA incorporation in the target protein.

4.1.1b AAV

Viral vectors are attractive delivery vehicles for transducing mammalian cells. It is possible to infect cells with precise copy numbers of viruses, making delivery of DNA much more homogenous than with transient transfection. Modern molecular biology techniques have rendered facile the production of high titers of recombinant viruses for use in complex experiments.

AAV is one of the most popular viral vectors for use in research and clinical applications. This virus exhibits low immunogenicity due to its small size, as well as low cytotoxicity because it cannot replicate without co-infection of a helper virus.[5–7] AAV is therefore an excellent choice for a DNA delivery vector due to this favorable biosafety profile for researchers and patients. Its single-stranded DNA genome is relatively stable, but it is small at approximately 5 kb, in comparison to other viruses used to deliver DNA to mammalian cells.[8] AAV's capsid is rigid and cannot tolerate inclusion of additional genetic material, so if an experiment requires the delivery of large amounts of genetic cargo, it will be necessary to use multiple rAAV vectors.[9,10]

One additional drawback to using AAV as a vector is particularly problematic for genetic code expansion experiments. Replication of AAV's single-stranded DNA genome relies on Rep proteins binding to the ITRs.[11,12] RNA sequences in the AAV genome can fold in on themselves to make hairpins, and Rep proteins can "mistake" these hairpins for the 5' ITR and truncate the genome during replication.[13,14] These truncations can excise key components of the rAAV cargo, and are much more likely to occur when the cargo contains tRNAs for use in stop codon suppression or sgRNAs for use in CRSIPR experiments. While AAV is an excellent vector for delivering proteins of interest to

mammalian cells, its tendency towards truncation with RNA cargos limits its use for genetic code expansion applications.

4.1.1c Lentivirus

Lentiviral vectors have increased in popularity as DNA delivery vehicles due to their large cargo capacity and ability to integrate into the mammalian genome. Retroviruses use reverse transcriptase to replicate their RNA genomes, so the DNA daughter strands can integrate into genomic DNA to produce stable cell lines. This feature is attractive for researchers because stably integrated genes can offer more homogenous and robust expression than can transient transfection. Since lentiviruses such as HIV-1 are highly pathogenic to humans, it is essential for researcher safety that all vectors used be replication-deficient. Recombinant lentivirus vectors are typically pseudotyped with VSV-G instead of the viral envelope protein Env to expand the viral tropism.[15] Third-generation lentivirus vectors split the replication-deficient HIV-1 genome into four plasmids to reduce the probability of recombination that could create replication-competent viruses.[16] While this third-generation system is safer for human researchers and patients receiving these vectors for gene therapies, it faces the same limitation of decreased transfection efficiency yielding lower viral titers than systems that rely on fewer plasmids.

Lentivirus has been used for genetic code expansion applications in mammalian cells. Expression of UAA-containing proteins appears to improve when the lentiviral vectors are used to stably integrate the aaRS/tRNA pair into the genome than when the cargo is delivered via transient transfection.,[17–19] but genomic integration can lead to disruption of essential host cell genes. Additionally, the risk of recombination in lentiviruses is still high with their single-stranded RNA genomes,[20] so it may be difficult

to reliably produce recombinant lentivirus containing tRNA cassettes to use for genetic code expansion.

4.1.1d Baculovirus

Baculovirus vectors solve the problems posed by transient transfection and by AAV and lentivirus vectors. As is the case with AAV and lentivirus, infection with baculovirus results in much more homogenous transgene expression than does transient transfection.[21] Baculovirus is an insect virus that has been pseudotyped with VSV-G to allow it to infect mammalian cells;[22–24] its genome is silent in human cells, so it is safe for researchers to use.[25] These features also make baculovirus a useful vector for *in vivo* applications.

Unlike AAV, baculovirus has a large genome with a potentially unlimited cargo capacity because of the mechanism of genome packaging involves continuous assembly of capsid proteins around the DNA until it is fully enclosed.[26,27] This massive cargo capacity allows researchers to produce recombinant baculovirus containing all of the genes needed for complex experiments, including genetic code expansion. Baculovirus's large, double-stranded DNA genome is quite stable and can carry highly repetitive cassettes of tRNAs without any evidence of recombination.[21] Because the virus genome itself is quite large (approximately 134 kb), it is not practical to clone transgenes directly into it.[26] Instead, transgenes can be cloned into shuttle plasmids which are transformed into DH10Bac cells to undergo recombination with a bacmid that contains the rest of the baculovirus genome.[28,29] Researchers can choose to transfect the shuttle plasmids directly into host cells, which is a convenient procedure, or to produce infective baculovirus, which will transduce cells more homogenously. These features (large and stable genome, safe use, and

flexibility of use) make baculovirus an ideal vector to deliver the genetic machinery needed for unnatural amino acid incorporation in mammalian cells.

4.1.2 A new baculovirus vector for the pyrrolysyl system

Previous work has produced baculovirus vectors for tyrosyl and pyrrolysyl systems to incorporate tyrosine and lysine derivatives in proteins in mammalian cells.[21,24] The baculovirus vector pAcBac3 in particular optimizes incorporation of O-methyltyrosine (Figure 4.1) over the original pAcBac2 system, with 20 copies of the suppressor tyrosyl tRNA (YtR) and UbiC promoter for the aaRS (Figure 4.2).[21] Since the pyrrolysyl system is popular for genetic code expansion applications in mammalian cells (due to the ease of evolving synthetases in bacteria that will be naturally orthogonal in eukaryotes),[30,31] it would be beneficial for the scientific community to have an optimized baculovirus vector for the pyrrolysyl system.

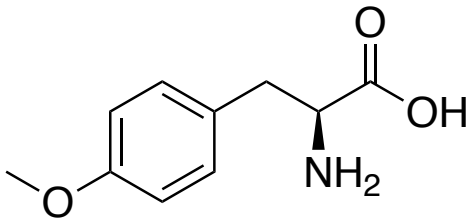


Figure 4.1. Structure of O-methyltyrosine. The baculovirus vector pAcBac3 was able to significantly improve the incorporation of O-methyltyrosine in proteins in mammalian cells.

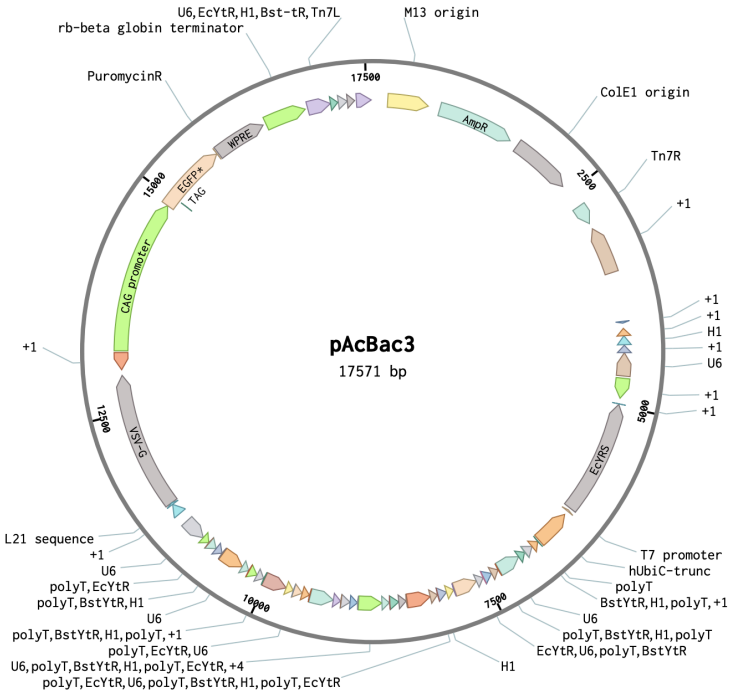


Figure 4.2. Map of pAcBac3 baculovirus vector.[21] 20 copies of the tRNA cassette (distributed over three locations within the vector, one of which contains 16 tandem repeats of the tRNA) rescues incorporation of O-methyltyrosine in mammalian cells.

As previously discussed, the VADER selection scheme has yielded tRNAs that install lysine variants in response to TAG stop codons much more efficiently.[32] Two of these variants, A2 and HTS25, have activities that improve the incorporation of AzK several-fold over that of the wild type PytR, in experiments where they are delivered via transient transfection (HTS25: unpublished data). Since infection by baculovirus results in much more homogenous host cell transduction than does transient transfection,[21] a baculovirus vector carrying these improved PytR variants would improve stop codon-containing transgene expression both by increased DNA delivery to the nucleus and by enhanced tRNA activity. This enhanced baculovirus vector will be highly beneficial to the scientific community.

Even though baculovirus vectors have the potential to be the standard vector for mammalian genetic code expansion, their use has not been broadly adopted by researchers in this field. The development of a baculovirus vector that contains evolved PytR variants, as well as genes for an evolved synthetase and stop codon-containing protein of interest that can be easily exchanged via molecular cloning, has the potential to popularize the use of these versatile vectors.

In this chapter, I demonstrate that delivering tRNAs by infecting cells with baculovirus allows researchers to deliver low copy numbers of tRNAs to cells and to see improved activity in an evolved PytR variant that is not evident when the cells receive more copies of the tRNAs. In addition, I develop an enhanced baculovirus vector using another, more efficient PytR variant. I determine the optimum promoter to use with the MbPylRS and whether a nuclear export sequence improves UAA incorporation in this system. Finally, I demonstrate that this vector can be used with a number of different evolved MbPylRS variants to incorporate a variety of UAAs into proteins in mammalian cells.

4.2 Results and Discussion

4.2.1 Comparing A2 PytR variant to wild type PytR

Previously, Kelemen et al. evolved a PytR variant A2 that outperforms the wild type PytR by 300%, as demonstrated via transient transfection in HEK-293T cells.[32] While this demonstrated improvement is considerable, it is possible that A2's actual improved efficiency would be even more evident were the tRNAs delivered at lower copy numbers. With transient transfection, cells that receive DNA may receive hundreds of copies of the tRNA, and the difference in performance between tRNAs may not be evident when there are so many copies of a less efficient tRNA. I generated baculovirus vectors to carry the

tRNAs so that I could precisely control the copy numbers delivered to cells, as well as ensure a homogenous distribution of infected cells.

While infective titer by FACS is more precise than by using the Clontech BacPAK™ Baculovirus Rapid Titer Kit, meaning that it is possible to deliver precise copy numbers of baculovirus to target cells, there is still variation in expression of the genes these viruses encode. To control for this variation in gene expression from the MTH-mCherry and tRNA containing viruses, I analyzed all GFP expression data by comparing it to the mCherry expression data for that particular treatment. Results are therefore given as normalized GFP and mCherry fluorescence in RFU, and also as GFP/mCherry ratios.

Expression of eGFP from the pAcBac-CMV-MbPylRS-CMV-GFPtag39 virus reached approximately 10% that of wild type eGFP expression from pAcBac-CMV-MbPylRS-CMV-GFPwt. For cells infected with baculovirus containing MTH-mCherry, mCherry expression increased approximately linearly as MOI increased, for both PytR and A2 virus (Figure 4.3). In cells infected with wild type PytR virus, GFP expression continued to increase with MOI up to the maximum MOI of 20 (Figure 4.4). For A2 virus, GFP expression plateaued at approximately 400 RFU with infection from 5-20 MOI, indicating that even at low copy numbers (20-80 copy numbers, with 4xtRNA per virus), this tRNA has maximized its ability to incorporate BocK at TAG codons in mammalian cells.

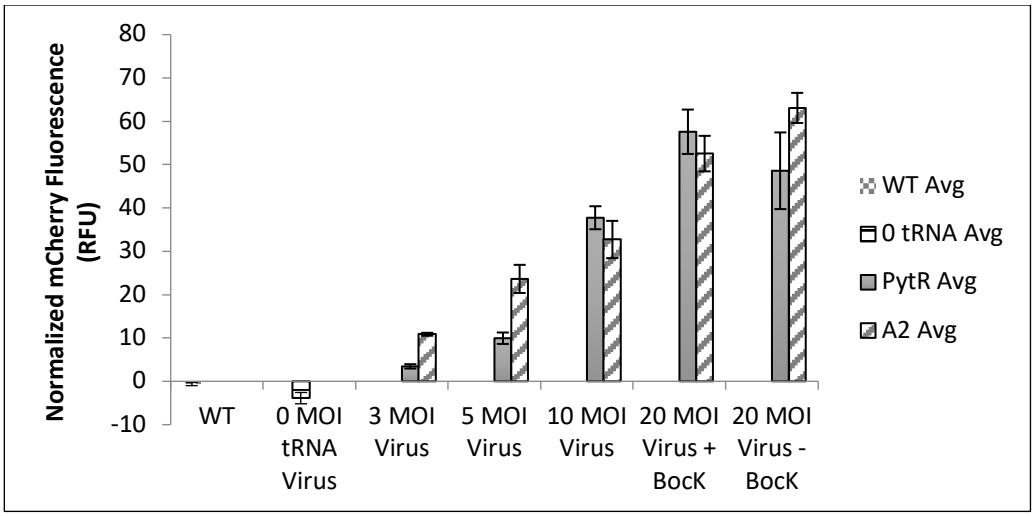
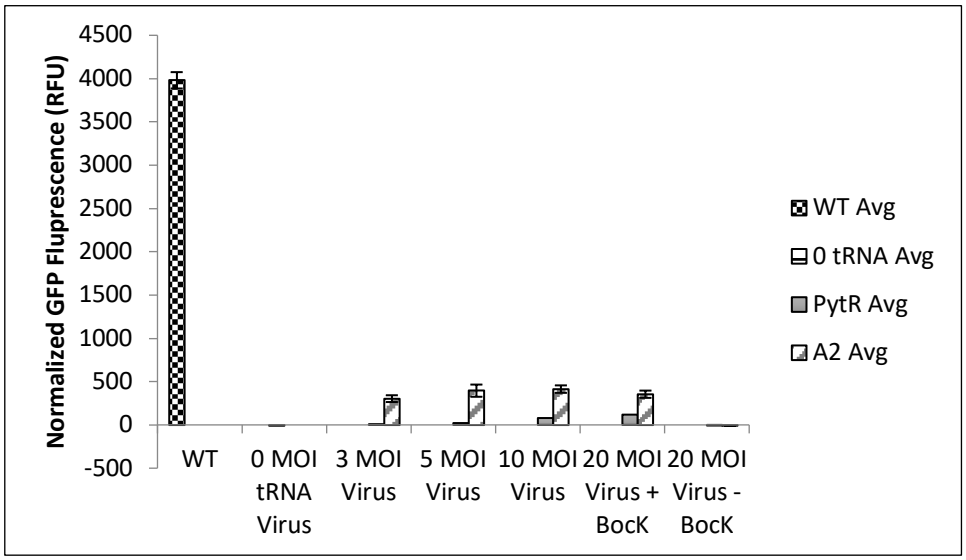
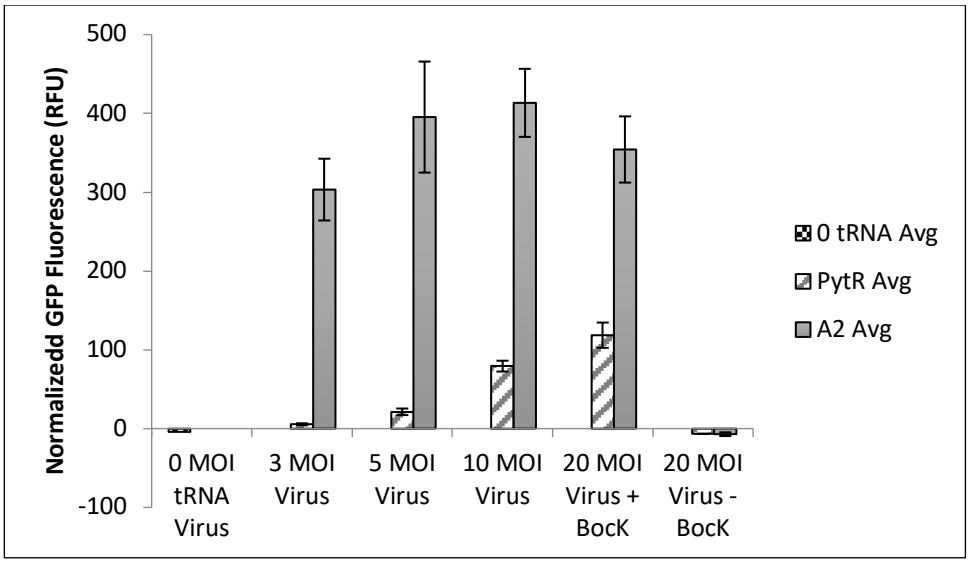


Figure 4.3. mCherry fluorescence of cells infected with baculovirus. HEK-293T cells were infected with 100 MOI of GFP-containing virus (either wild type GFP or GFP39tag) and varying MOIs of viruses containing 4 copies of a tRNA cassette (wild type PytR or A2) and mCherry expressed from the zinc metallothionine promoter. After 48 hours, cells were imaged and lysed, and mCherry and GFP fluorescence was taken.



(A)



(B)

Figure 4.4. GFP fluorescence of cells infected with baculovirus. HEK-293T cells were infected with 100 MOI of GFP-containing virus (either wild type GFP or GFP39tag) and varying MOIs of viruses containing 4 copies of a tRNA cassette (wild type PytR or A2) and mCherry expressed from the zinc metallothionine promoter. After 48 hours, cells were imaged and lysed, and mCherry and GFP fluorescence was taken. (A) GFP fluorescence of all treatments. (B) GFP fluorescence of cells that received GFP39tag virus and tRNA viruses.

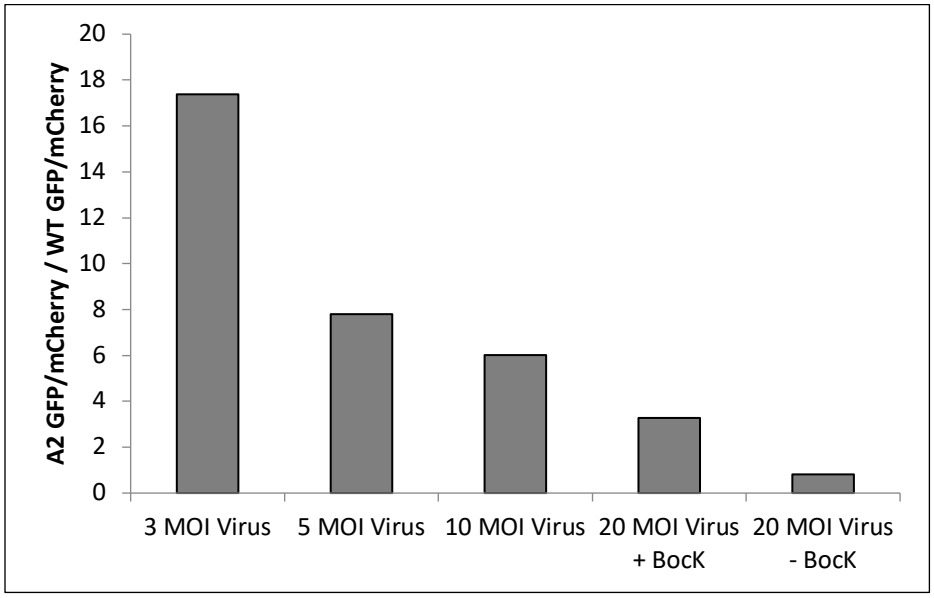


Figure 4.5. Comparison of GFP/mCherry ratios of A2 and wild type PytR. For each tRNA at each MOI, ratios were taken of average GFP values and average mCherry values. The ratio of these ratios (A2/wild type) at each MOI is provided here.

Comparing GFP/mCherry values for each virus at each MOI demonstrates the superiority of the A2 tRNA. The ratio between A2's and the wild type PytR's GFP/mCherry ratios is highest at MOI of 3 and decreases as MOI increases (Figure 4.5), which indicates that as copy number decreases, the heightened efficiency of this evolved tRNA becomes more important. At MOI of 3, the ratio between A2's and the wild type PytR's GFP/mCherry ratios is 17.4, which shows an improvement far beyond the 3-fold improvement previously suspected for A2 when the comparisons were made by delivering the tRNAs via transient transfection.

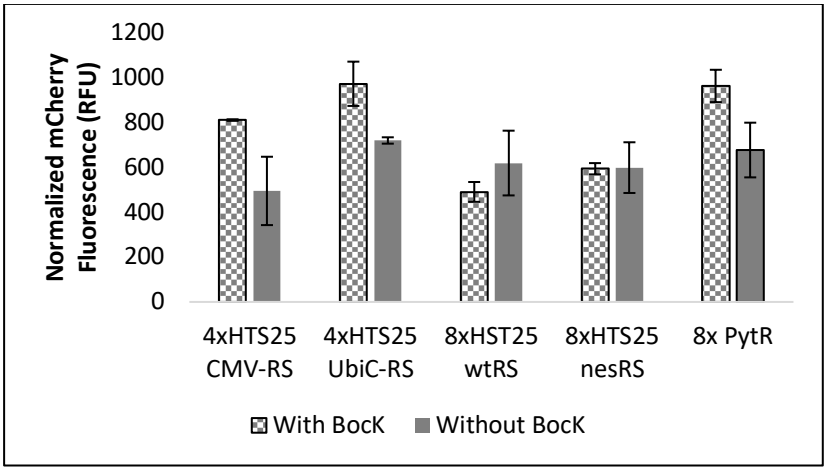
4.2.2 Designing an optimized PytR baculovorus vector

While infection via baculovirus allows for exquisite copy number control and homogenous DNA delivery, most applications for stop codon suppression in mammalian cells make use of transient transfection, as this method requires less time and labor than does preparing baculovirus vectors. Therefore, I set out to design a baculovirus vector that makes use of our highest-performing PytR, HTS25, that was developed using the VADER selection scheme. Each of these viral vectors contains an mCherry-TAG-GFP construct, which allows each vector to act as its own internal control for fluorescence expression.

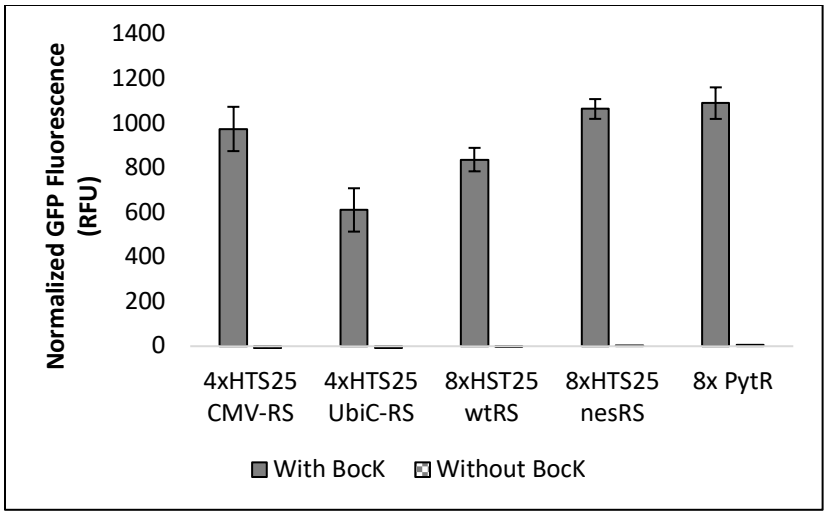
In the pAcBac3 vector optimized for TAG suppression and O-methyltyrosine (OMeY) incorporation, putting the synthetase under the control of the weaker promoter UbiC (as opposed to the strong promoter CMV) predictably decreased synthetase expression, but substantially improved the expression of eGFPtag39.[21] To determine which promoter would enable the expression of more GFP in this vector, I compared GFP/mCherry ratios in cells transfected with either pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-4xHTS25 or pAcBac-UbiC-MbPylRS-CAG-mCherry-TAG-GFP-

4xHTS25 (Figure 4.6). These vectors only contained one tRNA cassette due to cloning site restrictions after cloning in the different promoters. The GFP/mCherry ratio for the vector with the CMV promoter was 2.3x higher than that of the vector with the UbiC promoter, demonstrating that unlike with the tyrosyl vector, having the synthetase under the control of the stronger promoter in the pyrrolysyl vector improves stop codon suppression.

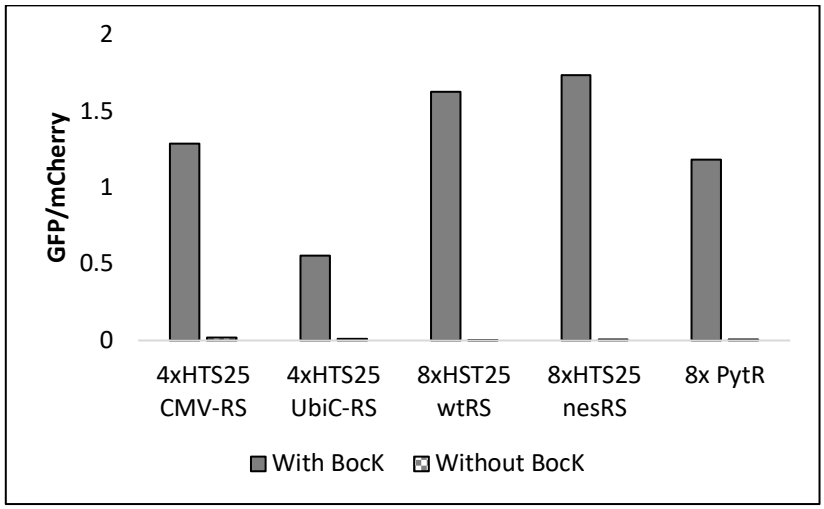
I cloned in an extra copy of the tRNA cassette into the vector with the CMV promoter, and made a similar construct using the wild type PytR, so that each vector had eight copies of its respective tRNA. The GFP/mCherry ratio for the 8xPytR vector was slightly lower than that for the 4xHTS25 vector, which demonstrates that even by transient transfection, which delivers high copy numbers of the DNA to cells, that cells needed twice as much of the wild type PytR to make as much GFP as the cells that received HTS25 did. The 8xHST25 vector performed 1.4x better than did the 8xPytR. It is likely that the differences between these vectors will be more extreme when they are packaged into baculoviruses and infected at low MOI.



(A)



(B)



(C)

Figure 4.6. Comparison of GFP/mCherry ratios of HTS25 and wild type PytR. HEK-293T cells were transfected with pAcBac vectors containing the mCherry-TAG-GFP cassette and the MbPylRS under control of either the CMV or the UbiC promoter. These vectors contained 4 or 8 copies of the HTS25 tRNA, or 8 copies of the wild type PytR. One of the 8xHTS25 vectors also contained an MbPylRs with a nuclear export sequence (nes) on its N-terminus. (A) mCherry fluorescence of lysate of infected cells. (B) GFP fluorescence of infected cell lysate. (C) Comparison of mCherry/GFP ratios between cells infected with each vector.

While the differences between the wild type PytR and the HTS25 variant are evident even by transfection, the nuclear export sequence on the N-terminus of the MbPylRS did not make a substantial improvement to BocK incorporation when these vectors were delivered by transient transfection. Many small differences between

performance of these synthetases may have masked because of the high copy numbers delivered to each cell in transfection experiments, so it is possible that the NES may offer a substantial improvement to BocK incorporation once baculovirus is produced using these vectors and used to infect mammalian cells.

4.2.3 Incorporating other optimized synthetases

There have been a variety of pyrrolysyl synthetase variants that have been optimized to charge lysine variants of various sizes and characteristics for different uses. I made pAcBac vectors with some of these MbPylRS variants to demonstrate the versatility of these vectors for incorporating such diverse UAAs.

The NBK RS developed by the Schultz lab can be used to incorporate photocaged lysine variants. I replaced the wild type MbPylRS with the NBK-RS to produce pAcBac-CMV-NBK-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25 and compared the incorporation of the photocaged lysine PCK (Figure 4.7) between cells transfected with this plasmid and its wild type MbPylRS counterpart. The GFP/mCherry ratios were 4.4x higher with the NBK-RS vector than with the wild type RS, and there was negligible GFP expression from either vector in the absence of PCK (Figures 4.8, 4.9). This NBK-RS pAcBac vector with the evolved tRNA cassettes is therefore a highly useful tool for incorporating photocaged lysine UAAs in mammalian cells.

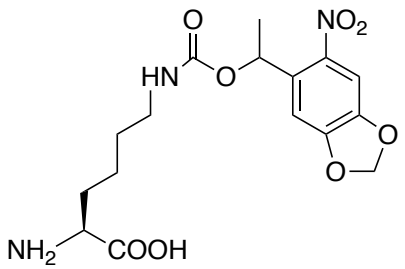


Figure 4.7. Structure of the photocaged lysine PCK.

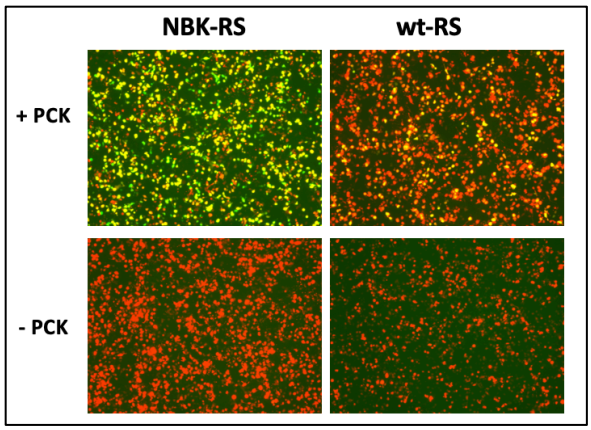
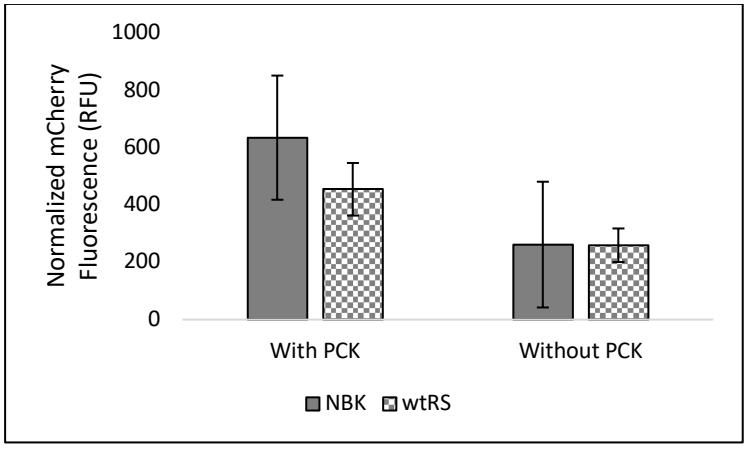
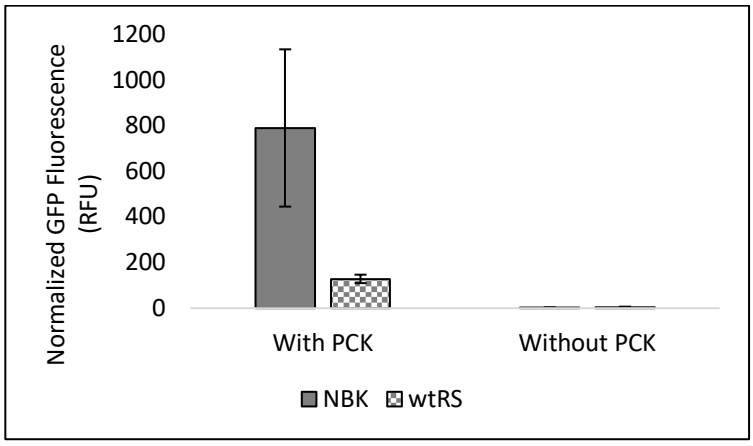


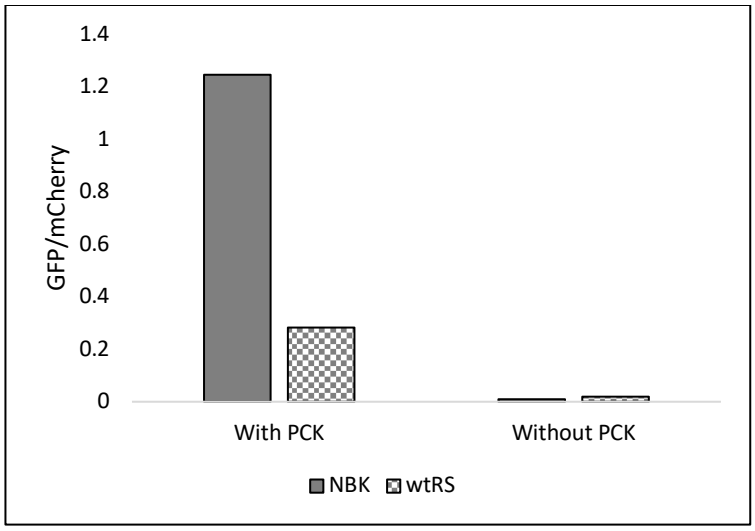
Figure 4.8. Comparison of NBK-MbPylRS and wild type MbPylRS to charge PCK. HEK-293T were transfected with pAcBac-CMV-NBK-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25 or pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25. Cells were imaged after 72 hours.



(A)



(B)



(C)

Figure 4.9. Comparison of GFP/mCherry ratios for NBK-RS and the wild type MbPylRS. HEK-293T cells were transfected with pAcBac vectors containing the mCherry-TAG-GFP cassette and either the wild type MbPylRS or the NBK-RS. Cells were imaged and lysed after 72 hours, then GFP and mCherry fluorescence were taken. (A) mCherry fluorescence of lysate of infected cells. (B) GFP fluorescence of infected cell lysate. (C) Comparison of mCherry/GFP ratios between cells infected with either the wild type synthetase or the NBK-RS.

Diazirine lysine (DiazK; Figure 4.10) is another photocaged amino acid for which an MbPylRS has been evolved. Unlike its NBK counterpart, pAcBac-CMV-DiazK-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25 did not result in higher GFP expression than the vector with the wild type MbPylRS (Figure 4.11). Rather, GFP/mCherry ratio for the wild type synthetase was 7.6x higher than that for the DiazK synthetase (Figure 4.12). It is possible that in the context of this vector, the DiazK synthetase was not expressed optimally for DiazK incorporation, and that using a different promoter for this gene would improve synthetase function.

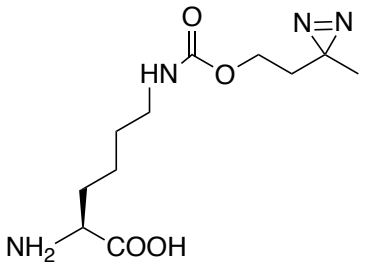


Figure 4.10. Structure of DiazK.

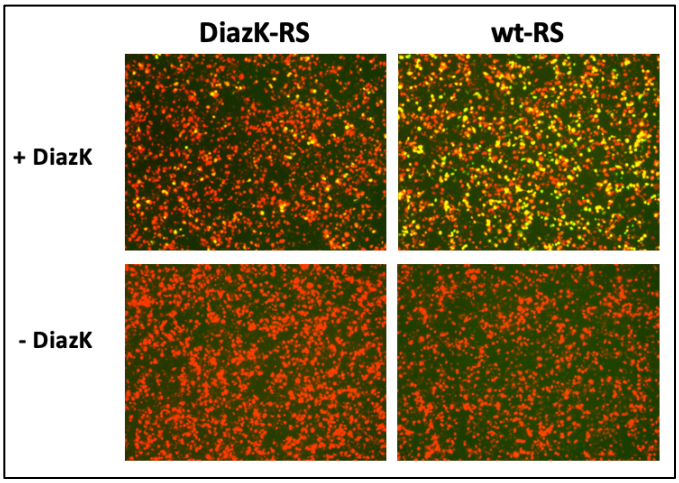
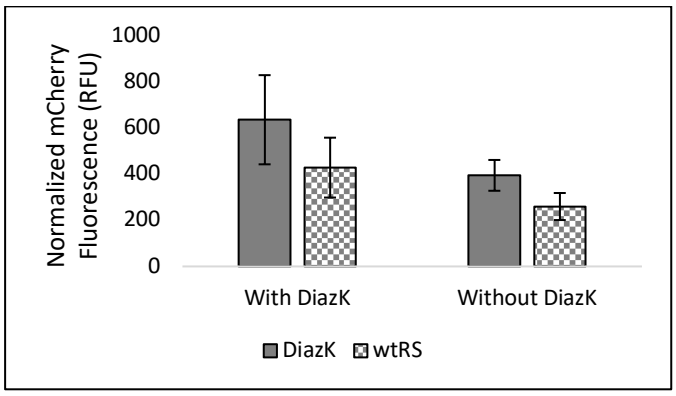
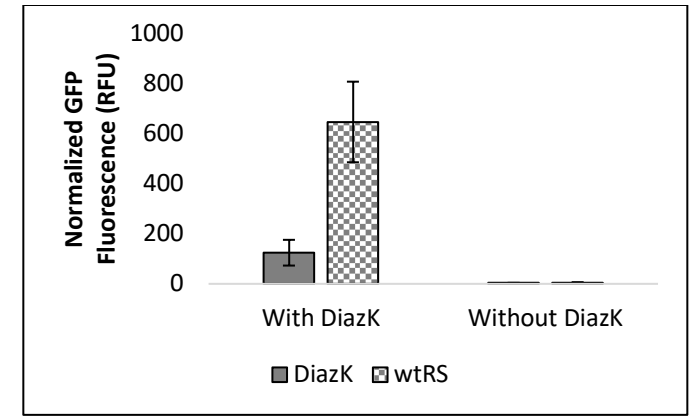


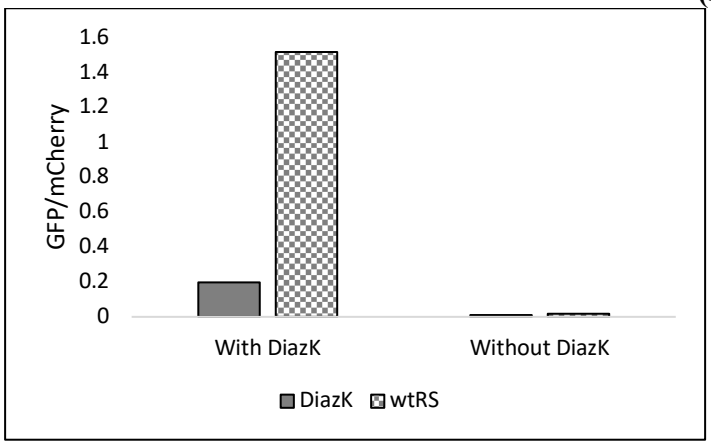
Figure 4.11. Comparison of DiazK-MbPylRS and wild type MbPylRS to charge DiazK. HEK-293T were transfected with pAcBac-CMV-DiazK-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25 or pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25. Cells were imaged after 72 hours.



(A)



(B)



(C)

Figure 4.12. Comparison of GFP/mCherry ratios for DiazK-RS and the wild type MbPylRS. HEK-293T cells were transfected with pAcBac vectors containing the mCherry-TAG-GFP cassette and either the wild type MbPylRS or the DiazK-RS. Cells were imaged and lysed after 72 hours, then GFP and mCherry fluorescence were taken. (A) mCherry fluorescence of lysate of infected cells. (B) GFP fluorescence of infected cell lysate. (C) Comparison of mCherry/GFP ratios between cells infected with either the wild type synthetase or the DiazK-RS.

N$^{\varepsilon}$-acetyl-L-lysine (AcK) is a UAA whose incorporation in mammalian cells at the point of translation mimics the posttranslational acetylation of lysine (Figure 4.13). There is an MbPylRS variant selected for incorporation of AcK. In this backbone, however, it did not incorporate AcK at a level that improved GFP expression more than did the wild type synthetase (Figure 4.14). Both synthetases incorporated GFP at a negligible level (Figure 4.15). The context of this vector may also require further optimization for use with this

synthetase, such as using a different promoter for the synthetase as the tyrosyl RS requires in pAcBac3.[21]

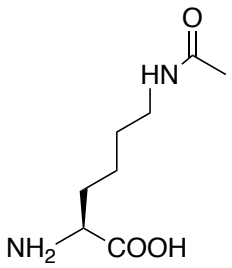


Figure 4.13. Structure of AcK.

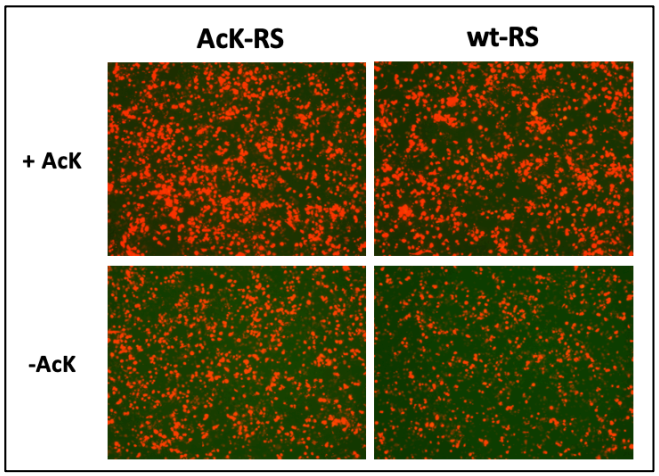
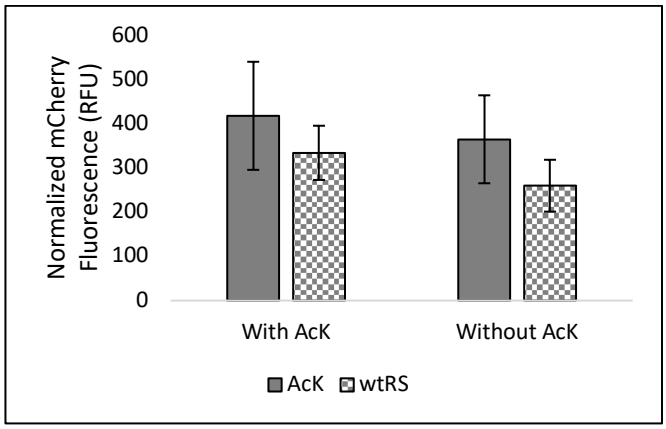
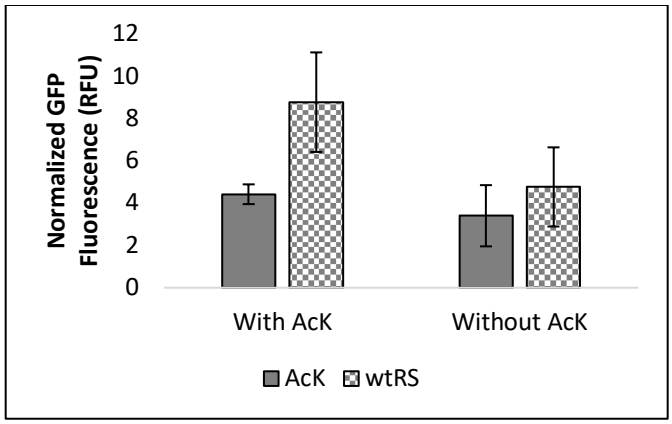


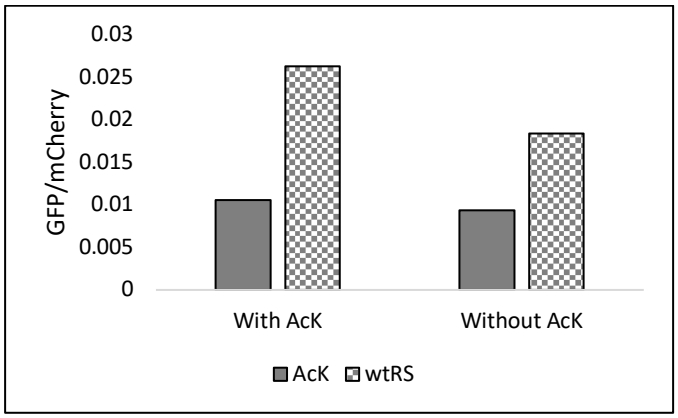
Figure 4.14. Comparison of AcK-MbPylRS and wild type MbPylRS to charge AcK. HEK-293T were transfected with pAcBac-CMV-AcK-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25 or pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25. Cells were imaged after 72 hours.



(A)



(B)



(C)

Figure 4.15. Comparison of GFP/mCherry ratios for AcK-RS and the wild type MbPylRS. HEK-293T cells were transfected with pAcBac vectors containing the mCherry-TAG-GFP cassette and either the wild type MbPylRS or the AcK-RS. Cells were imaged and lysed after 72 hours, then GFP and mCherry fluorescence were taken. (A) mCherry fluorescence of lysate of infected cells. (B) GFP fluorescence of infected cell lysate. (C) Comparison of mCherry/GFP ratios between cells infected with either the wild type synthetase or the AcK-RS.

There is a MbPylRS variant (the biotin-lysine variant, or Biotin-K) that is a polyspecific synthetase capable of charging a number of diverse lysine variants, including those too large for most synthetases. I compared the performance of this synthetase with that of the wild type MbPylRS at incorporating azidolysine (AzK; Figure 2.14) and cyclo-octyne lysine (SCOK; Figure 4.16). The wild type synthetase outperformed the Biotin-K synthetase at incorporating AzK into the mCherry-TAG-GFP construct (Figure 4.17), with a GFP/mCherry ratio 1.4x that of the Biotin-K's GFP/mCherry ratio. With SCOK,

however, the Biotin-K synthetase had a GFP/mCherry ratio an impressive 53x that of the wild type synthetase's GFP/mCherry ratio (Figure 4.18). The wild type synthetase did not charge SCOK well, as evidenced by the little GFP fluorescence in cells transfected with the wild type synthetase pAcBac construct and provided with SCOK. This wild type synthetase does not have an active site large enough to accommodate large amino acids such as SCOK, but the Biotin-K synthetase's evolved active site is. The Biotin-K synthetase did charge AzK, and did produce an appreciable amount of GFP, but its activity was not as high as the wild type synthetase's. This variation in performance is to be expected in a polyspecific synthetase, especially one with a substrate range as large as that of the Biotin-K synthetase, as it is not possible for a synthetase to optimally charge such different UAAs with equal efficiency.

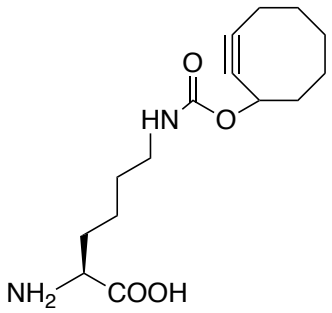


Figure 4.16. Structure of cyclo-octyne lysine (SCOK).

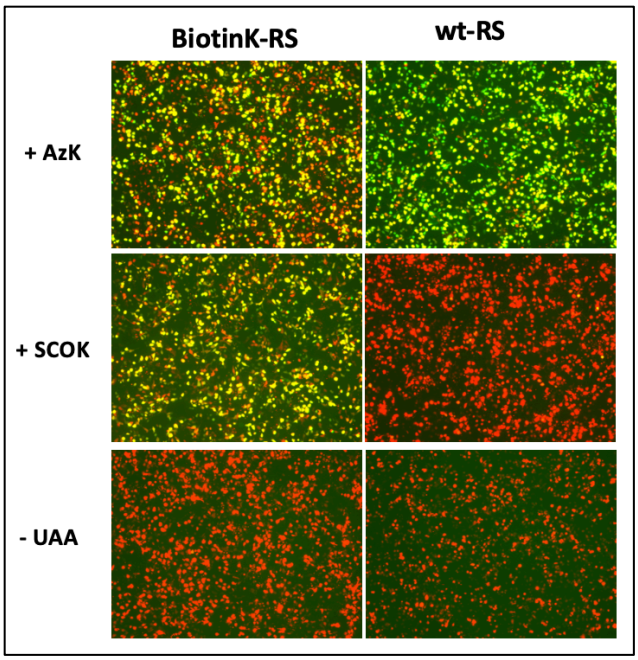
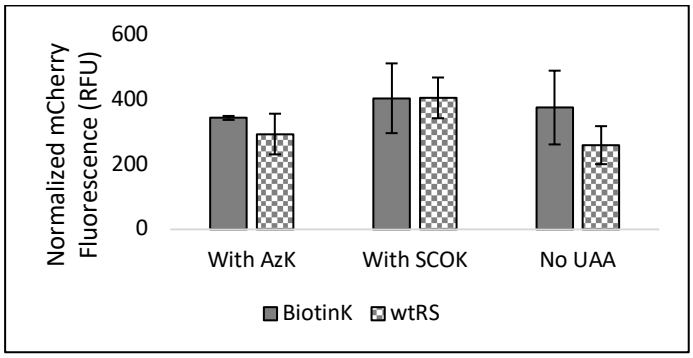
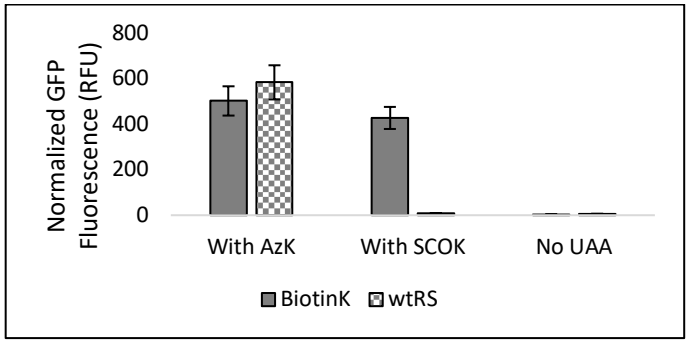


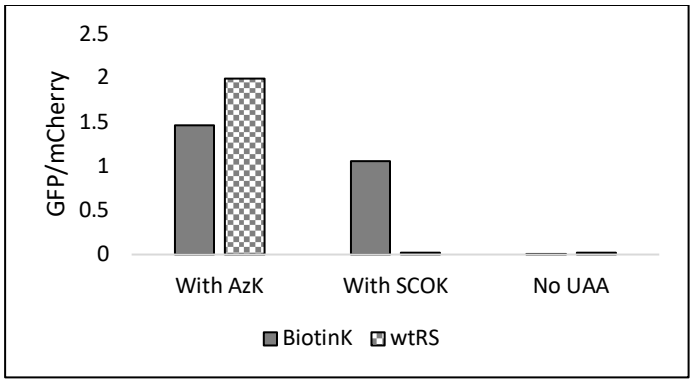
Figure 4.17. Comparison of BiotinK-MbPylRS and wild type MbPylRS to charge AzK and SCOK. HEK-293T were transfected with pAcBac-CMV-BiotinK-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25 or pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25. Cells were imaged after 72 hours.



(A)



(B)



(C)

Figure 4.18. Comparison of GFP/mCherry ratios for AcK-RS and the wild type MbPylRS. HEK-293T cells were transfected with pAcBac vectors containing the mCherry-TAG-GFP cassette and either the wild type MbPylRS or the AcK-RS. Cells were imaged and lysed after 72 hours, then GFP and mCherry fluorescence were taken. (A) mCherry fluorescence of lysate of infected cells. (B) GFP fluorescence of infected cell lysate. (C) Comparison of mCherry/GFP ratios between cells infected with either the wild type synthetase or the BiotinK-RS, for each UAA.

4.3 Conclusion

These experiments demonstrate the power of baculovirus vectors for incorporating UAAs into proteins in mammalian cells. These vectors, when made into P1 baculovirus, can deliver aaRS/tRNA pairs to mammalian cells with exquisite control over copy number. Transduction by virus allows the homogeneous distribution of DNA across the population of cells, which is useful in *in vivo* experiments, as well as delivery of low copy numbers of virus to each cell, which here allows the true elucidation of the improved efficiency of an evolved tRNA. Even though the superiority of this evolved tRNA is most evident in infection experiments, these pAcBac vectors are also useful when they are delivered via transient transfection. The NES appended to the N-terminus of the synthetase unfortunately does not boost stop codon suppression by a measurable amount with transfection, when many small differences in efficiencies between constructs are masked. HTS25's superiority over the wild type PytR in suppressing TAG codons is evident even in these transfection experiments, as eight copies of the wild type PytR are required for the same amount of stop

codon suppression as provided by four copies of HTS25. These vectors can be used for a variety of experiments, as the synthetase in the vector can be easily swapped for one specifically suited to the experiment. Overall, these pAcBac vectors present a useful tool for the chemical biology community to use in experiments involving lysine variants and TAG suppression in mammalian cells.

4.4 Acknowledgments

Dr. Yunan Zheng gave me access to her plasmid stocks and trained me to prepare recombinant baculovirus. Dr. Sarah Erickson gave me access to her plasmid stocks and stores of UAAs. Dr. Renpeng Gu gave me access to the stock of plasmid containing the DiazK-RS.

4.5 Materials and Methods

4.5.1. Strains and cell lines

Plasmid propagations were performed as described in Chapter 2 of this thesis. Sf9 cells and HEK-293T cells were maintained as described in Chapter 2.

4.5.2. Plasmids

All plasmid maps and sequences can be found in the Appendix.

The plasmids pAcBac-CMV-MbPylRS-CMV-GFPwt and pAcBac-CMV-MbPylRS-CMV-GFP39tag were propagated from Dr. Rachel Kelemen's stocks.

To construct pAcBac-4xPytRwt-MTH-mCherry and pAcBac-4xA2-MTH-mCherry, pAcBac-4xPytRwt and pAcBac-4xA2 (both from Dr. Kelemen's stocks) were cut with SbfI and NotI. An MTH-mCherry g-block was amplified with the primers MTH-mCherry-SbfI-F and MTH-mCherry-NotI-R, and this cassette was cloned into the pAcBac plasmids.

The mCherry-TAG-GFP construct consists of a wild type mCherry protein, followed by a flexible linker containing a stop codon (Gly-Gly-Ser-TAG-Gly-Gly-Ser), and then a wild type EGFP protein. To produce this construct, the N-terminus of the flexible linker was added to wild type EGFP using the primer NheI-linker-EGFP-F, with the reverse terminal primer GFP-EcoRI-6xHis-R. This fragment was cloned into pAcBac1-GFPwt using the EcoRI and NheI sites to produce the intermediate plasmid pAcBac1-GFPwt+linker. The mCherry component of the construct was amplified with the primers NheI-Kozak-mCherry-F and mCherry-linker-AvrII-R, which added a Kozak sequence to the 5' end of the DNA fragment and the other half of the linker to the C-terminus of mCherry. The intermediate plasmid pAcBac1-GFPwt+linker and this second DNA fragment were digested with NheI and AvrII, and the mCherry half of the construct was cloned into these sites, destroying the NheI and AvrII cut sites in the linker and producing the plasmid pAcBac-mCherry-TAG-GFP. The mCherry-TAG-GFP construct can be cut, intact, from the plasmid using the NheI and EcoRI cut sites for ease of future clonings.

To produce pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-4xHTS25 and pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-4xPytRwt, a number of intermediate plasmids were generated before arriving at the final constructs. I digested pIDT-4xHTS25 and pIDT-4xPytRwt with AvrII and NheI. I digested pAcBac-CMV-MbPylRS-WPRE-CAG-GFPtag39AvrIIdel-WPRE with AvrII and cloned the 4xHTS25 and 4xPytRwt cassettes in using those cut sites to generate the intermediate plasmids pAcBac-CMV-MbPylRS-CAG-GFPtagAvrIIdel-4xHTS25 and pAcBac-CMV-MbPylRS-CAG-GFPtagAvrIIdel-4xPytRwt. To add a second tRNA cassette to bring the total tRNA copy number in each plasmid to 8, I digested the intermediate plasmids with SpeI, which is

compatable with both AvrII and NheI, allowing me to produce the plasmids pAcBac-CMV-MbPylRS-CAG-GFPtagAvrIIdel-8xHTS25 and pAcBac-CMV-MbPylRS-CAG-GFPtagAvrIIdel-8xPytRwt.

It was necessary to replace the GFPtag construct with the mCherry-TAG-GFP construct so that there would be an internal fluorescent control in all experiments using these plasmids. SfiI sites were added to either end of the mCherry-TAG-GFP-WPRE construct using the primers Gibson-mCherry-F and Gibson-WPRE-R. The intermediate plasmids pAcBac-CMV-MbPylRS-CAG-GFPtagAvrIIdel-4xHTS25, pAcBac-CMV-MbPylRS-CAG-GFPtagAvrIIdel-4xPytRwt, pAcBac-CMV-MbPylRS-CAG-GFPtagAvrIIdel-8xHTS25, and pAcBac-CMV-MbPylRS-CAG-GFPtagAvrIIdel-8xPytRwt were all digested with SfiI to remove the GFPtag construct, and the mCherry-TAG-GFP-WPRE construct was added to each of these plasmids using Gibson assembly to produce pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-4xHTS25, pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-4xPytR, pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-8xHTS25, and pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-8xPytR.

pAcBac-UbiC-MbPylRS-CAG-mCherry-TAG-GFP-4xHTS25 were produced by digesting pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-4xHTS25 with SpeI and SacI, and using Gibson assembly with a UbiC fragment amplified with the primers Gibson-UbiC-F and Gibson-Ubic-R.

The nuclear export sequence was added to the N-terminus of the MbPylRS to produce NES-MbPylRS using the primers NheI-Lemke-NES-MbPylRS-F and MbPylRS-WPRE-XhoI-R. The point mutations to produce the Biotin-K MbPylRS were introduced

using the internal primers L274A-F, L274A-mut-R, V370R-F, V730R-mut-R, C313V-M315Y-F, and C313V-M315Y-mut-R; and the terminal primers MbPylRS-Nterm-F and MbPylRS-Cterm-R.

4.5.3 Unnatural amino acids

PCK was synthesized as previously described.[33] DiazK was purchased from SiChem (catalog number SC-8034). AcK was purchased from Thermo Fisher Scientific (catalog number AAJ613906). AzK was purchased from Iris Biotech (catalog number HAA1625). SCOK was purchased from SiChem (catalog number SC-8000).

4.5.4 Producing recombinant baculovirus

Recombinant baculoviruses were produced as described in detail in Chapter 2 of this thesis.

4.5.5 Comparing A2 PytR variant to wild type PytR

HEK-293T cells were infected with either pAcBac-CMV-MbPylRS-CMV-GFPwt or pAcBac-CMV-MbPylRS-CMV-GFPtag39 P1 virus at 100 MOI. Cells that received GFPtag39 virus received either pAcBac-4xPytRwt-MTH-mCherry or pAcBac-4xA2-MTH-mCherry P1 virus at the following MOIs: 0, 3, 5, 10, or 20. BocK was also added to the cell media of wells that received GFPtag39 and tRNA virus. No BocK controls wells received GFPtag39 virus and either tRNA virus at MOI 20.

4.5.6 Designing an optimized PytR baculovorus vector

To determine whether the CMV or UbiC promoter for the MbPylRS is optimum for enhancing the incorporation of lysine variants at TAG codons, pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-4xHTS25 was compared with pAcBac-UbiC-

MbPylRS-CAG-mCherry-TAG-GFP-4xHTS25 via transient transfection, with and without the addition of BocK to the cell media.

To demonstrate the improvement of the HTS25 PytR over the wild type PytR, I compared pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-8xPytR to pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25 via transient transfection, with and without the addition of BocK to the cell media. I also compared pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25 to pAcBac-CMV-NES-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25 via transient transfection to determine whether the Lemke nuclear export sequence improved the incorporation of BocK and therefore GFP expression.

4.5.7 Incorporating other optimized synthetases

I compared pAcBac-CMV-NBK-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25 to the same vector with the wild type synthetase via transient transfection with a photocaged lysine (PCK) as the substrate. I also tested pAcBac-CMV-DiazK-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25 with DiazK, pAcBac-CMV-AcK-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25 with AcK, and pAcBac-CMV-BiotinK-MbPylRS-CAG-mCherry-TAG-GFP-8xHST25 with azidolysine (AzK) and cyclo-octyne lysine (SCOK).

4.6 References

(1) Liu, C. C.; Schultz, P. G. *Annu. Rev. Biochem.* **2010**, *79* (1), 413.
(2) Cohen, R. N.; van der Aa, M. A. E. M.; Macaraeg, N.; Lee, A. P.; Szoka, F. C. *J. Control. Release* **2009**, *135* (2), 166.
(3) Carpentier, E.; Paris, S.; Kamen, A. A.; Durocher, Y. *J. Biotechnol.* **2007**, *128* (2), 268.
(4) Glover, D. J.; Leyton, D. L.; Moseley, G. W.; Jans, D. A. *J. Gene Med.* **2010**, *12* (1), 77.
(5) Atchison, R. W.; Casto, B. C.; Hammon, W. M. D. *Science (80-.).* **1965**, *149* (3685), 754.
(6) Samulski, R. J.; Muzyczka, N. *Annu. Rev. Virol.* **2014**, *1* (1), 427.
(7) Wang, D.; Tai, P. W. L.; Gao, G. *Nature Reviews Drug Discovery*. Nature Publishing Group May 1, 2019, pp 358–378.
(8) Srivastava, A.; Lusby, E. W.; Berns, K. I. *J. Virol.* **1983**, *45* (2).
(9) Nakai, H.; Storm, T. A.; Kay, M. A. *Nat. Biotechnol.* **2000**, *18* (5), 527.
(10) Rendahl, K. G.; Leff, S. E.; Otten, G. R.; Spratt, S. K.; Bohl, D.; Van Roey, M.; Donahue, B. A.; Cohen, L. K.; Mandel, R. J.; Danos, O.; Snyder, R. O. *Nat. Biotechnol.* **1998**, *16* (8), 757.
(11) Hauswirth, W. W.; Berns, K. I. *Virology* **1977**, *78* (2), 488.
(12) Im, D. S.; Muzyczka, N. *J. Virol.* **1992**, *66* (2).
(13) Xie, J.; Mao, Q.; Tai, P. W. L.; He, R.; Ai, J.; Su, Q.; Zhu, Y.; Ma, H.; Li, J.; Gong, S.; Wang, D.; Gao, Z.; Li, M.; Zhong, L.; Zhou, H.; Gao, G. *Mol. Ther.* **2017**, *25* (6), 1363.
(14) Tran, N. T.; Heiner, C.; Weber, K.; Weiand, M.; Wilmot, D.; Xie, J.; Wang, D.; Brown, A.; Manokaran, S.; Su, Q.; Zapp, M. L.; Gao, G.; Tai, P. W. L. *Mol. Ther. - Methods Clin. Dev.* **2020**, *18*, 639.
(15) Naldini, L.; Blömer, U.; Gallay, P.; Ory, D.; Mulligan, R.; Gage, F. H.; Verma, I. M.; Trono, D. *Science (80-.).* **1996**, *272* (5259), 263.
(16) Dull, T.; Zufferey, R.; Kelly, M.; Mandel, R. J.; Nguyen, M.; Trono, D.; Naldini, L. *J. Virol.* **1998**, *72* (11), 8463.
(17) Shen, B.; Xiang, Z.; Miller, B.; Louie, G.; Wang, W.; Noel, J. P.; Gage, F. H.; Wang, L. *Stem Cells* **2011**, *29* (8), 1231.
(18) Sakin, V.; Hanne, J.; Dunder, J.; Anders-Össwein, M.; Laketa, V.; Nikić, I.; Kräusslich, H. G.; Lemke, E. A.; Müller, B. *Cell Chem. Biol.* **2017**, *24* (5), 635.
(19) Zheng, Y.; Yu, F.; Wu, Y.; Si, L.; Xu, H.; Zhang, C.; Xia, Q.; Xiao, S.; Wang, Q.; He, Q.; Chen, P.; Wang, J.; Taira, K.; Zhang, L.; Zhou, D. *Nucleic Acids Res.* **2015**, *43* (11), 73.
(20) Chen, J.; Powell, D.; Hu, W.-S. *J. Virol.* **2006**, *80* (19), 9651.
(21) Zheng, Y.; Lewis, T. L.; Igo, P.; Polleux, F.; Chatterjee, A. *ACS Synth. Biol.* **2017**, *6* (1), 13.
(22) Barsoum, J.; Brown, R.; McKee, M.; Boyce, F. M. *Hum. Gene Ther.* **1997**, *8* (17), 2011.
(23) Kaikkonen, M. U.; Räty, J. K.; Airenne, K. J.; Wirth, T.; Heikura, T.; Ylä-Herttuala, S. *Gene Ther.* **2006**, *13* (4), 304.
(24) Chatterjee, A.; Xiao, H.; Bollong, M.; Ai, H. W.; Schultz, P. G. *Proc. Natl. Acad.*

Sci. U. S. A. **2013**, *110* (29), 11803.
(25) Kost, T. A.; Condreay, J. P. *Trends in Biotechnology*. Elsevier Current Trends April 1, 2002, pp 173–180.
(26) Ayres, M. D.; Howard, S. C.; Lopez-Ferber, M.; Lopez-Ferber, M.; Possee, R. D. *Virology* **1994**, *202* (2), 586.
(27) Cheshenko, N.; Krougliak, N.; Eisensmith, R. C.; Krougliak, V. A. *Gene Ther.* **2001**, *8* (11), 846.
(28) Jarvis, D. L. In *Methods in Enzymology*; Academic Press Inc., 2009; Vol. 463, pp 191–222.
(29) Kost, T. A.; Condreay, J. P.; Jarvis, D. L. *Nature Biotechnology*. Nature Publishing Group May 5, 2005, pp 567–575.
(30) Dumas, A.; Lercher, L.; Spicer, C. D.; Davis, B. G. *Chemical Science*. Royal Society of Chemistry January 1, 2015, pp 50–69.
(31) Italia, J. S.; Addy, P. S.; Wrobel, C. J. J.; Crawford, L. A.; Lajoie, M. J.; Zheng, Y.; Chatterjee, A. *Nat. Chem. Biol.* **2017**, *13* (4), 446.
(32) Kelemen, Rachel E.; Jewel, Delilah; Pasha, Muhammad; Cao, Xiaofu; Chatterjee, A. *Submiss.*
(33) Gautier, A.; Nguyen, D. P.; Lusic, H.; An, W.; Deiters, A.; Chin, J. W. *J. Am. Chem. Soc.* **2010**, *132* (12), 4086.

Chapter 5

A virus-based two-hybrid selection scheme to probe the mammalian interactome

5.1 Introduction

The prior chapters of this thesis focused on enhancing selection schemes for the directed evolution of tRNAs and optimizing viral vectors for genetic code expansion in mammalian cells. This chapter deals with using AAV-based selection schemes to develop a high-throughput two-hybrid assay for detecting binding partners in the mammalian proteome.

5.1.1 Two-hybrid selection

Many cellular processes are controlled by protein-protein interactions (PPIs). Typical cellular functions, such as growth and differentiation, or signal transduction, rely on PPIs.[1,2] Parasites and pathogens also manipulate PPIs to gain access to cells and replicate.[3] Understanding the interactome in mammalian cells can help researchers to better understand our physiology and how to treat diseases when the underlying causes involve interactions between proteins.

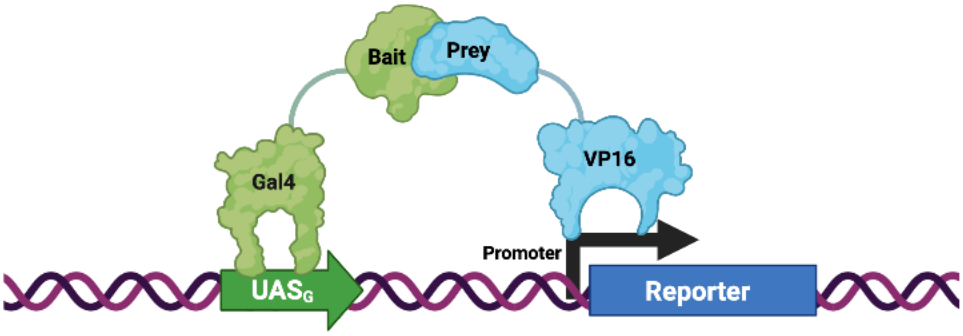


Figure 5.1. Overview of a two-hybrid assay. In a two-hybrid system, a DNA binding domain (typically Gal4) fused to a bait protein binds to the UAS_G region of the GAL1 promoter. The bait protein associates with the prey protein, which is fused to the VP16 transcription activator. VP16 binds to the promoter of the reporter gene, which is then transcribed. (Image designed using BioRender.)

In a two-hybrid experiment, cells are transformed with plasmids containing a DNA binding domain fused to a bait protein of interest, a transcription activator fused to a prey protein (or one member of a library of proteins), and a reporter gene that cannot be

expressed unless the DNA binding domain binds to an area upstream of the gene *and* the transcription activator is recruited. The DNA binding domain will bind the region upstream of the reporter gene, but the transcription activator will not be recruited unless the bait and prey proteins physically associate. Transcription of the reporter gene is thus tied to the interaction of the bait and prey proteins (Figure 5.1). Typically, the DNA binding domain used is the Gal4 DNA binding domain (the N-terminus of the GAL4 protein, which binds to the UAS_G in the GAL1 promoter).[4] The *Herpesvirus* protein VP16 is a popular transcription activator.[5]

5.1.2 Yeast two-hybrid systems

The most commonly used two-hybrid system is the yeast two-hybrid system.[6,7] Yeast cells exhibit a number of advantages over bacteria and mammalian cells for use in two-hybrid experiments,[5] although two-hybrid protocols exist for each of those cell types.[8–10] Transformation efficiency in yeast cells is good, and since yeast cells amplify and maintain plasmids, recovering plasmids from them after a screen or a selection is simple. Two-hybrid screens can be performed easily in yeast cells grown on X-gal with LacZ as the reporter gene, turning colonies blue if the bait and prey proteins associate. Selections are also straightforward using this system: growing an auxotrophic strain on media lacking a particular essential amino acid like histidine and using a gene such as *HIS3* as the reporter protein will only result in cell survival and colony growth if the bait and prey proteins interact. Since *Saccharomyces cerevisiae* is a eukaryote, mammalian bait and prey proteins are more likely to be translated and folded properly than if the selection process took place in bacteria. These features of the yeast two-hybrid system are attractive to researchers investigating the eukaryotic interactome.

While working with yeast is advantageous in many respects, yeast two-hybrid systems also present problems when working with the mammalian interactome. Even though yeast are eukaryotes, they are evolutionarily divergent from mammals, and many mammalian proteins will not fold correctly in yeast due to lack of appropriate chaperones. Additionally, certain PPIs may depend on post-translational modifications or other interactome-dependent changes that are not possible when the proteins are removed from their native context. Performing two-hybrid experiments directly in mammalian cells would resolve these challenges.

5.1.3. Mammalian two-hybrid systems

Mammalian two-hybrid systems exist that allow detection of protein binding partners in their native cellular contexts. In these systems, three plasmids are typically used to deliver the necessary genetic cargo to the cells for the assay: the Gal4-bait protein fusion, the VP16-prey protein fusion, and the reporter gene downstream of the UAS_G region of the GAL1 promoter.[10] This system is limited by the efficiency of transient transfection, as not all cells will receive all three plasmids necessary to undergo the two-hybrid assay. One solution could be a two-plasmid system in which the reporter gene and the bait/Gal4 binding domain are on the same plasmid and the VP16/prey fusion are on a different plasmid. Unfortunately, it is not possible to control DNA delivery in transient transfection across a large population in a single dish, so this method renders impossible the delivery of library of prey proteins so that each cell only receives one library member. Without the ability to deliver a single prey library member per cell, it is not possible to scan the proteome for binding partners in a two-hybrid assay.

A newer method using microarrays addresses this problem of extremely low-throughput pairwise selections. Cell array protein-protein interaction assay (CAPPIA) makes use of robotically printed microarrays and an autofluorescent reporter to deliver DNA individually to cells deposited on a slide.[11] This system permits testing of a library of prey proteins using transient transfection, which would be impossible in a larger plate where the researcher cannot control the delivery of individual plasmids. While CAPPIA addresses this problem, its implementation is also limited. While the authors of this method claim that CAPPIA is high throughput, they only tested a library of 160 prey proteins. That library size may be an improvement upon what an individual researcher can test manually, but under 200 proteins does not approach the size and complexity of the mammalian proteome that a researcher may want to screen. To have a truly high throughput library, it will be necessary to test several orders of magnitude more proteins that the CAPPIA system has been used for thus far. Additionally, performing CAPPIA requires the use of a robot, which not every lab has the means to purchase. A mammalian two-hybrid system should be as high-throughput and accessible as its yeast counterparts.

5.1.4. Developing high throughput two-hybrid systems for mammalian cells

In order to develop a truly high throughput mammalian two-hybrid system, it will be necessary to deliver the library of prey proteins to the host cells in a manner that ensures that each host cell receives only one library member. The delivery technique that works best for mammalian cells that meets that criterion is viral infection. As previously discussed, recombinant viruses are easy to produce and titer, which allows researchers to exert exquisite control over the copy number of viruses delivered to their host cells. Doing

so allows a researcher to perform a selection or screen in a large plate of cells, much like a bacterial or yeast selection, without the need for expensive robotics.

Recombinant AAV possesses a number of features that make it amenable to being used for two-hybrid selections in mammalian cells. It is possible to make VP2/VP3-only or VP3-only AAV capsids, but VP1 is required to produce infective virus.[12,13] It is also possible to knock out VP1 expression from the AAV2 genome and place transcription of VP1 under control of a separate promoter, which is a technique that has been used successfully to install UAAs into only VP1 and not VP2 or VP3 (unpublished data). Separating the genes for the proteins in this way can allow researchers to place VP1 transcription, and hence the production of infectious virus, under the control of a promoter in a two-hybrid system. VP1 and the packaging of infectious virus becomes the reporter for a prey library member binding to the bait protein. While AAV is a small virus with a limited cargo capacity,[14] the large majority of the human proteome is less than 1300 amino acids in length,[15] which is the upper limit for the size of a prey protein that can be fused to VP16 and have the gene for this fusion fit within the AAV capsid. Infectious virus will contain packaged DNA, which will include the gene for the prey protein whose association with the bait protein prompted transcription of VP1. The genetic cargo of these reporter viruses can be characterized to determine which library members are hits. rAAV can potentially be both an excellent delivery vector and a reporter for a high throughput mammalian two-hybrid system.

Screens such as CIPPIA and selections such as metabolics-based yeast two-hybrid systems are certainly useful for identifying protein binding partners, but this goal is not the only one that can be achieved with a selection system. Other selection schemes have

evolved promoters in bacteria and yeast to improve production of a protein of interest.[16–19] If a consistent bait and prey protein pair can be maintained, the sequences of promoters can be randomized and evolved to maximize reporter gene transcription.

Here, I present a two-hybrid selection scheme for use in mammalian cells. I demonstrate that existing two-hybrid assays using secreted alkaline phosphatase (SEAP) can be modified to use a fluorescent reporter. I then demonstrate that it is possible to place VP1 under control of the two-hybrid system such that production of infectious virus becomes the reporter. This system will allow researchers to develop a high throughput selection using AAV as both a vector to deliver library members and a reporter from which the sequence of a hit can be retrieved. I also demonstrate that this AAV-based two-hybrid system produces different quantities of virus depending on the strength of the promoter upstream of VP1. This AAV two-hybrid system can also therefore be used in the future to evolve minimal promoters for general synthetic biology uses.

5.2 Results and Discussion

5.2.1 Producing GFP with the use of two-hybrid binding partners

To test the generalized applicability of Matchmaker Mammalian Assay (Takara) protocol, I determined whether I could use a two-hybrid assay to express a reporter protein of interest (in this case, SEAP and eGFP). The Gal4 binding domain binds to GAL1 binding sites upstream of DNA to be transcribed, and the herpesvirus protein VP16 AD interacts with host transcription factors to induce transcription. With the Matchmaker assay, a fusion of the DNA binding domain GAL4 and the mouse p53 protein can bind to the fusion of the herpesvirus protein VP16 AD and the SV40 large T-antigen through the association of p53 and the T-antigen. To provide a control, VP16 AD is fused to the polyoma virus coat

protein (CP), which does not bind to p53. In addition to this test, in which GAL4 and VP16 must associate in the cell, I also provided a GAL4-VP16 fusion protein to determine the maximum level at which eGFP can be expressed using this system (when these proteins are constantly bound).

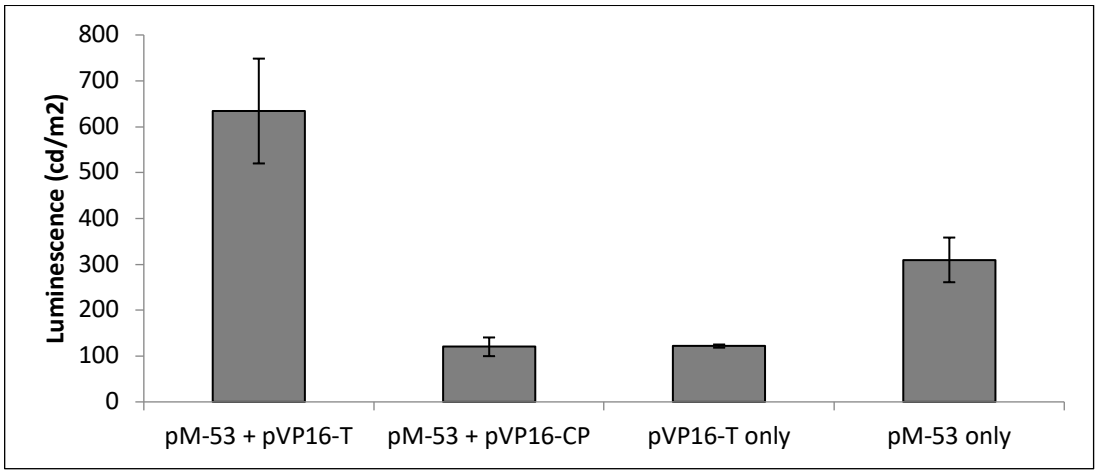
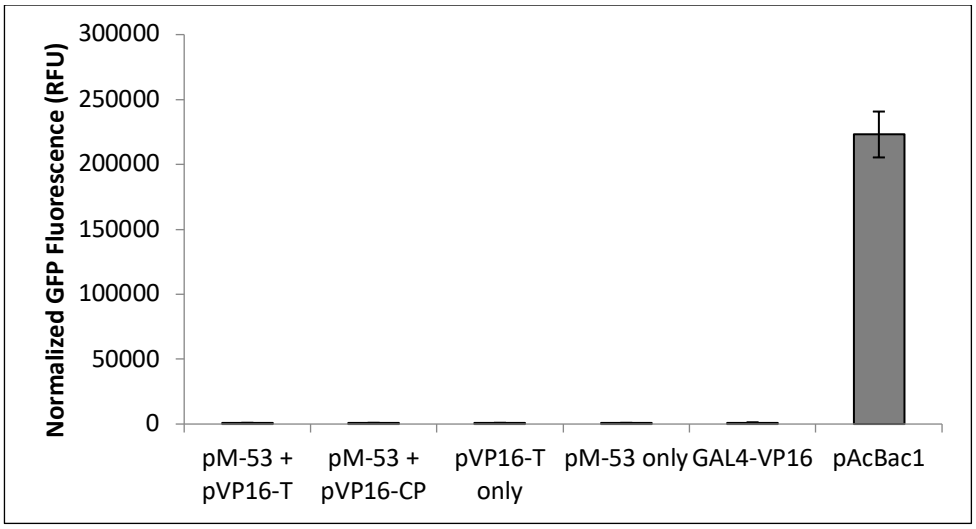
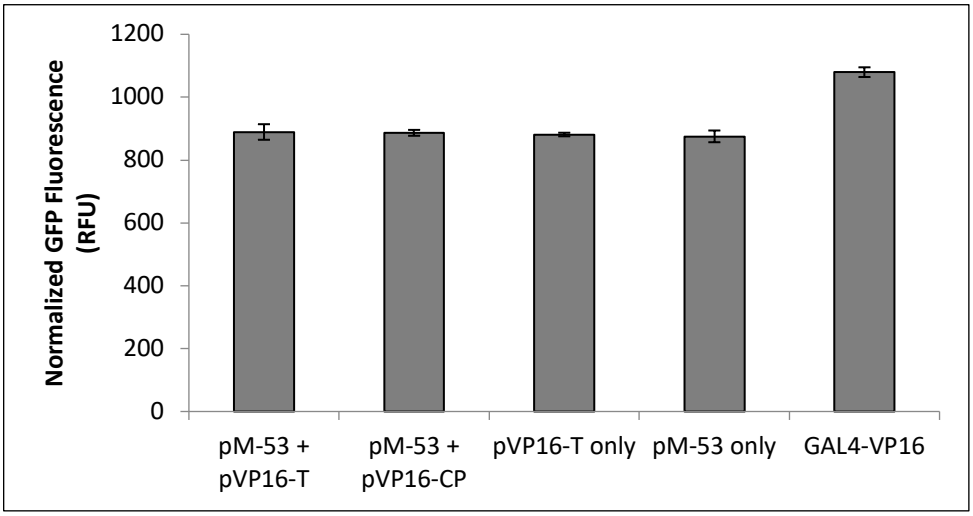


Figure 5.2. Luminescence of cells testing the two-hybrid Mammalian Matchmaker Assay. HEK-293T cells were transfected with plasmids to test this two-hybrid assay. After 72 hours, the SEAP detection assay was performed with CSPD as the substrate.

With the luminescence assay, cells produced much more SEAP when they received pM-53 and pVP16-T than when they received pM-53 and pVP16-CP or either pM-53 or pVP16-T alone (Figure 5.2). This result demonstrates that the two-hybrid system functions well in HEK-293T cells. While this system works well, luminescence assays are not as facile to perform as are fluorescence measurements, and so I replaced SEAP with GFP in the reporter plasmid to determine whether this system worked equally well for this reporter protein (Figure 5.3).



(A)



(B)

Figure 5.3. GFP fluorescence of cells using the two-hybrid system. SEAP was replaced with GFP in the Mammalian Matchmaker Assay, and this system was tested alongside a Gal4-VP16 fusion, with a constitutively expressed GFP as a control.

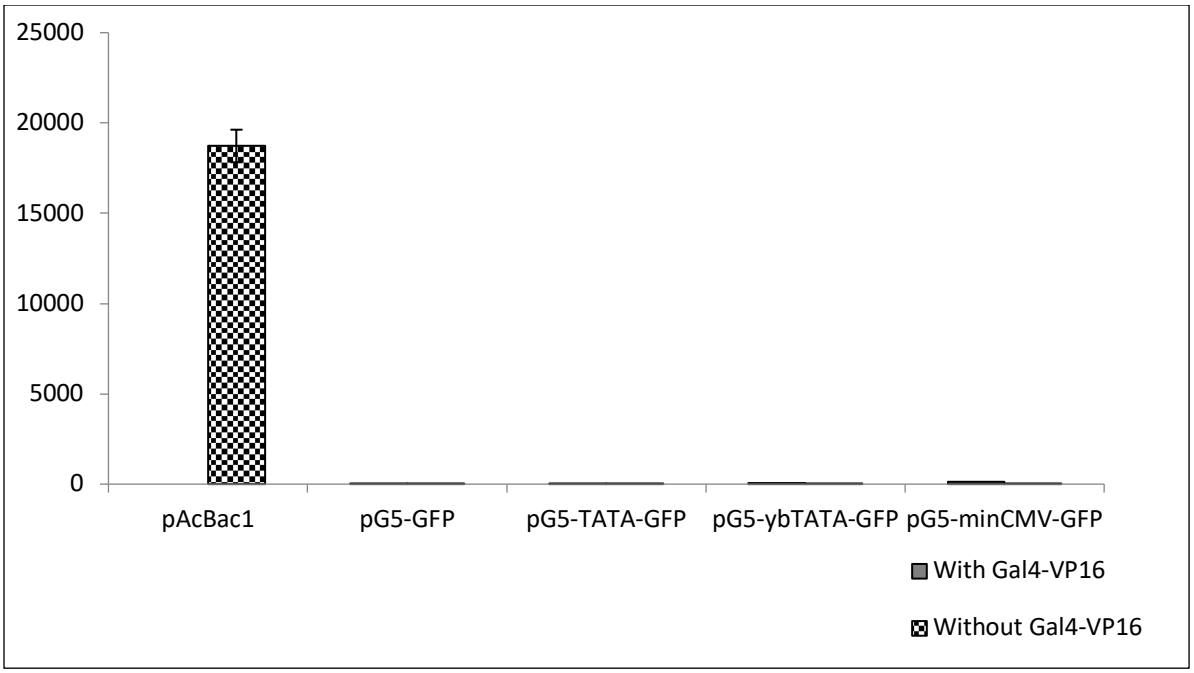
The vector pAcBac1-GFPwt acted as a control here, as transfection of this vector in mammalian cells results in high levels of expression of eGFP (Figure 5.3a). Transfection with the two-hybrid system of plasmids resulted in much lower levels of GFP production, as expected. Unlike the luminescence assay, there was not a large difference between the levels of GFP produced did not differ between treatments with pM-53 and pVP16-T, pM-53 and pVP16-CP, pM-53 alone, and pVP16-T alone (Figure 5.3b). However, the treatment

with the GAL4-VP16 fusion construct resulted in higher levels of GFP production, indicating that it is possible to differentiate between levels of association between the two-hybrid proteins based on this assay.

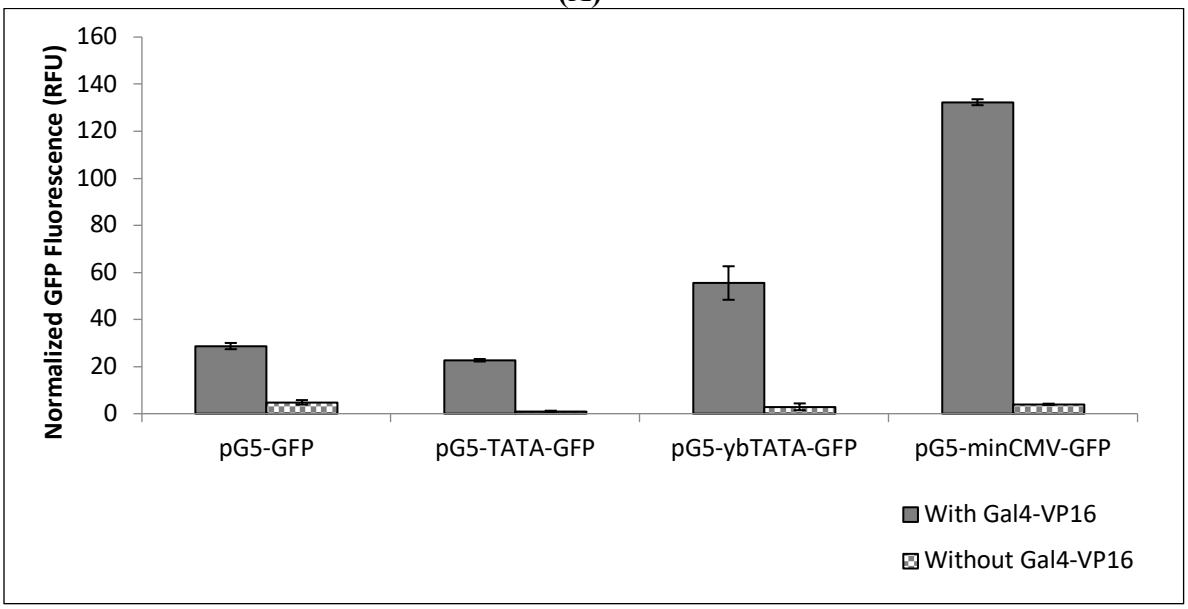
5.2.2 Testing the effects of minimal promoters on GFP production with a two-hybrid system

The original pG5 vector controls the transcription of the reporter protein (originally SEAP, which I swapped with GFP for convenience) with 5 consecutive GAL4 DNA binding domains and a Kozak sequence. To improve production of the reporter protein, I added different minimal promoters upstream of the Kozak sequence with the intention of improving transcription of the reporter protein gene. I tested 3 different minimal promoters: a TATA box, a YB-TATA box, and a minimal CMV synthetic promoter (see the appendix for promoter sequences). For these experiments, I used the GAL4-VP16 fusion construct to maximize transcription while testing which minimal promoter performed best.

In these experiments, the vector pAcBac1 acted as a control. Cells transfected with this plasmid produced much more GFP, which is expected because GFP is expressed constitutively from the CMV promoter in that vector. The YB-TATA and minCMV promoters performed better than the original construct, which lacked any promoter, as well as the TATA promoter (Figure 5.4). There was 4.6x as much fluorescence from cells which received GFP under the control of the minCMV promoter as from cells which received the original construct. Importantly, there was negligible fluorescence in the absence of the GAL4-VP16 fusion construct from all of the pG5-GFP plasmids tested. This system can be reliably used for two-hybrid selections, as even a small amount of leakiness from the promoter can cause false positives in a screen for protein binding partners.



(A)



(B)

Figure 5.4. Comparison of GFP fluorescence using minimal promoters. HEK-293T cells were transfected with Gal4-VP16 and GFP expressed from different minimal promoters. After 72 hours, cells were imaged and lysed, and GFP fluorescence was measured. (A) Including fluorescence from cells transfected with pAcBac1. (B) Only including fluorescence from cells using the two-hybrid system.

5.2.3 Producing rAAV dependent on two-hybrid selection

In order to create a system in which AAV can be used in a screen for mammalian protein binding partners, it is necessary for the production of AAV to be dependent on the association of Gal4 and VP16. To produce such a system, I used a construct in which VP1 is deleted from the genome of AAV and re-introduced under the control of the CMV promoter. In this case, I replaced GFP with VP1 in the pG5-minCV plasmid, so that transcription of VP1 was under the control of the five Gal4 binding domains and the synthetic minimal CMV promoter. In order for infectious AAV to be produced, the Gal4-VP16 construct must be present promote transcription of VP1, which is required for AAV to be infectious.

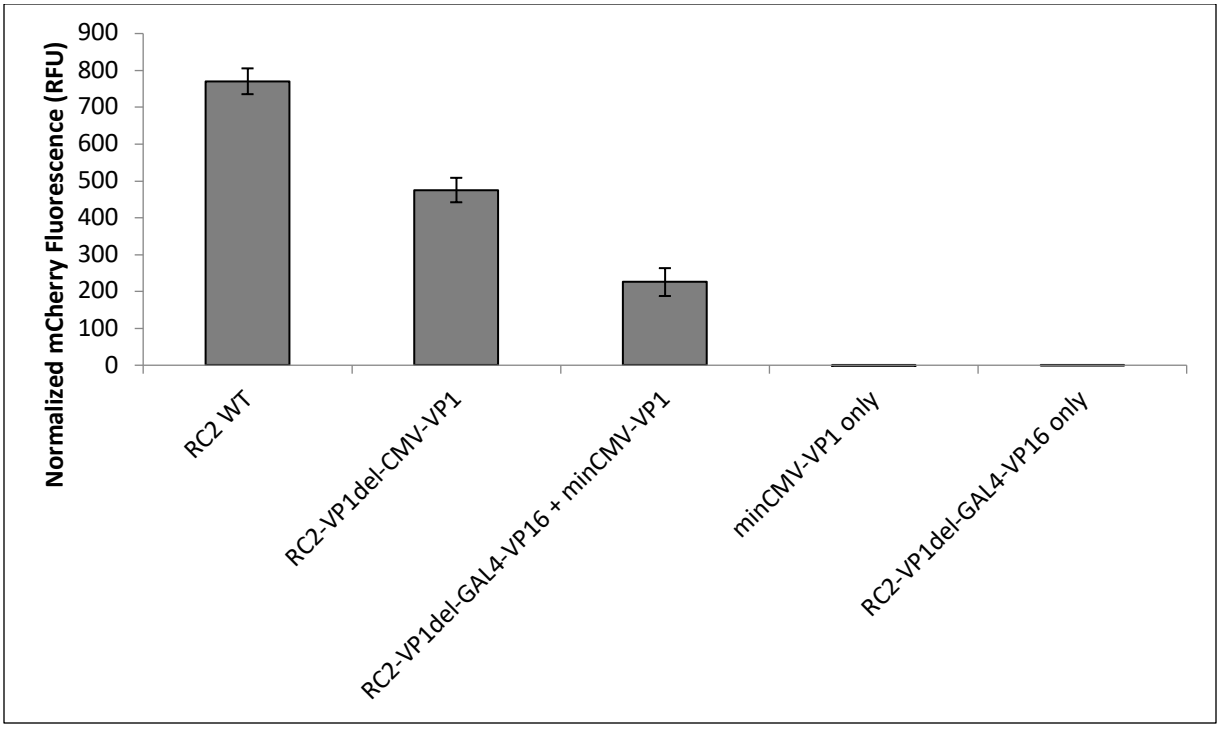


Figure 5.5. mCherry fluorescence of cells infected with rAAV-containing lysate. Cells were infected with rAAV produced by packaging cells to test the production of infectious virus when VP1 expression is controlled by Gal4-VP16. After 48 hours, these cells were imaged and lysed, and mCherry fluorescence was taken.

Production of AAV with the construct in which VP1 is deleted from the genome and then reintroduced on a CMV promoter (RC2-VP1del-CMV-VP1) was 61.7% of that produced from the wild type AAV2 genome (RC2wt), as measured by mCherry fluorescence of cells infected with lysate from HEK-293T cells packaging virus (Figure 5.5). When VP1 is provided on the pG5-minCMV-VP1 plasmid, and the rest of the AAV genome (RC2-VP1del) and GAL4-VP16 are provided on a different plasmid, mCherry fluorescence is 47.5% that from the RC2-VP1del-CMV-VP1 construct. This result indicates that a substantial quantity of output virus can be produced from this system.

There was negligible virus produced when only minCMV-VP1 was provided, as expected, since this treatment does not provide Rep, which is necessary for AAV DNA replication, or VP2 and VP3, which are necessary to produce the majority of the AAV capsid. Importantly, negligible infectious virus was produced when minCMV-VP1 was not provided, indicating that the production of VP1 from the minCMV promoter is indeed essential to produce infectious AAV and that the output of this virus can be used to determine whether GAL4 and VP16 are associated (Figure 5.6).

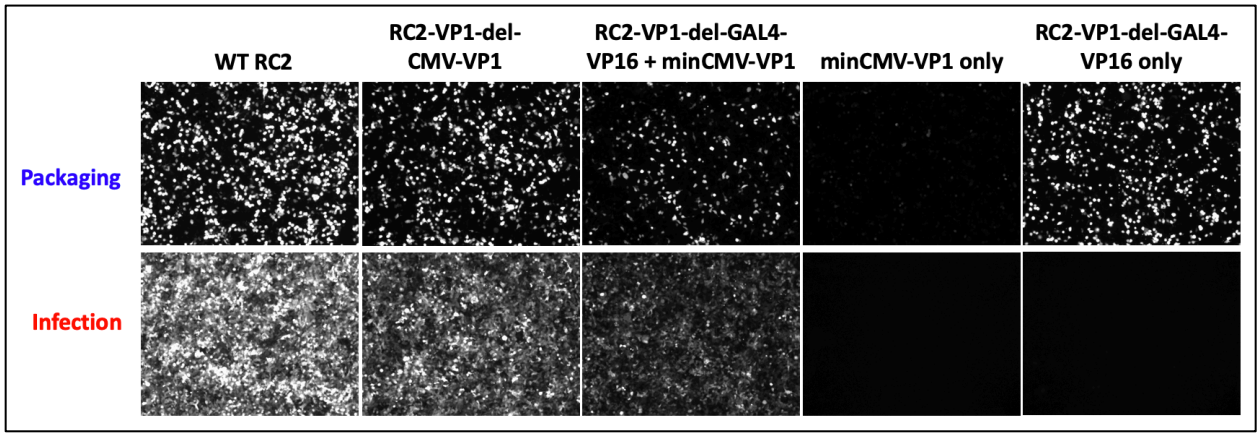


Figure 5.6. Cells packaging rAAV and cells infected with rAAV-containing lysate. Cells were infected with rAAV-mCherry and transfected plasmids to test the conditionality VP1 production dependent on Gal4-VP16 binding to the minCMV synthetic minimal

promoter. After 72 hours, the packaging cells were imaged and lysed, and new cells were infected with the lysate. These cells were imaged after 48 hours.

5.3 Conclusion

It is difficult to perform two-hybrid selections involving libraries of proteins in mammalian cells due to the technical challenge of being able to provide only one library member per cell. This virus-based system can circumvent that challenge and fill a gap in the available technology to enable high-throughput two-hybrid selections. Putting an essential viral protein, VP1, under the control of the GAL1 promoter enables infectious AAV to act as both a DNA delivery vector and the two-hybrid reporter. The platform presented here will be a valuable new application of directed evolution technology.

5.3.1 Future directions

This two-hybrid system, using VP1 as the output protein for the production of infectious AAV, has the potential to be a highly responsive selection system. One could provide a construct in which a bait protein is fused to GAL4 to cells via transient transfection, and also infect the cells with rAAV in which the genome is a library of proteins fused to VP16. With the two-hybrid selection system in which production of VP1 is tied to the association of GAL4 and VP16, infectious output virus would only be produced if the protein contained in the rAAV genome could bind to the bait protein. This system would allow researchers to scan the human proteome for binding partners of particular bait proteins of interest.

While transcription of VP1 from the minCMV promoter led to the production of a respectable quantity of AAV, the fluorescence output (and therefore likely infectious AAV production) was less than half that which was produced when VP1 was under the control of the larger CMV promoter. The difference between the activities of the typical CMV and

the minimal promoters was even more stark with GFP expressed constitutively or from the two-hybrid system. There appears to be a dynamic range in the activity of these promoters sufficient for the evolution of more active minimal promoters. The two-hybrid system, using infectious AAV as the reporter, would allow for this selection to be done with relative ease.

5.4. Materials and Methods

5.4.1 Strains and cell lines

Plasmid propagations were performed as described in Chapter 2 of this thesis. HEK-293T cells were maintained as described in Chapter 2.

5.4.2. Plasmids

All plasmid maps and sequences can be found in the Appendix.

pHelper, pAAV-mCherry+1xMmPytR, pIDTSMART-MbPylRS-RC2wt, pIDTSMART-MbPylRS-RC2-VP1del, and pIDTSMART-RC2-VP1del-CMV-VP1wt were propagated from Dr. Rachel Kelemen's stocks. pAcBac1-GFPwt was propagated from Dr. Yunan Zheng's stocks.

pG5-SEAP, pM53, pVP16-T, and pVP16-CP were part of the Matchmaker Mammalian Assay kit (Takara).

pECE-Gal4-VP16 was accessed from Addgene (plasmid #71728).

To produce pG5-GFP, eGFP was amplified from pAcBac1-GFPwt using the terminal primers XbaI-eGFP-R and SphI-Kozak-GFP-F, which also appended a Kozak sequence to the 5' end of the gene for GFP. pG5-SEAP was digested with XbaI and SphI, and GFP was cloned into the plasmid using these cut sites. The TATA sequence was appended to the 5' end of the GFP construct using the primer TATA-GFP-F. The YB

TATA sequence was added to the 5' end of the GFP construct using YB-TATA-GFP-F, and the minimal CMV promoter was added to the GFP construct using min-CMV-GFP-F. Each of these constructs was amplified using the reverse primer KpnI-GFP-R because pG5-GFP cannot be cut with XbaI due to DNA methylation that was not present in the context of pG5-SEAP. pG5-GFP was cut with KpnI and SphI, and the GFP constructs were cloned in to produce pG5-TATA-GFP, pG5-YBTATA-GFP, and pG5-minCMV-GFP.

The VP1-only construct was amplified out of pIDTSMART-RC2-VP1del-CMV-VP1wt using the terminal primers SphI-VP1-only-F and XbaI-VP1-only-R. The TATA, YB TATA, and minCMV sequences were appended to the 5' end of the VP1 construct using the respective forward primers SphI-TATA-VP1-only-F, SphI-YBTATA-VP1-only-F, and SphI-minCMV-VP1-only-F. These constructs were cloned into the pG5 backbone by digesting pG5-SEAP with SphI and XbaI and ligating in the digested VP1 constructs to produce pG5-VP1, pG5-TATA-VP1, pG5-YBTATA-VP1, and pG5-minCMV-VP1.

The Gal4-VP16 construct was amplified out of pECE-Gal4-VP16 using the terminal primers SbfI-Gal4-VP16-F and AvrII-Gal4-VP16-R. pIDTSMART-MbPylRS-RC2-VP1del was digested with SbfI and AvrII, and the Gal4-VP16 construct was cloned in to produce pIDTSMART-RC2-VP1del-Gal4-VP16.

5.4.3 Preparing recombinant AAV

Recombinant AAV was produced as described in detail in Chapter 2 of this thesis.

5.4.4 Producing GFP with the use of two-hybrid binding partners

The following plasmids were transfected into HEK-293T cells to test the Matchmaker luminescence assay: pG5-SEAP with pM-53 and pVP-16T, pG5-SEAP with pM-53 and pVP-16CP, pG5-SEAP with pM-53 only, and pG5-SEAP with pVP16-T only.

Cells were lysed with CelLytic™ M (Sigma-Aldrich) with 0.00001% Pierce Universal Nuclease. Lysate luminescence was measured using the GreatEscAPe™ SEAP Chemiluminescence Detection Kit with CSPD as the substrate.

To determine whether it was possible to drive production of a protein of interest by relying on two-hybrid selection, the following plasmids were transfected into HEK-293T cells: pG5-GFP with pM-53 and pVP-16T, pG5-GFP with pM-53 and pVP-16CP, pG5-GFP with pM-53 only, pG5-GFP with pVP16-T only, pG5-GFP with pECE-GAL4-VP16, and pAcBac1-GFPwt as a control. Cells were lysed with CelLytic™ M (Sigma-Aldrich) with 0.00001% Pierce Universal Nuclease, and GFP fluorescence was measured.

5.4.5 Testing the effects of different minimal promoters on GFP production with a two-hybrid system

The following plasmids were transfected into HEK-293T cells to test whether certain minimal promoters could enhance the production of GFP while still maintaining dependence on this two-hybrid system: pG5-GFP alone, pG5-GFP with pECE-GAL4-VP16, pG5-TATA-GFP alone, pG5-TATA-GFP with pECE-GAL4-VP16, pG5-YBTATA-GFP alone, pG5-YBTATA-GFP with pECE-GAL4-VP16, pG5-minCMV-GFP alone, pG5-minCMV-GFP with pECE-GAL4-VP16, and pAcBac1-GFPwt as a control. Cells were lysed with CelLytic™ M (Sigma-Aldrich) with 0.00001% Pierce Universal Nuclease, and GFP fluorescence was measured.

5.4.6 Producing rAAV dependent on two-hybrid selection

To determine whether it is possible to tie AAV production to this two-hybrid system, these plasmids were transfected into HEK-293T cells, along with pHelper: pIDTSMART-MbPylRS-RC2wt only, pIDTSMART-RC2-VP1del-CMV-VP1 only,

pIDTSMART-RC2-VP1del-GAL4-VP16 and pG5-minCMV-VP1, pG5-minCMV-VP1 only, and pIDTSMART-RC2-VP1del-GAL4-VP16 only. rAAV-mCherry-PytR was delivered via infection at 5 MOI, 4 hours before transfection. Cells were lysed with CelLytic™ M (Sigma-Aldrich) with 0.00001% Pierce Universal Nuclease. The resultant lysate was flash-frozen and stored at -80°C. Lysate was thawed and used to infect new HEK-293T cells with sodium butyrate to 1.0 mM. After 48 hours, cells were imaged and lysed, and mCherry fluorescence was measured.

5.5 References

(1) Loregian, A.; Palù, G. *J. Cell. Physiol.* **2005**, *204* (3), 750.
(2) Arkin, M. R.; Whitty, A. *Curr. Opin. Chem. Biol.* **2009**, *13* (3), 284.
(3) Loregian, A.; Marsden, H. S.; Palù, G. *Rev. Med. Virol.* **2002**, *12* (4), 239.
(4) Keegan, L.; Gill, G.; Ptashne, M. *Science (80-.).* **1986**, *231* (4739), 699.
(5) Fields, S.; Sternglanz, R. *Trends in Genetics*. Elsevier Current Trends August 1, 1994, pp 286–292.
(6) Fields, S.; Song, O. K. *Nature* **1989**, *340* (6230), 245.
(7) Fields, S. *Methods* **1993**, *5* (2), 116.
(8) Karimova, G.; Pidoux, J.; Ullmann, A.; Ladant, D. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (10), 5752.
(9) Finkel, T.; Duc, J.; Fearon, E. R.; Dang, C. V.; Tomaselli, G. F. *J. Biol. Chem.* **1993**, *268* (1), 5.
(10) Patrício, D.; Fardilha, M. *Drug Discov. Today* **2020**, *25* (4), 764.
(11) Fiebitz, A.; Nyarsik, L.; Haendler, B.; Hu, Y.-H.; Wagner, F.; Thamm, S.; Lehrach, H.; Janitz, M.; Vanhecke, D. *BMC Genomics* **2008**, *9*.
(12) Girod, A.; Wobus, C. E.; Zádori, Z.; Ried, M.; Leike, K.; Tijssen, P.; Kleinschmidt, J. A.; Hallek, M. *J. Gen. Virol.* **2002**, *83* (5), 973.
(13) Grieger, J. C.; Johnson, J. S.; Gurda-Whitaker, B.; Agbandje-McKenna, M.; Samulski, R. J. *J. Virol.* **2007**, *81* (15), 7833.
(14) Srivastava, A.; Lusby, E. W.; Berns, K. I. *J. Virol.* **1983**, *45* (2).
(15) Zhuang, Yonlong; Ma, Fei; Li-Ling, Jesse; Xu, Xiaofeng; Li, Y. *Mol. Biol. Evol.* **2003**, *20* (12), 1978.
(16) Yona, A. H.; Alm, E. J.; Gore, J. *Nat. Commun.* **2018**, *9* (1), 1.
(17) Yokobayashi, Y.; Weiss, R.; Arnold, F. H. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (26), 16587.
(18) Tyo, K. E. J.; Nevoigt, E.; Stephanopoulos, G. Academic Press, 2011; Vol. 497, pp 135–155.
(19) Alper, H.; Fischer, C.; Nevoigt, E.; Stephanopoulos, G. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (36), 12678.

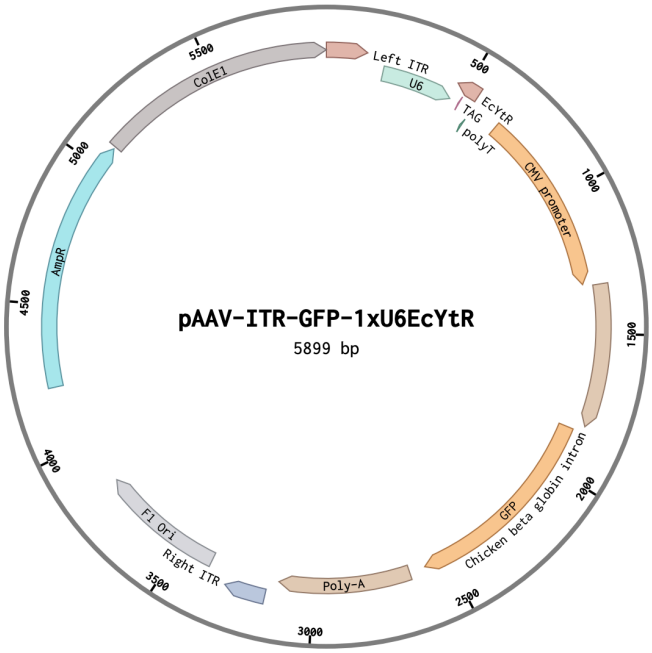
Appendix I

Plasmid maps and sequences

List of plasmid sequences

Chapter 2 Plasmids

pAAV-ITR-GFP+1xEcYtR



Sequence color-coding key

Feature	Color
Left ITR	text
U6 promoter	text
EcYtR	text
polyT	text
CMV promoter	text
Chicken beta globin intron	text
GFP	text
Poly-A	text
Right ITR	text
F1 Ori	text
Amp R	text
ColE1	text

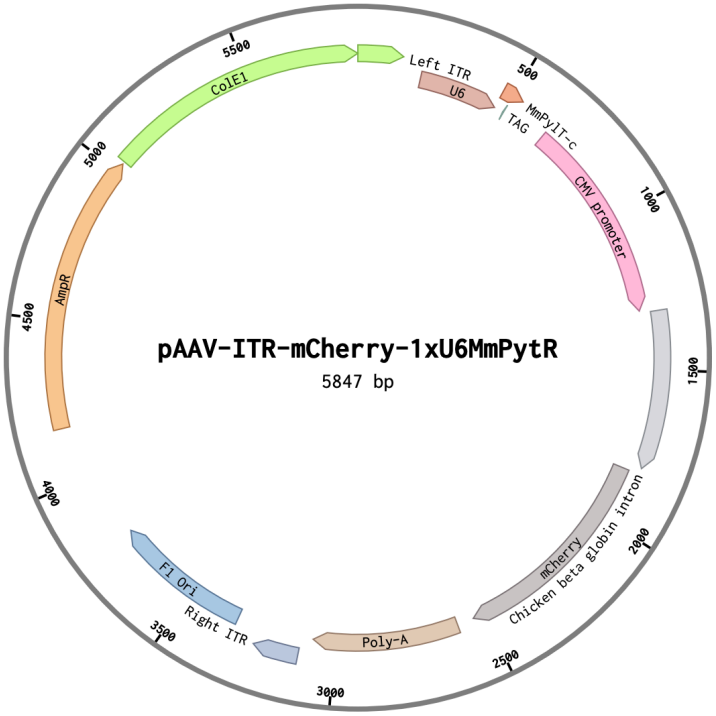
Sequence

cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgt
cgggcgacctttggtcgcccggcctcagtgagcgagcgagcgcgcagagagggagtggc
caactccatcactaggggttcctgcggccgcacgcgtctcgcggtccagtagtgatcga
cactgctcgatccgctcgcacccctaggtcgggcaggaagagggcctatttcccatgat

tccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttga
ctgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttggg
tagtttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttg
aaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacaccggtgggg
ttcccgagcggccaaagggagcagactctaaatctgccgtcacagacttcgaaggttcg
aatccttcccccaccatttttt
gctagcggatcgacgagagcagcgcgactggatctgt
cgcccgtctcaaacgcaaccctccggcggtcgcatatcattcaggacgagcctcagact
ccagcgtaac
acgcgtggagctagttattaatagtaatcaattacggggtcattagttc
atagcccatatatggagttccgcgttacataacttacggtaaatggcccgcctggctga
ccgcccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgtc
aatagggactttccattgacgtcaatgggtggagtatttacggtaaactgcccacttgg
cagtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaa
tggcccgcctggcattatgcccagtacatgaccttatgggactttcctacttggcagta
catctacgtattagtcatcgctattaccatggtgatgcggttttggcagtacatcaatg
ggcgtggatagcggtttgactcacggggatttccaagtctccaccccattgacgtcaat
gggagtttgttttgcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgcc
ccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcagagctcg
tttagtgaaccgtcagatcgcctggagacgccatccacgctgttttgacctccatagaa
gacaccgggaccgatccagcctcc
gcggatt
cgaatcccggccgggaacggtgcattgg
aacgcggattccccgtgccaagagtgacgtaagtaccgcctatagagtctataggccca
caaaaaatgctttcttctttaatatactttttgtttatcttatttctaatactttcc
ctaatctctttctttcagggcaataatgatacaatgtatcatgcctctttgcaccattc
taaagaataacagtgataatttctgggttaaggcaatagcaatatttctgcatataaat
atttctgcatataaattgtaactgatgtaagaggtttcatattgctaatagcagctaca
atccagctaccattctgcttttatttatggttgggataaggctggattattctgagtc
caagctaggccctttgctaatcatgttcatacctcttatcttcctcccacagctcctg
ggcaacgtgctggtctgtgtgctggcccatcactttggcaaagaattgggat
tcgaaca
tcgattgaattctga
atggtgagcaagggcgaggaggataacatggccatcatcaagga
gttcatgcgcttcaaggtgcacatggagggctccgtgaacggccacgagttcgagatcg
agggcgagggcgagggccgcccctacgagggcacccagaccgccaagctgaaggtgacc
aagggtggccccctgcccttcgcctgggacatcctgtcccctcagttcatgtacggctc
caaggcctacgtgaagcaccccgccgacatccccgactacttgaagctgtccttccccg
agggcttcaagtgggagcgcgtgatgaacttcgaggacggcggcgtggtgaccgtgacc
caggactcctccctgcaggacggcgagttcatctacaaggtgaagctgcgcggcaccaa
cttcccctccgacggccccgtaatgcagaagaagaccatgggctgggaggcctcctccg
agcggatgtaccccgaggacggcgccctgaagggcgagatcaagcagaggctgaagctg
aaggacggcggccactacgacgctgaggtcaagaccacctacaaggccaagaagcccgt
gcagctgcccggcgcctacaacgtcaacatcaagttggacatcacctcccacaacgagg
actacaccatcgtggaacagtacgaacgcgccgagggccgccactccaccggcggcatg
gacgagctgtacaag
ggatcctctagagtcgacctgcagaagcttgcctcgagcagcgc
tgctcgagagatct
acgggtggcatccctgtgacccctccccagtgcctctcctggccc
tggaagttgccactccagtgcccaccagccttgtcctaataaaattaagttgcatcatt
ttgtctgactaggtgtccttctataatattatggggtggaggggggtggtatggagcaa
ggggcaagttgggaagacaacctgtagggcctgcggggtctattgggaaccaagctgga
gtgcagtggcacaatcttggctcactgcaatctccgcctcctgggttcaagcgattctc
ctgcctcagcctcccgagttgttgggattccaggcatgcatgaccaggctcagctaatt
tttgtttttttggtagagacggggtttcaccatattggccaggctggtctccaactcct
aatctcaggtgatctacccaccttggcctcccaaattgctgggattacaggcgtgaacc

actgctcccttccctgtccttctgattttgtaggtaaccacgtgcggaccgagcggccg
caggaacccctagtgatggagttggccactccctctctgcgcgctcgctcgctcactga
ggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcg
agcgagcgcgcagctgcctgcagggcgcctgatgcggtattttctccttacgcatctg
tgcggtatttcacaccgcatacgtcaaagcaaccatagtacgcgccctgtagcggcgca
ttaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccct
agcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttcccc
gtcaagctctaaatcgggggctcccttagggttccgatttagtgctttacggcacctc
gaccccaaaaaacttgatttgggtgatggttcacgtagtgggccatcgccctgatagac
ggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgttccaaa
ctggaacaacactcaaccctatctcgggctattcttttgatttataagggattttgccg
atttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaa
caaaatattaacgtttacaatttatggtgcactctcagtacaatctgctctgatgccg
catagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgt
ctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtca
gaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctat
ttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcgg
ggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatcc
gctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatga
gtattcaacatttccgtgtcgcccttattccctttttgcggcattttgccttcctgtt
tttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacg
agtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccg
aagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcc
cgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgactt
ggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaat
tatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacg
atcggaggaccgaaggagctaaccgcttttttgcacaacatggggatcatgtaactcg
ccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacacca
cgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttact
ctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccact
tctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagc
gtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgta
gttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctga
gataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatac
tttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt
gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccc
cgtagaaaagatcaaaggatcttcttgagatccttttttttctgcgcgtaatctgctgct
tgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctacca
actcttttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttct
agtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcg
ctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccggg
ttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttc
gtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtg
agctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagc
ggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatct
ttatagtcctgtcgggtttcgccacctctgacttgagcgtcgattttgtgatgctcgt
caggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggcc
ttttgctggccttttgctcacatgt

pAAV- mCherry+1xMmPytR



Sequence color-coding key

Feature	Color
Left ITR	text
U6 promoter	text
MmPytR	text
polyT	text
CMV promoter	text
Chicken beta globin intron	text
mCherry	text
Poly-A	text
Right ITR	text
F1 Ori	text
Amp R	text
ColE1	text

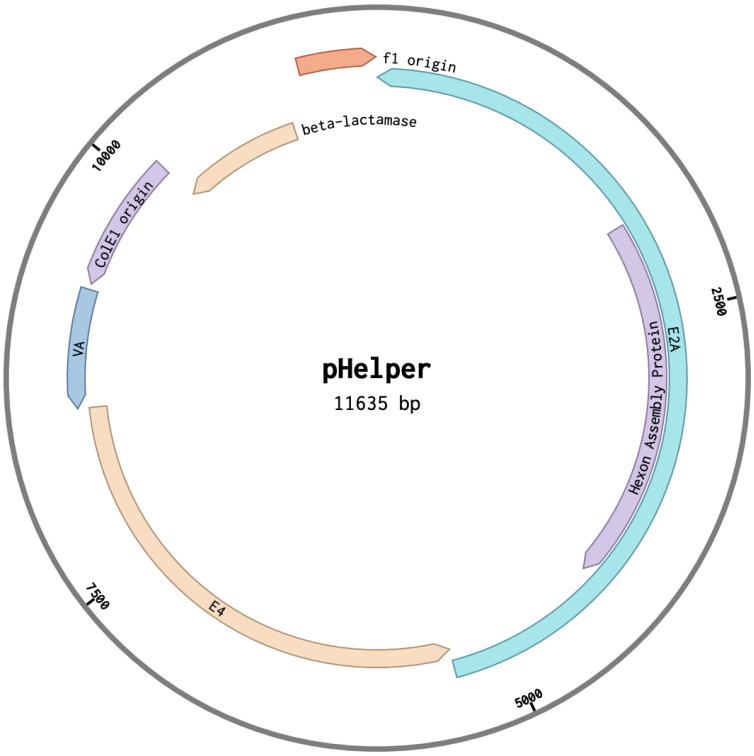
Sequence

cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgt
cgggcgacctttggtcgcccggcctcagtgagcgagcgagcgcgcagagagggagtggc
caactccatcactaggggttcctgcggccgcacgcgtctcgcggtccagtagtgatcga
cactgctcgatccgctcgcacccctaggtcgggcaggaagagggcctatttcccatgat
tccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttga
ctgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttggg

tagtttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttg
aaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacaccggaaacc
tgatcatgtagatcgaacggactctaaatccgttcagccgggttagattcccggggttt
ccgtttttt gctagcggatcgacgagagcagcgcgactggatctgtcgcccgtctcaaa
cgcaaccctccggcggtcgcatatcattcaggacgagcctcagactccagcgtaacacg
cgtggagctagttattaatagtaatcaattacggggtcattagttcatagcccatatat
ggagttccgcgttacataacttacggtaaatggcccgcctggctgaccgcccaacgacc
cccgcccattgacgtcaataatgacgtatgttcccatagtaacgtcaatagggactttc
cattgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtacatcaagt
gtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggcccgcctggc
attatgcccagtacatgaccttatgggactttcctacttggcagtacatctacgtatta
gtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggatagcg
gtttgactcacggggatttccaagtctccaccccattgacgtcaatgggagtttgtttt
gcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgccccattgacgcaaa
tgggcggtaggcgtgtacggtgggaggtctatataagcagagctcgtttagtgaaccgt
cagatcgcctggagacgccatccacgctgttttgacctccatagaagacaccgggaccg
atccagcctccgcggattcgaatcccggccgggaacggtgcattggaacgcggattccc
cgtgccaagagtgacgtaagtaccgcctatagagtctataggcccacaaaaaatgcttt
cttcttttaatatacttttttgtttatcttatttctaatactttccctaatctctttct
ttcagggcaataatgatacaatgtatcatgcctctttgcaccattctaaagaataacag
tgataatttctgggttaaggcaatagcaatatttctgcatataaatatttctgcatata
aattgtaactgatgtaagaggtttcatattgctaatagcagctacaatccagctaccat
tctgcttttatttatggttgggataaggctggattattctgagtccaagctaggccct
ttgctaatcatgttcatacctcttatcttcctcccacagctcctgggcaacgtgctgg
tctgtgtgctggcccatcactttggcaaagaattgggattcgaacatcgattgaattct
gaatggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctg
gacggcgacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccac
ctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctggc
ccaccctcgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgaccac
atgaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcac
catcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcg
acaccctggtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatc
ctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaa
gcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcg
tgcagctcgccgaccactaccagcagaacaccccatcggcgacggccccgtgctgctg
cccgacaaccactacctgagcacccagtccgccctgagcaaagaccccaacgagaagcg
cgatcacatggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacg
agctgtacaagtactcagatctcgagctcaagtaggatcctctagagtcgacctgcag
aagcttgcctcgagcagcgctgctcgagagatctacgggtggcatccctgtgacccctc
cccagtgcctctcctggccctggaagttgccactccagtgcccaccagccttgtcctaa
taaaattaagttgcatcattttgtctgactaggtgtccttctataatattatggggtgg
aggggggtggtatggagcaaggggcaagttgggaagacaacctgtagggcctgcggggt
ctattgggaaccaagctggagtgcagtggcacaatcttggctcactgcaatctccgcct
cctgggttcaagcgattctcctgcctcagcctcccgagttgttgggattccaggcatgc
atgaccaggctcagctaatttttgtttttttggtagagacggggtttcaccatattggc
caggctggtctccaactcctaatctcaggtgatctacccaccttggcctcccaaattgc
tgggattacaggcgtgaaccactgctcccttccctgtccttctgatttgtaggtaacc
acgtgcggaccgagcggccgcaggaaccctagtgatggagttggccactccctctctg

cgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgc
ccgggcggcctcagtgagcgagcgagcgcgcagctgcctgcaggggcgcctgatgcggt
attttctccttacgcatctgtgcggtatttcacaccgcatacgtcaaagcaaccatagt
acgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgacc
gctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttcctttctcgc
cacgttcgccggctttccccgtcaagctctaaatcgggggctccctttagggttccgat
ttagtgctttacggcacctcgaccccaaaaaacttgatttgggtgatggttcacgtagt
gggccatcgccctgatagacggtttttcgccctttgacgttggagtccacgttctttaa
tagtggactcttgttccaaactggaacaacactcaaccctatctcgggctattcttttg
atttataagggattttgccgatttcggcctattggttaaaaaatgagctgatttaacaa
aaatttaacgcgaattttaacaaaatattaacgtttacaatttatggtgcactctcag
tacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctg
acgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtc
tccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaa
gggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttaga
cgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaa
atacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataata
ttgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttttg
cggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgct
gaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagat
ccttgagagttttcgccccgaagaacgtttccaatgatgagcacttttaaagttctgc
tatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcata
cactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacgga
tggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcgg
ccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaac
atgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccatacc
aaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactat
taactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcg
gataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctga
taaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatg
gtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaa
cgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcaga
ccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaagga
tctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcg
ttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttt
tctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtt
tgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcag
ataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgt
agcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcg
ataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcgg
tcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccga
actgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaagg
cggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttcca
gggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcg
tcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcgg
cctttttacggttcctggccttttgctggccttttgctcacatgt

pHelper



Sequence color-coding key

Feature	Color
E2A	text
E4	text
VA	text
ColE1 ori	text
Beta-lactamase	text
F1 ori	text

Sequence

ggtacccaactccatgcttaacagtccccaggtacagcccaccctgcgtcgcaaccagg
aacagctctacagcttcctggagcgccactcgccctacttccgcagccacagtgcgcag
attaggagcgccacttctttttgtcacttgaaaaacatgtaaaaataatgtactaggag
acactttcaataaaggcaaatgttttatttgtacactctcgggtgattatttacccccc
caccccttgccgtctgcgccgtttaaaaatcaaaggggttctgccgcgcatcgctatgcg
ccactggcagggacacgttgcgatactggtgtttagtgctccacttaaactcaggcaca
accatccgcggcagctcggtgaagttttcactccacaggctgcgcaccatcaccaacgc
gtttagcaggtcgggcgccgatatcttgaagtcgcagttggggcctccgccctgcgcgc
gcgagttgcgatacacagggttgcagcactggaacactatcagcgccgggtggtgcacg
ctggccagcacgctcttgtcggagatcagatccgcgtccaggtcctccgcgttgctcag

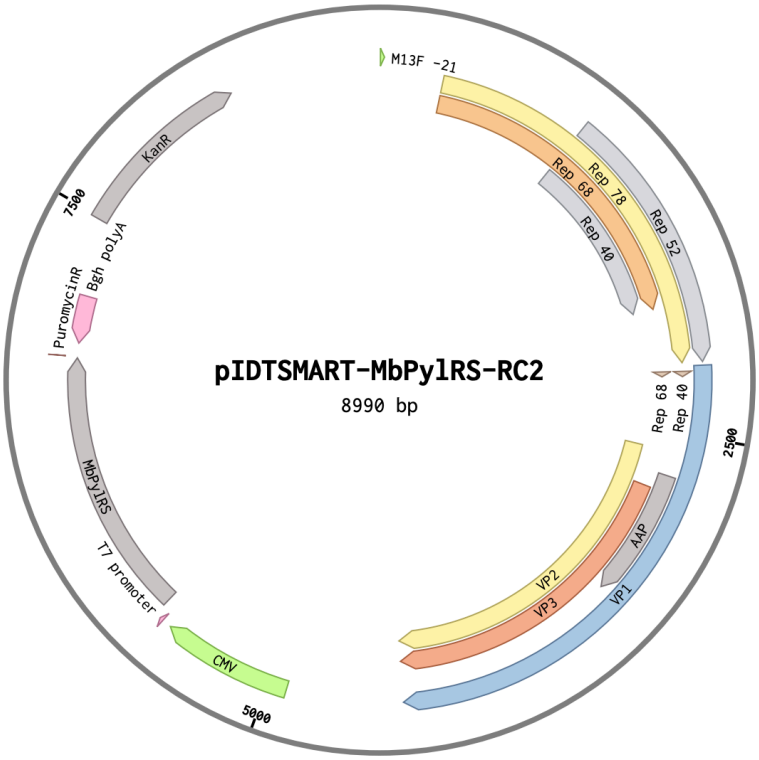
ggcgaacggagtcaactttggtagctgccttcccaaaaagggtgcatgcccaggctttg
agttgcactcgcaccgtagtggcatcagaaggtgaccgtgcccggtctgggcgttagga
tacagcgcctgcatgaaagccttgatctgcttaaaagccacctgagcctttgcgccttc
agagaagaacatgccgcaagacttgccggaaaactgattggccggacaggccgcgtcat
gcacgcagcaccttgcgtcggtgttggagatctgcaccacatttcggccccaccggttc
ttcacgatcttggccttgctagactgctccttcagcgcgcgctgcccgttttcgctcgt
cacatccatttcaatcacgtgctccttatttatcataatgctcccgtgtagacacttaa
gctcgccttcgatctcagcgcagcggtgcagccacaacgcgcagcccgtgggctcgtgg
tgcttgtaggttacctctgcaaacgactgcaggtacgcctgcaggaatcgccccatcat
cgtcacaaaggtcttgttgctggtgaaggtcagctgcaacccgcggtgctcctcgttta
gccaggtcttgcatacggccgccagagcttccacttggtcaggcagtagcttgaagttt
gcctttagatcgttatccacgtggtacttgtccatcaacgcgcgcgcagcctccatgcc
cttctcccacgcagacacgatcggcaggctcagcgggtttatcaccgtgctttcacttt
ccgcttcactggactcttccttttcctcttgcgtccgcataccccgcgccactgggtcg
tcttcattcagccgccgcaccgtgcgcttacctcccttgccgtgcttgattagcaccgg
tgggttgctgaaacccaccatttgtagcgccacatcttctctttcttcctcgctgtcca
cgatcacctctggggatggcgggcgctcgggcttgggagaggggcgcttctttttcttt
ttggacgcaatggccaaatccgccgtcgaggtcgatggccgcgggctgggtgtgcgcgg
caccagcgcatcttgtgacgagtcttcttcgtcctcggactcgagacgccgcctcagcc
gctttttgggggcgcgcggggaggcggcggcgacggcgacggggacgacacgtcctcc
atggttggtggacgtcgcgccgcaccgcgtccgcgctcgggggtggtttcgcgctgctc
ctcttcccgactggccatttccttctcctataggcagaaaaagatcatggagtcagtcg
agaaggaggacagcctaaccgccccctttgagttcgccaccaccgcctccaccgatgcc
gccaacgcgcctaccaccttccccgtcgaggcacccccgcttgaggaggaggaagtgat
tatcgagcaggacccaggttttgtaagcgaagacgacgaggatcgctcagtaccaacag
aggataaaaagcaagaccaggacgacgcagaggcaaacgaggaacaagtcgggcggggg
gaccaaaggcatggcgactacctagatgtgggagacgacgtgctgttgaagcatctgca
gcgccagtgcgccattatctgcgacgcgttgcaagagcgcagcgatgtgcccctcgcca
tagcggatgtcagccttgcctacgaacgccacctgttctcaccgcgcgtacccccaaa
cgccaagaaaacggcacatgcgagcccaacccgcgcctcaacttctaccccgtatttgc
cgtgccagaggtgcttgccacctatcacatctttttccaaaactgcaagataccctat
cctgccgtgccaaccgcagccgagcggacaagcagctggccttgcggcagggcgctgtc
atacctgatatcgcctcgctcgacgaagtgccaaaaatctttgagggtcttggacgcga
cgagaaacgcgcggcaaacgctctgcaacaagaaaacagcgaaaatgaaagtcactgtg
gagtgctggtggaacttgagggtgacaacgcgcgcctagccgtgctgaaacgcagcatc
gaggtcacccactttgcctacccggcacttaacctacccccaaggttatgagcacagt
catgagcgagctgatcgtgcgccgtgcacgacccctggagagggatgcaaacttgcaag
aacaaaccgaggagggcctacccgcagttggcgatgagcagctggcgcgctggcttgag
acgcgcgagcctgccgacttggaggagcgacgcaagctaatgatggccgcagtgcttgt
taccgtggagcttgagtgcatgcagcggttctttgctgacccggagatgcagcgcaagc
tagaggaaacgttgcactacacctttcgccagggctacgtgcgccaggcctgcaaaatt
tccaacgtggagctctgcaacctggtctcctaccttggaattttgcacgaaaaccgcct
cgggcaaaacgtgcttcattccacgctcaagggcgaggcgcgccgcgactacgtccgcg
actgcgtttacttatttctgtgctacacctggcaaacggccatgggcgtgtggcagcaa
tgcctggaggagcgcaacctaaaggagctgcagaagctgctaaagcaaaacttgaagga
cctatggacggccttcaacgagcgctccgtggccgcgcacctggcggacattatcttcc
ccgaacgcctgcttaaaaccctgcaacagggtctgccagacttcaccagtcaaagcatg
ttgcaaaactttaggaactttatcctagagcgttcaggaattctgcccgccacctgctg

tgcgcttcctagcgactttgtgcccattaagtaccgtgaatgccctccgccgctttggg
gtcactgctaccttctgcagctagccaactaccttgcctaccactccgacatcatggaa
gacgtgagcggtgacggcctactggagtgtcactgtcgctgcaacctatgcaccccgca
ccgctccctggtctgcaattcgcaactgcttagcgaaagtcaaattatcggtacctttg
agctgcagggtccctcgcctgacgaaaagtccgcggctccggggttgaaactcactccg
gggctgtggacgtcggcttaccttcgcaaatttgtacctgaggactaccacgcccacga
gattaggttctacgaagaccaatcccgcccgccaaatgcggagcttaccgcctgcgtca
ttacccagggccacatccttggccaattgcaagccatcaacaaagcccgccaagagttt
ctgctacgaaagggacgggggtttacctggacccccagtccggcgaggagctcaaccc
aatccccccgccgccgcagccctatcagcagccgcgggcccttgcttcccaggatggca
cccaaaaagaagctgcagctgccgccgccgccacccacggacgaggaggaatactggga
cagtcaggcagaggaggttttggacgaggaggaggagatgatggaagactgggacagcc
tagacgaagcttccgaggccgaagaggtgtcagacgaaacaccgtcaccctcggtcgca
ttcccctcgccggcgccccagaaattggcaaccgttcccagcatcgctacaacctccgc
tcctcaggcgccgccggcactgcctgttcgccgacccaaccgtagatgggacaccactg
gaaccagggccggtaagtctaagcagccgccgccgttagcccaagagcaacaacagcgc
caaggctaccgctcgtggcgcgggcacaagaacgccatagttgcttgcttgcaagactg
tgggggcaacatctccttcgcccgccgctttcttctctaccatcacggcgtggccttcc
cccgtaacatcctgcattactaccgtcatctctacagcccctactgcaccggcggcagc
ggcagcggcagcaacagcagcggtcacacagaagcaaaggcgaccggatagcaagactc
tgacaaagcccaagaaatccacagcggcggcagcagcaggaggaggagcgctgcgtctg
gcgcccaacgaacccgtatcgacccgcgagcttagaaataggatttttcccactctgta
tgctatatttcaacaaagcaggggccaagaacaagagctgaaaataaaaaacaggtctc
tgcgctccctcacccgcagctgcctgtatcacaaaagcgaagatcagcttcggcgcacg
ctggaagacgcggaggctctcttcagcaaatactgcgcgctgactcttaaggactagtt
tcgcgccctttctcaaatttaagcgcgaaaactacgtcatctccagcggccacacccgg
cgccagcacctgtcgtcagcgccattatgagcaaggaaattcccacgccctacatgtgg
agttaccagccacaaatgggacttgcggctggagctgcccaagactactcaacccgaat
aaactacatgagcgcgggaccccacatgatatcccgggtcaacggaatccgcgcccacc
gaaaccgaattctcctcgaacaggcggctattaccaccacacctcgtaataaccttaat
ccccgtagttggcccgctgccctggtgtaccaggaaagtcccgctcccaccactgtggt
acttcccagagacgcccaggccgaagttcagatgactaactcaggggcgcagcttgcgg
gcggctttcgtcacagggtgcggtcgcccgggcgttttagggcggagtaacttgcatgt
attgggaattgtagtttttttaaaatgggaagtgacgtatcgtgggaaaacggaagtga
agatttgaggaagttgtgggtttttttggctttcgtttctgggcgtaggttcgcgtgcgg
ttttctgggtgtttttttgtggactttaaccgttacgtcatttttttagtcctatatatac
tcgctctgtacttggccctttttacactgtgactgattgagctggtgccgtgtcgagtg
gtgtttttttaataggtttttttactggtaaggctgactgttatggctgccgctgtggaa
gcgctgtatgttgttctggagcgggagggtgctattttgcctaggcaggagggtttttc
aggtgtttatgtgtttttctctcctattaatttgttatacctcctatgggggctgtaa
tgttgtctctacgcctgcgggtatgtattcccccgggctatttcggtcgctttttagca
ctgaccgatgttaaccaacctgatgtgtttaccgagtcttacattatgactccggacat
gaccgaggaactgtcggtggtgctttttaatcacggtgaccagtttttttacggtcacg
ccggcatggccgtagtccgtcttatgcttataagggttgtttttcctgttgtaagacag
gcttctaatgtttaaatgtttttttttttgttatttttatttgtgtttaatgcaggaac
ccgcagacatgtttgagagaaaaatggtgtctttttctgtggtggttccggaacttacc
tgcctttatctgcatgagcatgactacgatgtgcttgctttttgcgcgaggctttgcc
tgatttttgagcagcaccttgcattttatatcgccgcccatgcaacaagcttacatag

gggctacgctggttagcatagctccgagtatgcgtgtcataatcagtgtgggttctttt
gtcatggttcctggcggggaagtggccgcgctggtccgtgcagacctgcacgattatgt
tcagctggccctgcgaagggacctacgggatcgcggtatttttgttaatgttccgcttt
tgaatcttatacaggtctgtgaggaacctgaatttttgcaatcatgattcgctgcttga
ggctgaaggtggagggcgctctggagcagatttttacaatggccggacttaatattcgg
gatttgcttagagacatattgataaggtggcgagatgaaaattatttgggcatggttga
aggtgctggaatgtttatagaggagattcaccctgaagggtttagcctttacgtccact
tggacgtgagggcagtttgccttttggaagccattgtgcaacatcttacaaatgccatt
atctgttctttggctgtagagtttgaccacgccaccggaggggagcgcgttcacttaat
agatcttcattttgaggttttggataatcttttggaataaaaaaaaaaaaacatggttc
ttccagctcttcccgctcctcccgtgtgtgactcgcagaacgaatgtgtaggttggctg
ggtgtggcttattctgcggtggtggatgttatcagggcagcggcgcatgaaggagttta
catagaacccgaagccaggggcgcctggatgctttgagagagtggatatactacaact
actacacagagcgagctaagcgacgagaccggagacgcagatctgtttgtcacgcccgc
acctggttttgcttcaggaaatatgactacgtccggcgttccatttggcatgacactac
gaccaacacgatctcggttgtctcggcgcactccgtacagtagggatcgcctacctcct
tttgagacagagacccgcgctaccatactggaggatcatccgctgctgcccgaatgtaa
cactttgacaatgcacaacgtgagttacgtgcgaggtcttccctgcagtgtgggattta
cgctgattcaggaatgggttgttccctgggatatggttctgacgcgggaggagcttgta
atcctgaggaagtgtatgcacgtgtgcctgtgttgtgccaacattgatatcatgacgag
catgatgatccatggttacgagtcctgggctctccactgtcattgttccagtcccggtt
ccctgcagtgcatagccggcgggcaggttttggccagctggtttaggatggtggtggat
ggcgccatgtttaatcagaggtttatatggtaccgggaggtggtgaattacaacatgcc
aaaagaggtaatgtttatgtccagcgtgtttatgaggggtcgccacttaatctacctgc
gcttgtggtatgatggccacgtgggttctgtggtccccgccatgagctttggatacagc
gccttgcactgtgggattttgaacaatattgtggtgctgtgctgcagttactgtgctga
tttaagtgagatcagggtgcgctgctgtgcccggaggacaaggcgtctcatgctgcggg
cggtgcgaatcatcgctgaggagaccactgccatgttgtattcctgcaggacggagcgg
cggcggcagcagtttattcgcgcgctgctgcagcaccaccgccctatcctgatgcacga
ttatgactctacccccatgtaggcgtggacttccccttcgccgcccgttgagcaaccgc
aagttggacagcagcctgtggctcagcagctggacagcgacatgaacttaagcgagctg
cccggggagtttattaatatcactgatgagcgtttggctcgacaggaaaccgtgtggaa
tataacacctaagaatatgtctgttacccatgatatgatgctttttaaggccagccggg
gagaaaggactgtgtactctgtgtgttgggagggaggtggcaggttgaatactagggtt
ctgtgagtttgattaaggtacggtgatcaatataagctatgtggtggtggggctatact
actgaatgaaaaatgacttgaaattttctgcaattgaaaaataaacacgttgaaacata
acatgcaacaggttcacgattctttattcctgggcaatgtaggagaaggtgtaagagtt
ggtagcaaaagtttcagtggtgtattttccactttcccaggaccatgtaaaagacatag
agtaagtgcttacctcgctagtttctgtggattcactagaatcgatgtaggatgttgcc
cctcctgacgcggtaggagaaggggagggtgccctgcatgtctgccgctgctcttgctc
ttgccgctgctgaggaggggggcgcatctgccgcagcaccggatgcatctgggaaaagc
aaaaaaggggctcgtccctgtttccggaggaatttgcaagcggggtcttgcatgacggg
gaggcaaacccccgttcgccgcagtccggccggcccgagactcgaaccgggggtcctgc
gactcaacccttggaaaataaccctccggctacagggagcgagccacttaatgctttcg
ctttccagcctaaccgcttacgccgcgcgcggccagtggccaaaaaagctagcgcagca
gccgccgcgcctggaaggaagccaaaaggagcgctcccccgttgtctgacgtcgcacac
ctgggttcgacacgcgggcggtaaccgcatggatcacggcggacggccggatccggggt
tcgaaccccggtcgtccgccatgataccttgcgaatttatccaccagaccacggaaga

gtgcccgcttacaggctctccttttgcacggtctagagcgtcaacgactgcgcacgcct
caccggccagagcgtcccgaccatggagcacttttgccgctgcgcaacatctggaacc
gcgtccgcgactttccgcgcgcctccaccaccgccgccggcatcacctggatgtccagg
tacatctacggattacgtcgacgtttaaaccatatgatcagctcactcaaaggcggtaa
tacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccag
caaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccc
ccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggac
tataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccgacc
ctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctca
tagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtg
tgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgag
tccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattag
cagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggct
acactagaagaacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaa
agagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgt
ttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatctttt
ctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgaga
ttatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaat
ctaaagtatatatgagtaacttggtctgacagttaccaatgcttaatcagtgaggcac
ctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtag
ataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgaga
cccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagc
gcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaa
gctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacagg
catcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgat
caaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcct
ccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcact
gcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtact
caaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtca
atacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacg
ttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaac
ccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtga
gcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttg
aatactcatactcttccttttcaatattattgaagcatttatcagggttattgtctca
tgagcggatacatatttgaatgtatttagaaaaataaacaaatagggttccgcgcaca
tttccccgaaaagtgccacctaaattgtaagcgttaatattttgttaaaattcgcgtta
aatttttgttaaatcagctcatttttaaccaataggccgaaatcggcaaaatcccttta
taaatcaaaagaatagaccgagatagggttgagtgttgttccagtttggaacaagagtc
cactattaaagaacgtggactccaacgtcaaagggcgaaaaaccgtctatcagggcgat
ggcccactacgtgaaccatcaccctaatcaagtttttggggtcgaggtgccgtaaagc
actaaatcggaaccctaaagggagcccccgatttagagcttgacggggaaagccggcga
acgtggcgagaaaggaagggaagaaagcgaaaggagcgggcgctagggcgctggcaagt
gtagcggtcacgctgcgcgtaaccaccacacccgccgcgcttaatgcgccgctacaggg
cgcgatggatcc

pIDT-MbPylRS-RC2



Sequence color-coding key

Feature	Color
Rep	text
Cap	text
CMV promoter	text
T7 promoter	text
MbPylRS	text
Bgh polyA	text
KanR	text

Sequence: Wild type RC2

cccgtgtaaaacgacggccagtttatctagtcagcttgattctagctgatcgtggaccg
gaaggtgagccagtgagttgattgcagtccagttacgctggagtctgaggctcgtcctg
aatgatatgcgaccgccggagggttgcgtttgagacgggcgacagatccagtcgcgctg
ctctcgtcgatccgctagggcggccgctctagaactagtggatcccccggaagatcaga
agttcctattccgaagttcctattctctagaaagtataggaacttctgatctgcgcagc
cgccatgccgggggttttacgagattgtgattaaggtccccagcgaccttgacgagcatc
tgcccggcatttctgacagctttgtgaactgggtggccgagaaggaatgggagttgccg
ccagattctgacatggatctgaatctgattgagcaggcacccctgaccgtggccgagaa
gctgcagcgcgactttctgacggaatggcgccgtgtgagtaaggccccggaggcccttt

tctttgtgcaatttgagaagggagagagctacttccacatgcacgtgctcgtggaaacc
accggggtgaaatccatggttttgggacgtttcctgagtcagattcgcgaaaaactgat
tcagagaatttaccgcgggatcgagccgactttgccaaactggttcgcggtcacaaaga
ccagaaatggcgccggaggcgggaacaaggtggtggatgagtgctacatccccaattac
ttgctccccaaaacccagcctgagctccagtgggcgtggactaatatggaacagtattt
aagcgcctgtttgaatctcacggagcgtaaacggttggtggcgcagcatctgacgcacg
tgtcgcagacgcaggagcagaacaaagagaatcagaatcccaattctgatgcgccggtg
atcagatcaaaaacttcagccaggtacatggagctggtcgggtggctcgtggacaaggg
gattacctcggagaagcagtggatccaggaggaccaggcctcatacatctccttcaatg
cggcctccaactcgcggtcccaaatcaaggctgccttggacaatgcgggaaagattatg
agcctgactaaaaccgcccccgactacctggtgggccagcagcccgtggaggacatttc
cagcaatcggatttataaaattttggaactaaacgggtacgatccccaatatgcggctt
ccgtctttctgggatgggccacgaaaaagttcggcaagaggaacaccatctggctgttt
gggcctgcaactaccgggaagaccaacatcgcggaggccatagcccacactgtgccctt
ctacgggtgcgtaaactggaccaatgagaactttcccttcaacgactgtgtcgacaaga
tggtgatctggtgggaggaggggaagatgaccgccaaggtcgtggagtcggccaaagcc
attctcggaggaagcaaggtgcgcgtggaccagaaatgcaagtcctcggcccagataga
cccgactcccgtgatcgtcacctccaacaccaacatgtgcgccgtgattgacgggaact
caacgaccttcgaacaccagcagccgttgcaagaccggatgttcaaatttgaactcacc
cgccgtctggatcatgactttgggaaggtcaccaagcaggaagtcaaagacttttccg
gtgggcaaaggatcacgtggttgaggtggagcatgaattctacgtcaaaaagggtggag
ccaagaaaagacccgcccccagtgacgcagatataagtgagcccaaacgggtgcgcgag
tcagttgcgcagccatcgacgtcagacgcggaagcttcgatcaactacgcagacaggta
ccaaaacaaatgttctcgtcacgtgggcatgaatctgatgctgtttccctgcagacaat
gcgagagaatgaatcagaattcaaatatctgcttcactcacggacagaaagactgttta
gagtgctttcccgtgtcagaatctcaacccgtttctgtcgtcaaaaaggcgtatcagaa
actgtgctacattcatcatatcatgggaaaggtgccagacgcttgcactgcctgcgatc
tggtcaatgtggatttggatgactgcatctttgaacaataaatgatttaaatcaggtat
ggctgccgatggttatcttccagattggctcgaggacactctctctgaaggaataagac
agtggtggaagctcaaacctggcccaccaccaccaaagcccgcagagcggcataaggac
gacagcaggggtcttgtgcttcctgggtacaagtacctcggacccttcaacggactcga
caagggagagccggtcaacgaggcagacgccgcggccctcgagcacgacaaagcctacg
accggcagctcgacagcggagacaacccgtacctcaagtacaaccacgccgacgcggag
tttcaggagcgccttaaagaagatacgtcttttgggggcaacctcggacgagcagtctt
ccaggcgaaaaagagggttcttgaacctctgggcctggttgaggaacctgttaagacgg
ctccgggaaaaaagaggccggtagagcactctcctgtggagccagactcctcctcggga
accggaaaggcgggccagcagcctgcaagaaaaagattgaattttggtcagactggaga
cgcagactcagtacctgacccccagcctctcggacagccaccagcagccccctctggtc
tgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataacgagggc
gccgacggagtgggtaattcctcgggaaattggcattgcgattccacatggatgggcga
cagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaaccacctct
acaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttggctacagc
acccttgggggtattttgacttcaacagattccactgccacttttcaccacgtgactg
gcaaagactcatcaacaacaactggggattccgacccaagagactcaacttcaagctct
ttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgccaataac
cttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacgtcctcgg
ctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtgccacagt
atggatacctcaccctgaacaacgggagtcaggcagtaggacgctcttcattttactgc

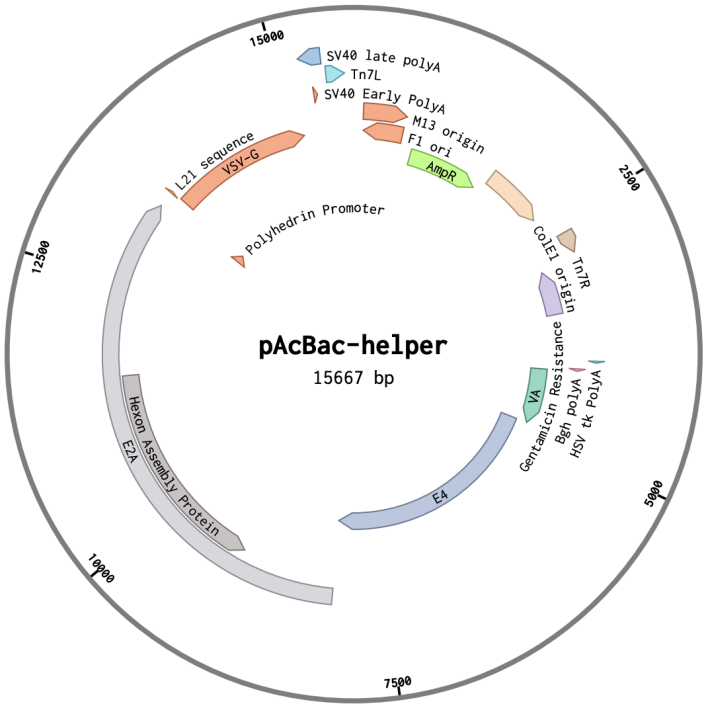
ctggagtactttccttctcagatgctgcgtaccggaaacaactttaccttcagctacac
ttttgaggacgttcctttccacagcagctacgctcacagccagagtctggaccgtctca
tgaatcctctcatcgaccagtacctgtattacttgagcagaacaaacactccaagtgga
accaccacgcagtcaaggcttcagttttctcaggccggagcgagtgacattcgggacca
gtctaggaactggcttcctggaccctgttaccgccagcagcgagtatcaaagacatctg
cggataacaacaacagtgaatactcgtggactggagctaccaagtaccacctcaatggc
agagactctctggtgaatccgggcccggccatggcaagccacaaggacgatgaagaaaa
gtttttcctcagagcggggttctcatctttgggaagcaaggctcagagaaaacaaatg
tggacattgaaaaggtcatgattacagacgaagaggaaatcaggacaaccaatcccgtg
gctacggagcagtatggttctgtatctaccaacctccagagaggcaacagacaagcagc
taccgcagatgtcaacacacaaggcgttcttccaggcatggtctggcaggacagagatg
tgtaccttcaggggcccatctgggcaaagattccacacacggacggacattttcacccc
tctcccctcatgggtggattcggacttaaacaccctcctccacagattctcatcaagaa
caccccggtacctgcgaatccttcgaccaccttcagtgcggcaaagtttgcttccttca
tcacacagtactccacgggacaggtcagcgtggagatcgagtgggagctgcagaaggaa
aacagcaaacgctggaatcccgaaattcagtacacttccaactacaacaagtctgttaa
tgtggactttactgtggacactaatggcgtgtattcagagcctcgccccattggcacca
gatacctgactcgtaatctgtaattgcttgttaatcaataaaccgtttaattcgtttca
gttgaactttggtctctgcgtatttctttcttatctagtttccatggctacgtagataa
gtagcatggcgggttaatcattaactacagcccgggcgtttaaacagcgggcggagggg
tggagtcgtgacgtgaattacgtcatagggttagggaggtcctgtattagaggtcacgt
gagtgttttgcgacattttgcgacaccatgtggtctcgctgggggggggggcccgagtg
agcacgcagggtctccattttgaagcgggaggtttgaacgagcgctggcgcgctcactg
gccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgcct
tgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcc
cttcccatgcatcggccgcaaatacctgcaggatccgttttgcgctgcttcgcgatgta
cgggccagatatacgcgttgacattgattattgactagttattaatagtaatcaattac
ggggtcattagttcatagcccatatatggagttccgcgttacataacttacggtaaatg
gcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtatgtt
cccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacggta
aactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgacg
tcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt
cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttg
gcagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacc
ccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgt
cgtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtcta
tataagcagagctctctggctaactagagaacccactgcttactggcttatcgaaatta
atacgactcactatagggagacccaagctggctagcgccaccatggataaaaaaccatt
agatgttttaatatctgcgaccgggctctggatgtccaggactggcacgctccacaaaa
tcaagcaccatgaggtctcaagaagtaaaatatacattgaaatggcgtgtggagaccat
cttgttgtgaataattccaggagttgtagaacagccagagcattcagacatcataagta
cagaaaaacctgcaaacgatgtagggtttcggacgaggatatcaataattttctcacaa
gatcaaccgaaagcaaaaacagtgtgaaagttagggtagtttctgctccaaaggtcaaa
aaagctatgccgaaatcagtttcaagggctccgaagcctctggaaaattctgtttctgc
aaaggcatcaacgaacacatccagatctgtaccttcgcctgcaaaatcaactccaaatt
cgtctgttcccgcatcggctcctgctccttcacttacaagaagccagcttgatagggtt
gaggctctcttaagtccagaggataaaatttctctaaatatggcaaagcctttcaggga
acttgagcctgaacttgtgacaagaagaaaaaacgattttcagcggctctataccaatg

atagagaagactacctcggtaaactcgaacgtgatattacgaaatttttcgtagaccgg
ggttttctggagataaagtctcctatccttattccggcggaatacgtggagagaatggg
tattaataatgatactgaactttcaaaacagatcttccgggtggataaaaatctctgct
tgaggccaatgcttgccccgactctttacaactatctgcgaaaactcgataggatttta
ccaggcccaataaaaatttttcgaagtcggaccttgttaccggaaagagtctgacggcaa
agagcacctggaagaatttactatggtgaacttctgtcagatgggttcgggatgtactc
gggaaaatcttgaagctctcatcaaagagtttctggactatctggaaatcgacttcgaa
atcgtaggagattcctgtatggtctttggggatactcttgatataatgcacggggacct
ggagctttcttcggcagtcgtcgggccagtttctcttgatagagaatggggtattgaca
aaccatggataggtgcaggttttggtcttgaacgcttgctcaaggttatgcacggcttt
aaaaacattaagagggcatcaaggtccgaatcttactataatgggatttcaaccaatct
gtaagaattcaacgcgttaagtcgactttaactcgagtctagagggcccgtttaaaccc
gctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccccc
gtgccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgagga
aattgcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcagg
acagcaaggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctct
atggcttctgaggcggaaagaaccctaggggtgcgagcggatcgagcagtgtcgatcac
tactggaccgcgagctgtgctgcgacccgtgatcttacggcattatacgtatgatcggt
ccacgatcagctagattatctagtcagcttgatgtcatagctgtttcctgaggctcaat
actgaccatttaaatcatacctgacctccatagcagaaagtcaaaagcctccgaccgga
ggcttttgacttgatcggcacgtaagaggttccaactttcaccataatgaaataagatc
actaccgggcgtatttttgagttatcgagattttcaggagctaaggaagctaaaatga
gccatattcaacgggaaacgtcttgcttgaagccgcgattaaattccaacatggatgct
gatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtgcgacaatcta
tcgattgtatgggaagcccgatgcgccagagttgtttctgaaacatggcaaaggtagcg
ttgccaatgatgttacagatgagatggtcaggctaaactggctgacggaatttatgcct
cttccgaccatcaagcattttatccgtactcctgatgatgcatggttactcaccactgc
gatcccagggaaaacagcattccaggtattagaagaatatcctgattcaggtgaaaata
ttgttgatgcgctggcagtgttcctgcgccggttgcattcgattcctgtttgtaattgt
ccttttaacggcgatcgcgtatttcgtctcgctcaggcgcaatcacgaatgaataacgg
tttggttggtgcgagtgattttgatgacgagcgtaatggctggcctgttgaacaagtct
ggaaagaaatgcataaactcttgccattctcaccggattcagtcgtcactcatggtgat
ttctcacttgataaccttatttttgacgaggggaaattaataggttgtattgatgttgg
acgagtcggaatcgcagaccgataccaggatcttgccatcctatggaactgcctcggtg
agttttctccttcattacagaaacggcttttcaaaaatatggtattgataatcctgat
atgaataaattgcagtttcacttgatgctcgatgagttttctaatgaggacctaaatg
taatcacctggctcaccttcgggtgggcctttctgcgttgctggcgttttccataggc
tccgccccccctgacgagcatcacaaaaatcgatgctcaagtcagaggtggcgaaacccg
acaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgt
tccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgc
tttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctg
ggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcg
tcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaaca
ggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaac
tacggctacactagaagaacagtatttggtatctgcgctctgctgaagccagttacctc
ggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttt
ttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttga
ttttctaccgaagaaaggccca

The sequence for RC2-454TAG is identical except the sequence for *Cap* is:

atggctgccgatggttatcttccagattggctcgaggacactctctctgaaggaataag
acagtggtggaagctcaaacctggcccaccaccaccaaagcccgcagagcggcataagg
acgacagcaggggtcttgtgcttcctgggtacaagtacctcggacccttcaacggactc
gacaagggagagccggtcaacgaggcagacgccgcggccctcgagcacgacaaagccta
cgaccggcagctcgacagcggagacaacccgtacctcaagtacaaccacgccgacgcgg
agtttcaggagcgccttaaagaagatacgtcttttgggggcaacctcggacgagcagtc
ttccaggcgaaaaagagggttcttgaacctctgggcctggttgaggaacctgttaagac
ggctccgggaaaaaagaggccggtagagcactctcctgtggagccagactcctcctcgg
gaaccggaaaggcgggccagcagcctgcaagaaaaagattgaattttggtcagactgga
gacgcagactcagtacctgacccccagcctctcggacagccaccagcagccccctctgg
tctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataacgagg
gcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatggatgggc
gacagagtcatcaccaccagcTAGcgaacctgggccctgcccacctacaacaaccacct
ctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttggctaca
gcacccccttgggggtattttgacttcaacagattccactgccacttttcaccacgtgac
tggcaaagactcatcaacaacaactggggattccgacccaagagactcaacttcaagct
ctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgccaata
accttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacgtcctc
ggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtgccaca
gtatggatacctcaccctgaacaacgggagtcaggcagtaggacgctcttcatttttact
gcctggagtactttccttctcagatgctgcgtaccggaaacaactttaccttcagctac
acttttgaggacgttcctttccacagcagctacgctcacagccagagtctggaccgtct
catgaatcctctcatcgaccagtacctgtattacttgagcagaacaaacactccaagtg
gaaccaccacgcagtcaaggcttcagttttctcaggccggagcgagtgacattcgggac
cagtctaggaactggcttcctggaccctgttaccgccagcagcgagtatcaaagacatc
tgcggataacaacaacagtgaatactcgtggactggagctaccaagtaccacctcaatg
gcagagactctctggtgaatccgggcccggccatggcaagccacaaggacgatgaagaa
aagtttttcctcagagcggggttctcatctttgggaagcaaggctcagagaaaacaaa
tgtggacattgaaaaggtcatgattacagacgaagaggaaatcaggacaaccaatcccg
tggctacggagcagtatggttctgtatctaccaacctccagagaggcaacagacaagca
gctaccgcagatgtcaacacacaaggcgttcttccaggcatggtctggcaggacagaga
tgtgtaccttcaggggcccatctgggcaaagattccacacacggacggacattttcacc
cctctcccctcatgggtggattcggacttaaacaccctcctccacagattctcatcaag
aacaccccggtacctgcgaatccttcgaccaccttcagtgcggcaaagtttgcttcctt
catcacacagtactccacgggacaggtcagcgtggagatcgagtgggagctgcagaagg
aaaacagcaaacgctggaatcccgaaattcagtacacttccaactacaacaagtctgtt
aatgtggactttactgtggacactaatggcgtgtattcagagcctcgccccattggcac
cagatacctgactcgtaatctgtaa

pAcBac-Helper



Sequence color-coding key

Feature	Color
M13 ori	text
AmpR	text
ColE1 ori	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
VA	text
E4	text
E2A	text
Polyhedrin promoter	text
L21 sequence	text
VSV-G	text
SV40 late polyA	text
Tn7L	text

Sequence

ttctctgtcacagaatgaaaatttttctgtcatctcttcgttattaatgtttgtaattg
actgaatatcaacgcttatttgcagcctgaatggcgaatgggacgcgccctgtagcggc

gcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgc
cctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttc
cccgtcaagctctaaatcgggggctcccttagggttccgatttagtgctttacggcac
ctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgata
gacggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgttcc
aaactggaacaacactcaaccctatctcggtctattcttttgatttataagggattttg
ccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattt
taacaaaatattaacgtttacaatttcaggtggcacttttcggggaaatgtgcgcggaa
cccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataa
ccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccg
tgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaa
cgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaa
ctggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaat
gatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggc
aagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcacca
gtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccat
aaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaagg
agctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaa
ccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaat
ggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaac
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggccctt
ccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtat
cattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacgg
ggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactg
attaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaa
acttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgacca
aaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaa
ggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaacc
accgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaagg
taactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagtta
ggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgtt
accagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgat
agttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagc
ttggagcgaacgacctacaccgaactgagatacctacagcgtgagcattgagaaagcgc
cacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacag
gagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcggg
tttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcct
atggaaaaacgccagcaacgcggcctttttacggttcctggccttttgctggccttttg
ctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgccttt
gagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcga
ggaagcggaagagcgcctgatgcggtattttctccttacgcatctgtgcggtatttcac
accgcagaccagccgcgtaacctggcaaaatcggttacggttgagtaataaatggatgc
cctgcgtaagcgggtgtgggcggacaataaagtcttaaactgaacaaaatagatctaaa
ctatgacaataaagtcttaaactagacagaatagttgtaaactgaaatcagtccagtta
tgctgtgaaaaagcatactggacttttgttatggctaaagcaaactcttcattttctga
agtgcaaattgcccgtcgtattaaagaggggcgtggccaagggcatggtaaagactata
ttcgcggcgttgtgacaatttaccgaacaactccgcggccgggaagccgatctcggctt
gaacgaattgttaggtggcggtacttgggtcgatatcaaagtgcatcacttcttcccgt

atgcccaactttgtatagagagccactgcgggatcgtcaccgtaatctgcttgcacgta
gatcacataagcaccaagcgcgttggcctcatgcttgaggagattgatgagcgcggtgg
caatgccctgcctccggtgctcgccggagactgcgagatcatagatatagatctcacta
cgcggctgctcaaacctgggcagaacgtaagccgcgagagcgccaacaaccgcttcttg
gtcgaaggcagcaagcgcgatgaatgtcttactacggagcaagttcccgaggtaatcgg
agtccggctgatgttgggagtaggtggctacgtctccgaactcacgaccgaaaagatca
agagcagcccgcatggatttgacttggtcagggccgagcctacatgtgcgaatgatgcc
catacttgagccacctaactttgttttagggcgactgccctgctgcgtaacatcgttgc
tgctgcgtaacatcgttgctgctccataacatcaaacatcgacccacggcgtaacgcgc
ttgctgcttggatgcccgaggcatagactgtacaaaaaaacagtcataacaagccatga
aaaccgccactgcgccgttaccaccgctgcgttcggtcaaggttctggaccagttgcgt
gagcgcatacgctacttgcattacagtttacgaaccgaacaggcttatgtcaactgggt
tcgtgccttcatccgtttccacggtgtgcgtcacccggcaaccttgggcagcagcgaag
tcgaggcatttctgtcctggctggcgaacgagcgcaaggtttcggtctccacgcatcgt
caggcattggcggccttgctgttcttctacggcaaggtgctgtgcacggatctgccctg
gcttcaggagatcggtagacctcggccgtcgcggcgcttgccggtggtgctgaccccgg
atgaagtggttcgcatcctcggttttctggaaggcgagcatcgtttgttcgcccaggac
tctagctatagttctagtggttggctacgtacccgtagtggctatggcagggcttgcgc
ttaatgcgccgctacagggcgcgtggggataccccctagagccccagctggttctttcc
gcctcagaagccatagagcccaccgcatccccagcatgccgtaatccgtagatgtacct
ggacatccaggtgatgccggcggcggtggtggaggcgcgcggaaagtcgcggacgcggt
tccagatgttgcgcagcggcaaaaagtgctccatggtcgggacgctctggccggtgagg
cgtgcgcagtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactct
tccgtggtctggtggataaattcgcaagggtatcatggcggacgaccggggttcgaacc
ccggatccggccgtccgccgtgatccatgcggttaccgcccgcgtgtcgaacccaggtg
tgcgacgtcagacaacgggggagcgctccttttggcttccttccaggcgcggcggctgc
tgcgctagctttttggccactggccgcgcgcggcgtaagcggttaggctggaaagcga
aagcattaagtggctcgctccctgtagccggagggttattttccaagggttgagtcgca
ggaccccggttcgagtctcgggccggccggactgcggcgaacggggttgcctcccc
gtcatgcaagaccccgcttgcaaattcctccggaaacagggacgagccccttttttgct
ttcccagatgcatccggtgctgcggcagatgcgcccccctcctcagcagcggcaagag
caagagcagcggcagacatgcagggcaccctcccctctcctaccgcgtcaggaggggc
aacatcctacatcgattctagtgaatccacagaaactagcgaggtaagcacttactcta
tgtcttttacatggtcctgggaaagtggaaaatacaccactgaaactttgctaccaac
tcttacaccttctcctacattgcccaggaataaagaatcgtgaacctgttgcatgttat
gtttcaacgtgtttatttttcaattgcagaaaatttcaagtcatttttcattcagtagt
atagccccaccaccacatagcttatattgatcaccgtaccttaatcaaactcacagaac
cctagtattcaacctgccacctccctcccaacacacagagtacacagtcctttctcccc
ggctggccttaaaaagcatcatatcatgggtaacagacatattcttaggtgttatattc
cacacggtttcctgtcgagccaaacgctcatcagtgatattaataaactccccgggcag
ctcgcttaagttcatgtcgctgtccagctgctgagccacaggctgctgtccaacttgcg
gttgctcaacgggcggcgaaggggaagtccacgcctacatgggggtagagtcataatcg
tgcatcaggatagggcggtggtgctgcagcagcgcgcgaataaactgctgccgccgccg
ctccgtcctgcaggaatacaacatggcagtggtctcctcagcgatgattcgcaccgccc
gcagcatgagacgccttgtcctccgggcacagcagcgcaccctgatctcacttaaatca
gcacagtaactgcagcacagcaccacaatattgttcaaaatcccacagtgcaaggcgct
gtatccaaagctcatggcggggaccacagaacccacgtggccatcataccacaagcgca
ggtagattaagtggcgacccctcataaacacgctggacataaacattacctctttggc

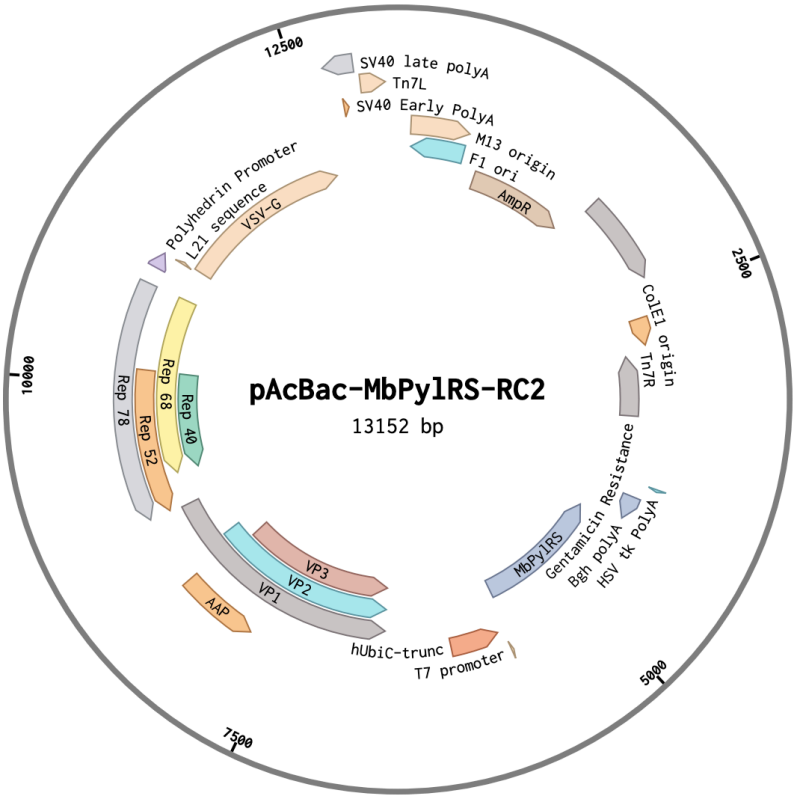
atgttgtaattcaccacctcccggtaccatataaacctctgattaaacatggcgccatc
caccaccatcctaaaccagctggccaaaacctgcccgccggctatgcactgcagggaac
cgggactggaacaatgacagtggagagcccaggactcgtaaccatggatcatcatgctc
gtcatgatatcaatgttggcacaacacaggcacacgtgcatacacttcctcaggattac
aagctcctcccgcgtcagaaccatatcccagggaacaacccattcctgaatcagcgtaa
atcccacactgcagggaagacctcgcacgtaactcacgttgtgcattgtcaaagtgtta
cattcgggcagcagcggatgatcctccagtatggtagcgcgggtctctgtctcaaaagg
aggtaggcgatccctactgtacggagtgcgccgagacaaccgagatcgtgttggtcgta
gtgtcatgccaaatggaacgccggacgtagtcatatttcctgaagcaaaaccaggtgcg
ggcgtgacaaacagatctgcgtctccggtctcgtcgcttagctcgctctgtgtagtagt
tgtagtatatccactctctcaaagcatccaggcgccccctggcttcgggttctatgtaa
actccttcatgcgccgctgccctgataacatccaccaccgcagaataagccacacccag
ccaacctacacattcgttctgcgagtcacacacgggaggagcgggaagagctggaagaa
ccatgttttttttttttattccaaaagattatccaaaacctcaaaatgaagatctatt
aagtgaacgcgctcccctccggtggcgtggtcaaactctacagccaaagaacagataat
ggcatttgtaagatgttgcacaatggcttccaaaaggcaaactgccctcacgtccaagt
ggacgtaaaggctaaaccccttcagggtgaatctcctctataaacattccagcaccttca
accatgcccaaataattttcatctcgccaccttatcaatatgtctctaagcaaatcccg
aatattaagtccggccattgtaaaaatctgctccagagcgccctccaccttcagcctca
agcagcgaatcatgattgcaaaaattcaggttcctcacagacctgtataagattcaaaa
gcggaacattaacaaaaataccgcgatcccgtaggtcccttcgcagggccagctgaaca
taatcgtgcaggtctgcacggaccagcgcggccacttccccgccaggaaccatgacaaa
agaacccacactgattatgacacgcatactcggagctatgctaaccagcgtagccccta
tgtaagcttgttgcatgggcggcgatataaaatgcaaggtgctgctcaaaaaatcaggc
aaagcctcgcgcaaaaaagcaagcacatcgtagtcatgctcatgcagataaaggcaggt
aagttccggaaccaccacagaaaaagacaccatttttctctcaaacatgtctgcgggtt
cctgcattaaacacaaaataaaataacaaaaaaaaaaaacatttaaacattagaagcctg
tcttacaacaggaaaaacaacccttataagcataagacggactacggccatgccggcgt
gaccgtaaaaaaactggtcaccgtgattaaaaagcaccaccgacagttcctcggtcatg
tccggagtcataatgtaagactcggtaaacacatcaggttggttaacatcggtcagtgc
taaaaagcgaccgaaatagcccgggggaatacatacccgcaggcgtagagacaacatta
cagcccccataggaggtataacaaaattaataggagagaaaaacacataaacacctgaa
aaaccctcctgcctaggcaaaatagcaccctcccgctccagaacaacatacagcgcttc
cacagcggcagccataacagtcagccttaccagtaaaaaaacctattaaaaaacaccac
tcgacacggcaccagctcaatcagtcacagtgtaaaaagggccaagtacagagcgagta
tatataggactaaaaaatgacgtaacggttaaagtccacaaaaaacacccagaaaaccg
cacgcgaacctacgcccagaaacgaaagccaaaaaacccacaacttcctcaaatcttca
cttccgttttcccacgatacgtcacttcccattttaaaaaaactacaattcccaataca
tgcaagttactccgccctaaaacgcccgggcgaccgcaccctgtgacgaaagccgcccg
caagctgcgcccctgagttagtcatctgaacttcggcctgggcgtctctgggaagtacc
acagtggtgggagcgggactttcctggtacaccagggcagcgggccaactacggggatt
aaggttattacgaggtgtggtggtaatagccgcctgttcgaggagaattcggtttcggt
gggcgcggattccgttgacccgggatatcatgtggggtcccgcgctcatgtagtttatt
cgggttgagtagtcttgggcagctccagccgcaagtcccatttgtggctggtaactcca
catgtagggcgtgggaatttccttgctcataatggcgctgacgacaggtgctggcgccg
ggtgtggccgctggagatgacgtagttttcgcgcttaaatttgagaaagggcgcgaaac
tagtccttaagagtcagcgcgcagtatttgctgaagagagcctccgcgtcttccagcgt
gcgccgaagctgatcttcgcttttgtgatacaggcagctgcgggtgagggagcgcagag

acctgtttttattttcagctcttgttcttggcccctgctttgttgaaatatagcatac
agagtgggaaaaatcctatttctaagctcgcgggtcgatacgggttcgttgggcgccag
acgcagcgctcctcctcctgctgctgccgccgctgtggatttcttgggctttgtcagag
tcttgctatccggtcgcctttgcttctgtgtgaccgctgctgttgctgccgctgccgct
gccgccggtgcagtaggggctgtagagatgacggtagtaatgcaggatgttacgggggа
aggccacgccgtgatggtagagaagaaagcggcgggcgaaggagatgttgcccccacag
tcttgcaagcaagcaactatggcgttcttgtgcccgcgccacgagcggtagccttggcg
ctgttgttgctcttgggctaacggcggcggctgcttagacttaccggccctggttccag
tggtgtcccatctacggttgggtcggcgaacaggcagtgccggcggcgcctgaggagcg
gaggttgtagcgatgctgggaacggttgccaatttctggggcgccggcgaggggaatgc
gaccgagggtgacggtgtttcgtctgacacctcttcggcctcggaagcttcgtctaggc
tgtcccagtcttccatcatctcctcctcctcgtccaaaacctcctctgcctgactgtcc
cagtattcctcctcgtccgtgggtggcggcggcggcagctgcagcttctttttgggtgc
catcctgggaagcaagggcccgcggctgctgatagggctgcggcggcgggggattggg
ttgagctcctcgccggactgggggtccaggtaaaccccccgtcccttтcgtagcagaaa
ctcttggcgggctttgttgatggcttgcaattggccaaggatgtggccctgggtaatga
cgcaggcggtaagctccgcatttggcgggcgggattggtcttcgtagaacctaatctcg
tgggcgtggtagtcctcaggtacaaatttgcgaaggtaagccgacgtccacagccccgg
agtgagtttcaaccccggagccgcggacttttcgtcaggcgagggaccctgcagctcaa
aggtaccgataatttgactttcgctaagcagttgcgaattgcagaccagggagcggtgc
ggggtgcataggttgcagcgacagtgacactccagtaggccgtcaccgctcacgtcttc
catgatgtcggagtggtaggcaaggtagttggctagctgcagaaggtagcagtgacccc
aaagcggcggagggcattcacggtacttaatgggcacaaagtcgctaggaagcgcacag
caggtggcgggcagaattcctgaacgctctaggataaagttcctaaagttttgcaacat
gctttgactggtgaagtctggcagaccctgttgcagggtttaagcaggcgttcgggga
agataatgtccgccaggtgcgcggccacggagcgctcgttgaaggccgtccataggtcc
ttcaagttttgctttagcagcttctgcagctcctttaggttgcgctcctccaggcattg
ctgccacacgcccatggccgtttgccaggtgtagcacagaaataagtaaacgcagtcgc
ggacgtagtcgcggcgcgcctcgcccttgagcgtggaatgaagcacgttttgcccgagg
cggttttcgtgcaaaattccaaggtaggagaccaggttgcagagctccacgttggaaat
tttgcaggcctggcgcacgtagccctggcgaaaggtgtagtgcaacgtttcctctagct
tgcgctgcatctccgggtcagcaaagaaccgctgcatgcactcaagctccacggtaaca
agcactgcggccatcattagcttgcgtcgctcctccaagtcggcaggctcgcgcgtctc
aagccagcgcgccagctgctcatcgccaactgcgggtaggccctcctcggtttgttctt
gcaagtttgcatccctctccaggggtcgtgcacggcgcacgatcagctcgctcatgact
gtgctcataaccttgggggtaggttaagtgccgggtaggcaaagtgggtgacctcgat
gctgcgtttcagcacggctaggcgcgcgttgtcaccctcaagttccaccagcactccac
agtgactttcattttcgctgttttcttgttgcagagcgtttgccgcgcgtttctcgtcg
cgtccaagaccctcaaagattttggcacttcgtcgagcgaggcgatatcaggtatgac
agcgccctgccgcaaggccagctgcttgtccgctcggctgcggttggcacggcaggata
ggggtatcttgcagttttggaaaaagatgtgataggtggcaagcacctctggcacggca
aatacggggtagaagttgaggcgcgggttgggctcgcatgtgccgttttcttggcgttt
ggggggtacgcgcggtgagaacaggtggcgttcgtaggcaaggctgacatccgctatgg
cgaggggcacatcgctgcgctcttgcaacgcgtcgcagataatggcgcactggcgctgc
agatgcttcaacagcacgtcgtctcccacatctaggtagtcgccatgcctttggtcccc
ccgcccgacttgttcctcgtttgcctctgcgtcgtcctggtcttgcttttatcctctg
ttggtactgagcgatcctcgtcgtcttcgcttacaaaacctgggtcctgctcgataatc
acttcctcctcctcaagcgggggtgcctcgacggggaaggtggtaggcgcgttggcggc

atcggtggaggcggtggtggcgaactcaaaggggcggttaggctgtcctccttctcga
ctgactccatgatctttttctgcctataggagaaggaaatggccagtcgggaagaggag
cagcgcgaaaccacccccgagcgcggacgcggtgcggcgcgacgtccaccaaccatgga
ggacgtgtcgtccccgtcgccgtcgccgccgcctccccgcgcgcccccaaaaaagcggc
tgaggcggcgtctcgagtccgaggacgaagaagactcgtcacaagatgcgctggtgccg
cgcacacccagcccgcggccatcgacctcgacggcggatttggccattgcgtccaaaaa
gaaaaagaagcgcccctctcccaagcccgagcgcccgccatccccagaggtgatcgtgg
acagcgaggaagaaagagaagatgtggcgctacaaatggtgggtttcagcaacccaccg
gtgctaatcaagcacggcaagggaggtaagcgcacggtgcggcggctgaatgaagacga
cccagtggcgcggggtatgcggacgcaagaggaaaaggaagagtccagtgaagcggaaa
gtgaaagcacggtgataaacccgctgagcctgccgatcgtgtctgcgtgggagaagggc
atggaggctgcgcgcgcgttgatggacaagtaccacgtggataacgatctaaaggcaaa
cttcaagctactgcctgaccaagtggaagctctggcggccgtatgcaagacctggctaa
acgaggagcaccgcgggttgcagctgaccttcaccagcaacaagacctttgtgacgatg
atggggcgattcctgcaggcgtacctgcagtcgtttgcagaggtaacctacaagcacca
cgagcccacgggctgcgcgttgtggctgcaccgctgcgctgagatcgaaggcgagctta
agtgtctacacgggagcattatgataaataaggagcacgtgattgaaatggatgtgacg
agcgaaaacgggcagcgcgcgctgaaggagcagtctagcaaggccaagatcgtgaagaa
ccggtggggccgaaatgtggtgcagatctccaacaccgacgcaaggtgctgcgtgcatg
acgcggcctgtccggccaatcagttttccggcaagtcttgcggcatgttcttctctgaa
ggcgcaaaggctcaggtggcttttaagcagatcaaggctttcatgcaggcgctgtatcc
taacgcccagaccgggcacggtcaccttctgatgccactacggtgcgagtgcaactcaa
agcctgggcatgcacccttttgggaaggcagctaccaaagttgactccgttcgccctg
agcaacgcggaggacctggacgcggatctgatctccgacaagagcgtgctggccagcgt
gcaccacccggcgctgatagtgttccagtgctgcaaccctgtgtatcgcaactcgcgcg
cgcagggcggaggccccaactgcgacttcaagatatcggcgcccgacctgctaaacgcg
ttggtgatggtgcgcagcctgtggagtgaaaacttcaccgagctgccgcggatggttgt
gcctgagtttaagtggagcactaaacaccagtatcgcaacgtgtccctgccagtggcgc
atagcgatgcgcggcagaaccccttgatttttaaacggcgcagacggcaagggtgggg
ggtaaataatcacccgagagtgtacaaataaaaacatttgcctttattgaaagtgtctc
ctagtacattatttttacatgttttcaagtgacaaaaagaagtggcgctcctaatctg
cgcactgtggctgcggaagtagggcgagtggcgctccaggaagctgtagagctgttcct
ggttgcgacgcagggtgggctgtacctggggactgttaagcatggagttgggtaccgcg
gccgcggtccgtatactccggaatattaatagatcatggagataattaaaatgataacc
atctcgcaaataaataagtattttactgttttcgtaacagttttgtaataaaaaaacct
ataaatattccggattattcataccgtcccaccatcgggcgcgaactcctaaaaaaccg
ccaccatgaagtgccttttgtacttagccttttttattcattggggtgaattgcaagttc
accatagttttccacacaaccaaaaaggaaactggaaaaatgttccttctaattacca
ttattgcccgtcaagctcagatttaaattggcataatgacttaataggcacagccttac
aagtcaaaatgcccaagagtcacaaggctattcaagcagacggttggatgtgtcatgct
tccaaatgggtcactacttgtgatttccgctggtatggaccgaagtatataacacattc
catccgatccttcactccatctgtagaacaatgcaaggaaagcattgaacaaacgaaac
aaggaacttggctgaatccaggcttccctcctcaaagttgtggatatgcaactgtgacg
gatgccgaagcagtgattgtccaggtgactcctcaccatgtgctggttgatgaatacac
aggagaatgggttgattcacagttcatcaacggaaaatgcagcaattacatatgcccca
ctgtccataactctacaacctggcattctgactataaggtcaaagggctatgtgattct
aacctcatttccatggacatcaccttcttctcagaggacggagagctatcatccctggg
aaaggagggcacagggttcagaagtaactactttgcttatgaaactggaggcaaggcct

gcaaaatgcaatactgcaagcattggggagtcagactcccatcaggtgtctggttcgag
atggctgataaggatctctttgctgcagccagattccctgaatgcccagaagggtcaag
tatctctgctccatctcagacctcagtggatgtaagtctaattcaggacgttgagagga
tcttggattattccctctgccaagaaacctggagcaaaatcagagcgggtcttccaatc
tctccagtggatctcagctatcttgctcctaaaaacccaggaaccggtcctgctttcac
cataatcaatggtaccctaaaatactttgagaccagatacatcagagtcgatattgctg
ctccaatcctctcaagaatggtcggaatgatcagtggaactaccacagaaagggaactg
tgggatgactgggcaccatatgaagacgtggaaattggacccaatggagttctgaggac
cagttcaggatataagtttcctttatacatgattggacatggtatgttggactccgatc
ttcatcttagctcaaaggctcaggtgttcgaacatcctcacattcaagacgctgcttcg
caacttcctgatgatgagagtttatttttggtgatactgggctatccaaaaatccaat
cgagcttgtagaaggttggttcagtagttggaaaagctctattgcctctttttcttta
tcatagggttaatcattggactattcttggttctccgagttggtatccatctttgcatt
aaattaaagcacaccaagaaaagacagatttatacagacatagagatgaaccgacttgg
aaagtgataaggccaggccggccaagcttgtcgagaagtactagaggatcataatcagc
cataccacatttgtagaggttttacttgctttaaaaaacctcccacacctccccctgaa
cctgaaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttataatg
gttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcactgcat
tctagttgtggtttgtccaaactcatcaatgtatcttatcatgtctggatctgatcact
gcttgagcctaggagatccgaaccagataagtgaaatctagttccaaactattttgtca
ttttttaattttcgtattagcttacgacgctacacccagttcccatctattttgtcactc
ttccctaaataatccttaaaaactccatttccacccctcccagttcccaactattttgt
ccgcccacagcggggcatttttcttcctgttatgttttttaatcaaacatcctgccaact
ccatgtgacaaaccgtcatcttcggctacttt

pAcBac-MbPylRS-RC2



Sequence color-coding key

Feature	Color
M13 ori	text
AmpR	text
ColE1 ori	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
MbPylRS	text
T7 promoter	text
hUbiC-trunc	text
Cap	text
Rep	text
Polyhedrin promoter	text
L21 sequence	text
VSV-G	text
SV40 late PolyA	text
Tn7L	text

Sequence: wild type RC2

ttctctgtcacagaatgaaaatttttctgtcatctcttcgttattaatgtttgtaattg
actgaatatcaacgcttatttgcagcctgaatggcgaatgggacgcgccctgtagcggc
gcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgc
cctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttc
cccgtcaagctctaaatcgggggctcccttagggttccgatttagtgctttacggcac
ctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgata
gacggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgttcc
aaactggaacaacactcaaccctatctcggtctattcttttgatttataagggatttg
ccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattt
taacaaaatattaacgtttacaatttcaggtggcacttttcggggaaatgtgcgcggaa
cccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataa
ccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccg
tgtcgcccttattccctttttgcggcattttgccttcctgttttgctcacccagaaa
cgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaa
ctggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaat
gatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggc
aagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcacca
gtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccat
aaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaagg
agctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaa
ccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaat
ggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaac
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggccctt
ccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtat
cattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacgg
ggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactg
attaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaa
acttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgacca
aaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaa
ggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaacc
accgctaccagcggtggtttgtttgccggatcaagagctaccaactcttttccgaagg
taactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagtta
ggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgtt
accagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgat
agttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagc
ttggagcgaacgacctacaccgaactgagatacctacagcgtgagcattgagaaagcgc
cacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacag
gagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcggg
tttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcct
atggaaaaacgccagcaacgcggcctttttacggttcctggccttttgctggccttttg
ctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgccttt
gagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcga
ggaagcggaagagcgcctgatgcggtattttctccttacgcatctgtgcggtatttcac
accgcagaccagccgcgtaacctggcaaaatcggttacggttgagtaataaatggatgc
cctgcgtaagcgggtgtgggcggacaataaagtcttaaactgaacaaaatagatctaaa
ctatgacaataaagtcttaaactagacagaatagttgtaaactgaaatcagtccagtta

tgctgtgaaaaagcatactggacttttgttatggctaaagcaaactcttcattttctga
agtgcaaattgcccgtcgtattaaagaggggcgtggccaagggcatggtaaagactata
ttcgcggcgttgtgacaatttaccgaacaactccgcggccgggaagccgatctcggctt
gaacgaattgttaggtggcggtacttgggtcgatatcaaagtgcatcacttcttcccgt
atgcccaactttgtatagagagccactgcgggatcgtcaccgtaatctgcttgcacgta
gatcacataagcaccaagcgcgttggcctcatgcttgaggagattgatgagcgcggtgg
caatgccctgcctccggtgctcgccggagactgcgagatcatagatatagatctcacta
cgcggctgctcaaacctgggcagaacgtaagccgcgagagcgccaacaaccgcttcttg
gtcgaaggcagcaagcgcgatgaatgtcttactacggagcaagttcccgaggtaatcgg
agtccggctgatgttgggagtaggtggctacgtctccgaactcacgaccgaaaagatca
agagcagcccgcatggatttgacttggtcagggccgagcctacatgtgcgaatgatgcc
catacttgagccacctaactttgttttagggcgactgccctgctgcgtaacatcgttgc
tgctgcgtaacatcgttgctgctccataacatcaaacatcgacccacggcgtaacgcgc
ttgctgcttggatgcccgaggcatagactgtacaaaaaaacagtcataacaagccatga
aaaccgccactgcgccgttaccaccgctgcgttcggtcaaggttctggaccagttgcgt
gagcgcatacgctacttgcattacagtttacgaaccgaacaggcttatgtcaactgggt
tcgtgccttcatccgtttccacggtgtgcgtcacccggcaaccttgggcagcagcgaag
tcgaggcatttctgtcctggctggcgaacgagcgcaaggtttcggtctccacgcatcgt
caggcattggcggccttgctgttcttctacggcaaggtgctgtgcacggatctgccctg
gcttcaggagatcggtagacctcggccgtcgcggcgcttgccggtggtgctgaccccgg
atgaagtggttcgcatcctcggttttctggaaggcgagcatcgtttgttcgcccaggac
tctagctatagttctagtggttggctacgtacccgtagtggctatggcagggcttgcgc
ttaatgcgccgctacagggcgcgtggggataccccctagagccccagctggttctttcc
gcctcagaagccatagagcccaccgcatccccagcatgcctgctattgtcttcccaatc
ctccccttgctgtcctgccccaccccaccccccagaatagaatgacacctactcagac
aatgcgatgcaatttcctcattttattaggaaaggacagtgggagtggcaccttccagg
gtcaaggaaggcacgggggaggggcaaacaacagatggctggcaactagaaggcacagt
cgaggctgatcagcgggtttaaacgggccctctagactcgagttaaagtcgacttaacg
cgttgaattcttacagattggttgaaatcccattatagtaagattcggaccttgatgcc
ctcttaatgtttttaaagccgtgcataaccttgagcaagcgttcaagaccaaaacctgc
acctatccatggtttgtcaataccccattctctatcaagagaaactggcccgacgactg
ccgaagaaagctccaggtccccgtgcattatatcaagagtatccccaaagaccatacag
gaatctcctacgatttcgaagtcgatttccagatagtccagaaactctttgatgagagc
ttcaagattttcccgagtacatcccgaacccatctgacagaagttcaccatagtaaatt
cttccaggtgctctttgccgtcagactctttccggtaacaaggtccgacttcgaaaatt
tttattgggcctggtaaaatcctatcgagttttcgcagatagttgtaaagagtcggggc
aagcattggcctcaagcagagatttttatccacccggaagatctgttttgaaagttcag
tatcattattaatacccattctctccacgtattccgccggaataaggataggagacttt
atctccagaaaaccccggtctacgaaaaatttcgtaatatcacgttcgagtttaccgag
gtagtcttctctatcattggtatagagccgctgaaaatcgtttttttcttcttgtcacaa
gttcaggctcaagttccctgaaaggctttgccatatttagagaaattttatcctctgga
cttaagagagcctcaaccctatcaagctggcttcttgtaagtgaaggagcaggagccga
tgcgggaacagacgaatttggagttgattttgcaggcgaaggtacagatctggatgtgt
tcgttgatgcctttgcagaaacagaattttccagaggcttcggagcccttgaaactgat
ttcggcatagcttttttgaccttttggagcagaaactaccctaactttcacactgttttt
gctttcggttgatcttgtgagaaaattattgatatcctcgtccgaaaccctacatcgtt
tgcaggtttttctgtacttatgatgtctgaatgctctggctgttctacaactcctggaa
ttattcacaacaagatggtctccacacgccatttcaatgtatattttacttcttgagac

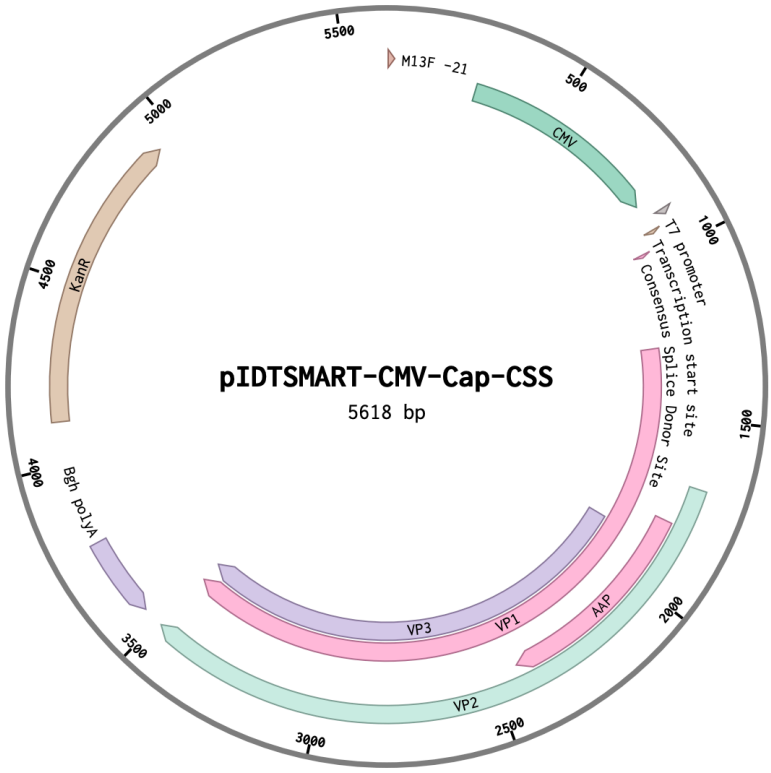
ctcatggtgcttgattttgtggagcgtgccagtcctggacatccagagcccggtcgcag
atattaaaacatctaatggtttttatccatggtggcgctagccagcttgggtctccct
atagtgagtcgtattaatttcgataagccagtaagcagtgggttctctagttagccaga
gagctctagaccaagtgacgatcacagcgatccacaaacaagaaccgcgacccaaatcc
cggctgcgacggaactagctgtgccacacccggcgcgtccttatataatcatcggcgtt
caccgccccacggagatccctccgcagaatcgccgagaagggactactttcctcgcct
gttccgctctctggaaagaaaaccagtgccctagagtcacccaagtcccgtcctaaaat
gtccttctgctgatactggggttctaaggccgagtcttatgagcagcgggccgctgtcc
tgagcgtccgggcggaaggatcaggacgctcgctgcgcccttcgtctgacgtggcagcg
ctcgccgtgaggaggggggcgcccgcgggaggcgccaaaacccggcgcggaggccttcg
aacggccactagtcaataatcaatgtcaacgcgtatatctggcccgtacatcgcgaagc
agcgcaaaacggatcctgcaggtatttgcggccgatgcatgggaagggcgatcggtgcg
ggcctcttcgctattacgccagctggcgaaaggggatgtgctgcaaggcgattaagtt
gggtaacgccagggttttcccagtcacgacgttgtaaaacgacggccagtgagcgcgcc
agcgctcgttcaaacctcccgcttcaaaatggagaccctgcgtgctcactcgggccccc
ccccccagcgagaccacatggtgtcgcaaaatgtcgcaaaacactcacgtgacctctaa
tacaggacctccctaaccctatgacgtaattcacgtcacgactccacccctccgcccgc
tgtttaaacgcccgggctgtagttaatgattaacccgccatgctacttatctacgtagc
catggaaactagataagaaagaaatacgcagagaccaaagttcaactgaaacgaattaa
acggtttattgattaacaagcaattacagattacgagtcaggtatctggtgccaatggg
gcgaggctctgaatacacgccattagtgtccacagtaaagtccacattaacagacttgt
tgtagttggaagtgtactgaatttcgggattccagcgtttgctgttttccttctgcagc
tcccactcgatctccacgctgacctgtcccgtggagtactgtgtgatgaaggaagcaaa
ctttgccgcactgaaggtggtcgaaggattcgcaggtaccggggtgttcttgatgagaa
tctgtggaggagggtgtttaagtccgaatccacccatgaggggagaggggtgaaaatgt
ccgtccgtgtgtggaatctttgcccagatgggcccctgaaggtacacatctctgtcctg
ccagaccatgcctggaagaacgccttgtgtgttgacatctgcggtagctgcttgtctgt
tgcctctctggaggttggtagatacagaaccatactgctccgtagccacgggattggtt
gtcctgatttcctcttcgtctgtaatcatgacctttcaatgtccacatttgttttctc
tgagccttgcttcccaaagatgagaaccccgctctgaggaaaaaactttcttcatcgt
ccttgtggcttgccatggccgggcccggattcaccagagagtctctgccattgaggtgg
tacttggtagctccagtccacgagtattcactgttgttgttatccgcagatgtctttga
tactcgctgctggcggtaacagggtccaggaagccagttcctagactggtcccgaatgt
cactcgctccggcctgagaaaactgaagccttgactgcgtggtggttccacttggagtg
tttgttctgctcaagtaatacaggtactggtcgatgagaggattcatgagacggtccag
actctggctgtgagcgtagctgctgtggaaaggaacgtcctcaaaagtgtagctgaagg
taaagttgtttccggtacgcagcatctgagaaggaaagtactccaggcagtaaaatgaa
gagcgtcctactgcctgactcccgttgttcagggtgaggtatccatactgtggcaccat
gaagacgtctgctgggaacggcgggaggcatccttgatgcgccgagccgaggacgtacg
ggagctggtactccgagtcagtaaacacctgaaccgtgctggtaaggttattggcaatc
gtcgtcgtaccgtcattctgcgtgacctctttgacttgaatgttaaagagcttgaagtt
gagtctcttgggtcggaatccccagttgttgttgatgagtctttgccagtcacgtggtg
aaaagtggcagtggaatctgttgaagtcaaaataccccaaggggtgctgtagccaaag
tagtgattgtcgttcgaggctcctgattggctggaaatttgtttgtagaggtggttgtt
gtaggtgggcagggcccaggttcgggtgctggtggtgatgactctgtcgcccatccatg
tggaatcgcaatgccaatttcccgaggaattacccactccgtcggcgccctcgttattg
tctgccattggtgcgccactgcctgtagccatcgtattagttcccagaccagaggggc
tgctggtggctgtccgagaggctgggggtcaggtactgagtctgcgtctccagtctgac

caaaattcaatctttttcttgcaggctgctggcccgcctttccggttcccgaggaggag
tctggctccacaggagagtgctctaccggcctctttttcccggagccgtcttaacagg
ttcctcaaccaggcccagaggttcaagaaccctctttttcgcctggaagactgctcgtc
cgaggttgcccccaaaagacgtatcttctttaaggcgctcctgaaactccgcgtcggcg
tggttgtacttgaggtacgggttgtctccgctgtcgagctgccggtcgtaggctttgtc
gtgctcgagggccgcggcgtctgcctcgttgaccggctctcccttgtcgagtccgttga
agggtccgaggtacttgtacccaggaagcacaagacccctgctgtcgtccttatgccgc
tctgcgggctttggtggtggtgggccaggtttgagcttccaccactgtcttattccttc
agagagagtgtcctcgagccaatctggaagataaccatcggcagccatacctgatttaa
atcatttattgttcaaagatgcagtcatccaaatccacattgaccagatcgcaggcagt
gcaagcgtctggcacctttcccatgatatgatgaatgtagcacagtttctgatacgcct
ttttgacgacagaaacgggttgagattctgacacgggaaagcactctaaacagtctttc
tgtccgtgagtgaagcagatatttgaattctgattcattctctcgcattgtctgcaggg
aaacagcatcagattcatgcccacgtgacgagaacatttgttttggtacctgtctgcgt
agttgatcgaagcttccgcgtctgacgtcgatggctgcgcaactgactcgcgcacccgt
ttgggctcacttatatctgcgtcactgggggcgggtcttttcttggctccacccttttt
gacgtagaattcatgctccacctcaaccacgtgatcctttgcccaccggaaaaagtctt
tgacttcctgcttggtgaccttcccaaagtcatgatccagacggcgggtgagttcaaat
ttgaacatccggtcttgcaacggctgctggtgttcgaaggtcgttgagttcccgtcaat
cacggcgcacatgttggtgttggaggtgacgatcacgggagtcgggtctatctgggccg
aggacttgcatttctggtccacgcgcaccttgcttcctccgagaatggctttggccgac
tccacgaccttggcggtcatcttcccctcctcccaccagatcaccatcttgtcgacaca
gtcgttgaagggaaagttctcattggtccagtttacgcacccgtagaagggcacagtgt
gggctatggcctccgcgatgttggtcttcccggtagttgcaggcccaaacagccagatg
gtgttcctcttgccgaacttttcgtggcccatcccagaaagacggaagccgcatattg
gggatcgtacccgtttagttccaaaatttataaatccgattgctggaaatgtcctcca
cgggctgctggcccaccaggtagtcgggggcggttttagtcaggctcataatctttccc
gcattgtccaaggcagccttgatttgggaccgcgagttggaggccgcattgaaggagat
gtatgaggcctggtcctcctggatccactgcttctccgaggtaatccccttgtccacga
gccacccgaccagctccatgtacctggctgaagttttgatctgatcaccggcgcatca
gaattgggattctgattctctttgttctgctcctgcgtctgcgacacgtgcgtcagatg
ctgcgccaccaaccgtttacgctccgtgagattcaaacaggcgcttaaatactgttcca
tattagtccacgcccactggagctcaggctgggttttggggagcaagtaattggggatg
tagcactcatccaccaccttgttcccgcctccggcgccatttctggtctttgtgaccgc
gaaccagtttggcaaagtcggctcgatcccgcggtaaattctctgaatcagttttcgc
gaatctgactcaggaaacgtcccaaaaccatggatttcaccccggtggtttccacgagc
acgtgcatgtggaagtagctctctcccttctcaaattgcacaaagaaaagggcctccgg
ggccttactcacacggcgccattccgtcagaaagtcgcgctgcagcttctcggccacgg
tcaggggtgcctgctcaatcagattcagatccatgtcagaatctggcggcaactcccat
tccttctcggccacccagttcacaaagctgtcagaaatgccgggcagatgctcgtcaag
gtcgctggggaccttaatcacaatctcgtaaaaccccggcatggcggctgcgcagatca
gaagttcctatactttctagagaataggaacttcggaataggaacttctgatcttccgg
gggatccactagttctagagcggccgcggtccgtatactccggaatattaatagatcat
ggagataattaaaatgataaccatctcgcaaataaataagtattttactgttttcgtaa
cagttttgtaataaaaaaacctataaatattccggattattcataccgtcccaccatcg
ggcgcgaactcctaaaaaaaccgccaccatgaagtgcctttgtacttagccttttttatt
cattggggtgaattgcaagttcaccatagttttccacacaaccaaaaaggaaactgga
aaaatgttccttctaattaccattattgcccgtcaagctcagatttaaattggcataat

gacttaataggcacagccttacaagtcaaaatgcccaagagtcacaaggctattcaagc
agacggttggatgtgtcatgcttccaaatgggtcactacttgtgatttccgctggtatg
gaccgaagtatataacacattccatccgatccttcactccatctgtagaacaatgcaag
gaaagcattgaacaaacgaaacaaggaacttggctgaatccaggcttccctcctcaaag
ttgtggatatgcaactgtgacggatgccgaagcagtgattgtccaggtgactcctcacc
atgtgctggttgatgaatacacaggagaatgggttgattcacagttcatcaacggaaaa
tgcagcaattacatatgccccactgtccataactctacaacctggcattctgactataa
ggtcaaagggctatgtgattctaacctcatttccatggacatcaccttcttctcagagg
acggagagctatcatccctgggaaaggagggcacagggttcagaagtaactactttgct
tatgaaactggaggcaaggcctgcaaaatgcaatactgcaagcattggggagtcagact
cccatcaggtgtctggttcgagatggctgataaggatctctttgctgcagccagattcc
ctgaatgcccagaagggtcaagtatctctgctccatctcagacctcagtggatgtaagt
ctaattcaggacgttgagaggatcttggattattccctctgccaagaaacctggagcaa
aatcagagcgggtcttccaatctctccagtggatctcagctatcttgctcctaaaaacc
caggaaccggtcctgctttcaccataatcaatggtaccctaaaatactttgagaccaga
tacatcagagtcgatattgctgctccaatcctctcaagaatggtcggaatgatcagtgg
aactaccacagaaagggaactgtgggatgactgggcaccatatgaagacgtggaaattg
gacccaatggagttctgaggaccagttcaggatataagtttcctttatacatgattgga
catggtatgttggactccgatcttcatcttagctcaaaggctcaggtgttcgaacatcc
tcacattcaagacgctgcttcgcaacttcctgatgatgagagtttatttttggtgata
ctgggctatccaaaaatccaatcgagcttgtagaaggttggttcagtagttggaaaagc
tctattgcctctttttctttatcatagggttaatcattggactattcttggttctccg
agttggtatccatctttgcattaaattaaagcacaccaagaaaagacagatttatacag
acatagagatgaaccgacttggaaagtgataaggccaggccggccaagcttgtcgagaa
gtactagaggatcataatcagccataccacatttgtagaggttttacttgctttaaaaa
acctcccacacctcccccctgaacctgaaacataaaatgaatgcaattgttgttgttaac
ttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaa
taaagcatttttttcactgcattctagttgtggtttgtccaaactcatcaatgtatctt
atcatgtctggatctgatcactgcttgagcctaggagatccgaaccagataagtgaaat
ctagttccaaactattttgtcatttttaattttcgtattagcttacgacgctacaccca
gttcccatctattttgtcactcttccctaaataatccttaaaaactccatttccacccc
tcccagttcccaactattttgtccgcccacagcggggcatttttcttcctgttatgttt
ttaatcaaacatcctgccaactccatgtgacaaaccgtcatcttcggctacttt

The sequence for RC2-454TAG in this plasmid is identical except the sequence for *Cap* has the same substitution as in pIDT-MbPylRS-RC2-454TAG.

pIDTSMART-CMV-CAP-CSS



Sequence color-coding key

Feature	Color
CMV promoter	text
T7 promoter	text
Transcription start site	text
Consensus splice donor site	text
Cap	text
Bgh polyA	text
KanR	text

Sequence: wild type Cap

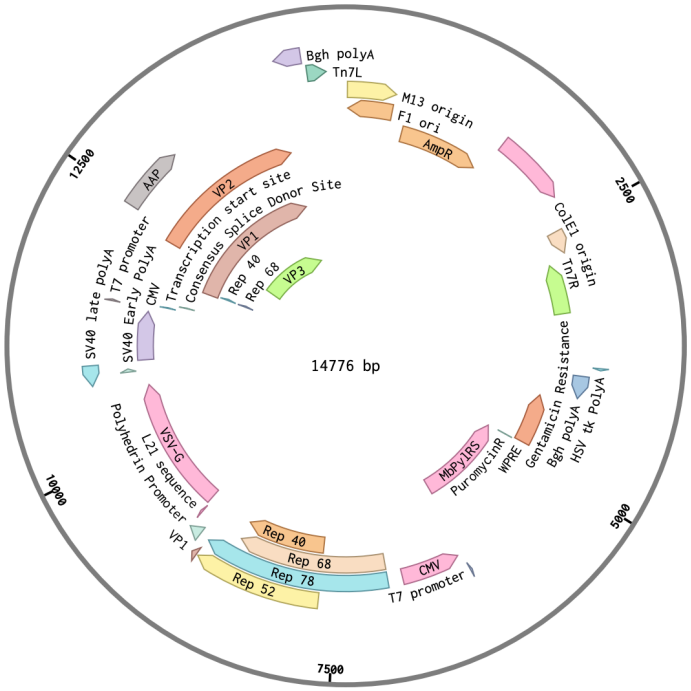
cccgtgtaaaacgacggccagtttatctagtcagcttgattctagctgatcgtggaccg
gaaggtgagccagtgagttgattgcagtccagttacgctggagtctgaggctcgtcctg
aatgatatgcgaccgccggagggttgcgtttgagacgggcgacagatccagtcgcgctg
ctctcgtcgatccgctagggcggccgcaaatacctgcaggatccgttttgcgctgcttc
gcgatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaa
tcaattacggggtcattagttcatagcccatatatggagttccgcgttacataacttac
ggtaaatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatga
cgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactat
ttacggtaaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccc

tattgacgtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttat
gggactttcctacttggcagtacatctacgtattagtcatcgctattaccatggtgatg
cggttttggcagtacatcaatgggcgtggatagcggtttgactcacggggatttccaag
tctccaccccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttc
caaaatgtcgtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgg
gaggtctatataagcagagctctctggctaactagagaacccactgcttactggcttat
cgaaattaatacgactcactatagggagacccaagctggctagcagttgcgcagccatc
gacgtcagacgcggaagcttcgatcaactacgcagacaggtaagtaaacaaatgttctc
gtcacgtgggcatgaatctgatgctgtttccctgcagacaatgcgagagaatgaatcag
aattcaaatatctgcttcactcacggacagaaagactgtttagagtgctttcccgtgtc
agaatctcaacccgtttctgtcgtcaaaaaggcgtatcagaaactgtgctacattcatc
atatcatgggaaaggtgccagacgcttgcactgcctgcgatctggtcaatgtggatttg
gatgactgcatctttgaacaataaatgatttaaatcaggtatggctgccgatggttatc
ttccagattggctcgaggacactctctctgaaggaataagacagtggtggaagctcaaa
cctggcccaccaccaccaaagcccgcagagcggcataaggacgacagcaggggtcttgt
gcttcctgggtacaagtacctcggacccttcaacggactcgacaagggagagccggtca
acgaggcagacgccgcggcccctcgagcacgacaaagcctacgaccggcagctcgacagc
ggagacaacccgtacctcaagtacaaccacgccgacgcggagtttcaggagcgccttaa
agaagatacgtcttttgggggcaacctcggacgagcagtcttccaggcgaaaaagaggg
ttcttgaacctctgggcctggttgaggaacctgttaagacggctccgggaaaaaagagg
ccggtagagcactctcctgtggagccagactcctcctcgggaaccggaaaggcgggcca
gcagcctgcaagaaaaagattgaattttggtcagactggagacgcagactcagtacctg
acccccagcctctcggacagccaccagcagccccctctggtctgggaactaatacgatg
gctacaggcagtggcgcaccaatggcagacaataacgagggcgccgacggagtgggtaa
ttcctcgggaaattggcattgcgattccacatggatgggcgacagagtcatcaccacca
gcacccgaacctgggccctgcccacctacaacaaccacctctacaaacaaatttccagc
caatcaggagcctcgaacgacaatcactactttggctacagcacccttgggggtattt
tgacttcaacagattccactgccacttttcaccacgtgactggcaaagactcatcaaca
acaactggggattccgacccaagagactcaacttcaagctctttaacattcaagtcaaa
gaggtcacgcagaatgacggtacgacgacgattgccaataaccttaccagcacggttca
ggtgtttactgactcggagtaccagctcccgtacgtcctcggctcggcgcatcaaggat
gcctcccgccgttcccagcagacgtcttcatggtgccacagtatggatacctcaccctg
aacaacgggagtcaggcagtaggacgctcttcattttactgcctggagtactttccttc
tcagatgctgcgtaccggaaacaactttaccttcagctacacttttgaggacgttcctt
tccacagcagctacgctcacagccagagtctggaccgtctcatgaatcctctcatcgac
cagtacctgtattacttgagcagaacaaacactccaagtggaaccaccacgcagtcaag
gcttcagttttctcaggccggagcgagtgacattcgggaccagtctaggaactggcttc
ctggaccctgttaccgccagcagcgagtatcaaagacatctgcggataacaacaacagt
gaatactcgtggactggagctaccaagtaccacctcaatggcagagactctctggtgaa
tccgggcccggccatggcaagccacaaggacgatgaagaaaagtttttttcctcagagcg
gggttctcatctttgggaagcaaggctcagagaaaacaaatgtggacattgaaaaggtc
atgattacagacgaagaggaaatcaggacaaccaatcccgtggctacggagcagtatgg
ttctgtatctaccaacctccagagaggcaacagacaagcagctaccgcagatgtcaaca
cacaaggcgttcttccaggcatggtctggcaggacagagatgtgtaccttcaggggccc
atctgggcaaagattccacacacggacggacattttcacccctctcccctcatgggtgg
attcggacttaaacaccctcctccacagattctcatcaagaacaccccggtacctgcga
atccttcgaccaccttcagtgcggcaaagtttgcttccttcatcacacagtactccacg
ggacaggtcagcgtggagatcgagtgggagctgcagaaggaaaacagcaaacgctggaa

tcccgaaattcagtacacttccaactacaacaagtctgttaatgtggactttactgtgg
acactaatggcgtgtattcagagcctcgccccattggcaccagatacctgactcgtaat
ctgtaattgctgtcgactttaactcgagtctagagggcccgtttaaacccgctgatcag
cctcgactgtgccttctagttgccagccatctgttgtttgcccctccccgtgccttcc
ttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgcatc
gcattgtctgagtaggtgtcattctattctgggggtggggtggggcaggacagcaagg
gggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggcttct
gaggcggaaagaaccctaggggtgcgagcggatcgagcagtgtcgatcactactggacc
gcgagctgtgctgcgacccgtgatcttacggcattatacgtatgatcggtccacgatca
gctagattatctagtcagcttgatgtcatagctgtttcctgaggctcaatactgaccat
ttaaatcatacctgacctccatagcagaaagtcaaaagcctccgaccggaggcttttga
cttgatcggcacgtaagaggttccaactttcaccataatgaaataagatcactaccggg
cgtatttttgagttatcgagattttcaggagctaaggaagctaaaatgagccatattc
aacgggaaacgtcttgcttgaagccgcgattaaattccaacatggatgctgatttatat
gggtataaatgggctcgcgataatgtcgggcaatcaggtgcgacaatctatcgattgta
tgggaagcccgatgcgccagagttgtttctgaaacatggcaaaggtagcgttgccaatg
atgttacagatgagatggtcaggctaaactggctgacggaatttatgcctcttccgacc
atcaagcattttatccgtactcctgatgatgcatggttactcaccactgcgatcccagg
gaaaacagcattccaggtattagaagaatatcctgattcaggtgaaaatattgttgatg
cgctggcagtgttcctgcgccggttgcattcgattcctgtttgtaattgtccttttaac
ggcgatcgcgtatttcgtctcgctcaggcgcaatcacgaatgaataacggtttggttgg
tgcgagtgattttgatgacgagcgtaatggctggcctgttgaacaagtctggaaagaaa
tgcataaactcttgccattctcaccggattcagtcgtcactcatggtgatttctcactt
gataaccttatttttgacgaggggaaattaataggttgtattgatgttggacgagtcgg
aatcgcagaccgataccaggatcttgccatcctatggaactgcctcggtgagttttctc
cttcattacagaaacggctttttcaaaaatatggtattgataatcctgatatgaataaa
ttgcagtttcacttgatgctcgatgagttttctaatgaggacctaaatgtaatcacct
ggctcaccttcgggtgggcctttctgcgttgctggcgttttccataggctccgccccc
ctgacgagcatcacaaaaatcgatgctcaagtcagaggtggcgaaacccgacaggacta
taaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccct
gccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcata
gctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtg
cacgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtc
caacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagca
gagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctac
actagaagaacagtatttggtatctgcgctctgctgaagccagttacctcggaaaaaga
gttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttg
caagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgattttctacc
gaagaaaggccca

The sequence for RC2-454TAG in this plasmid is identical except the sequence for *Cap* has the same substitution as in pIDT-MbPylRS-RC2-454TAG.

pAcBac1-Rep-CMV-Cap-CSS-CMV-MbPylRS



Sequence color-coding key

Feature	Color
M13 ori	text
AmpR	text
ColE1 ori	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
WPRE	text
MbPylRS	text
T7 promoter	text
CMV promoter	text
Rep	text
Polyhedrin promoter	text
L21 sequence	text
VSV-G	text
SV40 late polyA	text
CMV promoter	text
T7 promoter	text
Transcription start site	text
Consensus Splice Donor Site	text
Cap	text

Bgh polyA	text
Tn7L	text

Sequence: wild type Cap

ctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggt
tacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttct
tcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctc
cctttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgattaggg
tgatggttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttgg
agtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatc
tcggtctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaa
tgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgtttacaattt
caggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaata
cattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattg
aaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcgg
cattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaa
gatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatcct
tgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctat
gtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacac
tattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatgg
catgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggcca
acttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatg
ggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaa
cgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaa
ctggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggat
aaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataa
atctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggta
agccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacga
aatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagacca
agtttactcatatatactttagattgatttaaaacttcattttaatttaaaaggatct
aggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttc
cactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttttct
gcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgc
cggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagata
ccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagc
accgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgata
agtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcg
ggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaact
gagatacctacagcgtgagcattgagaaagcgccacgcttcccgaagggagaaaggcgg
acaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccaggg
ggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcg
atttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcct
ttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcag
ccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcctgatgcggt
atttctccttacgcatctgtgcggtatttcacaccgcagaccagccgcgtaacctggc
aaaatcggttacggttgagtaataaatggatgccctgcgtaagcgggtgtgggcggaca

ataaagtcttaaactgaacaaaatagatctaaactatgacaataaagtcttaaactaga
cagaatagttgtaaactgaaatcagtccagttatgctgtgaaaaagcatactggacttt
tgttatggctaaagcaaactcttcattttctgaagtgcaaattgcccgtcgtattaaag
aggggcgtggccaagggcatggtaaagactatattcgcggcgttgtgacaatttaccga
acaactccgcggccgggaagccgatctcggcttgaacgaattgttaggtggcggtactt
gggtcgatatcaaagtgcatcacttcttcccgtatgcccaactttgtatagagagccac
tgcgggatcgtcaccgtaatctgcttgcacgtagatcacataagcaccaagcgcgttgg
cctcatgcttgaggagattgatgagcgcggtggcaatgccctgcctccggtgctcgccg
gagactgcgagatcatagatatagatctcactacgcggctgctcaaacctgggcagaac
gtaagccgcgagagcgccaacaaccgcttcttggtcgaaggcagcaagcgcgatgaatg
tcttactacggagcaagttcccgaggtaatcggagtccggctgatgttgggagtaggtg
gctacgtctccgaactcacgaccgaaaagatcaagagcagcccgcatggatttgacttg
gtcagggccgagcctacatgtgcgaatgatgcccatacttgagccacctaactttgttt
tagggcgactgccctgctgcgtaacatcgttgctgctgcgtaacatcgttgctgctcca
taacatcaaacatcgacccacggcgtaacgcgcttgctgcttggatgcccgaggcatag
actgtacaaaaaaacagtcataacaagccatgaaaaccgccactgcgccgttaccaccg
ctgcgttcggtcaaggttctggaccagttgcgtgagcgcatacgctacttgcattacag
tttacgaaccgaacaggcttatgtcaactgggttcgtgccttcatccgtttccacggtg
tgcgtcacccggcaaccttgggcagcagcgaagtcgaggcatttctgtcctggctggcg
aacgagcgcaaggtttcggtctccacgcatcgtcaggcattggcggccttgctgttctt
ctacggcaaggtgctgtgcacggatctgccctggcttcaggagatcggtagacctcggc
cgtcgcggcgcttgccggtggtgctgaccccggatgaagtggttcgcatcctcggtttt
ctggaaggcgagcatcgtttgttcgcccaggactctagctatagttctagtggttggct
acgtacccgtagtggctatggcagggcttgcgcttaatgcgccgctacagggcgcgtgg
ggataccccctagagccccagctggttctttccgcctcagaagccatagagcccaccgc
atccccagcatgcctgctattgtcttcccaatcctccccttgctgtcctgccccaccc
cacccccagaatagaatgacacctactcagacaatgcgatgcaatttcctcattttat
taggaaaggacagtgggagtggcaccttccagggtcaaggaaggcacgggggaggggca
aacaacagatggctggcaactagaaggcacagtcgaggctgatcagcgggtttaaacgg
gccctctagactcgagttaaagtcgacgcggggaggcggcccaaagggagatccgactc
gtctgagggcgaaggcgaagacgcggaagaggccgcagagccggcagcaggccgcggga
aggaaggtccgctggattgagggccgaagggacgtagcagaaggacgtcccgcgcagaa
tccaggtggcaacacaggcgagcagccaaggaaaggacgatgatttccccgacaacacc
acggaattgtcagtgcccaacagccgagcccctgtccagcagcgggcaaggcaggcggc
gatgagttccgccgtggcaatagggaggggaaagcgaaagtcccggaaaggagctgac
aggtggtggcaatgccccaaccagtgggggttgcgtcagcaaacacagtgcacaccacg
ccacgttgcctgacaacgggccacaactcctcataaagagacagcaaccaggatttata
caaggaggagaaaatgaaagccatacgggaagcaatagcatgatacaaaggcattaaag
cagcgtatccacatagcgtaaaaggagcaacatagttaagaataccagtcaatctttca
caaattttgtaatccagaggttgattgtcgacttaacgcgttgaattcttacagattgg
ttgaaatcccattatagtaagattcggaccttgatgccctcttaatgttttaaagccg
tgcataaccttgagcaagcgttcaagaccaaaacctgcacctatccatggtttgtcaat
accccattctctatcaagagaaactggcccgacgactgccgaagaaagctccaggtccc
cgtgcattatatcaagagtatccccaaagaccatacaggaatctcctacgatttcgaag
tcgatttccagatagtccagaaactctttgatgagagcttcaagattttcccgagtaca
tcccgaacccatctgacagaagttcaccatagtaaattcttccaggtgctctttgccgt
cagactctttccggtaacaaggtccgacttcgaaaatttttattgggcctggtaaaatc
ctatcgagtttcgcagatagttgtaaagagtcggggcaagcattggcctcaagcagag

attttatccacccggaagatctgttttgaaagttcagtatcattattaatacccattc
tctccacgtattccgccggaataaggataggagactttatctccagaaaaccccggtct
acgaaaaatttcgtaatatcacgttcgagtttaccgaggtagtcttctctatcattggt
atagagccgctgaaaatcgtttttttcttcttgtcacaagttcaggctcaagttccctga
aaggctttgccatatttagagaaattttatcctctggacttaagagagcctcaacccta
tcaagctggcttcttgtaagtgaaggagcaggagccgatgcgggaacagacgaatttgg
agttgattttgcaggcgaaggtacagatctggatgtgttcgttgatgcctttgcagaaa
cagaatttccagaggcttcggagcccttgaaactgatttcggcatagcttttttgacc
tttggagcagaaactaccctaactttcacactgtttttgctttcggttgatcttgtgag
aaaattattgatatcctcgtccgaaaccctacatcgtttgcaggttttctgtacttat
gatgtctgaatgctctggctgttctacaactcctggaattattcacaacaagatggtct
ccacacgccatttcaatgtatattttacttcttgagacctcatggtgcttgattttgtg
gagcgtgccagtcctggacatccagagcccggtcgcagatattaaaacatctaatggtt
ttttatccatggtggcgctagccagcttgggtctccctatagtgagtcgtattaatttc
gataagccagtaagcagtgggttctctagttagccagagagctctgcttatatagacct
cccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacatttt
ggaaagtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggag
acttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaacc
gcatcaccatggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcc
cataaggtcatgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaat
aggggggcgtacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgta
aatagtccacccattgacgtcaatggaaagtccctattggcgttactatgggaacatac
gtcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttac
cgtaagttatgtaacgcggaactccatatatgggctatgaactaatgaccccgtaattg
attactattaataactagtcaataatcaatgtcaacgcgtatatctggcccgtacatcg
cgaagcagcgcaaaacggatcctgcagggctctagaactagtggatccccgggaagatc
agaagttcctattccgaagttcctattctctagaaagtataggaacttctgatctgcgc
agccgccatgccggggttttacgagattgtgattaaggtccccagcgaccttgacgagc
atctgcccggcatttctgacagctttgtgaactgggtggccgagaaggaatgggagttg
ccgccagattctgacatggatctgaatctgattgagcaggcacccctgaccgtggccga
gaagctgcagcgcgactttctgacggaatggcgccgtgtgagtaaggccccggaggccc
ttttctttgtgcaatttgagaagggagagagctacttccacatgcacgtgctcgtggaa
accaccggggtgaaatccatggttttgggacgtttcctgagtcagattcgcgaaaaact
gattcagagaatttaccgcgggatcgagccgactttgccaaactggttcgcggtcacaa
agaccagaaatggcgccggaggcgggaacaaggtggtggatgagtgctacatccccaat
tacttgctccccaaaacccagcctgagctccagtgggcgtggactaatatggaacagta
tttaagcgcctgtttgaatctcacggagcgtaaacggttggtggcgcagcatctgacgc
acgtgtcgcagacgcaggagcagaacaaagagaatcagaatcccaattctgatgcgccg
gtgatcagatcaaaaacttcagccaggtacatggagctggtcgggtggctcgtggacaa
ggggattacctcggagaagcagtggatccaggaggaccaggcctcatacatctccttca
atgcggcctccaactcgcggtcccaaatcaaggctgccttggacaatgcgggaaagatt
atgagcctgactaaaaccgcccccgactacctggtgggccagcagcccgtggaggacat
ttccagcaatcggatttataaaattttggaactaaacgggtacgatccccaatatgcgg
cttccgtctttctgggatgggccacgaaaaagttcggcaagaggaacaccatctggctg
tttgggcctgcaactaccgggaagaccaacatcgcggaggccatagcccacactgtgcc
cttctacgggtgcgtaaactggaccaatgagaactttcccttcaacgactgtgtcgaca
agatggtgatctggtgggaggaggggaagatgaccgccaaggtcgtggagtcggccaaa
gccattctcggaggaagcaaggtgcgcgtggaccagaaatgcaagtcctcggcccagat

agacccgactcccgtgatcgtcacctccaacaccaacatgtgcgccgtgattgacggga
actcaacgaccttcgaacaccagcagccgttgcaagaccggatgttcaaatttgaactc
acccgccgtctggatcatgactttgggaaggtcaccaagcaggaagtcaaagacttttt
ccggtgggcaaaggatcacgtggttgaggtggagcatgaattctacgtcaaaaagggtg
gagccaagaaaagacccgcccccagtgacgcagatataagtgagcccaaacgggtgcgc
gagtcagttgcgcagccatcgacgtcagacgcggaagcttcgatcaactacgcagacag
gtaccaaaacaaatgttctcgtcacgtgggcatgaatctgatgctgtttccctgcagac
aatgcgagagaatgaatcagaattcaaatatctgcttcactcacggacagaaagactgt
ttagagtgctttcccgtgtcagaatctcaacccgtttctgtcgtcaaaaaggcgtatca
gaaactgtgctacattcatcatatcatgggaaaggtgccagacgcttgcactgcctgcg
atctggtcaatgtggatttggatgactgcatctttgaacaataaatgatttaaatcagg
tatggctgccgatggttatcttccagattggctcgaggacactctctctgagcggccgc
ggtccgtatactccggaatattaatagatcatggagataattaaaatgataaccatctc
gcaaataaataagtattttactgttttcgtaacagttttgtaataaaaaaacctataaa
tattccggattattcataccgtcccaccatcgggcgcgaactcctaaaaaaccgccacc
atgaagtgccttttgtacttagccttttattcattggggtgaattgcaagttcaccat
agttttccacacaaccaaaaaggaaactggaaaaatgttccttctaattaccattatt
gcccgtcaagctcagatttaaattggcataatgacttaataggcacagccttacaagtc
aaaatgcccaagagtcacaaggctattcaagcagacggttggatgtgtcatgcttccaa
atgggtcactacttgtgatttccgctggtatggaccgaagtatataacacattccatcc
gatccttcactccatctgtagaacaatgcaaggaaagcattgaacaaacgaaacaagga
acttggctgaatccaggcttccctcctcaaagttgtggatatgcaactgtgacggatgc
cgaagcagtgattgtccaggtgactcctcaccatgtgctggttgatgaatacacaggag
aatgggttgattcacagttcatcaacggaaaatgcagcaattacatatgccccactgtc
cataactctacaacctggcattctgactataaggtcaaagggctatgtgattctaacct
catttccatggacatcaccttcttctcagaggacggagagctatcatccctgggaaagg
agggcacagggttcagaagtaactactttgcttatgaaactggaggcaaggcctgcaaa
atgcaatactgcaagcattggggagtcagactcccatcaggtgtctggttcgagatggc
tgataaggatctctttgctgcagccagattccctgaatgcccagaagggtcaagtatct
ctgctccatctcagacctcagtggatgtaagtctaattcaggacgttgagaggatcttg
gattattccctctgccaagaaacctggagcaaaatcagagcgggtcttccaatctctcc
agtggatctcagctatcttgctcctaaaaacccaggaaccggtcctgctttcaccataa
tcaatggtaccctaaaatactttgagaccagatacatcagagtcgatattgctgctcca
atcctctcaagaatggtcggaatgatcagtggaactaccacagaaagggaactgtggga
tgactgggcaccatatgaagacgtggaaattggacccaatggagttctgaggaccagtt
caggatataagtttcctttatacatgattggacatggtatgttggactccgatcttcat
cttagctcaaaggctcaggtgttcgaacatcctcacattcaagacgctgcttcgcaact
tcctgatgatgagagtttatttttggtgatactgggctatccaaaaatccaatcgagc
ttgtagaaggttggttcagtagttggaaaagctctattgcctctttttctttatcata
gggttaatcattggactattcttggttctccgagttggtatccatctttgcattaaatt
aaagcacaccaagaaaagacagatttatacagacatagagatgaaccgacttggaaagt
gataaggccaggccggccaagcttgtcgagaagtactagaggatcataatcagccatac
cacatttgtagaggttttacttgctttaaaaaacctcccacacctcccctgaacctga
aacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttataatggttac
aaataaagcaatagcatcacaaatttcacaaataaagcatttttttcactgcattctag
ttgtggtttgtccaaactcatcaatgtatcttatcatgtctggatctgatcactgcttg
agcctagttattaatagtaatcaattacggggtcattagttcatagcccatatatggag
ttccgcgttacataacttacggtaaatggcccgcctggctgaccgcccaacgacccccg

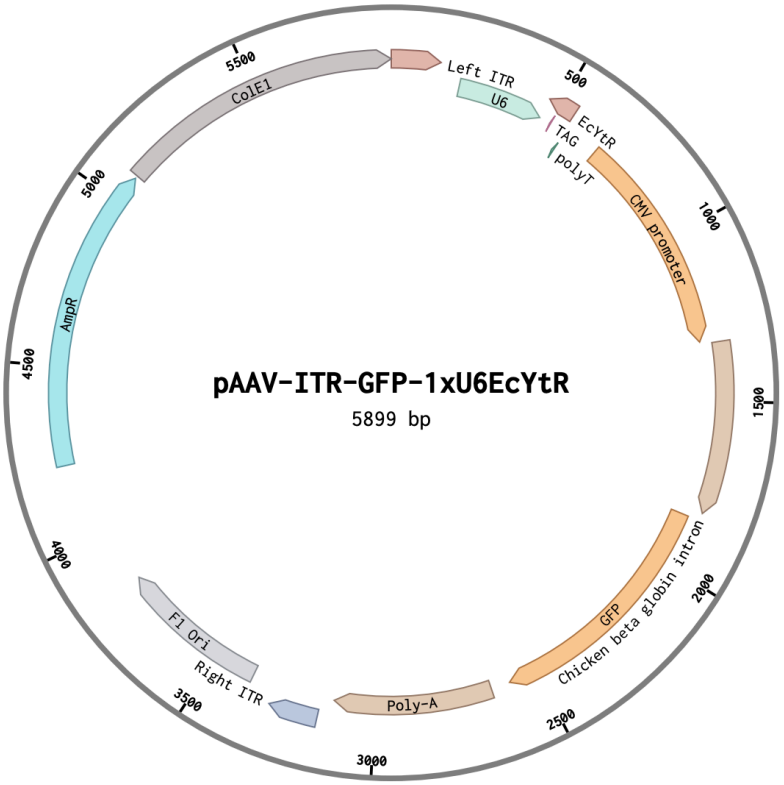
cccattgacgtcaataatgacgtatgttcccatagtaacgccaatagggactttccatt
gacgtcaatgggtggactatttacggtaaactgcccacttggcagtacatcaagtgtat
catatgccaagtacgccccctattgacgtcaatgacggtaaatggcccgcctggcatta
tgcccagtacatgaccttatgggactttcctacttggcagtacatctacgtattagtca
tcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggatagcggttt
gactcacggggatttccaagtctccaccccattgacgtcaatgggagtttgttttggca
ccaaaatcaacgggactttccaaaatgtcgtaacaactccgccccattgacgcaaatgg
gcggtaggcgtgtacggtgggaggtctatataagcagagctctctggctaactagagaa
cccactgcttactggcttatcgaaattaatacgactcactatagggagacccaagctgg
ctagcagttgcgcagccatcgacgtcagacgcggaagcttcgatcaactacgcagacag
gtaagtaaacaaatgttctcgtcacgtgggcatgaatctgatgctgtttccctgcagac
aatgcgagagaatgaatcagaattcaaatatctgcttcactcacggacagaaagactgt
ttagagtgctttcccgtgtcagaatctcaacccgtttctgtcgtcaaaaaggcgtatca
gaaactgtgctacattcatcatatcatgggaaaggtgccagacgcttgcactgcctgcg
atctggtcaatgtggatttggatgactgcatctttgaacaataaatgatttaaatcagg
tatggctgccgatggttatcttccagattggctcgaggacactctctctgaaggaataa
gacagtggtggaagctcaaacctggcccaccaccaccaaagcccgcagagcggcataag
gacgacagcaggggtcttgtgcttcctgggtacaagtacctcggacccttcaacggact
cgacaagggagagccggtcaacgaggcagacgccgcggccctcgagcacgacaaagcct
acgaccggcagctcgacagcggagacaacccgtacctcaagtacaaccacgccgacgcg
gagtttcaggagcgccttaaagaagatacgtcttttgggggcaacctcggacgagcagt
cttccaggcgaaaaagagggttcttgaacctctgggcctggttgaggaacctgttaaga
cggctccgggaaaaaagaggccggtagagcactctcctgtggagccagactcctcctcg
ggaaccggaaaggcgggccagcagcctgcaagaaaaagattgaattttggtcagactgg
agacgcagactcagtacctgaccccagcctctcggacagccaccagcagccccctctg
gtctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataacgag
ggcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatggatggg
cgacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaaccacc
tctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttggctac
agcacccctttgggggtattttgacttcaacagattccactgccacttttcaccacgtga
ctggcaaagactcatcaacaacaactggggattccgacccaagagactcaacttcaagc
tctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgccaat
aaccttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacgtcct
cggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtgccac
agtatggatacctcaccctgaacaacgggagtcaggcagtaggacgctcttcatttac
tgcctggagtactttcctttctcagatgctgcgtaccggaaacaactttaccttcagcta
cacttttgaggacgttcctttccacagcagctacgctcacagccagagtctggaccgtc
tcatgaatcctctcatcgaccagtacctgtattacttgagcagaacaaacactccaagt
ggaaccaccacgcagtcaaggcttcagttttctcaggccggagcgagtgacattcggga
ccagtctaggaactggcttcctggaccctgttaccgccagcagcgagtatcaaagacat
ctgcggataacaacaacagtgaatactcgtggactggagctaccaagtaccacctcaat
ggcagagactctctggtgaatccgggcccggccatggcaagccacaaggacgatgaaga
aaagttttttcctcagagcggggttctcatctttgggaagcaaggctcagagaaaacaa
atgtggacattgaaaaggtcatgattacagacgaagaggaaatcaggacaaccaatccc
gtggctacggagcagtatggttctgtatctaccaacctccagagaggcaacagacaagc
agctaccgcagatgtcaacacacaaggcgttcttccaggcatggtctggcaggacagag
atgtgtaccttcagggcccatctgggcaaagattccacacacggacggacattttcac
ccctctcccctcatgggtggattcggacttaaacaccctcctccacagattctcatcaa

gaacaccccggtacctgcgaatccttcgaccaccttcagtgcggcaaagtttgcttcct
tcatcacacagtactccacgggacaggtcagcgtggagatcgagtgggagctgcagaag
gaaaacagcaaacgctggaatcccgaaattcagtacacttccaactacaacaagtctgt
taatgtggactttactgtggacactaatggcgtgtattcagagcctcgccccattggca
ccagatacctgactcgtaatctgtaattgctgtcgactttaactcgagtctagagggcc
cgtttaaacccgctgatcagcctcgactgtgccttctagttgccagccatctgttgttt
gcccctcccccgtgccttccttgaccctggaaggtgccactcccactgtcctttcctaa
taaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctgggggggtgg
ggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatg
cggtgggctctatggcttctgaggcggaaagaaccctaggagatccgaaccagataagt
gaaatctagttccaaactattttgtcatttttaattttcgtattagcttacgacgctac
acccagttcccatctattttgtcactcttccctaaataatccttaaaaactccatttcc
acccctcccagttcccaactattttgtccgcccacagcggggcatttttcttcctgtta
tgttttaatcaaacatcctgccaactccatgtgacaaaccgtcatcttcggctacttt
ttctctgtcacagaatgaaaatttttctgtcatctcttcgttattaatgtttgtaattg
actgaatatcaacgcttatttgcagc

The sequence for RC2-454TAG in this plasmid is identical except the sequence for *Cap* has the same substitution as in pIDT-MbPylRS-RC2-454TAG.

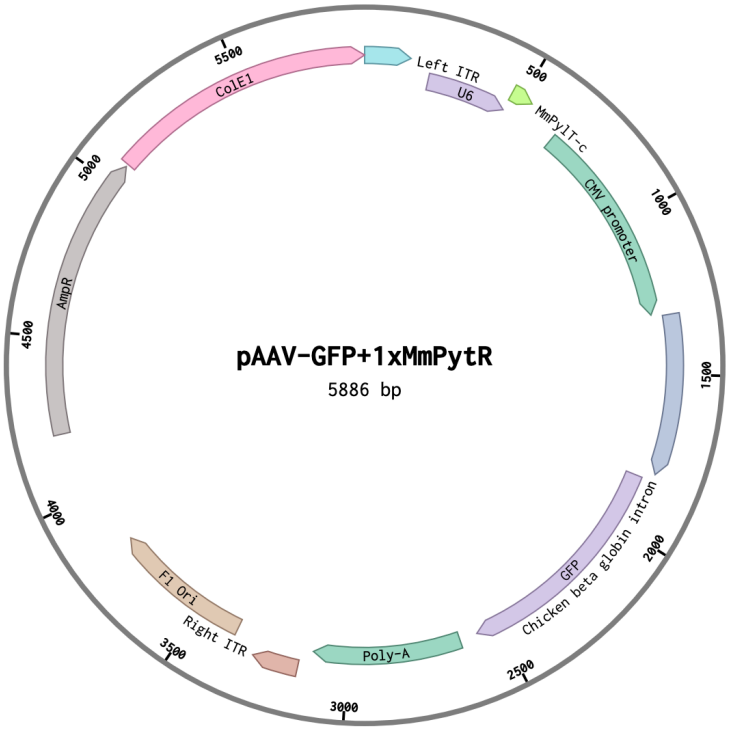
Chapter 3 Plasmids

pAAV-ITR-GFP+1xEcYtR



The sequence for this plasmid is identical to the pAAV-ITR-GFP-1xEcYtR found in the Chapter 2 section of this appendix.

pAAV-GFP+1xMmPytR



Sequence color-coding key

Feature	Color
Left ITR	text
U6 promoter	text
MmPytR	text
polyT	text
CMV promoter	text
Chicken beta globin intron	text
GFP	text
Poly-A	text
Right ITR	text
F1 Ori	text
Amp R	text
ColE1	text

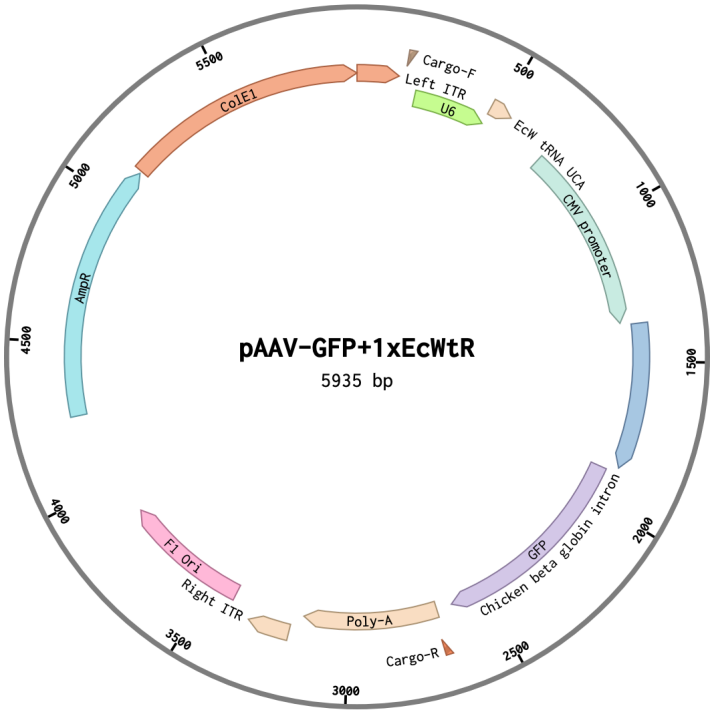
Sequence

cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgt
cgggcgacctttggtcgcccggcctcagtgagcgagcgagcgcgcagagagggagtggc
caactccatcactaggggttcctgcggccgcacgcgtctcgcggtccagtagtgatcga
cactgctcgatccgctcgaccccctaggtcgggcaggaagagggcctatttcccatgat
tccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttga

ctgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttggg
tagtttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttg
aaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacaccggaaacc
tgatcatgtagatcgaacggactctaaatccgttcagccgggttagattcccggggttt
ccgtttttttgctagcggatcgacgagagcagcgcgactggatctgtcgcccgtctcaaa
cgcaaccctccggcggtcgcatatcattcaggacgagcctcagactccagcgtaacacg
cgtggagctagttattaatagtaatcaattacggggtcattagttcatagcccatatat
ggagttccgcgttacataacttacggtaaatggcccgcctggctgaccgcccaacgacc
cccgcccattgacgtcaataatgacgtatgttcccatagtaacgtcaatagggactttc
cattgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtacatcaagt
gtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggcccgcctggc
attatgcccagtacatgaccttatgggactttcctacttggcagtacatctacgtatta
gtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggatagcg
gtttgactcacggggatttccaagtctccaccccattgacgtcaatgggagtttgtttt
gcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgccccattgacgcaaa
tgggcggtaggcgtgtacggtgggaggtctatataagcagagctcgtttagtgaaccgt
cagatcgcctggagacgccatccacgctgttttgacctccatagaagacaccgggaccg
atccagcctccgcggattcgaatcccggccgggaacggtgcattggaacgcggattccc
cgtgccaagagtgacgtaagtaccgcctatagagtctataggcccacaaaaaatgcttt
cttcttttaatatacttttttgtttatcttatttctaatactttccctaatctctttct
ttcagggcaataatgatacaatgtatcatgcctctttgcaccattctaaagaataacag
tgataatttctgggttaaggcaatagcaatatttctgcatataaatatttctgcatata
aattgtaactgatgtaagaggtttcatattgctaatagcagctacaatccagctaccat
tctgcttttatttatggttgggataaggctggattattctgagtccaagctaggccct
tttgctaatcatgttcatacctcttatcttcctcccacagctcctgggcaacgtgctgg
tctgtgtgctggcccatcactttggcaaagaattgggattcgaacatcgattgaattct
gaatggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctg
gacggcgacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccac
ctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctggc
ccaccctcgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgaccac
atgaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcac
catcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcg
acaccctggtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatc
ctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaa
gcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcg
tgcagctcgccgaccactaccagcagaacaccccatcggcgacggccccgtgctgctg
cccgacaaccactacctgagcacccagtccgccctgagcaaagaccccaacgagaagcg
cgatcacatggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacg
agctgtacaagtactcagatctcgagctcaagtaggatcctctagagtcgacctgcag
aagcttgcctcgagcagcgctgctcgagagatctacgggtggcatccctgtgacccctc
cccagtgcctctcctggccctggaagttgccactccagtgcccaccagccttgtcctaa
taaaattaagttgcatcattttgtctgactaggtgtccttctataatattatggggtgg
aggggggtggtatggagcaaggggcaagttgggaagacaacctgtagggcctgcggggt
ctattgggaaccaagctggagtgcagtggcacaatcttggctcactgcaatctccgcct
cctgggttcaagcgattctcctgcctcagcctcccgagttgttgggattccaggcatgc
atgaccaggctcagctaatttttgtttttttggtagagacggggtttcaccatattggc
caggctggtctccaactcctaatctcaggtgatctacccaccttggcctcccaaattgc
tgggattacaggcgtgaaccactgctcccttccctgtccttctgatttgtaggtaacc

acgtgcggaccgagcggccgcaggaacccctagtgatggagttggccactccctctctg
cgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgc
ccgggcggcctcagtgagcgagcgagcgcgcagctgcctgcagggcgcctgatgcggt
attttctccttacgcatctgtgcggtatttcacaccgcatacgtcaaagcaaccatagt
acgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgacc
gctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttcctttctcgc
cacgttcgccggctttccccgtcaagctctaaatcgggggctccctttagggttccgat
ttagtgctttacggcacctcgaccccaaaaaacttgatttgggtgatggttcacgtagt
gggccatcgccctgatagacggtttttcgccctttgacgttggagtccacgttctttaa
tagtggactcttgttccaaactggaacaacactcaaccctatctcgggctattcttttg
atttataagggattttgccgatttcggcctattggttaaaaaatgagctgatttaacaa
aaatttaacgcgaattttaacaaaatattaacgtttacaatttatggtgcactctcag
tacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctg
acgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtc
tccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaa
gggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttaga
cgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaa
atacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataata
ttgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttttg
cggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgct
gaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagat
ccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgc
tatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcata
cactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacgga
tggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcgg
ccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaac
atgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccatacc
aaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactat
taactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcg
gataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctga
taaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatg
gtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaa
cgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcaga
ccaagtttactcatatatactttagattgatttaaaacttcattttaatttaaaaggga
tctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcg
ttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttt
tctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtt
tgccggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcag
ataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgt
agcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcg
ataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcgg
tcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccga
actgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaagg
cggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttcca
gggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcg
tcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcgg
cctttttacggttcctggccttttgctggccttttgctcacatgt

pAAV-GFP+1xEcWtR



Sequence color-coding key

Feature	Color
Left ITR	text
U6 promoter	text
EcWtR	text
polyT	text
CMV promoter	text
Chicken beta globin intron	text
GFP	text
Poly-A	text
Right ITR	text
F1 Ori	text
Amp R	text
ColE1	text

Sequence

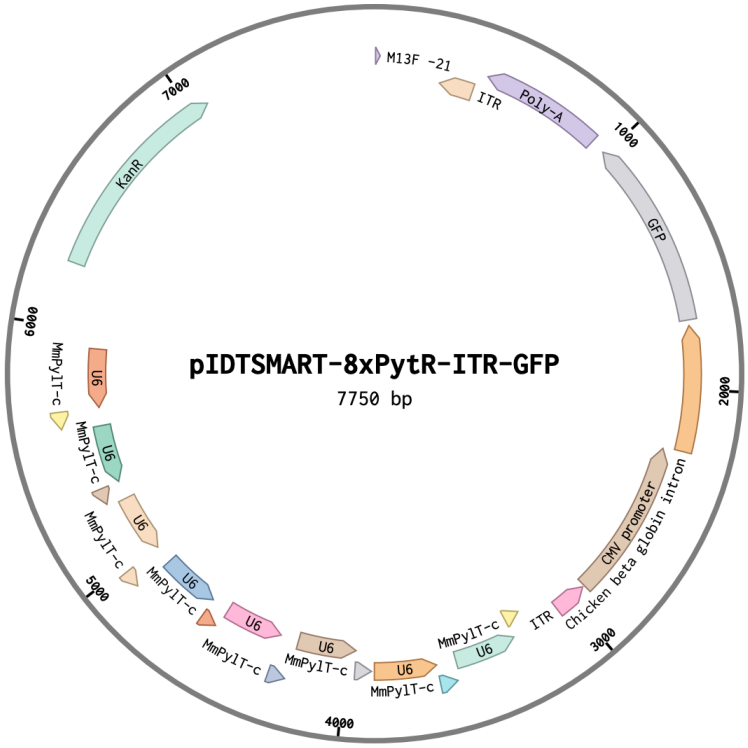
cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgt
cgggcgacctttggtcgcccggcctcagtgagcgagcgagcgcgcagagagggagtggc
caactccatcactaggggttcctgcggccgcacgcgtctcgcggtccagtagtgatcga
cactgctcgatccgctcgcacccctaggtcgggcaggaagagggcctatttcccatgat
tccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttga
ctgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttggg

tagtttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttg
aaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacaccaggggcg
tagttcaattggtagagcaccggtcttcaaaaccgggtgttgggagttcgagtctctcc
gcccctgccatttttgctagcggatcgacgagagcagcgcgactggatctgtcgcccg
tctcaaacgcaaccctccggcggtcgcatatcattcaggacgagcctcagactccagcg
taacacgcgtggagctagttattaatagtaatcaattacggggtcattagttcatagcc
catatatggagttccgcgttacataacttacggtaaatggcccgcctggctgaccgccc
aacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgtcaatagg
gactttccattgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtac
atcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggccc
gcctggcattatgcccagtacatgaccttatgggactttcctacttggcagtacatcta
cgtattagtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtg
gatagcggtttgactcacggggatttccaagtctccaccccattgacgtcaatgggagt
ttgttttgcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgccccattg
acgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcagagctcgtttagt
gaaccgtcagatcgcctggagacgccatccacgctgttttgacctccatagaagacacc
gggaccgatccagcctccgcggattcgaatcccggccgggaacggtgcattggaacgcg
gattccccgtgccaagagtgacgtaagtaccgcctatagagtctataggcccacaaaaa
atgctttcttctttaatatactttttgtttatcttatttctaatactttccctaatc
tctttctttcagggcaataatgatacaatgtatcatgcctctttgcaccattctaaaga
ataacagtgataatttctgggttaaggcaatagcaatatttctgcatataaatatttct
gcatataaattgtaactgatgtaagaggtttcatattgctaatagcagctacaatccag
ctaccattctgcttttatttatggttgggataaggctggattattctgagtccaagct
aggccctttgctaatcatgttcatacctcttatcttcctcccacagctcctgggcaac
gtgctggtctgtgtgctggcccatcactttggcaaagaattgggattcgaacatcgatt
gaattctgaatggtgagcaagggcgaggaggataacatggccatcatcaaggagttcat
gcgcttcaaggtgcacatggagggctccgtgaacggccacgagttcgagatcgagggcg
agggcgagggccgcccctacgagggcacccagaccgccaagctgaaggtgaccaagggt
ggcccctgcccttcgcctgggacatcctgtcccctcagttcatgtacggctccaaggc
ctacgtgaagcaccccgccgacatccccgactacttgaagctgtccttccccgagggct
tcaagtgggagcgcgtgatgaacttcgaggacggcggcgtggtgaccgtgacccaggac
tcctccctgcaggacggcgagttcatctacaaggtgaagctgcgcggcaccaacttccc
ctccgacggccccgtaatgcagaagaagaccatgggctgggaggcctcctccgagcgga
tgtaccccgaggacggcgccctgaagggcgagatcaagcagaggctgaagctgaaggac
ggcggccactacgacgctgaggtcaagaccacctacaaggccaagaagcccgtgcagct
gcccggcgcctacaacgtcaacatcaagttggacatcacctcccacaacgaggactaca
ccatcgtggaacagtacgaacgcgccgagggccgccactccaccggcggcatggacgag
ctgtacaaggatcctctagagtcgacctgcagaagcttgcctcgagcagcgctgctcg
agagatctacgggtggcatccctgtgacccctccccagtgcctctcctggccctggaag
ttgccactccagtgcccaccagccttgtcctaataaaattaagttgcatcattttgtct
gactaggtgtccttctataatattatggggtggaggggggtggtatggagcaaggggca
agttgggaagacaacctgtagggcctgcggggtctattgggaaccaagctggagtgcag
tggcacaatcttggctcactgcaatctccgcctcctgggttcaagcgattctcctgcct
cagcctcccgagttgttgggattccaggcatgcatgaccaggctcagctaatttttgtt
tttttggtagagacggggtttcaccatattggccaggctggtctccaactcctaatctc
aggtgatctacccaccttggcctcccaaattgctgggattacaggcgtgaaccactgct
cccttccctgtccttctgattttgtaggtaaccacgtgcggaccgagcggccgcaggaa
cccctagtgatggagttggccactccctctctgcgcgctcgctcgctcactgaggccgg

gcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgag
cgcgcagctgcctgcagggcgcctgatgcggtattttctccttacgcatctgtgcggt
atttcacaccgcatacgtcaaagcaaccatagtacgcgccctgtagcggcgcattaagc
gcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccctagcgcc
cgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaag
ctctaaatcgggggctcccttagggttccgatttagtgctttacggcacctcgacccc
aaaaaacttgatttgggtgatggttcacgtagtgggccatcgccctgatagacggtttt
tcgccctttgacgttggagtccacgttctttaatagtggactcttgttccaaactggaa
caacactcaaccctatctcgggctattcttttgatttataagggattttgccgatttcg
gcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaat
attaacgtttacaatttatggtgcactctcagtacaatctgctctgatgccgcatagt
taagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctc
ccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggtt
ttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttat
aggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcggggaaat
gtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcat
gagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattc
aacatttccgtgtcgcccttattccctttttgcggcattttgccttcctgtttttgct
cacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtggg
ttacatcgaactggatctcaacagcggtaagatccttgagagtttcgccccgaagaac
gttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtatt
gacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttga
gtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgca
gtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcgga
ggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttga
tcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgc
ctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagct
tcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcg
ctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggt
ctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatc
tacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagatagg
tgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttaga
ttgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataat
ctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtaga
aaagatcaaaggatcttcttgagatccttttttctgcgcgtaatctgctgcttgcaaa
caaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctt
tttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgta
gccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgc
taatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggac
tcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcac
acagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctat
gagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagg
gtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatag
tcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcagggg
ggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgc
tggccttttgctcacatgt

The mCherry sequence for the mCherry-containing version of this plasmid is:
atggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctgga
cggcgacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccacct
acggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctggccc
accctcgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgaccacat
gaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcacca
tcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgac
accctggtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcct
ggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaagc
agaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtg
cagctcgccgaccactaccagcagaacaccccatcggcgacggccccgtgctgctgcc
cgacaaccactacctgagcacccagtccgccctgagcaaagaccccaacgagaagcgcg
atcacatggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacgag
ctgtacaagtactcagatctcgagctcaagtag

pIDTSMART-8xPytR-ITR-GFP



Sequence color-coding key

Feature	Color
ITR	text
Poly-A	text
GFP	text
Chicken beta globin intron	text
CMV promoter	text
ITR	text
MmPylT-c	text
U6	text
KanR	text

Sequence

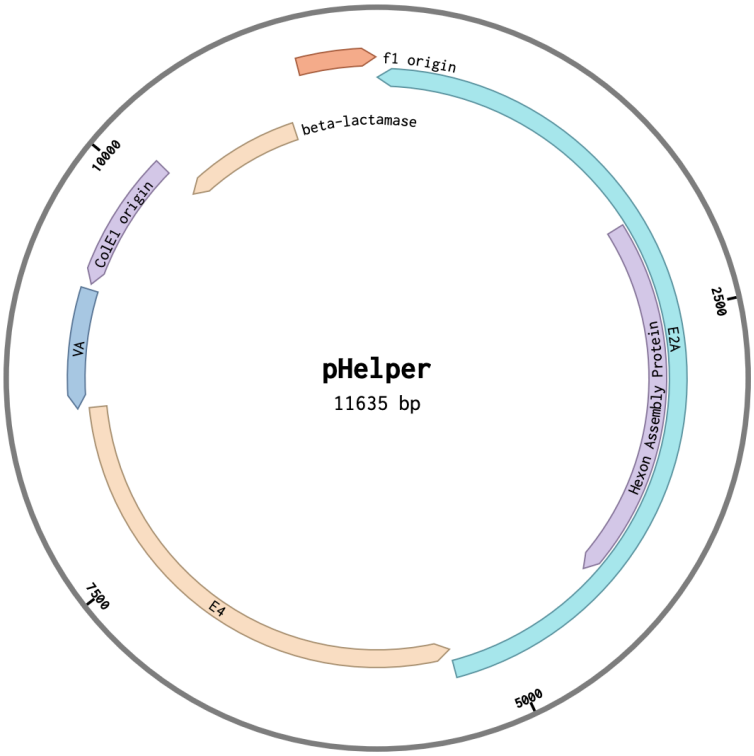
cccgtgtaaaacgacggccagtttatctagtcagcttgattctagctgatcgtggaccg
gaaggtgagccagtgagttgattgcagtccagttacgctggagtctgaggctcgtcctg
aatgatatgcgaccgccggagggttgcgtttgagacgggcgacagatccagtcgcgctg
ctctcgtcgatccgctaggactatggttgctttgacgtatgcggtgtgaaataccgcac
agatgcgtaaggagaaaataccgcatcaggcgcccctgcaggcagctgcgcgctcgctc
gctcactgaggccgcccgggcaaagcccgggcgtcgggcgacctttggtcgcccggcct
cagtgagcgagcgagcgcgcagagagggagtggccaactccatcactaggggttcctgc
ggccgctcggtccgcacgtggttacctacaaaatcagaaggacagggaagggagcagtg

gttcacgcctgtaatcccagcaatttgggaggccaaggtgggtagatcacctgagatta
ggagttggagaccagcctggccaatatggtgaaaccccgtctctaccaaaaaaacaaaa
attagctgagcctggtcatgcatgcctggaatcccaacaactcgggaggctgaggcagg
agaatcgcttgaacccaggaggcggagattgcagtgagccaagattgtgccactgcact
ccagcttggttcccaatagaccccgcaggccctacaggttgtcttcccaacttgcccct
tgctccataccacccccctccaccccataatattatagaaggacacctagtcagacaaa
atgatgcaacttaattttattaggacaaggctggtgggcactggagtggcaacttccag
ggccaggagaggcactggggaggggtcacagggatgccacccgtagatctctcgagcag
cgctgctcgaggcaagcttctgcaggtcgactctagaggatccctacttgagctcgaga
tctgagtacttgtacagctcgtccatgccgagagtgatcccggcggcggtcacgaactc
cagcaggaccatgtgatcgcgcttctcgttggggtctttgctcagggcggactgggtgc
tcaggtagtggttgtcgggcagcagcacggggccgtcgccgatgggggtgttctgctgg
tagtggtcggcgagctgcacgctgccgtcctcgatgttgtggcggatcttgaagttcac
cttgatgccgttcttctgcttgtcggccatgatatagacgttgtggctgttgtagttgt
actccagcttgtgccccaggatgttgccgtcctccttgaagtcgatgcccttcagctcg
atgcggttcaccagggtgtcgccctcgaacttcacctcggcgcgggtcttgtagttgcc
gtcgtccttgaagaagatggtgcgctcctggacgtagccttcgggcatggcggacttga
agaagtcgtgctgcttcatgtggtcggggtagcggctgaagcactgcacgccgtaggtc
agggtggtcacgagggtgggccagggcacgggcagcttgccggtggtgcagatgaactt
cagggtcagcttgccgtaggtggcatcgccctcgccctcgccggacacgctgaacttgt
ggccgtttacgtcgccgtccagctcgaccaggatgggcaccaccccggtgaacagctcc
tcgcccttgctcaccattcagaattcaatcgatgttcgaatcccaattctttgccaaag
tgatgggccagcacacagaccagcacgttgcccaggagctgtgggaggaagataagagg
tatgaacatgattagcaaaagggcctagcttggactcagaataatccagccttatccca
accataaaataaaagcagaatggtagctggattgtagctgctattagcaatatgaaacc
tcttacatcagttacaatttatatgcagaaatatttatatgcagaaatattgctattgc
cttaacccagaaattatcactgttattctttagaatggtgcaaagaggcatgatacatt
gtatcattattgccctgaaagaaagagattagggaaagtattagaaataagataaacaa
aaaagtatattaaaagaagaaagcatttttgtgggcctatagactctataggcggtac
ttacgtcactcttggcacggggaatccgcgttccaatgcaccgttcccggccgggattc
gaatccgcggaggctggatcggtcccggtgtcttctatggaggtcaaaacagcgtggat
ggcgtctccaggcgatctgacggttcactaaacgagctctgcttatatagacctcccac
cgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacattttggaaa
gtcccgttgattttggtgcaaaacaaactcccattgacgtcaatggggtggagacttgg
aaatccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaaccgcatca
ccatggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcccataag
gtcatgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaataggggg
cgtacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgtaaatact
ccacccattgacgtcaatggaaagtccctattgacgttactatgggaacatacgtcatt
attgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttaccgtaag
ttatgtaacgcggaactccatatatgggctatgaactaatgaccccgtaattgattact
attaataactagctccacgcgtgcggccgcaggaacccctagtgatggagttggccact
ccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgccc
gggctttgcccgggcggcctcagtgagcgagcgagcgcgcagctgcctgcaggacatgt
gagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgttttcc
cataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcg
aaacccgagctagcaaaaacggaaaccccgggaatctaacccggctgaacggatttag
agtccgttcgatctacatgatcaggtttccggtgtttcgtcctttccacaagatatata

aagccaagaaatcgaaatactttcaagttacggtaagcatatgatagtccattttaaaa
cataattttaaaactgcaaactacccaagaaattattactttctacgtcacgtattttg
tactaatatctttgtgtttacagtcaaattaattctaattatctctctaacagccttgt
atcgtatatgcaaatatgaaggaatcatgggaaataggccctcttcctgcccgacctag
caaaaaacggaaaccccgggaatctaacccggctgaacggatttagagtccgttcgatc
tacatgatcaggtttccggtgtttcgtcctttccacaagatatataaagccaagaaatc
gaaatactttcaagttacggtaagcatatgatagtccattttaaaacataatttaaaa
ctgcaaactacccaagaaattattactttctacgtcacgtattttgtactaatatcttt
gtgtttacagtcaaattaattctaattatctctctaacagccttgtatcgtatatgcaa
atatgaaggaatcatgggaaataggccctcttcctgcccgacctagcaaaaaacggaaa
ccccgggaatctaacccggctgaacggatttagagtccgttcgatctacatgatcaggt
ttccggtgtttcgtcctttccacaagatatataaagccaagaaatcgaaatactttcaa
gttacggtaagcatatgatagtccattttaaaacataatttaaaactgcaaactaccc
aagaaattattactttctacgtcacgtattttgtactaatatctttgtgtttacagtca
attaattctaattatctctctaacagccttgtatcgtatatgcaaatatgaaggaatc
atgggaaataggccctcttcctgcccgacctagcaaaaaacggaaaccccgggaatcta
acccggctgaacggatttagagtccgttcgatctacatgatcaggtttccggtgtttcg
tcctttccacaagatatataaagccaagaaatcgaaatactttcaagttacggtaagca
tatgatagtccattttaaaacataatttaaaactgcaaactacccaagaaattattac
tttctacgtcacgtattttgtactaatatctttgtgtttacagtcaaattaattctaat
tatctctctaacagccttgtatcgtatatgcaaatatgaaggaatcatgggaaataggc
cctcttcctgcccgacctagcaaaaaacggaaaccccgggaatctaacccggctgaacg
gatttagagtccgttcgatctacatgatcaggtttccggtgtttcgtcctttccacaag
atatataaagccaagaaatcgaaatactttcaagttacggtaagcatatgatagtccat
tttaaaacataatttaaaactgcaaactacccaagaaattattactttctacgtcacg
tattttgtactaatatctttgtgtttacagtcaaattaattctaattatctctctaaca
gccttgtatcgtatatgcaaatatgaaggaatcatgggaaataggccctcttcctgccc
gacctagcaaaaaacggaaaccccgggaatctaacccggctgaacggatttagagtccg
ttcgatctacatgatcaggtttccggtgtttcgtcctttccacaagatatataaagcca
agaaatcgaaatactttcaagttacggtaagcatatgatagtccattttaaaacataat
tttaaaactgcaaactacccaagaaattattactttctacgtcacgtattttgtactaa
tatctttgtgtttacagtcaaattaattctaattatctctctaacagccttgtatcgta
tatgcaaatatgaaggaatcatgggaaataggccctcttcctgcccgacctagcaaaaa
acggaaaccccgggaatctaacccggctgaacggatttagagtccgttcgatctacatg
atcaggtttccggtgtttcgtcctttccacaagatatataaagccaagaaatcgaaata
ctttcaagttacggtaagcatatgatagtccattttaaaacataatttaaaactgcaa
actacccaagaaattattactttctacgtcacgtattttgtactaatatctttgtgttt
acagtcaaattaattctaattatctctctaacagccttgtatcgtatatgcaaatatga
aggaatcatgggaaataggccctcttcctgcccgacctagcaaaaaacggaaaccccgg
gaatctaacccggctgaacggatttagagtccgttcgatctacatgatcaggtttccgg
tgtttcgtcctttccacaagatatataaagccaagaaatcgaaatactttcaagttacg
gtaagcatatgatagtccattttaaaacataatttaaaactgcaaactacccaagaaa
ttattactttctacgtcacgtattttgtactaatatctttgtgtttacagtcaaattaa
ttctaattatctctctaacagccttgtatcgtatatgcaaatatgaaggaatcatggga
aataggccctcttcctgcccgacctaggggtgcgagcggatcgagcagtgtcgatcact
actggaccgcgagctgtgctgcgacccgtgatcttacggcattatacgtatgatcggtc
cacgatcagctagattatctagtcagcttgatgtcatagctgtttcctgaggctcaata
ctgaccatttaaatcatacctgacctccatagcagaaagtcaaaagcctccgaccggag

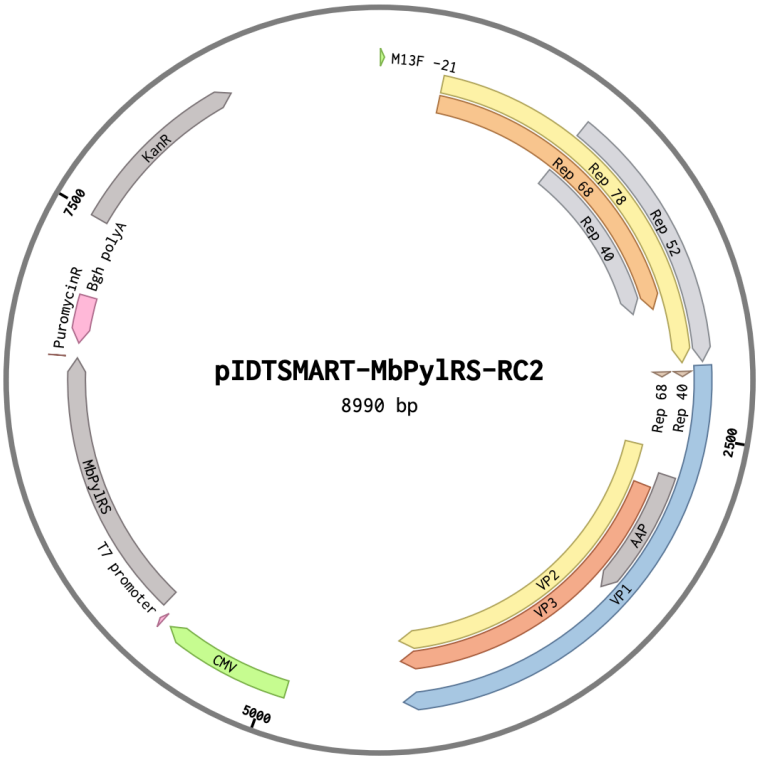
gcttttgacttgatcggcacgtaagaggttccaactttcaccataatgaaataagatca
ctaccgggcgtatttttgagttatcgagatttcaggagctaaggaagctaaaatgag
ccatattcaacgggaaacgtcttgcttgaagccgcgattaaattccaacatggatgctg
atttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtgcgacaatctat
cgattgtatgggaagcccgatgcgccagagttgtttctgaaacatggcaaaggtagcgt
tgccaatgatgttacagatgagatggtcaggctaaactggctgacggaatttatgcctc
ttccgaccatcaagcattttatccgtactcctgatgatgcatggttactcaccactgcg
atcccagggaaaacagcattccaggtattagaagaatatcctgattcaggtgaaaatat
tgttgatgcgctggcagtgttcctgcgccggttgcattcgattcctgtttgtaattgtc
cttttaacggcgatcgcgtatttcgtctcgctcaggcgcaatcacgaatgaataacggt
ttggttggtgcgagtgattttgatgacgagcgtaatggctggcctgttgaacaagtctg
gaaagaaatgcataaactcttgccattctcaccggattcagtcgtcactcatggtgatt
tctcacttgataaccttatttttgacgaggggaaattaataggttgtattgatgttgga
cgagtcggaatcgcagaccgataccaggatcttgccatcctatggaactgcctcggtga
gttttctccttcattacagaaacggctttttcaaaaatatggtattgataatcctgata
tgaataaattgcagtttcacttgatgctcgatgagtttttctaatgaggacctaaatgt
aatcacctggctcaccttcgggtgggcctttctgcgttgctggcgttttccataggct
ccgccccctgacgagcatcacaaaaatcgatgctcaagtcagaggtggcgaaacccga
caggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgtt
ccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgct
ttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgg
gctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgt
cttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacag
gattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaact
acggctacactagaagaacagtatttggtatctgcgctctgctgaagccagttacctcg
gaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttt
tttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgat
tttctaccgaagaaaggccca

pHelper



The sequence for this plasmid is identical to the pHelper found in the Chapter 2 section of this appendix.

pIDTSMART-MbPylRS-RC2



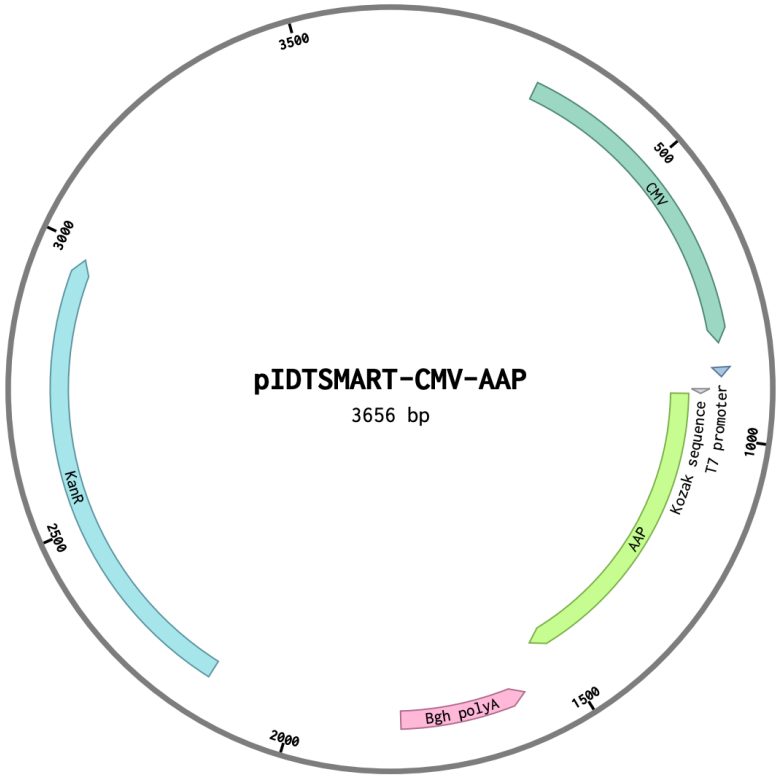
The sequence for this plasmid is identical to the pIDTSMART-MbPylRS-RC2wt and pIDTSMART-MbPylRS-RC2-454TAG found in the Chapter 2 section of this appendix.

The sequence for the *Cap* gene for pIDTSMART-MbPylRS-RC2-AAPstop60 is:

atggctgccgatggttatcttccagattggctcgaggacactctctctgaaggaataag
acagtggtggaagctcaaacctggcccaccaccaccaaagcccgcagagcggcataagg
acgacagcaggggtcttgtgcttcctgggtacaagtacctcggacccttcaacggactc
gacaagggagagccggtcaacgaggcagacgccgcggccctcgagcacgacaaagccta
cgaccggcagctcgacagcggagacaacccgtacctcaagtacaaccacgccgacgcgg
agtttcaggagcgccttaaagaagatacgtcttttgggggcaacctcggacgagcagtc
ttccaggcgaaaaagagggttcttgaacctctgggcctggttgaggaacctgttaagac
ggctccgggaaaaaagaggccggtagagcactctcctgtggagccagactcctcctcgg
gaaccggaaaggcgggccagcagcctgcaagaaaaagattgaattttggtcagactgga
gacgcagactcagtacctgacccccagcctctcggacagccaccagcagccccctctgg
tctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataacgagg
gcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatggatgggc
gacagagtGatcaccaccagcacccgaacctgggccctgcccacctacaacaaccacct
ctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttggctaca
gcacccctgggggtattttgacttcaacagattccactgccacttttcaccacgtgac
tggcaaagactcatcaacaacaactggggattccgacccaagagactcaacttcaagct
ctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgccaata
accttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacgtcctc
ggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtgccaca
gtatggatacctcaccctgaacaacgggagtcaggcagtaggacgctcttcattttact

gcctggagtactttccttctcagatgctgcgtaccggaaacaactttaccttcagctac
acttttgaggacgttcctttccacagcagctacgctcacagccagagtctggaccgtct
catgaatcctctcatcgaccagtacctgtattacttgagcagaacaaacactccaagtg
gaaccaccacgcagtcaaggcttcagtttctcaggccggagcgagtgacattcgggac
cagtctaggaactggcttcctggaccctgttaccgccagcagcgagtatcaaagacatc
tgcggataacaacaacagtgaatactcgtggactggagctaccaagtaccacctcaatg
gcagagactctctggtgaatccgggcccggccatggcaagccacaaggacgatgaagaa
aagtttttcctcagagcggggttctcatctttgggaagcaaggctcagagaaaacaaa
tgtggacattgaaaaggtcatgattacagacgaagaggaaatcaggacaaccaatcccg
tggctacggagcagtatggttctgtatctaccaacctccagagaggcaacagacaagca
gctaccgcagatgtcaacacacaaggcgttcttccaggcatggtctggcaggacagaga
tgtgtaccttcaggggcccatctgggcaaagattccacacacggacggacattttcacc
cctctcccctcatgggtggattcggacttaaacaccctcctccacagattctcatcaag
aacaccccggtacctgcgaatccttcgaccaccttcagtgcggcaaagtttgcttcctt
catcacacagtactccacgggacaggtcagcgtggagatcgagtgggagctgcagaagg
aaaacagcaaacgctggaatcccgaaattcagtacacttccaactacaacaagtctgtt
aatgtggactttactgtggacactaatggcgtgtattcagagcctcgccccattggcac
cagatacctgactcgtaatctgtaa

pIDTSMART-CMV-AAP



Sequence color-coding key

Feature	Color
CMV	text
T7 promoter	text
Kozak sequence	text
AAP	text
Bgh polyA	text
KanR	text

Sequence

cccgtgtaaaacgacggccagtttatctagtcagcttgattctagctgatcgtggaccg
gaaggtgagccagtgagttgattgcagtccagttacgctggagtctgaggctcgtcctg
aatgatatgcgaccgccggagggttgcgtttgagacgggcgacagatccagtcgcgctg
ctctcgtcgatccgctagggcggccgcaaatacctgcaggatccgttttgcgctgcttc
gcgatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtat
agttcatagcccatatatggagttccgcgttacataacttacggtaaatggcccgcctg
gctgaccgcccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagta
acgccaatagggactttccattgacgtcaatgggtggactatttacggtaaactgccca
cttggcagtacatcaagtgtatcatatgccaagtacgcccctattgacgtcaatgacg
gtaaatggcccgcctggcattatgcccagtacatgaccttatgggactttcctacttgg

cagtacatctacgtattagtcatcgctattaccatggtgatgcggttttggcagtacat
caatgggcgtggatagcggtttgactcacggggatttccaagtctccaccccattgacg
tcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaac
tccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcag
agctctctggctaactagagaacccactgcttactggcttatcgaaattaatacgactc
actatagggagacccaagctggctagcgccgccaccctggagacgcagactcagtacct
gacccccagcctctcggacagccaccagcagccccctctggtctgggaactaatacgat
ggctacaggcagtggcgcaccaatggcagacaataacgagggcgccgacggagtgggta
attcctcgggaaattggcattgcgattccacatggatgggcgacagagtcatcaccacc
agcacccgaacctgggccctgcccacctacaacaaccacctctacaaacaaatttccag
ccaatcaggagcctcgaacgacaatcactactttggctacagcacccttgggggtatt
ttgacttcaacagattccactgccacttttcaccacgtgactggcaaagactcatcaac
aacaactggggattccgacccaagagactcaacttcaagctctttaacattcaagtcaa
agaggtcacgcagaatgacggtacgacgacgattgccaataaccttaccagcacggttc
aggtgtttactgactcggagtaccagctcccgtacgtcctcggctcggcgcatcaagga
tgcctcccgccgttcccagcagacgtcttcatggtgccacagtatggatacctcaccct
gactcgagtcgagtctagagggcccgtttaaacccgctgatcagcctcgactgtgcctt
ctagttgccagccatctgttgtttgcccctccccgtgccttccttgaccctggaaggt
gccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattgtctgagtag
gtgtcattctattctgggggggtggggtggggcaggacagcaagggggaggattgggaag
acaatagcaggcatgctggggatgcggtgggctctatggcttctgaggcggaaagaacc
ctaggggtgcgagcggatcgagcagtgtcgatcactactggaccgcgagctgtgctgcg
acccgtgatcttacggcattatacgtatgatcggtccacgatcagctagattatctagt
cagcttgatgtcatagctgtttcctgaggctcaatactgaccatttaaatcatacctga
cctccatagcagaaagtcaaaagcctccgaccggaggcttttgacttgatcggcacgta
agaggttccaactttcaccataatgaaataagatcactaccgggcgtatttttgagtt
atcgagattttcaggagctaaggaagctaaaatgagccatattcaacgggaaacgtctt
gcttgaagccgcgattaaattccaacatggatgctgatttatatgggtataaatgggct
cgcgataatgtcgggcaatcaggtgcgacaatctatcgattgtatgggaagcccgatgc
gccagagttgtttctgaaacatggcaaaggtagcgttgccaatgatgttacagatgaga
tggtcaggctaaactggctgacggaatttatgcctcttccgaccatcaagcattttatc
cgtactcctgatgatgcatggttactcaccactgcgatcccagggaaaacagcattcca
ggtattagaagaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcc
tgcgccggttgcattcgattcctgtttgtaattgtcctttaacggcgatcgcgtattt
cgtctcgctcaggcgcaatcacgaatgaataacggtttggttggtgcgagtgattttga
tgacgagcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaactcttgc
cattctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttt
gacgaggggaaattaataggttgtattgatgttggacgagtcggaatcgcagaccgata
ccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacagaaac
ggcttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttcacttg
atgctcgatgagtttttctaatgaggacctaaatgtaatcacctggctcaccttcgggt
gggcctttctgcgttgctggcgtttttccataggctccgccccctgacgagcatcaca
aaaatcgatgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcg
tttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggata
cctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggt
atctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgtt
cagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagaca
cgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtag

gcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacagta
tttggtatctgcgctctgctgaagccagttacctcggaaaaagagttggtagctcttga
tccggcaaacaaaccaccgctggtagcggtggttttttgtttgcaagcagcagattac
gcgcagaaaaaaaggatctcaagaagatcctttgatttctaccgaagaaaggccca

The sequence for the *AAP* gene for pIDTSMART-CMV-AAP-L12TAG is:
ctggagacgcagactcagtacctgacccccagctagtcggacagccaccagcagccccc
tctggtctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataa
cgagggcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatgga
tgggcgacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaac
cacctctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttgg
ctacagcacccctttgggggtattttgacttcaacagattccactgccacttttcaccac
gtgactggcaaagactcatcaacaacaactggggattccgacccaagagactcaacttc
aagctctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgc
caataaccttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacg
tcctcggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtg
ccacagtatggatacctcaccctga

The sequence for the *AAP* gene for pIDTSMART-CMV-AAP-T78TAG is:
ctggagacgcagactcagtacctgacccccagcctctcggacagccaccagcagccccc
tctggtctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataa
cgagggcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatgga
tgggcgacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaTAGac
cacctctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttgg
ctacagcacccctttgggggtattttgacttcaacagattccactgccacttttcaccac
gtgactggcaaagactcatcaacaacaactggggattccgacccaagagactcaacttc
aagctctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgc
caataaccttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacg
tcctcggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtg
ccacagtatggatacctcaccctga

The sequence for the *AAP* gene for pIDTSMART-CMV-AAP-T97TAG is:
ctggagacgcagactcagtacctgacccccagcctctcggacagccaccagcagccccc
tctggtctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataa
cgagggcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatgga
tgggcgacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaac
cacctctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactTAGttgg
ctacagcacccctttgggggtattttgacttcaacagattccactgccacttttcaccac
gtgactggcaaagactcatcaacaacaactggggattccgacccaagagactcaacttc
aagctctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgc
caataaccttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacg
tcctcggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtg
ccacagtatggatacctcaccctga

The sequence for the *AAP* gene for pIDTSMART-CMV-AAP-T110TAG is:
ctggagacgcagactcagtacctgacccccagcctctcggacagccaccagcagccccc
tctggtctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataa
cgagggcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatgga

tgggcgacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaac
cacctctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttgg
ctacagcacccttgggggtattttgacttcaTAGgattccactgccacttttcaccac
gtgactggcaaagactcatcaacaacaactggggattccgacccaagagactcaacttc
aagctctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgc
caataaccttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacg
tcctcggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtg
ccacagtatggatacctcaccctga

The sequence for the *AAP* gene for pIDTSMART-CMV-AAP-T124TAG is:
ctggagacgcagactcagtacctgacccccagcctctcggacagccaccagcagccccc
tctggtctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataa
cgagggcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatgga
tgggcgacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaac
cacctctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttgg
ctacagcacccttgggggtattttgacttcaacagattccactgccacttttcaccac
gtgactggcaaagacTAGtcaacaacaactggggattccgacccaagagactcaacttc
aagctctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgc
caataaccttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacg
tcctcggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtg
ccacagtatggatacctcaccctga

The sequence for the *AAP* gene for pIDTSMART-CMV-AAP-T177TAG is:
ctggagacgcagactcagtacctgacccccagcctctcggacagccaccagcagccccc
tctggtctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataa
cgagggcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatgga
tgggcgacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaac
cacctctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttgg
ctacagcacccttgggggtattttgacttcaacagattccactgccacttttcaccac
gtgactggcaaagactcatcaacaacaactggggattccgacccaagagactcaacttc
aagctctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgc
caataaccttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtTAG
tcctcggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtg
ccacagtatggatacctcaccctga

The sequence for the *AAP* gene for pIDTSMART-CMV-AAP-T97TGA is:
ctggagacgcagactcagtacctgacccccagcctctcggacagccaccagcagccccc
tctggtctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataa
cgagggcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatgga
tgggcgacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaac
cacctctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactTGAttgg
ctacagcacccttgggggtattttgacttcaacagattccactgccacttttcaccac
gtgactggcaaagactcatcaacaacaactggggattccgacccaagagactcaacttc
aagctctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgc
caataaccttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacg
tcctcggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtg
ccacagtatggatacctcaccctga

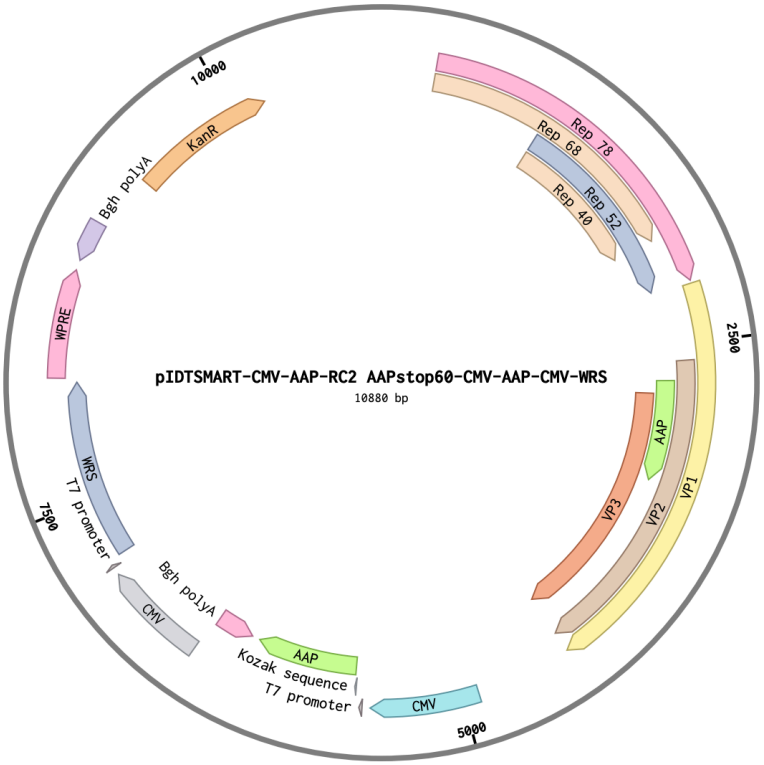
The sequence for the *AAP* gene for pIDTSMART-CMV-AAP-T110TGA is:

```
ctggagacgcagactcagtacctgacccccagcctctcggacagccaccagcagccccc
tctggtctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataa
cgagggcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatgga
tgggcgacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaac
cacctctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttgg
ctacagcacccttgggggtattttgacttcaTGAgattccactgccacttttcaccac
gtgactggcaaagactcatcaacaacaactggggattccgacccaagagactcaacttc
aagctctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgc
caataaccttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacg
tcctcggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtg
ccacagtatggatacctcaccctga
```

The sequence for the *AAP* gene for pIDTSMART-CMV-AAP-T124TGA is:

```
ctggagacgcagactcagtacctgacccccagcctctcggacagccaccagcagccccc
tctggtctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataa
cgagggcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatgga
tgggcgacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaac
cacctctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttgg
ctacagcacccttgggggtattttgacttcaacagattccactgccacttttcaccac
gtgactggcaaagacTGAtcaacaacaactggggattccgacccaagagactcaacttc
aagctctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgc
caataaccttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacg
tcctcggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtg
ccacagtatggatacctcaccctga
```

pIDTSMART-RC2-AAPstop60-CMV-AAP-CMV-WRS



Sequence color-coding key

Feature	Color
Rep	text
Cap – AAPstop60	text
CMV	text
T7 promoter	text
Kozak sequence	text
AAP	text
Bgh polyA	text
CMV	text
T7 promoter	text
WRS	text
WPRE	text
Bgh polyA	text
KanR	text

Sequence for AAP-97TGA

cccgtgtaaaacgacggccagtttatctagtcagcttgattctagctgatcgtggaccg
gaaggtgagccagtgagttgattgcagtccagttacgctggagtctgaggctcgtcctg

aatgatatgcgaccgccggagggttgcgtttgagacgggcgacagatccagtcgcgctg
ctctcgtcgatccgctagggcggccgctctagaactagtggatcccccggaagatcaga
agttcctattccgaagttcctattctctagaaagtataggaacttctgatctgcgcagc
cgccatgccgggtttacgagattgtgattaaggtccccagcgaccttgacgagcatc
tgcccggcatttctgacagctttgtgaactgggtggccgagaaggaatgggagttgccg
ccagattctgacatggatctgaatctgattgagcaggcacccctgaccgtggccgagaa
gctgcagcgcgactttctgacggaatggcgccgtgtgagtaaggccccggaggcccttt
tctttgtgcaatttgagaagggagagagctacttccacatgcacgtgctcgtggaaacc
accggggtgaaatccatggttttgggacgtttcctgagtcagattcgcgaaaaactgat
tcagagaatttaccgcgggatcgagccgactttgccaaactggttcgcggtcacaaaga
ccagaaatggcgccggaggcgggaacaaggtggtggatgagtgctacatccccaattac
ttgctccccaaaacccagcctgagctccagtgggcgtggactaatatggaacagtattt
aagcgcctgtttgaatctcacggagcgtaaacggttggtggcgcagcatctgacgcacg
tgtcgcagacgcaggagcagaacaaagagaatcagaatcccaattctgatgcgccggtg
atcagatcaaaaacttcagccaggtacatggagctggtcgggtggctcgtggacaaggg
gattacctcggagaagcagtggatccaggaggaccaggcctcatacatctccttcaatg
cggcctccaactcgcggtcccaaatcaaggctgccttggacaatgcgggaaagattatg
agcctgactaaaaccgcccccgactacctggtgggccagcagcccgtggaggacatttc
cagcaatcggatttataaaattttggaactaaacgggtacgatccccaatatgcggctt
ccgtctttctgggatgggccacgaaaaagttcggcaagaggaacaccatctggctgttt
gggcctgcaactaccgggaagaccaacatcgcggaggccatagcccacactgtgccctt
ctacgggtgcgtaaactggaccaatgagaactttcccttcaacgactgtgtcgacaaga
tggtgatctggtgggaggaggggaagatgaccgccaaggtcgtggagtcggccaaagcc
attctcggaggaagcaaggtgcgcgtggaccagaaatgcaagtcctcggcccagataga
cccgactcccgtgatcgtcacctccaacaccaacatgtgcgccgtgattgacgggaact
caacgaccttcgaacaccagcagccgttgcaagaccggatgttcaaatttgaactcacc
cgccgtctggatcatgactttgggaaggtcaccaagcaggaagtcaaagacttttccg
gtgggcaaaggatcacgtggttgaggtggagcatgaattctacgtcaaaaagggtggag
ccaagaaaagacccgcccccagtgacgcagatataagtgagcccaaacgggtgcgcgag
tcagttgcgcagccatcgacgtcagacgcggaagcttcgatcaactacgcagacaggta
ccaaaacaaatgttctcgtcacgtgggcatgaatctgatgctgtttccctgcagacaat
gcgagagaatgaatcagaattcaaatatctgcttcactcacggacagaaagactgttta
gagtgctttcccgtgtcagaatctcaacccgtttctgtcgtcaaaaaggcgtatcagaa
actgtgctacattcatcatatcatgggaaaggtgccagacgcttgcactgcctgcgatc
tggtcaatgtggatttggatgactgcatctttgaacaataaatgatttaaatcaggtat
ggctgccgatggttatcttccagattggctcgaggacactctctctgaaggaataagac
agtggtggaagctcaaacctggcccaccaccaccaaagcccgcagagcggcataaggac
gacagcaggggtcttgtgcttcctgggtacaagtacctcggacccttcaacggactcga
caagggagagccggtcaacgaggcagacgccgcggccctcgagcacgacaaagcctacg
accggcagctcgacagcggagacaacccgtacctcaagtacaaccacgccgacgcggag
tttcaggagcgccttaaagaagatacgtcttttgggggcaacctcggacgagcagtctt
ccaggcgaaaaagagggttcttgaacctctgggcctggttgaggaacctgttaagacgg
ctccgggaaaaaagaggccggtagagcactctcctgtggagccagactcctcctcggga
accggaaaggcgggccagcagcctgcaagaaaaagattgaattttggtcagactggaga
cgcagactcagtacctgacccccagcctctcggacagccaccagcagccccctctggtc
tgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataacgagggc
gccgacggagtgggtaattcctcgggaaattggcattgcgattccacatggatgggcga
cagagtgatcaccaccagcacccgaacctgggccctgcccacctacaacaaccacctct

acaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttggctacagc
acccccttgggggtattttgacttcaacagattccactgccacttttcaccacgtgactg
gcaaagactcatcaacaacaactggggattccgacccaagagactcaacttcaagctct
ttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgccaataac
cttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacgtcctcgg
ctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtgccacagt
atggatacctcaccctgaacaacgggagtcaggcagtaggacgctcttcattttactgc
ctggagtactttccttctcagatgctgcgtaccggaaacaactttaccttcagctacac
ttttgaggacgttcctttccacagcagctacgctcacagccagagtctggaccgtctca
tgaatcctctcatcgaccagtacctgtattacttgagcagaacaaacactccaagtgga
accaccacgcagtcaaggcttcagttttctcaggccggagcgagtgacattcgggacca
gtctaggaactggcttcctggaccctgttaccgccagcagcgagtatcaaagacatctg
cggataacaacaacagtgaatactcgtggactggagctaccaagtaccacctcaatggc
agagactctctggtgaatccgggcccggccatggcaagccacaaggacgatgaagaaaa
gtttttttcctcagagcggggttctcatctttgggaagcaaggctcagagaaaacaaatg
tggacattgaaaaggtcatgattacagacgaagaggaaatcaggacaaccaatcccgtg
gctacggagcagtatggttctgtatctaccaacctccagagaggcaacagacaagcagc
taccgcagatgtcaacacacaaggcgttcttccaggcatggtctggcaggacagagatg
tgtaccttcaggggcccatctgggcaaagattccacacacggacggacattttcacccc
tctcccctcatgggtggattcggacttaaacaccctcctccacagattctcatcaagaa
caccccggtacctgcgaatccttcgaccaccttcagtgcggcaaagtttgcttccttca
tcacacagtactccacgggacaggtcagcgtggagatcgagtgggagctgcagaaggaa
aacagcaaacgctggaatcccgaaattcagtacacttccaactacaacaagtctgttaa
tgtggactttactgtggacactaatggcgtgtattcagagcctcgccccattggcacca
gatacctgactcgtaatctgtaattgcttgttaatcaataaaccgtttaattcgtttca
gttgaactttggtctctgcgtatttctttcttatctagtttccatggctacgtagataa
gtagcatggcgggttaatcattaactacagcccgggcgtttaaacagcgggcggagggg
tggagtcgtgacgtgaattacgtcatagggttagggaggtcctgtattagaggtcacgt
gagtgttttgcgacattttgcgacaccatgtggtctcgctgggggggggggcccgagtg
agcacgcagggtctccattttgaagcgggaggtttgaacgagcgctggcgcgctcactg
gccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgcct
tgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcc
cttcccatgcatcggccgcaaatacctgcaggatccgttttgcgctgcttcgcgatgta
cgggccagatatacgcgttgacattgattattgactagttattaatagtaatcaattac
ggggtcattagttcatagcccatatatggagttccgcgttacataacttacggtaaatg
gcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtatgtt
cccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacggta
aactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgacg
tcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt
cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttg
gcagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacc
ccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgt
cgtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtcta
tataagcagagctctctggctaactagagaacccactgcttactggcttatcgaaatta
atacgactcactatagggagacccaagctggctagcgccgccacctggagacgcagac
tcagtacctgacccccagcctctcggacagccaccagcagccccctctggtctgggaac
taatacgatggctacaggcagtggcgcaccaatggcagacaataacgagggcgccgacg
gagtgggtaattcctcgggaaattggcattgcgattccacatggatgggcgacagagtc

atcaccaccagcacccgaacctgggccctgcccacctacaacaaccacctctacaaaca
aatttccagccaatcaggagcctcgaacgacaatcactgattggctacagcacccctg
ggggtattttgacttcaacagattccactgccacttttcaccacgtgactggcaaagac
tcatcaacaacaactggggattccgacccaagagactcaacttcaagctctttaacatt
caagtcaaagaggtcacgcagaatgacggtacgacgacgattgccaataaccttaccag
cacggttcaggtgtttactgactcggagtaccagctcccgtacgtcctcggctcggcgc
atcaaggatgcctcccgccgttcccagcagacgtcttcatggtgccacagtatggatac
ctcaccctgactcgagtcgagtctagagggcccgtttaaacccgctgatcagcctcgac
tgtgccttctagttgccagccatctgttgtttgccccctcccccgtgccttccttgaccc
tggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattgt
ctgagtaggtgtcattctattctgggggggtggggtggggcaggacagcaaggggaggga
ttgggaagacaatagcaggcatgctggggatgcggtgggctctatggcttctgaggcgg
aaagaaccctaggttattaatagtaatcaattacggggtcattagttcatagcccatat
atggagttccgcgttacataacttacggtaaatggcccgcctggctgaccgcccaacga
cccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaatagggactt
tccattgacgtcaatgggtggactatttacggtaaactgcccacttggcagtacatcaa
gtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggcccgcctg
gcattatgcccagtacatgaccttatgggactttcctacttggcagtacatctacgtat
tagtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggatag
cggtttgactcacggggatttccaagtctccaccccattgacgtcaatgggagtttgtt
ttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgccccattgacgc
aaatgggcggtaggcgtgtacggtgggaggtctatataagcagagctctctggctaact
agagaacccactgcttactggcttatcgaaattaatacgactcactatagggagaccca
agctggctagcgccaccatgactaagcccatcgtttttgctggcgcacagccctcaggt
gaattgaccattggtaactacatgggtgcgctgcgtcagtgggtaaacatgcaggatga
ctaccattgcatttactgtatcgttgaccaacacgcgatcaccgtgcgccaggatgcac
agaagctgcgtaaagcgacgctggatacgctggccttgtatctggcttgtggtatcgat
cctgagaaaagcaccatttttgttcagtcccacgtgccggaacatgcacagttaggctg
ggcactgaactgctatacctacttcggcgaactgagtcgcatgacgcagtttaaagata
aatctgcgcgttatgccgagaacatcaacgctggtctgtttgactatccggtgctgatg
gcagcggacatcctgctgtatcaaactaatctgggtccttgtggtgaagaccagaaaca
gcacctcgaactgagccgcgatattgcccagcgtttcaacgcgctgtatggcgagatct
ttaaggtgccggagccgtttattccgaaatctggcgcgcgcgtaatgtcgctgctggag
ccgaccaagaagatgtccaagtctgacgataatcgcaataacgttatcggcctgctgga
agatccgaaatcggtagtgaagaaaatcaaacgtgcggtcactgactccgacgagccgc
cggtagttcgctacgatgtgcagaacaaagcgggcgtttccaacctgttggatatcctt
tcagcggtaacgggccagagcatcccagaactggaaaaacagttcgaaggcaagatgta
tggtcatctgaaaggtgaagtggctgatgccgtttccggtatgctgactgaattgcagg
aacgctatcaccgtttccgcaacgatgaagccttcctgcaacaggtgatgaaagatggc
gcggaaaaagccagcgcgcacgcttcccgtacgctaaaagcggtgtacgaagcgattgg
ttttgtggcgaagccgtaagaattcaacgcgttaagtcgacaatcaacctctggattac
aaaatttgtgaaagattgactggtattcttaactatgttgctccttttacgctatgtgg
atacgctgctttaatgcctttgtatcatgctattgcttcccgtatggctttcattttct
cctccttgtataaatcctggttgctgtctctttatgaggagttgtggcccgttgtcagg
caacgtggcgtggtgtgcactgtgtttgctgacgcaacccccactggttggggcattgc
caccacctgtcagctcctttccgggactttcgctttccccctccctattgccacggcgg
aactcatcgccgcctgccttgcccgctgctggacaggggctcggctgttgggcactgac
aattccgtggtgttgtcggggaaatcatcgtcctttccttggctgctcgcctgtgttgc

cacctggattctgcgcgggacgtccttctgctacgtcccttcggccctcaatccagcgg
accttccttcccgcggcctgctgccggctctgcggcctcttccgcgtcttcgccttcgc
cctcagacgagtcggatctccctttgggccgcctccccgcgtcgactttaactcgagtc
tagagggcccgtttaaacccgctgatcagcctcgactgtgccttctagttgccagccat
ctgttgtttgcccctcccccgtgccttccttgaccctggaaggtgccactcccactgtc
ctttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattct
ggggggtggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatg
ctggggatgcggtgggctctatggactaggggtgcgagcggatcgagcagtgtcgatca
ctactggaccgcgagctgtgctgcgacccgtgatcttacggcattatacgtatgatcgg
tccacgatcagctagattatctagtcagcttgatgtcatagctgtttcctgaggctcaa
tactgaccatttaaatcatacctgacctccatagcagaaagtcaaaagcctccgaccgg
aggcttttgacttgatcggcacgtaagaggttccaactttcaccataatgaaataagat
cactaccgggcgtatttttgagttatcgagattttcaggagctaaggaagctaaaatg
agccatattcaacgggaaacgtcttgcttgaagccgcgattaaattccaacatggatgc
tgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtgcgacaatct
atcgattgtatgggaagcccgatgcgccagagttgtttctgaaacatggcaaaggtagc
gttgccaatgatgttacagatgagatggtcaggctaaactggctgacggaatttatgcc
tcttccgaccatcaagcattttatccgtactcctgatgatgcatggttactcaccactg
cgatcccagggaaaacagcattccaggtattagaagaatatcctgattcaggtgaaaat
attgttgatgcgctggcagtgttcctgcgccggttgcattcgattcctgtttgtaattg
tccttttaacggcgatcgcgtatttcgtctcgctcaggcgcaatcacgaatgaataacg
gtttggttggtgcgagtgattttgatgacgagcgtaatggctggcctgttgaacaagtc
tggaaagaaatgcataaactcttgccattctcaccggattcagtcgtcactcatggtga
tttctcacttgataaccttatttttgacgaggggaaattaataggttgtattgatgttg
gacgagtcggaatcgcagaccgataccaggatcttgccatcctatggaactgcctcggt
gagttttctccttcattacagaaacggctttttcaaaaatatggtattgataatcctga
tatgaataaattgcagtttcacttgatgctcgatgagttttctaatgaggacctaaat
gtaatcacctggctcaccttcgggtgggcctttctgcgttgctggcgttttccatagg
ctccgccccctgacgagcatcacaaaaatcgatgctcaagtcagaggtggcgaaaccc
gacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctg
ttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcg
ctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagct
gggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatc
gtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaac
aggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaa
ctacggctacactagaagaacagtatttggtatctgcgctctgctgaagccagttacct
cggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtt
ttttgtttgcaagcagcagattacgcgcagaaaaaaggatctcaagaagatcctttg
attttctaccgaagaaaggccca

The sequence for the *AAP* gene for pIDTSMART-RC2-AAPstop60-CMV-AAP110TGA-CMV-WRS is:

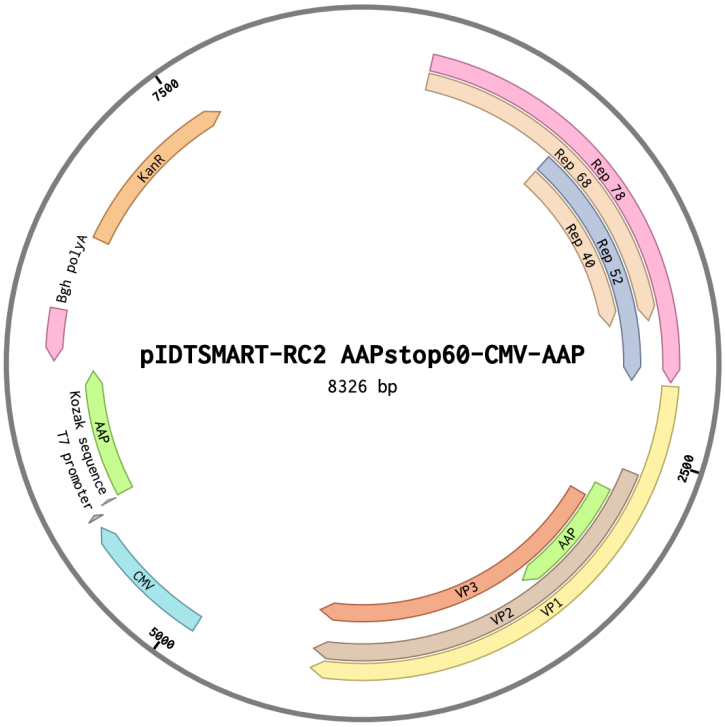
ctggagacgcagactcagtacctgacccccagcctctcggacagccaccagcagccccc
tctggtctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataa
cgagggcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatgga
tgggcgacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaac
cacctctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttgg
ctacagcacccttgggggtattttgacttcaTGAgattccactgccacttttcaccac

gtgactggcaaagactcatcaacaacaactggggattccgacccaagagactcaacttc
aagctctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgc
caataaccttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacg
tcctcggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtg
ccacagtatggatacctcaccctga

The sequence for the *AAP* gene for pIDTSMART-RC2-AAPstop60-CMV-AAP124TGA-CMV-WRS is:

ctggagacgcagactcagtacctgacccccagcctctcggacagccaccagcagccccc
tctggtctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataa
cgagggcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatgga
tgggcgacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaac
cacctctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttgg
ctacagcacccttgggggtattttgacttcaacagattccactgccacttttcaccac
gtgactggcaaagacTGAtcaacaacaactggggattccgacccaagagactcaacttc
aagctctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgc
caataaccttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacg
tcctcggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtg
ccacagtatggatacctcaccctga

pIDTSMART-RC2-AAPstop60-CMV-AAP



Sequence color-coding key

Feature	Color
Rep	text
Cap – AAPstop60	text
CMV	text
T7 promoter	text
Kozak sequence	text
AAP	text
Bgh polyA	text
KanR	text

Sequence for wild type AAP

cccgtgtaaaacgacggccagtttatctagtcagcttgattctagctgatcgtggaccg
gaaggtgagccagtgagttgattgcagtccagttacgctggagtctgaggctcgtcctg
aatgatatgcgaccgccggagggttgcgtttgagacgggcgacagatccagtcgcgctg
ctctcgtcgatccgctagggcggccgctctagaactagtggatcccccggaagatcaga
agttcctattccgaagttcctattctctagaaagtataggaacttctgatctgcgcagc
cgccatgccggggttttacgagattgtgattaaggtccccagcgaccttgacgagcatc
tgcccggcatttctgacagctttgtgaactgggtggccgagaaggaatgggagttgccg
ccagattctgacatggatctgaatctgattgagcaggcacccctgaccgtggccgagaa
gctgcagcgcgactttctgacggaatggcgccgtgtgagtaaggccccggaggcccttt
tctttgtgcaatttgagaagggagagagctacttccacatgcacgtgctcgtggaaacc

accggggtgaaatccatggttttgggacgtttcctgagtcagattcgcgaaaaactgat
tcagagaatttaccgcgggatcgagccgactttgccaaactggttcgcggtcacaaaga
ccagaaatggcgccggaggcgggaacaaggtggtggatgagtgctacatccccaattac
ttgctccccaaaacccagcctgagctccagtgggcgtggactaatatggaacagtattt
aagcgcctgtttgaatctcacggagcgtaaacggttggtggcgcagcatctgacgcacg
tgtcgcagacgcaggagcagaacaaagagaatcagaatcccaattctgatgcgccggtg
atcagatcaaaaacttcagccaggtacatggagctggtcgggtggctcgtggacaaggg
gattacctcggagaagcagtggatccaggaggaccaggcctcatacatctccttcaatg
cggcctccaactcgcggtcccaaatcaaggctgccttggacaatgcgggaaagattatg
agcctgactaaaaccgcccccgactacctggtgggccagcagcccgtggaggacatttc
cagcaatcggatttataaaattttggaactaaacgggtacgatccccaatatgcggctt
ccgtctttctgggatgggccacgaaaaagttcggcaagaggaacaccatctggctgttt
gggcctgcaactaccgggaagaccaacatcgcggaggccatagcccacactgtgccctt
ctacgggtgcgtaaactggaccaatgagaactttcccttcaacgactgtgtcgacaaga
tggtgatctggtgggaggaggggaagatgaccgccaaggtcgtggagtcggccaaagcc
attctcggaggaagcaaggtgcgcgtggaccagaaatgcaagtcctcggcccagataga
cccgactcccgtgatcgtcacctccaacaccaacatgtgcgccgtgattgacgggaact
caacgaccttcgaacaccagcagccgttgcaagaccggatgttcaaatttgaactcacc
cgccgtctggatcatgactttgggaaggtcaccaagcaggaagtcaaagacttttccg
gtgggcaaaggatcacgtggttgaggtggagcatgaattctacgtcaaaaagggtggag
ccaagaaaagacccgcccccagtgacgcagatataagtgagcccaaacgggtgcgcgag
tcagttgcgcagccatcgacgtcagacgcggaagcttcgatcaactacgcagacaggta
ccaaaacaaatgttctcgtcacgtgggcatgaatctgatgctgtttccctgcagacaat
gcgagagaatgaatcagaattcaaatatctgcttcactcacggacagaaagactgttta
gagtgctttcccgtgtcagaatctcaacccgtttctgtcgtcaaaaaggcgtatcagaa
actgtgctacattcatcatatcatgggaaaggtgccagacgcttgcactgcctgcgatc
tggtcaatgtggatttggatgactgcatctttgaacaataaatgatttaaatcaggtat
ggctgccgatggttatcttccagattggctcgaggacactctctctgaaggaataagac
agtggtggaagctcaaacctggcccaccaccaccaaagcccgcagagcggcataaggac
gacagcaggggtcttgtgcttcctgggtacaagtacctcggacccttcaacggactcga
caagggagagccggtcaacgaggcagacgccgcggccctcgagcacgacaaagcctacg
accggcagctcgacagcggagacaacccgtacctcaagtacaaccacgccgacgcggag
tttcaggagcgccttaaagaagatacgtcttttgggggcaacctcggacgagcagtctt
ccaggcgaaaaagagggttcttgaacctctgggcctggttgaggaacctgttaagacgg
ctccgggaaaaaagaggccggtagagcactctcctgtggagccagactcctcctcggga
accggaaaggcgggccagcagcctgcaagaaaaagattgaattttggtcagactggaga
cgcagactcagtacctgacccccagcctctcggacagccaccagcagccccctctggtc
tgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataacgagggc
gccgacggagtgggtaattcctcgggaaattggcattgcgattccacatggatgggcga
cagagtgatcaccaccagcacccgaacctgggccctgcccacctacaacaaccacctct
acaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttggctacagc
acccccttgggggtattttgacttcaacagattccactgccacttttcaccacgtgactg
gcaaagactcatcaacaacaactggggattccgacccaagagactcaacttcaagctct
ttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgccaataac
cttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacgtcctcgg
ctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtgccacagt
atggatacctcaccctgaacaacgggagtcaggcagtaggacgctcttcattttactgc
ctggagtactttccttctcagatgctgcgtaccggaaacaactttaccttcagctacac

ttttgaggacgttcctttccacagcagctacgctcacagccagagtctggaccgtctca
tgaatcctctcatcgaccagtacctgtattacttgagcagaacaaacactccaagtgga
accaccacgcagtcaaggcttcagttttctcaggccggagcgagtgacattcgggacca
gtctaggaactggcttcctggaccctgttaccgccagcagcgagtatcaaagacatctg
cggataacaacaacagtgaatactcgtggactggagctaccaagtaccacctcaatggc
agagactctctggtgaatccgggcccggccatggcaagccacaaggacgatgaagaaaa
gtttttttcctcagagcggggttctcatctttgggaagcaaggctcagagaaaacaaatg
tggacattgaaaaggtcatgattacagacgaagaggaaatcaggacaaccaatcccgtg
gctacggagcagtatggttctgtatctaccaacctccagagaggcaacagacaagcagc
taccgcagatgtcaacacacaaggcgttcttccaggcatggtctggcaggacagagatg
tgtaccttcaggggcccatctgggcaaagattccacacacggacggacattttcacccc
tctcccctcatgggtggattcggacttaaacaccctcctccacagattctcatcaagaa
caccccggtacctgcgaatccttcgaccaccttcagtgcggcaaagtttgcttccttca
tcacacagtactccacgggacaggtcagcgtggagatcgagtgggagctgcagaaggaa
aacagcaaacgctggaatcccgaaattcagtacacttccaactacaacaagtctgttaa
tgtggactttactgtggacactaatggcgtgtattcagagcctcgccccattggcacca
gatacctgactcgtaatctgtaattgcttgttaatcaataaaccgtttaattcgtttca
gttgaactttggtctctgcgtatttctttcttatctagtttccatggctacgtagataa
gtagcatggcgggttaatcattaactacagcccgggcgtttaaacagcgggcggagggg
tggagtcgtgacgtgaattacgtcatagggttagggaggtcctgtattagaggtcacgt
gagtgttttgcgacattttgcgacaccatgtggtctcgctgggggggggggcccgagtg
agcacgcagggtctccattttgaagcgggaggtttgaacgagcgctggcgcgctcactg
gccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgcct
tgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcc
cttcccatgcatcggccgcaaatacctgcaggatccgttttgcgctgcttcgcgatgta
cgggccagatatacgcgttgacattgattattgactagttattaatagtaatcaattac
ggggtcattagttcatagcccatatatggagttccgcgttacataacttacggtaaatg
gcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtatgtt
cccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacggta
aactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgacg
tcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt
cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttg
gcagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacc
ccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgt
cgtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtcta
tataagcagagctctctggctaactagagaacccactgcttactggcttatcgaaatta
atacgactcactatagggagacccaagctggctagcgccgccacctggagacgcagac
tcagtacctgacccccagcctctcggacagccaccagcagccccctctggtctgggaac
taatacgatggctacaggcagtggcgcaccaatggcagacaataacgagggcgccgacg
gagtgggtaattcctcgggaaattggcattgcgattccacatggatgggcgacagagtc
atcaccaccagcacccgaacctgggccctgcccacctacaacaaccacctctacaaaca
aatttccagccaatcaggagcctcgaacgacaatcactactttggctacagcacccctt
ggggtattttgacttcaacagattccactgccacttttcaccacgtgactggcaaaga
ctcatcaacaacaactggggattccgacccaagagactcaacttcaagctctttaacat
tcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgccaataaccttacca
gcacggttcaggtgtttactgactcggagtaccagctcccgtacgtcctcggctcggcg
catcaaggatgcctcccgccgttcccagcagacgtcttcatggtgccacagtatggata
cctcaccctgactcgagtcgagtctagagggcccgtttaaacccgctgatcagcctcga

ctgtgccttctagttgccagccatctgttgtttgcccctccccgtgccttccttgacc
ctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattg
tctgagtaggtgtcattctattctgggggtggggtggggcaggacagcaaggggaagg
attgggaagacaatagcaggcatgctggggatgcggtgggctctatggcttctgaggcg
gaaagaaccctaggggtgcgagcggatcgagcagtgtcgatcactactggaccgcgagc
tgtgctgcgacccgtgatcttacggcattatacgtatgatcggtccacgatcagctaga
ttatctagtcagcttgatgtcatagctgtttcctgaggctcaatactgaccatttaaat
catacctgacctccatagcagaaagtcaaaagcctccgaccggaggcttttgacttgat
cggcacgtaagaggttccaactttcaccataatgaaataagatcactaccgggcgtatt
ttttgagttatcgagatttcaggagctaaggaagctaaaatgagccatattcaacggg
aaacgtcttgcttgaagccgcgattaaattccaacatggatgctgatttatatgggtat
aaatgggctcgcgataatgtcgggcaatcaggtgcgacaatctatcgattgtatgggaa
gcccgatgcgccagagttgtttctgaaacatggcaaaggtagcgttgccaatgatgtta
cagatgagatggtcaggctaaactggctgacggaatttatgcctcttccgaccatcaag
cattttatccgtactcctgatgatgcatggttactcaccactgcgatcccagggaaaac
agcattccaggtattagaagaatatcctgattcaggtgaaaatattgttgatgcgctgg
cagtgttcctgcgccggttgcattcgattcctgtttgtaattgtccttttaacggcgat
cgcgtatttcgtctcgctcaggcgcaatcacgaatgaataacggtttggttggtgcgag
tgattttgatgacgagcgtaatggctggcctgttgaacaagtctggaaagaaatgcata
aactcttgccattctcaccggattcagtcgtcactcatggtgatttctcacttgataac
cttatttttgacgaggggaaattaataggttgtattgatgttggacgagtcggaatcgc
agaccgataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcat
tacagaaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcag
tttcacttgatgctcgatgagtttttctaatgaggacctaaatgtaatcacctggctca
ccttcgggtgggcctttctgcgttgctggcgttttccataggctccgccccctgacg
agcatcacaaaaatcgatgctcaagtcagaggtggcgaaacccgacaggactataaaga
taccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccgaccctgccgct
taccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcac
gctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaa
ccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaaccc
ggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcga
ggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactaga
agaacagtatttggtatctgcgctctgctgaagccagttacctcggaaaaagagttggt
agctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagca
gcagattacgcgcagaaaaaaggatctcaagaagatcctttgattttctaccgaagaa
aggccca

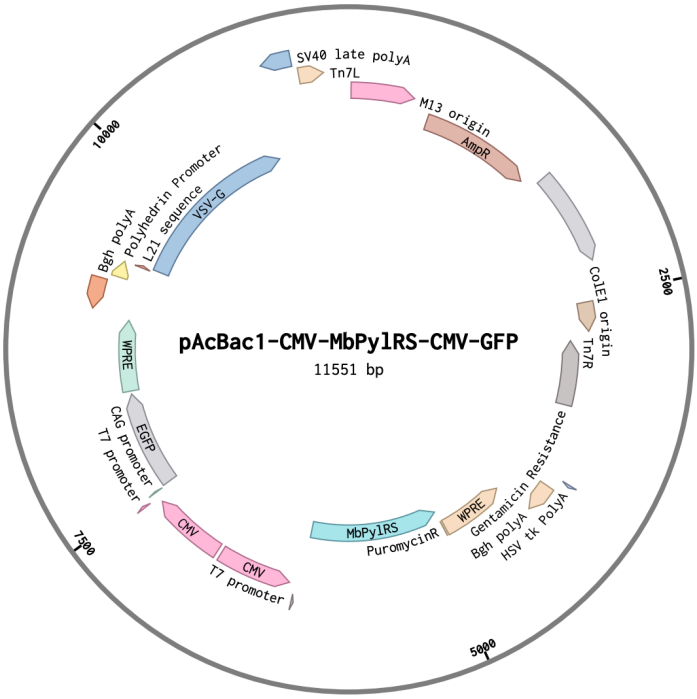
The sequence for the *AAP* gene for pIDTSMART-RC2-AAPstop60-CMV-AAP-97TGA is:

ctggagacgcagactcagtacctgacccccagcctctcggacagccaccagcagccccc
tctggtctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataa
cgagggcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatgga
tgggcgacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaac
cacctctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactTGAttgg
ctacagcacccctttgggggtattttgacttcaacagattccactgccacttttcaccac
gtgactggcaaagactcatcaacaacaactggggattccgacccaagagactcaacttc
aagctctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgc

caataaccttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacg
tcctcggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtg
ccacagtatggatacctcaccctga

Chapter 4 Plasmids

pAcBac-CMV-MbPylRS-CMV-GFP



Sequence color-coding key

Feature	Color
M13 origin	text
AmpR	text
ColE1 origin	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
WPRE	text
MbPylRS	text
T7 promoter	text
CMV promoter	text
CMV Promoter	text
T7 promoter	text
EGFP	text
WPRE	text
Bgh polyA	text
Polyhedrin Promoter	text

L21 sequence	text
VSV-G	text
SV40 late polyA	text
Tn7L	text

Sequence for wild type EGFP

ctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggt
tacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttct
tcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctc
cctttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgattaggg
tgatggttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttgg
agtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatc
tcggtctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaa
tgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgtttacaattt
caggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaata
cattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattg
aaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcgg
catttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaa
gatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatcct
tgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctat
gtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacac
tattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatgg
catgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggcca
acttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatg
ggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaa
cgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaa
ctggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggat
aaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataa
atctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggta
agccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacga
aatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagacca
agtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatct
aggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttc
cactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttct
gcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgc
cggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcagata
ccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagc
accgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgata
agtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcg
ggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaact
gagatacctacagcgtgagcattgagaaagcgccacgcttcccgaagggagaaaggcgg
acaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccaggg
ggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcg
atttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcct
ttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcag
ccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcctgatgcggt

attttctccttacgcatctgtgcggtatttcacaccgcagaccagccgcgtaacctggc
aaaatcggttacggttgagtaataaatggatgccctgcgtaagcgggtgtgggcggaca
ataaagtcttaaactgaacaaaatagatctaaactatgacaataaagtcttaaactaga
cagaatagttgtaaactgaaatcagtccagttatgctgtgaaaaagcatactggacttt
tgttatggctaaagcaaactcttcattttctgaagtgcaaattgcccgtcgtattaaag
aggggcgtggccaagggcatggtaaagactatattcgcggcgttgtgacaatttaccga
acaactccgcggccgggaagccgatctcggcttgaacgaattgttaggtggcggtactt
gggtcgatatcaaagtgcatcacttcttcccgtatgcccaactttgtatagagagccac
tgcgggatcgtcaccgtaatctgcttgcacgtagatcacataagcaccaagcgcgttgg
cctcatgcttgaggagattgatgagcgcggtggcaatgccctgcctccggtgctcgccg
gagactgcgagatcatagatatagatctcactacgcggctgctcaaacctgggcagaac
gtaagccgcgagagcgccaacaaccgcttcttggtcgaaggcagcaagcgcgatgaatg
tcttactacggagcaagttcccgaggtaatcggagtccggctgatgttgggagtaggtg
gctacgtctccgaactcacgaccgaaaagatcaagagcagcccgcatggatttgacttg
gtcagggccgagcctacatgtgcgaatgatgcccatacttgagccacctaactttgttt
tagggcgactgccctgctgcgtaacatcgttgctgctgcgtaacatcgttgctgctcca
taacatcaaacatcgacccacggcgtaacgcgcttgctgcttggatgcccgaggcatag
actgtacaaaaaaacagtcataacaagccatgaaaaccgccactgcgccgttaccaccg
ctgcgttcggtcaaggttctggaccagttgcgtgagcgcatacgctacttgcattacag
tttacgaaccgaacaggcttatgtcaactgggttcgtgccttcatccgtttccacggtg
tgcgtcacccggcaaccttgggcagcagcgaagtcgaggcatttctgtcctggctggcg
aacgagcgcaaggtttcggtctccacgcatcgtcaggcattggcggccttgctgttctt
ctacggcaaggtgctgtgcacggatctgccctggcttcaggagatcggtagacctcggc
cgtcgcggcgcttgccggtggtgctgaccccggatgaagtggttcgcatcctcggtttt
ctggaaggcgagcatcgtttgttcgcccaggactctagctatagttctagtggttggct
acgtacccgtagtggctatggcagggcttgcgcttaatgcgccgctacagggcgcgtgg
ggataccccctagagccccagctggttctttccgcctcagaagccatagagcccaccgc
atccccagcatgcctgctattgtcttcccaatcctccccttgctgtcctgccccaccc
cacccccagaatagaatgacacctactcagacaatgcgatgcaatttcctcattttat
taggaaaggacagtgggagtggcaccttccagggtcaaggaaggcacgggggaggggca
aacaacagatggctggcaactagaaggcacagtcgaggctgatcagcgggtttaaacgg
gccctctagactcgagttaaagtcgtcgcggggaggcggcccaaagggagatccgactc
gtctgagggcgaaggcgaagacgcggaagaggccgcagagccggcagcaggccgcggga
aggaaggtccgctggattgagggccgaagggacgtagcagaaggacgtcccgcgcagaa
tccaggtggcaacacaggcgagcagccaaggaaaggacgatgatttccccgacaacacc
acggaattgtcagtgcccaacagccgagcccctgtccagcagcgggcaaggcaggcggc
gatgagttccgccgtggcaatagggaggggggaaagcgaaagtcccggaaaggagctgac
aggtggtggcaatgccccaaccagtgggggttgcgtcagcaaacacagtgcacaccacg
ccacgttgcctgacaacgggccacaactcctcataaagagacagcaaccaggatttata
caaggaggagaaaatgaaagccatacgggaagcaatagcatgatacaaaggcattaaag
cagcgtatccacatagcgtaaaaggagcaacatagttaagaataccagtcaatctttca
caaattttgtaatccagaggttgattgtcgtcttaacgcgttgaattcttacagattgg
ttgaaatcccattatagtaagattcggaccttgatgccctcttaatgttttaaagccg
tgcataaccttgagcaagcgttcaagaccaaaacctgcacctatccatggtttgtcaat
accccattctctatcaagagaaactggcccgacgactgccgaagaaagctccaggtccc
cgtgcattatatcaagagtatccccaaagaccatacaggaatctcctacgatttcgaag
tcgatttccagatagtccagaaactctttgatgagagcttcaagattttcccgagtaca
tcccgaacccatctgacagaagttcaccatagtaaattcttccaggtgctctttgccgt

cagactctttccggtaacaaggtccgacttcgaaaatttttattgggcctggtaaaatc
ctatcgagttttcgcagatagttgtaaagagtcggggcaagcattggcctcaagcagag
atttttatccacccggaagatctgttttgaaagttcagtatcattattaatacccattc
tctccacgtattccgccggaataaggataggagactttatctccagaaaaccccggtct
acgaaaaatttcgtaatatcacgttcgagtttaccgaggtagtcttctctatcattggt
atagagccgctgaaaatcgtttttttcttcttgtcacaagttcaggctcaagttccctga
aaggctttgccatatttagagaaattttatcctctggacttaagagagcctcaacccta
tcaagctggcttcttgtaagtgaaggagcaggagccgatgcgggaacagacgaatttgg
agttgattttgcaggcgaaggtacagatctggatgtgttcgttgatgcctttgcagaaa
cagaattttccagaggcttcggagcccttgaaactgatttcggcatagcttttttgacc
tttggagcagaaactaccctaactttcacactgtttttgctttcggttgatcttgtgag
aaaattattgatatcctcgtccgaaaccctacatcgtttgcaggttttctgtacttat
gatgtctgaatgctctggctgttctacaactcctggaattattcacaacaagatggtct
ccacacgccatttcaatgtatattttacttcttgagacctcatggtgcttgattttgtg
gagcgtgccagtcctggacatccagagcccggtcgcagatattaaaacatctaatggtt
ttttatccatggtggcgctagtcagcttgggtctccctatagtgagtcgtattaatttc
gataagccagtaagcagtgggttctctagttagccagagagctctgcttatatagacct
cccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacatttt
ggaaagtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggag
acttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaacc
gcatcaccatggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcc
cataaggtcatgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaat
aggggggcgtacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgta
aatagtccacccattgacgtcaatggaaagtccctattggcgttactatgggaacatac
gtcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttac
cgtaagttatgtaacgcggaactccatatatgggctatgaactaatgaccccgtaattg
attactattaataactagaataatgctagcaataatcctgcaggattattctagttatt
aatagtaatcaattacggggtcattagttcatagcccatatatggagttccgcgttaca
taacttacggtaaatggcccgcctggctgaccgcccaacgacccccgcccattgacgtc
aataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatggg
tggactatttacggtaaactgcccacttggcagtacatcaagtgtatcatatgccaagt
acgccccctattgacgtcaatgacggtaaatggcccgcctggcattatgcccagtacat
gaccttatgggactttcctacttggcagtacatctacgtattagtcatcgctattacca
tggtgatgcggttttggcagtacatcaatgggcgtggatagcggtttgactcacgggga
tttccaagtctccaccccattgacgtcaatgggagtttgttttggcaccaaaatcaacg
ggactttccaaaatgtcgtaacaactccgccccattgacgcaaatgggcggtaggcgtg
tacggtgggaggtctatataagcagagctgtcccgggaactagagaacccactgcttac
tggcttatcgaaattaatacgactcactatagggagacccggtaagctggatagcgccg
ccaccatggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgag
ctggacggcgacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgc
cacttacggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccct
ggcccaccctcgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgac
cacatgaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcg
caccatcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagg
gcgacaccctggtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaac
atcctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccga
caagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggca
gcgtgcagctcgccgaccactaccagcagaacaccccatcggcgacggccccgtgctg

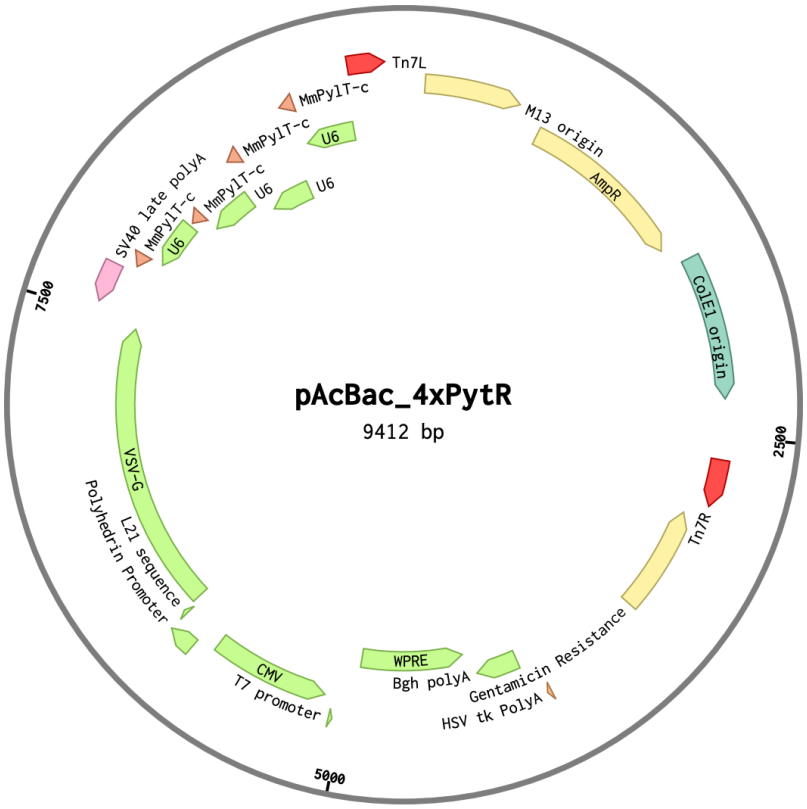
ctgcccgacaaccactacctgagcacccagtccgccctgagcaaagaccccaacgagaa
gcgcgatcacatggtcctgctggagttcgtgaccgccgccgggatcactctcggcatgg
acgagctgtacaagggcccttcgaacaaaaactcatctcagaagaggatctgaatatg
cataccggtcatcatcaccatcaccattgataaaaattcaacgcgttaagtcgacaatc
aacctctggattacaaaatttgtgaaagattgactggtattcttaactatgttgctcct
tttacgctatgtggatacgctgctttaatgcctttgtatcatgctattgcttcccgtat
ggctttcattttctcctccttgtataaatcctggttgctgtctctttatgaggagttgt
ggcccgttgtcaggcaacgtggcgtggtgtgcactgtgtttgctgacgcaacccccact
ggttggggcattgccaccacctgtcagctcctttccgggactttcgctttcccctccc
tattgccacggcggaactcatcgccgcctgccttgcccgctgctggacaggggctcggc
tgttgggcactgacaattccgtggtgttgtcggggaaatcatcgtcctttccttggctg
ctcgcctgtgttgccacctggattctgcgcgggacgtccttctgctacgtcccttcggc
cctcaatccagcggaccttccttcccgcggcctgctgccggctctgcggcctcttccgc
gtcttcgccttcgccctcagacgagtcggatctcccttgggccgcctccccgcgacga
ctttaactcgagtctagagggcccgtttaaacccgctgatcagcctcgactgtgccttc
tagttgccagccatctgttgtttgcccctcccccgtgccttccttgaccctggaaggtg
ccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattgtctgagtagg
tgtcattctattctgggggggtggggtggggcaggacagcaagggggaggattgggaaga
caatagcaggcatgctggggatgcggtgggctctatggattattgcggccgcggtccgt
atactccggaatattaatagatcatggagataattaaaatgataaccatctcgcaaata
aataagtattttactgttttcgtaacagtttgtaataaaaaaacctataaatattccg
gattattcataccgtcccaccatcgggcgcgaactcctaaaaaaccgccaccatgaagt
gcttttgtacttagccttttattcattggggtgaattgcaagttcaccatagttttt
ccacacaaccaaaaaggaaactggaaaaatgttccttctaattaccattattgcccgtc
aagctcagatttaaattggcataatgacttaataggcacagccttacaagtcaaaatgc
ccaagagtcacaaggctattcaagcagacggttggatgtgtcatgcttccaaatgggtc
actacttgtgatttccgctggtatggaccgaagtatataacacattccatccgatcctt
cactccatctgtagaacaatgcaaggaaagcattgaacaaacgaaacaaggaacttggc
tgaatccaggcttccctcctcaaagttgtggatatgcaactgtgacggatgccgaagca
gtgattgtccaggtgactcctcaccatgtgctggttgatgaatacacaggagaatgggt
tgattcacagttcatcaacggaaaatgcagcaattacatatgccccactgtccataact
ctacaacctggcattctgactataaggtcaaagggctatgtgattctaacctcatttcc
atggacatcaccttcttctcagaggacggagagctatcatccctgggaaaggagggcac
agggttcagaagtaactactttgcttatgaaactggaggcaaggcctgcaaaatgcaat
actgcaagcattggggagtcagactcccatcaggtgtctggttcgagatggctgataag
gatctctttgctgcagccagattccctgaatgcccagaagggtcaagtatctctgctcc
atctcagacctcagtggatgtaagtctaattcaggacgttgagaggatcttggattatt
ccctctgccaagaaacctggagcaaaatcagagcgggtcttccaatctctccagtggat
ctcagctatcttgctcctaaaaacccaggaaccggtcctgctttcaccataatcaatgg
taccctaaaatactttgagaccagatacatcagagtcgatattgctgctccaatcctct
caagaatggtcggaatgatcagtggaactaccacagaaagggaactgtgggatgactgg
gcaccatatgaagacgtggaaattggacccaatggagttctgaggaccagttcaggata
taagtttcctttatacatgattggacatggtatgttggactccgatcttcatcttagct
caaaggctcaggtgttcgaacatcctcacattcaagacgctgcttcgcaacttcctgat
gatgagagtttatttttggtgatactgggctatccaaaaatccaatcgagcttgtaga
aggttggttcagtagttggaaaagctctattgcctctttttctttatcatagggttaa
tcattggactattcttggttctccgagttggtatccatctttgcattaaattaaagcac
accaagaaaagacagatttatacagacatagagatgaaccgacttggaaagtgataagg

ccaggccggccaagcttgtcgagaagtactagaggatcataatcagccataccacattt
gtagaggttttacttgctttaaaaaacctcccacacctccccctgaacctgaaacataa
aatgaatgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataaa
gcaatagcatcacaaatttcacaaataaagcatttttttcactgcattctagttgtggt
ttgtccaaactcatcaatgtatcttatcatgtctggatctgatcactgcttgagcctag
gagatccgaaccagataagtgaaatctagttccaaactattttgtcatttttaattttc
gtattagcttacgacgctacacccagttcccatctattttgtcactcttccctaaataa
tccttaaaaactccatttccacccctcccagttcccaactattttgtccgcccacagcg
gggcatttttcttcctgttatgttttttaatcaaacatcctgccaactccatgtgacaaa
ccgtcatcttcggctacttttttctctgtcacagaatgaaaatttttctgtcatctcttc
gttattaatgtttgtaattgactgaatatcaacgcttatttgcagc

The sequence for GFP-39tag is:

atggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctgga
cggcgacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccactt
agggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctggccc
accctcgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgaccacat
gaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcacca
tcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgac
accctggtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcct
ggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaagc
agaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtg
cagctcgccgaccactaccagcagaacaccccatcggcgacggccccgtgctgctgcc
cgacaaccactacctgagcacccagtccgccctgagcaaagaccccaacgagaagcgcg
atcacatggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacgag
ctgtacaaggggcccttcgaacaaaaactcatctcagaagaggatctgaatatgcatac
cggtcatcatcaccatcaccattgataa

pAcBac-4xPytR



Sequence color-coding key

Feature	Color
M13 origin	text
AmpR	text
ColE1 origin	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
WPRE	text
T7 promoter	text
CMV promoter	text
Polyhedrin Promoter	text
L21 sequence	text
VSV-G	text
SV40 late polyA	text
MmPytR	text
U6	text
Tn7L	text

Sequence of 4xPytR-A2 variant

ttctctgtcacagaatgaaaatttttctgtcatctcttcgttattaatgtttgtaattg
actgaatatcaacgcttatttgcagcctgaatggcgaatgggacgcgccctgtagcggc
gcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgc
cctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttc
cccgtcaagctctaaatcgggggctcccttagggttccgatttagtgctttacggcac
ctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgata
gacggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgttcc
aaactggaacaacactcaaccctatctcggtctattcttttgatttataagggattttg
ccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattt
taacaaaatattaacgtttacaatttcaggtggcacttttcggggaaatgtgcgcggaa
cccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataa
ccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccg
tgtcgcccttattccctttttgcggcattttgccttcctgttttgctcacccagaaa
cgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaa
ctggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaat
gatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggc
aagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcacca
gtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccat
aaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaagg
agctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaa
ccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaat
ggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaac
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggccctt
ccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtat
cattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacgg
ggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactg
attaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaa
acttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgacca
aaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaa
ggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaacc
accgctaccagcggtggtttgtttgccggatcaagagctaccaactcttttccgaaggg
taactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagtta
ggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgtt
accagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgat
agttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagc
ttggagcgaacgacctacaccgaactgagatacctacagcgtgagcattgagaaagcgc
cacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacag
gagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcggg
tttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcct
atggaaaaacgccagcaacgcggcctttttacggttcctggccttttgctggccttttg
ctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgccttt
gagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcga
ggaagcggaagagcgcctgatgcggtattttctccttacgcatctgtgcggtatttcac
accgcagaccagccgcgtaacctggcaaaatcggttacggttgagtaataaatggatgc
cctgcgtaagcgggtgtgggcggacaataaagtcttaaactgaacaaaatagatctaaa
ctatgacaataaagtcttaaactagacagaatagttgtaaactgaaatcagtccagtta

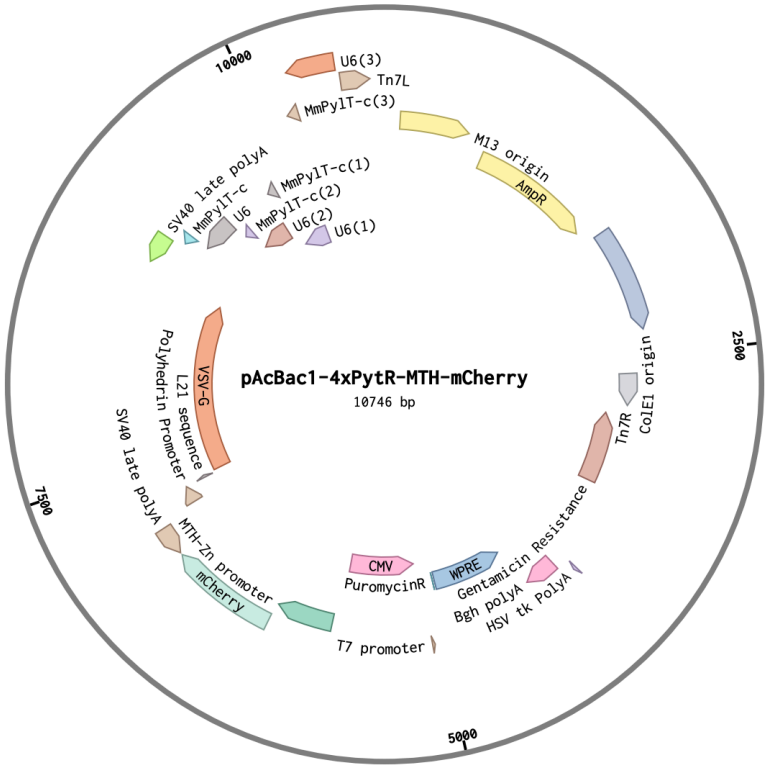
tgctgtgaaaaagcatactggacttttgttatggctaaagcaaactcttcattttctga
agtgcaaattgcccgtcgtattaaagaggggcgtggccaagggcatggtaaagactata
ttcgcggcgttgtgacaatttaccgaacaactccgcggccgggaagccgatctcggctt
gaacgaattgttaggtggcggtacttgggtcgatatcaaagtgcatcacttcttcccgt
atgcccaactttgtatagagagccactgcgggatcgtcaccgtaatctgcttgcacgta
gatcacataagcaccaagcgcgttggcctcatgcttgaggagattgatgagcgcggtgg
caatgccctgcctccggtgctcgccggagactgcgagatcatagatatagatctcacta
cgcggctgctcaaacctgggcagaacgtaagccgcgagagcgccaacaaccgcttcttg
gtcgaaggcagcaagcgcgatgaatgtcttactacggagcaagttcccgaggtaatcgg
agtccggctgatgttgggagtaggtggctacgtctccgaactcacgaccgaaaagatca
agagcagcccgcatggatttgacttggtcagggccgagcctacatgtgcgaatgatgcc
catacttgagccacctaactttgttttagggcgactgccctgctgcgtaacatcgttgc
tgctgcgtaacatcgttgctgctccataacatcaaacatcgacccacggcgtaacgcgc
ttgctgcttggatgcccgaggcatagactgtacaaaaaaacagtcataacaagccatga
aaaccgccactgcgccgttaccaccgctgcgttcggtcaaggttctggaccagttgcgt
gagcgcatacgctacttgcattacagtttacgaaccgaacaggcttatgtcaactgggt
tcgtgccttcatccgtttccacggtgtgcgtcacccggcaaccttgggcagcagcgaag
tcgaggcatttctgtcctggctggcgaacgagcgcaaggtttcggtctccacgcatcgt
caggcattggcggccttgctgttcttctacggcaaggtgctgtgcacggatctgccctg
gcttcaggagatcggtagacctcggccgtcgcggcgcttgccggtggtgctgaccccgg
atgaagtggttcgcatcctcggttttctggaaggcgagcatcgtttgttcgcccaggac
tctagctatagttctagtggttggctacgtacccgtagtggctatggcagggcttgcgc
ttaatgcgccgctacagggcgcgtggggataccccctagagccccagctggttctttcc
gcctcagaagccatagagcccaccgcatccccagcatgcctgctattgtcttcccaatc
ctccccccttgctgtcctgccccaccccaccccccagaatagaatgacacctactcagac
aatgcgatgcaatttcctcattttattaggaaaggacagtgggagtggcaccttccagg
gtcaaggaaggcacgggggaggggcaaacaacagatggctggcaactagaaggcacagt
cgaggctgatcagcgggtttaaacgggccctctagactcgagttaaagtcgacgcgggg
aggcggcccaaagggagatccgactcgtctgagggcgaaggcgaagacgcggaagaggc
cgcagagccggcagcaggccgcgggaaggaaggtccgctggattgagggccgaagggac
gtagcagaaggacgtcccgcgcagaatccaggtggcaacacaggcgagcagccaaggaa
aggacgatgatttccccgacaacaccacggaattgtcagtgcccaacagccgagcccct
gtccagcagcgggcaaggcaggcggcgatgagttccgccgtggcaatagggaggggaa
agcgaaagtcccggaaaggagctgacaggtggtggcaatgccccaaccagtgggggttg
cgtcagcaaacacagtgcacaccacgccacgttgcctgacaacgggccacaactcctca
taaagagacagcaaccaggatttatacaaggaggagaaaatgaaagccatacgggaagc
aatagcatgatacaaaggcattaaagcagcgtatccacatagcgtaaaaggagcaacat
agttaagaataccagtcaatctttcacaaattttgtaatccagaggttgattgtcgact
taacgcgttgaattccaccacactggactagtggatccgagctcggtaccaagcttaag
tttaaacgctagccagcttgggtctccctatagtgagtcgtattaatttcgataagcca
gtaagcagtgggttctctagttagccagagagctctgcttatatagacctcccaccgta
cacgcctaccgcccatttgcgtcaatggggcggagttgttacgacattttggaaagtcc
cgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggagacttggaaa
tccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaaccgcatcacca
tggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcccataaggtc
atgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaataggggcgt
acttggcatatgatacacttgatgtactgccaagtgggcagtttaccgtaaatagtcca
cccattgacgtcaatggaaagtccctattggcgttactatgggaacatacgtcattatt

gacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttaccgtaagtta
tgtaacgcggaactccatatatgggctatgaactaatgaccccgtaattgattactatt
aataactagtcaataatcaatgtcaacgcgtatatctggcccgtacatcgcgaagcagc
gcaaaacggatcctgcaggtatttgcggccgcggtccgtatactccggaatattaatag
atcatggagataattaaaatgataaccatctcgcaaataaataagtattttactgtttt
cgtaacagttttgtaataaaaaaacctataaatattccggattattcataccgtcccac
catcgggcgcgaactcctaaaaaaccgccaccatgaagtgccttttgtacttagccttt
ttattcattggggtgaattgcaagttcaccatagttttccacacaaccaaaaaggaaa
ctggaaaaatgttccttctaattaccattattgcccgtcaagctcagatttaaattggc
ataatgacttaataggcacagccttacaagtcaaaatgcccaagagtcacaaggctatt
caagcagacggttggatgtgtcatgcttccaaatgggtcactacttgtgatttccgctg
gtatggaccgaagtatataacacattccatccgatccttcactccatctgtagaacaat
gcaaggaaagcattgaacaaacgaaacaaggaacttggctgaatccaggcttccctcct
caaagttgtggatatgcaactgtgacggatgccgaagcagtgattgtccaggtgactcc
tcaccatgtgctggttgatgaatacacaggagaatgggttgattcacagttcatcaacg
gaaaatgcagcaattacatatgccccactgtccataactctacaacctggcattctgac
tataaggtcaaagggctatgtgattctaacctcatttccatggacatcaccttcttctc
agaggacggagagctatcatccctgggaaaggagggcacagggttcagaagtaactact
ttgcttatgaaactggaggcaaggcctgcaaaatgcaatactgcaagcattggggagtc
agactcccatcaggtgtctggttcgagatggctgataaggatctctttgctgcagccag
attccctgaatgcccagaagggtcaagtatctctgctccatctcagacctcagtggatg
taagtctaattcaggacgttgagaggatcttggattattccctctgccaagaaacctgg
agcaaaatcagagcgggtcttccaatctctccagtggatctcagctatcttgctcctaa
aaacccaggaaccggtcctgctttcaccataatcaatggtaccctaaaatactttgaga
ccagatacatcagagtcgatattgctgctccaatcctctcaagaatggtcggaatgatc
agtggaactaccacagaaagggaactgtgggatgactgggcaccatatgaagacgtgga
aattggacccaatggagttctgaggaccagttcaggatataagtttcctttatacatga
ttggacatggtatgttggactccgatcttcatcttagctcaaaggctcaggtgttcgaa
catcctcacattcaagacgctgcttcgcaacttcctgatgatgagagtttatttttgg
tgatactgggctatccaaaaatccaatcgagcttgtagaaggttggttcagtagttgga
aaagctctattgcctctttttctttatcatagggttaatcattggactattcttggtt
ctccgagttggtatccatctttgcattaaattaaagcacaccaagaaaagacagattta
tacagacatagagatgaaccgacttggaaagtgataaggccaggccggccaagcttgtc
gagaagtactagaggatcataatcagccataccacatttgtagaggttttacttgcttt
aaaaaacctcccacacctccccctgaacctgaaacataaaatgaatgcaattgttgttg
ttaacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttc
acaaataaagcatttttttcactgcattctagttgtggtttgtccaaactcatcaatgt
atcttatcatgtctggatctgatcactgcttgagcctagcaaaaaacgaggaccccggg
aatctaacccggctgaacggatttagagtccgttcgatctacatgatcaggtccccggt
gtttcgtcctttccacaagatatataaagccaagaaatcgaaatactttcaagttacgg
taagcatatgatagtccatttaaaacataatttaaaactgcaaactacccaagaaat
tattactttctacgtcacgtattttgtactaatatctttgtgtttacagtcaaattaat
tctaattatctctctaacagccttgtatcgtatatgcaaatatgaaggaatcatgggaa
ataggccctcttcctgcccgaggtacccctagcaaaaaacgaggaccccgggaatctaa
cccggctgaacggatttagagtccgttcgatctacatgatcaggtccccggtgtttcgt
cctttccacaagatatataaagccaagaaatcgaaatactttcaagttacggtaagcat
atgatagtccatttaaaacataatttaaaactgcaaactacccaagaaattattact
ttctacgtcacgtattttgtactaatatctttgtgtttacagtcaaattaattctaatt

atctctctaacagccttgtatcgtatatgcaaatatgaaggaatcatgggaaataggcc
ctcttcctgcccgaggtacccctagcaaaaaacgaggaccccgggaatctaacccggct
gaacggatttagagtccgttcgatctacatgatcaggtccccggtgtttcgtcctttcc
acaagatatataaagccaagaaatcgaaatactttcaagttacggtaagcatatgatag
tccattttaaaacataattttaaaactgcaaactacccaagaaattattactttctacg
tcacgtattttgtactaatatctttgtgtttacagtcaaattaattctaattatctctc
taacagccttgtatcgtatatgcaaatatgaaggaatcatgggaaataggccctcttcc
tgcccgaggtacccctagcaaaaaacgaggaccccgggaatctaacccggctgaacgga
tttagagtccgttcgatctacatgatcaggtccccggtgtttcgtcctttccacaagat
atataaagccaagaaatcgaaatactttcaagttacggtaagcatatgatagtccattt
taaaacataattttaaaactgcaaactacccaagaaattattactttctacgtcacgta
ttttgtactaatatctttgtgtttacagtcaaattaattctaattatctctctaacagc
cttgtatcgtatatgcaaatatgaaggaatcatgggaaataggccctcttcctgcccga
ggtaccctaggagatccgaaccagataagtgaaatctagttccaaactattttgtcat
ttttaattttcgtattagcttacgacgctacacccagttcccatctattttgtcactct
tccctaaataatccttaaaaactccatttccacccctcccagttcccaactattttgtc
cgcccacagcggggcatttttcttcctgttatgtttttaatcaaacatcctgccaactc
catgtgacaaaccgtcatcttcggctacttt

The sequence for the wild type PytR is:
ggaaacctgatcatgtagatcgaacggactctaaatccgttcagccgggttagattccc
ggggtttccg

pAcBac-4xPytR-MTH-mCherry



Sequence color-coding key

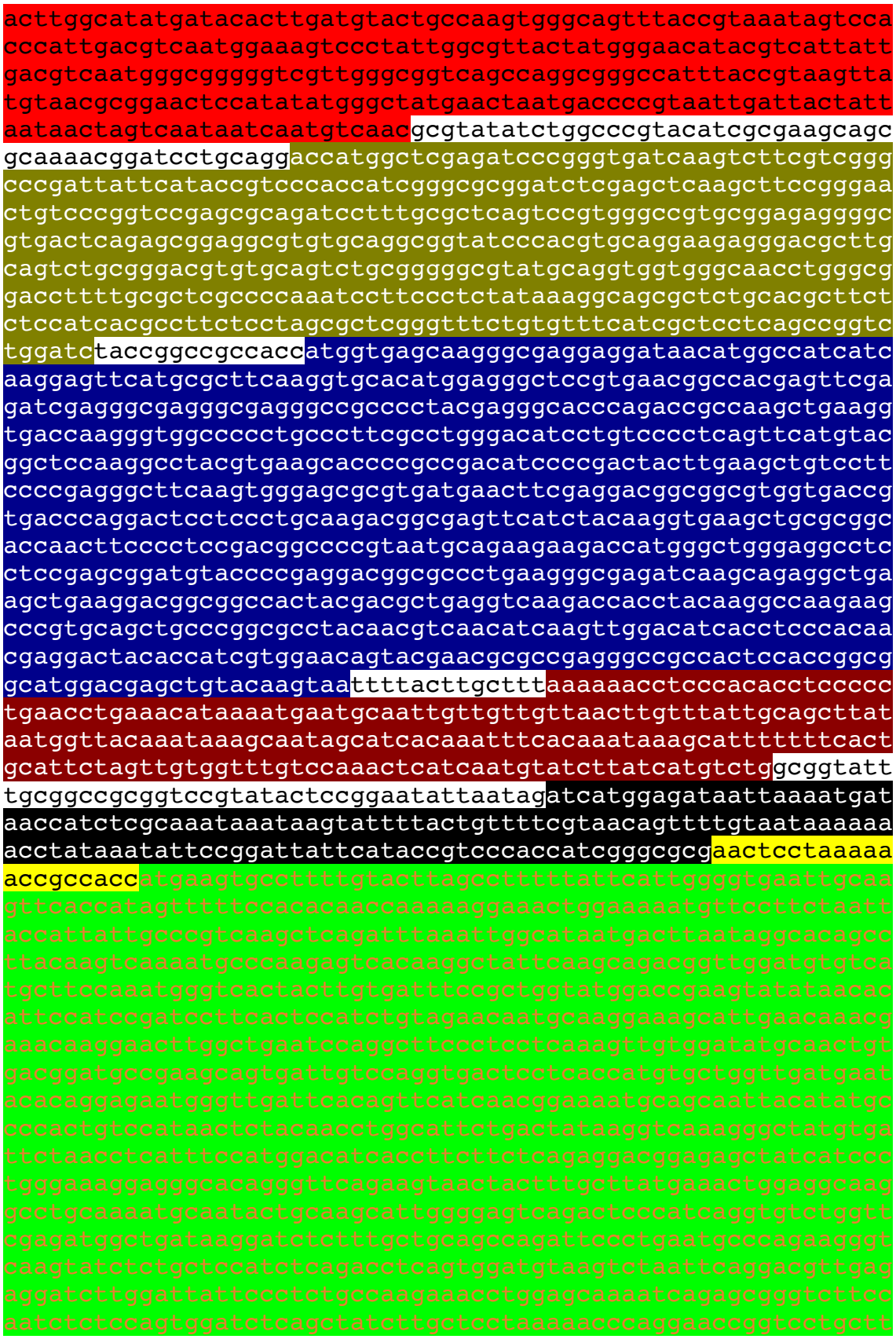
Feature	Color
M13 origin	text
AmpR	text
ColE1 origin	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
WPRE	text
T7 promoter	text
CMV promoter	text
MTH-Zn promoter	text
mCherry	text
SV40 late polyA	text
Polyhedrin Promoter	text
L21 sequence	text
VSV-G	text
SV40 late polyA	text
MmPytR	text
U6	text

Tn7L	text

Sequence of 4xPytR-A2 variant

ttctctgtcacagaatgaaaatttttctgtcatctcttcgttattaatgtttgtaattg
actgaatatcaacgcttatttgcagcctgaatggcgaatgggacgcgccctgtagcggc
gcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgc
cctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttc
cccgtcaagctctaaatcgggggctcccttagggttccgatttagtgctttacggcac
ctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgata
gacggttttcgccctttgacgttggagtccacgttctttaatagtggactcttgttcc
aaactggaacaacactcaaccctatctcggtctattcttttgatttataagggatttg
ccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattt
taacaaaatattaacgtttacaatttcaggtggcacttttcggggaaatgtgcgcggaa
cccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataa
ccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccg
tgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaa
cgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaa
ctggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaat
gatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggc
aagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcacca
gtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccat
aaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaagg
agctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaa
ccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaat
ggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaac
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggccctt
ccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtat
cattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacgg
ggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactg
attaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaa
acttcattttaatttaaaaggatctaggtgaagatccttttgataatctcatgacca
aaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaa
ggatcttcttgagatccttttttttctgcgcgtaatctgctgcttgcaaacaaaaaaacc
accgctaccagcggtggtttgtttgccggatcaagagctaccaactcttttttccgaagg
taactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagtta
ggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgtt
accagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgat
agttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagc
ttggagcgaacgacctacaccgaactgagatacctacagcgtgagcattgagaaagcgc
cacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacag
gagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcggg
tttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcct
atggaaaaacgccagcaacgcggcctttttacggttcctggccttttgctggccttttg
ctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgccttt
gagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcga
ggaagcggaagagcgcctgatgcggtattttctccttacgcatctgtgcggtatttcac
accgcagaccagccgcgtaacctggcaaaatcggttacggttgagtaataaatggatgc

cctgcgtaagcgggtgtgggcggacaataaagtcttaaactgaacaaaatagatctaaa
ctatgacaataaagtcttaaactagacagaatagttgtaaactgaaatcagtccagtta
tgctgtgaaaaagcatactggacttttgttatggctaaagcaaactcttcattttctga
agtgcaaattgcccgtcgtattaaagaggggcgtggccaagggcatggtaaagactata
ttcgcggcgttgtgacaatttaccgaacaactccgcggccgggaagccgatctcggctt
gaacgaattgttaggtggcggtacttgggtcgatatcaaagtgcatcacttcttcccgt
atgcccaactttgtatagagagccactgcgggatcgtcaccgtaatctgcttgcacgta
gatcacataagcaccaagcgcgttggcctcatgcttgaggagattgatgagcgcggtgg
caatgccctgcctccggtgctcgccggagactgcgagatcatagatatagatctcacta
cgcggctgctcaaacctgggcagaacgtaagccgcgagagcgccaacaaccgcttcttg
gtcgaaggcagcaagcgcgatgaatgtcttactacggagcaagttcccgaggtaatcgg
agtccggctgatgttgggagtaggtggctacgtctccgaactcacgaccgaaaagatca
agagcagcccgcatggatttgacttggtcagggccgagcctacatgtgcgaatgatgcc
catacttgagccacctaactttgttttagggcgactgccctgctgcgtaacatcgttgc
tgctgcgtaacatcgttgctgctccataacatcaaacatcgacccacggcgtaacgcgc
ttgctgcttggatgcccgaggcatagactgtacaaaaaaacagtcataacaagccatga
aaaccgccactgcgccgttaccaccgctgcgttcggtcaaggttctggaccagttgcgt
gagcgcatacgctacttgcattacagtttacgaaccgaacaggcttatgtcaactgggt
tcgtgccttcatccgtttccacggtgtgcgtcacccggcaaccttgggcagcagcgaag
tcgaggcatttctgtcctggctggcgaacgagcgcaaggtttcggtctccacgcatcgt
caggcattggcggccttgctgttcttctacggcaaggtgctgtgcacggatctgccctg
gcttcaggagatcggtagacctcggccgtcgcggcgcttgccggtggtgctgaccccgg
atgaagtggttcgcatcctcggttttctggaaggcgagcatcgtttgttcgcccaggac
tctagctatagttctagtggttggctacgtacccgtagtggctatggcagggcttgcgc
ttaatgcgccgctacagggcgcgtggggataccccctagagccccagctggttctttcc
gcctcagaagccatagagcccaccgcatccccagcatgcctgctattgtcttcccaatc
ctccccttgctgtcctgccccaccccaccccccagaatagaatgacacctactcagac
aatgcgatgcaatttcctcattttattaggaaaggacagtgggagtggcaccttccagg
gtcaaggaaggcacgggggagggcaaacaacagatggctggcaactagaaggcacagt
cgaggctgatcagcgggtttaaacgggccctctagactcgagttaaagtcgacgcgggg
aggcggcccaaagggagatccgactcgtctgagggcgaaggcgaagacgcggaagaggc
cgcagagccggcagcaggccgcgggaaggaaggtccgctggattgagggccgaagggac
gtagcagaaggacgtcccgcgcagaatccaggtggcaacacaggcgagcagccaaggaa
aggacgatgatttccccgacaacaccacggaattgtcagtgcccaacagccgagcccct
gtccagcagcgggcaaggcaggcggcgatgagttccgccgtggcaatagggaggggaa
agcgaaagtcccggaaaggagctgacaggtggtggcaatgccccaaccagtgggggttg
cgtcagcaaacacagtgcacaccacgccacgttgcctgacaacgggccacaactcctca
taaagagacagcaaccaggatttatacaaggaggagaaaatgaaagccatacgggaagc
aatagcatgatacaaaggcattaaagcagcgtatccacatagcgtaaaaggagcaacat
agttaagaataccagtcaatctttcacaaattttgtaatccagaggttgattgtcgact
taacgcgttgaattccaccacactggactagtggatccgagctcggtaccaagcttaag
tttaaacgctagccagcttgggtctccctatagtgagtcgtattaatttcgataagcca
gtaagcagtgggttctctagttagccagagagctctgcttatatagacctcccaccgta
cacgcctaccgcccatttgcgtcaatggggcggagttgttacgacattttggaaagtcc
cgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggagacttggaaa
tccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaaccgcatcacca
tggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcccataaggtc
atgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaataggggggcgt

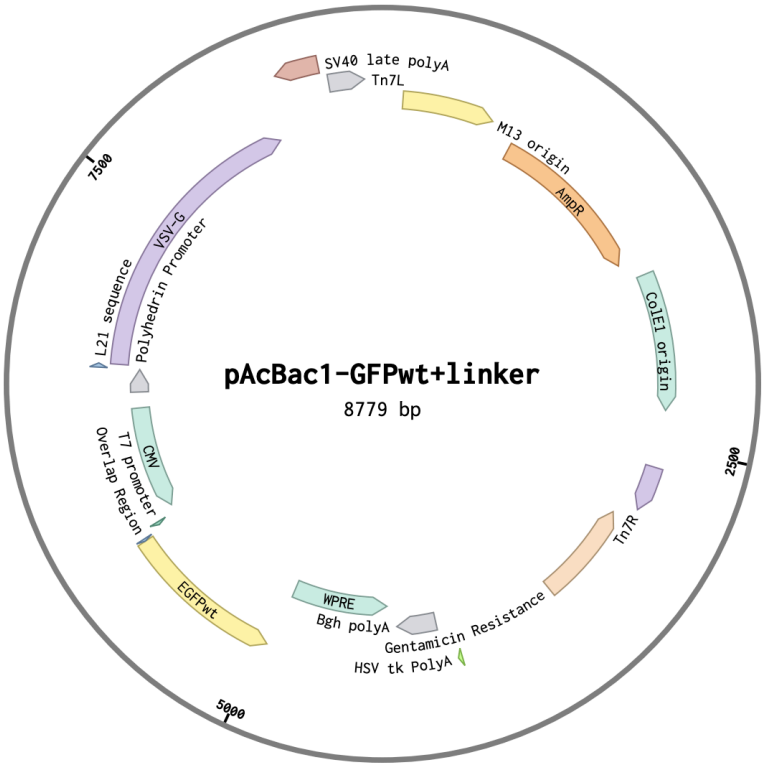

acttggcatatgatacacttgatgtactgccaagtgggcagtttaccgtaaatagtcca
cccattgacgtcaatggaaagtccctattggcgttactatgggaacatacgtcattatt
gacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttaccgtaagtta
tgtaacgcggaactccatatatgggctatgaactaatgaccccgtaattgattactatt
aataactagtcaataatcaatgtcaacgcgtatatctggcccgtacatcgcgaagcagc
gcaaaacggatcctgcaggaccatggctcgagatcccgggtgatcaagtcttcgtcggg
cccgattattcataccgtcccaccatcgggcgcggatctcgagctcaagcttccgggaa
ctgtcccggtccgagcgcagatcctttgcgctcagtccgtgggccgtgcggagaggggc
gtgactcagagcggaggcgtgtgcaggcggtatcccacgtgcaggaagagggacgcttg
cagtctgcgggacgtgtgcagtctgcgggggcgtatgcaggtggtgggcaacctgggcg
gacctttgcgctcgccccaaatccttccctctataaaggcagcgctctgcacgcttct
ctccatcacgccttctcctagcgctcgggtttctgtgtttcatcgctcctcagccggtc
tggatctaccggccgccaccatggtgagcaagggcgaggaggataacatggccatcatc
aaggagttcatgcgcttcaaggtgcacatggagggctccgtgaacggccacgagttcga
gatcgagggcgagggcgagggccgcccctacgagggcacccagaccgccaagctgaagg
tgaccaagggtggcccccctgcccttcgcctgggacatcctgtcccctcagttcatgtac
ggctccaaggcctacgtgaagcaccccgccgacatccccgactacttgaagctgtcctt
ccccgagggcttcaagtgggagcgcgtgatgaacttcgaggacggcggcgtggtgaccg
tgacccaggactcctccctgcaagacggcgagttcatctacaaggtgaagctgcgcggc
accaacttcccctccgacggccccgtaatgcagaagaagaccatgggctgggaggcctc
ctccgagcggatgtaccccgaggacggcgccctgaagggcgagatcaagcagaggctga
agctgaaggacggcggccactacgacgctgaggtcaagaccacctacaaggccaagaag
cccgtgcagctgcccggcgcctacaacgtcaacatcaagttggacatcacctcccacaa
cgaggactacaccatcgtggaacagtacgaacgcgccgagggccgccactccaccggcg
gcatggacgagctgtacaagtaattttacttgctttaaaaaacctcccacacctccccc
tgaacctgaaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttat
aatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcact
gcattctagttgtggtttgtccaaactcatcaatgtatcttatcatgtctggcggtatt
tgcggccgcggtccgtatactccggaatattaatagatcatggagataattaaaatgat
aaccatctcgcaaataaataagtattttactgttttcgtaacagttttgtaataaaaaa
acctataaatattccggattattcataccgtcccaccatcgggcgcgaactcctaaaaa
accgccaccatgaagtgcctttgtacttagccttttattcattggggtgaattgcaa
gttcaccatagtttttccacacaaccaaaaaggaaactggaaaaatgttccttctaatt
accattattgcccgtcaagctcagatttaaattggcataatgacttaataggcacagcc
ttacaagtcaaaatgcccaagagtcacaaggctattcaagcagacggttggatgtgtca
tgcttccaaatgggtcactacttgtgatttccgctggtatggaccgaagtatataacac
attccatccgatccttcactccatctgtagaacaatgcaaggaaagcattgaacaaacg
aaacaaggaacttggctgaatccaggcttccctcctcaaagttgtggatatgcaactgt
gacggatgccgaagcagtgattgtccaggtgactcctcaccatgtgctggttgatgaat
acacaggagaatgggttgattcacagttcatcaacggaaaatgcagcaattacatatgc
cccactgtccataactctacaacctggcattctgactataaggtcaaagggctatgtga
ttctaacctcatttccatggacatcaccttcttctcagaggacggagagctatcatccc
tgggaaaggagggcacagggttcagaagtaactactttgcttatgaaactggaggcaag
gcctgcaaaatgcaatactgcaagcattggggagtcagactcccatcaggtgtctggtt
cgagatggctgataaggatctctttgctgcagccagattccctgaatgcccagaagggt
caagtatctctgctccatctcagacctcagtggatgtaagtctaattcaggacgttgag
aggatcttggattattccctctgccaagaaacctggagcaaaatcagagcgggtcttcc
aatctctccagtggatctcagctatcttgctcctaaaaacccaggaaccggtcctgctt

tcaccataatcaatggtaccctaaaatactttgagaccagatacatcagagtcgatatt
gctgctccaatcctctcaagaatggtcggaatgatcagtggaactaccacagaaaggga
actgtgggatgactgggcaccatatgaagacgtggaaattggacccaatggagttctga
ggaccagttcaggatataagtttcctttatacatgattggacatggtatgttggactcc
gatcttcatcttagctcaaaggctcaggtgttcgaacatcctcacattcaagacgctgc
ttcgcaacttcctgatgatgagagtttatttttggtgatactgggctatccaaaaatc
caatcgagcttgtagaaggttggttcagtagttggaaaagctctattgcctctttttc
tttatcatagggttaatcattggactattcttggttctccgagttggtatccatctttg
cattaaattaaagcacaccaagaaaagacagatttatacagacatagagatgaaccgac
ttggaaagtgataaggccaggccggccaagcttgtcgagaagtactagaggatcataat
cagccataccacatttgtagaggttttacttgctttaaaaaacctcccacacctccccc
tgaacctgaaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttat
aatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcact
gcattctagttgtggtttgtccaaactcatcaatgtatcttatcatgtctggatctgat
cactgcttgagcctagcaaaaaacgaggaccccgggaatctaacccggctgaacggatt
tagagtccgttcgatctacatgatcaggtccccggtgtttcgtcctttccacaagatat
ataaagccaagaaatcgaaatactttcaagttacggtaagcatatgatagtccatttta
aaacataattttaaaactgcaaactacccaagaaattattactttctacgtcacgtatt
ttgtactaatatctttgtgtttacagtcaaattaattctaattatctctctaacagcct
tgtatcgtatatgcaaatatgaaggaatcatgggaaataggccctcttcctgcccgagg
taccctagcaaaaaacgaggaccccgggaatctaacccggctgaacggatttagagtc
cgttcgatctacatgatcaggtccccggtgtttcgtcctttccacaagatatataaagc
caagaaatcgaaatactttcaagttacggtaagcatatgatagtccattttaaaacata
attttaaaactgcaaactacccaagaaattattactttctacgtcacgtattttgtact
aatatctttgtgtttacagtcaaattaattctaattatctctctaacagccttgtatcg
tatatgcaaatatgaaggaatcatgggaaataggccctcttcctgcccgaggtaccct
agcaaaaaacgaggaccccgggaatctaacccggctgaacggatttagagtccgttcga
tctacatgatcaggtccccggtgtttcgtcctttccacaagatatataaagccaagaaa
tcgaaatactttcaagttacggtaagcatatgatagtccattttaaaacataattttaa
aactgcaaactacccaagaaattattactttctacgtcacgtattttgtactaatatct
ttgtgtttacagtcaaattaattctaattatctctctaacagccttgtatcgtatatgc
aaatatgaaggaatcatgggaaataggccctcttcctgcccgaggtaccctagcaaaa
aacgaggaccccgggaatctaacccggctgaacggatttagagtccgttcgatctacat
gatcaggtccccggtgtttcgtcctttccacaagatatataaagccaagaaatcgaaat
actttcaagttacggtaagcatatgatagtccattttaaaacataattttaaaactgca
aactacccaagaaattattactttctacgtcacgtattttgtactaatatctttgtgtt
tacagtcaaattaattctaattatctctctaacagccttgtatcgtatatgcaaatatg
aaggaatcatgggaaataggccctcttcctgcccgaggtaccctaggagatccgaacc
agataagtgaaatctagttccaaactattttgtcatttttaattttcgtattagcttac
gacgctacacccagttcccatctattttgtcactcttccctaaataatccttaaaaact
ccatttccacccctcccagttcccaactattttgtccgcccacagcggggcatttttct
tcctgttatgtttttaatcaaacatcctgccaactccatgtgacaaaccgtcatcttcg
gctacttt

The sequence for the wild type PytR is:

ggaaacctgatcatgtagatcgaacggactctaaatccgttcagccgggttagattccc
ggggtttccg

pAcBac1-GFPwt+linker



Sequence color-coding key

Feature	Color
M13 origin	text
AmpR	text
ColE1 origin	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
WPRE	text
EGFPwt	text
Overlap region	text
T7 promoter	text
CMV promoter	text
Polyhedrin Promoter	text
L21 sequence	text
VSV-G	text
SV40 late polyA	text
Tn7L	text

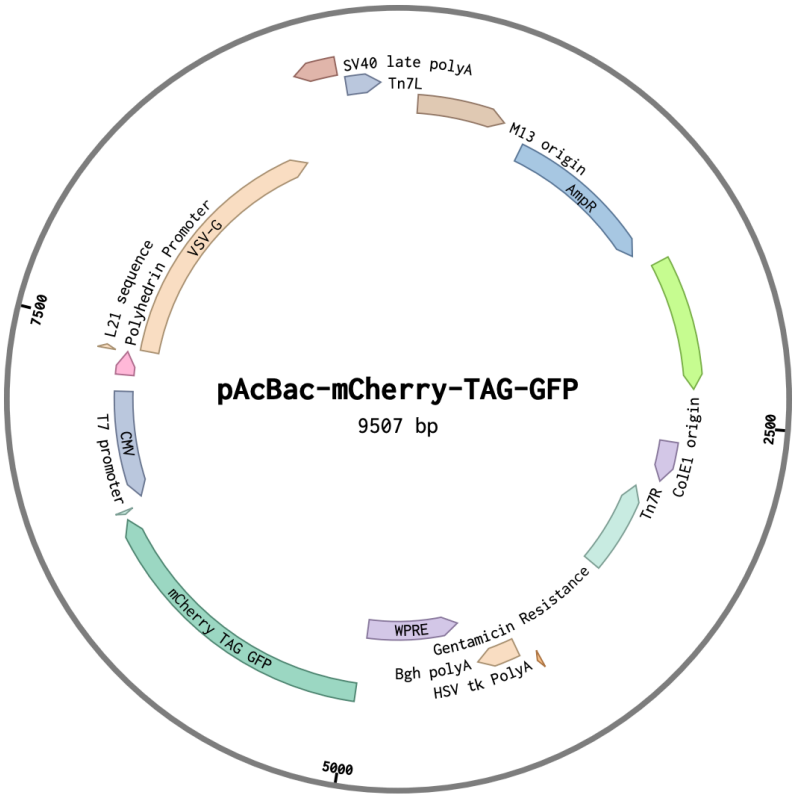
ttctctgtcacagaatgaaaatttttctgtcatctcttcgttattaatgtttgtaattg
actgaatatcaacgcttatttgcagcctgaatggcgaatgggacgcgccctgtagcggc
gcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgc
cctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttc
cccgtcaagctctaaatcgggggctcccttagggttccgatttagtgctttacggcac
ctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgata
gacggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgttcc
aaactggaacaacactcaaccctatctcggtctattcttttgatttataagggatttg
ccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattt
taacaaaatattaacgtttacaatttcaggtggcacttttcggggaaatgtgcgcggaa
cccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataa
ccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccg
tgtcgcccttattccctttttgcggcattttgccttcctgtttttgctcacccagaaa
cgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaa
ctggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaat
gatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggc
aagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcacca
gtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccat
aaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaagg
agctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaa
ccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaat
ggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaac
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggccctt
ccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtat
cattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacgg
ggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactg
attaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaa
acttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgacca
aaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaa
ggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaacc
accgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaagg
taactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagtta
ggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgtt
accagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgat
agttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagc
ttggagcgaacgacctacaccgaactgagatacctacagcgtgagcattgagaaagcgc
cacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacag
gagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcggg
tttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcct
atggaaaaacgccagcaacgcggcctttttacggttcctggccttttgctggccttttg
ctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgccttt
gagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcga
ggaagcggaagagcgcctgatgcggtattttctccttacgcatctgtgcggtatttcac
accgcagaccagccgcgtaacctggcaaaatcggttacggttgagtaataaatggatgc
cctgcgtaagcgggtgtgggcggacaataaagtcttaaactgaacaaaatagatctaaa
ctatgacaataaagtcttaaactagacagaatagttgtaaactgaaatcagtccagtta

tgctgtgaaaaagcatactggacttttgttatggctaaagcaaactcttcattttctga
agtgcaaattgcccgtcgtattaaagaggggcgtggccaagggcatggtaaagactata
ttcgcggcgttgtgacaatttaccgaacaactccgcggccgggaagccgatctcggctt
gaacgaattgttaggtggcggtacttgggtcgatatcaaagtgcatcacttcttcccgt
atgcccaactttgtatagagagccactgcgggatcgtcaccgtaatctgcttgcacgta
gatcacataagcaccaagcgcgttggcctcatgcttgaggagattgatgagcgcggtgg
caatgccctgcctccggtgctcgccggagactgcgagatcatagatatagatctcacta
cgcggctgctcaaacctgggcagaacgtaagccgcgagagcgccaacaaccgcttcttg
gtcgaaggcagcaagcgcgatgaatgtcttactacggagcaagttcccgaggtaatcgg
agtccggctgatgttgggagtaggtggctacgtctccgaactcacgaccgaaaagatca
agagcagcccgcatggatttgacttggtcagggccgagcctacatgtgcgaatgatgcc
catacttgagccacctaactttgttttagggcgactgccctgctgcgtaacatcgttgc
tgctgcgtaacatcgttgctgctccataacatcaaacatcgacccacggcgtaacgcgc
ttgctgcttggatgcccgaggcatagactgtacaaaaaaacagtcataacaagccatga
aaaccgccactgcgccgttaccaccgctgcgttcggtcaaggttctggaccagttgcgt
gagcgcatacgctacttgcattacagtttacgaaccgaacaggcttatgtcaactgggt
tcgtgccttcatccgtttccacggtgtgcgtcacccggcaaccttgggcagcagcgaag
tcgaggcatttctgtcctggctggcgaacgagcgcaaggtttcggtctccacgcatcgt
caggcattggcggccttgctgttcttctacggcaaggtgctgtgcacggatctgccctg
gcttcaggagatcggtagacctcggccgtcgcggcgcttgccggtggtgctgaccccgg
atgaagtggttcgcatcctcggttttctggaaggcgagcatcgtttgttcgcccaggac
tctagctatagttctagtggttggcctacgtacccgtagtggctatggcagggcttgcg
cttaatgcgccgctacagggcgcgtggggataccccctagagccccagctggttctttc
cgcctcagaagccatagagcccaccgcatccccagcatgcctgctattgtcttcccaat
cctccccttgctgtcctgccccaccccaccccccagaatagaatgacacctactcaga
caatgcgatgcaatttcctcattttattaggaaaggacagtgggagtggcaccttccag
ggtcaaggaaggcacgggggagggcaaacaacagatggctggcaactagaaggcacag
tcgaggctgatcagcgggtttaaacgggccctctagactcgagttaaagtcgacgcggg
gaggcggcccaaagggagatccgactcgtctgagggcgaaggcgaagacgcggaagagg
ccgcagagccggcagcaggccgcgggaaggaaggtccgctggattgagggccgaaggga
cgtagcagaaggacgtcccgcgcagaatccaggtggcaacacaggcgagcagccaagga
aaggacgatgatttccccgacaacaccacggaattgtcagtgcccaacagccgagcccc
tgtccagcagcgggcaaggcaggcggcgatgagttccgccgtggcaatagggaggggga
aagcgaaagtcccggaaaggagctgacaggtggtggcaatgccccaaccagtgggggtt
gcgtcagcaaacacagtgcacaccacgccacgttgcctgacaacgggccacaactcctc
ataaagagacagcaaccaggatttatacaaggaggagaaaatgaaagccatacgggaag
caatagcatgatacaaaggcattaaagcagcgtatccacatagcgtaaaaggagcaaca
tagttaagaataccagtcaatctttcacaaattttgtaatccagaggttgattgtcgac
ttaacgcgttggaattcctcattaatggtgatggtgatgatgaccggtatgcatattca
gatcctcttctgagatgagttttgttcgaagggccccttgtacagctcgtccatgccg
agagtgatcccggcggcggtcacgaactccagcaggaccatgtgatcgcgcttctcgtt
ggggtctttgctcagggcggactgggtgctcaggtagtggttgtcgggcagcagcacgg
ggccgtcgccgatgggggtgttctgctggtagtggtcggcgagctgcacgctgccgtcc
tcgatgttgtggcggatcttgaagttcaccttgatgccgttcttctgcttgtcggccat
gatatagacgttgtggctgttgtagttgtactccagcttgtgccccaggatgttgccgt
cctccttgaagtcgatgcccttcagctcgatgcggttcaccagggtgtcgccctcgaac
ttcacctcggcgcgggtcttgtagttgccgtcgtccttgaagaagatggtgcgctcctg
gacgtagccttcgggcatggcggacttgaagaagtcgtgctgcttcatgtggtcggggt

agcggctgaagcactgcacgccgtaggtcagggtggtcacgagggtgggccagggcacg
ggcagcttgccggtggtgcagatgaacttcagggtcagcttgccgtaggtggcatcgcc
ctcgccctcgccggacacgctgaacttgtggccgtttacgtcgccgtccagctcgacca
ggatgggcaccaccccggtgaacagctcctcgcccttgctcacggagccgccctgctag
ccagcttgggtctccctatagtgagtcgtattaatttcgataagccagtaagcagtggg
ttctctagttagccagagagctctgcttatatagacctcccaccgtacacgcctaccgc
ccatttgcgtcaatggggcggagttgttacgacattttggaaagtcccgttgattttgg
tgccaaaacaaactcccattgacgtcaatggggtggagacttggaaatccccgtgagtc
aaaccgctatccacgcccattgatgtactgccaaaaccgcatcaccatggtaatagcga
tgactaatacgtagatgtactgccaagtaggaaagtcccataaggtcatgtactgggca
taatgccaggcgggccatttaccgtcattgacgtcaataggggggcgtacttggcatatg
atacacttgatgtactgccaagtgggcagtttaccgtaaatagtccacccattgacgtc
aatggaaagtccctattggcgttactatgggaacatacgtcattattgacgtcaatggg
cgggggtcgttgggcggtcagccaggcgggccatttaccgtaagttatgtaacgcggaa
ctccatatatgggctatgaactaatgaccccgtaattgattactattaataactagtca
ataatcaatgtcaacgcgtatatctggcccgtacatcgcgaagcagcgcaaaacggatc
ctgcaggtatttgcggccgcggtccgtatactccggaatattaatagatcatggagata
attaaaatgataaccatctcgcaaataaataagtattttactgttttcgtaacagtttt
gtaataaaaaaacctataaatattccggattattcataccgtcccaccatcgggcgcga
actcctaaaaaaccgccacc atgaagtgcctttgtacttagccttttattcattggg
gtgaattgcaagttcaccatagttttccacacaaccaaaaaggaaactggaaaaatgt
tccttctaattaccattattgcccgtcaagctcagatttaaattggcataatgacttaa
taggcacagccttacaagtcaaaatgcccaagagtcacaaggctattcaagcagacggt
tggatgtgtcatgcttccaaatgggtcactacttgtgatttccgctggtatggaccgaa
gtatataacacattccatccgatccttcactccatctgtagaacaatgcaaggaaagca
ttgaacaaacgaaacaaggaacttggctgaatccaggcttccctcctcaaagttgtgga
tatgcaactgtgacggatgccgaagcagtgattgtccaggtgactcctcaccatgtgct
ggttgatgaatacacaggagaatgggttgattcacagttcatcaacggaaaatgcagca
attacatatgccccactgtccataactctacaacctggcattctgactataaggtcaaa
gggctatgtgattctaacctcatttccatggacatcaccttcttctcagaggacggaga
gctatcatccctgggaaaggagggcacagggttcagaagtaactactttgcttatgaaa
ctggaggcaaggcctgcaaaatgcaatactgcaagcattggggagtcagactcccatca
ggtgtctggttcgagatggctgataaggatctctttgctgcagccagattccctgaatg
cccagaagggtcaagtatctctgctccatctcagacctcagtggatgtaagtctaattc
aggacgttgagaggatcttggattattccctctgccaagaaacctggagcaaaatcaga
gcgggtcttccaatctctccagtggatctcagctatcttgctcctaaaaacccaggaac
cggtcctgctttcaccataatcaatggtaccctaaaatactttgagaccagatacatca
gagtcgatattgctgctccaatcctctcaagaatggtcggaatgatcagtggaactacc
acagaaagggaactgtgggatgactgggcaccatatgaagacgtggaaattggacccaa
tggagttctgaggaccagttcaggatataagtttcctttatacatgattggacatggta
tgttggactccgatcttcatcttagctcaaaggctcaggtgttcgaacatcctcacatt
caagacgctgcttcgcaacttcctgatgatgagagtttatttttggtgatactgggct
atccaaaaatccaatcgagcttgtagaaggttggttcagtagttggaaaagctctattg
cctctttttctttatcatagggttaatcattggactattcttggttctccgagttggt
atccatctttgcattaaattaaagcacaccaagaaaagacagatttatacagacataga
gatgaaccgacttggaaagtgataaggccaggccggccaagcttgtcgagaagtactag
aggatcataatcagccataccacatttgtagaggttttacttgctttaaaaaacctccc
acacctcccctgaacctgaaacataaaatgaatgcaattgttgttgttaacttgtttta

ttgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagca
ttttttcactgcattctagttgtggtttgtccaaactcatcaatgtatcttatcatgt
ctggatctgatcactgcttgagcctaggagatccgaaccagataagtgaaatctagttc
caaactattttgtcatttttaattttcgtattagcttacgacgctacacccagttccca
tctattttgtcactcttccctaaataatccttaaaaactccatttccacccctcccagt
tcccaactattttgtccgcccacagcggggcatttttcttcctgttatgttttaatca
aacatcctgccaactccatgtgacaaaccgtcatcttcggctacttt

pAcBac-mCherry-TAG-GFP



Sequence color-coding key

Feature	Color
M13 origin	text
AmpR	text
ColE1 origin	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
WPRE	text
mCherry-TAG-GFP	text
T7 promoter	text
CMV promoter	text
Polyhedrin Promoter	text
L21 sequence	text
VSV-G	text
SV40 late polyA	text
Tn7L	text

Sequence

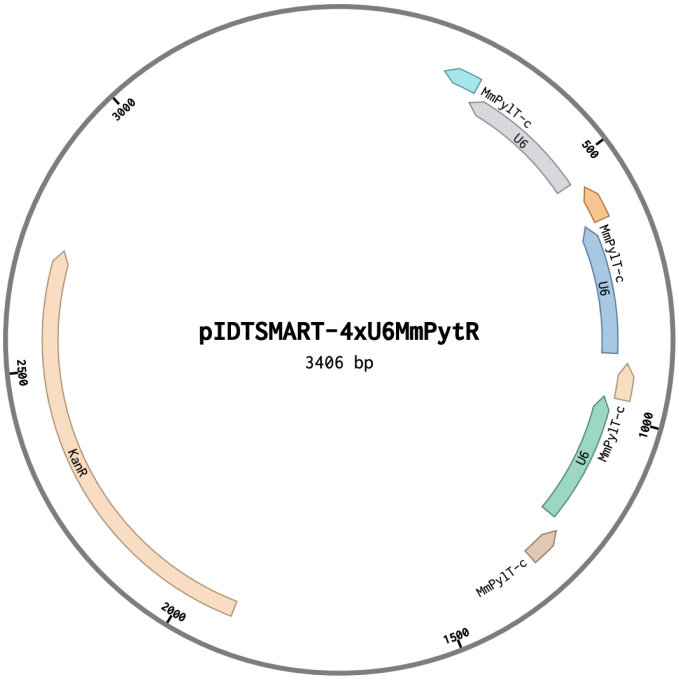
ttctctgtcacagaatgaaaatttttctgtcatctcttcgttattaatgtttgtaattg
actgaatatcaacgcttatttgcagcctgaatggcgaatgggacgcgccctgtagcggc
gcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgc
cctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttc
cccgtcaagctctaaatcgggggctcccttagggttccgatttagtgctttacggcac
ctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgata
gacggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgttcc
aaactggaacaacactcaaccctatctcggtctattcttttgatttataagggattttg
ccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattt
taacaaaatattaacgtttacaatttcaggtggcacttttcggggaaatgtgcgcggaa
cccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataa
ccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccg
tgtcgcccttattccctttttgcggcattttgccttcctgttttgctcacccagaaa
cgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaa
ctggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaat
gatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggc
aagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcacca
gtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccat
aaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaagg
agctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaa
ccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaat
ggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaac
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggccctt
ccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtat
cattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacgg
ggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactg
attaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaa
acttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgacca
aaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaa
ggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaacc
accgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaagg
taactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagtta
ggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgtt
accagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgat
agttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagc
ttggagcgaacgacctacaccgaactgagatacctacagcgtgagcattgagaaagcgc
cacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacag
gagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcggg
tttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcct
atggaaaaacgccagcaacgcggcctttttacggttcctggccttttgctggccttttg
ctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgccttt
gagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcga
ggaagcggaagagcgcctgatgcggtattttctccttacgcatctgtgcggtatttcac
accgcagaccagccgcgtaacctggcaaaatcggttacggttgagtaataaatggatgc
cctgcgtaagcgggtgtgggcggacaataaagtcttaaactgaacaaaatagatctaaa
ctatgacaataaagtcttaaactagacagaatagttgtaaactgaaatcagtccagtta

tgctgtgaaaaagcatactggacttttgttatggctaaagcaaactcttcattttctga
agtgcaaattgcccgtcgtattaaagaggggcgtggccaagggcatggtaaagactata
ttcgcggcgttgtgacaatttaccgaacaactccgcggccgggaagccgatctcggctt
gaacgaattgttaggtggcggtacttgggtcgatatcaaagtgcatcacttcttcccgt
atgcccaactttgtatagagagccactgcgggatcgtcaccgtaatctgcttgcacgta
gatcacataagcaccaagcgcgttggcctcatgcttgaggagattgatgagcgcggtgg
caatgccctgcctccggtgctcgccggagactgcgagatcatagatatagatctcacta
cgcggctgctcaaacctgggcagaacgtaagccgcgagagcgccaacaaccgcttcttg
gtcgaaggcagcaagcgcgatgaatgtcttactacggagcaagttcccgaggtaatcgg
agtccggctgatgttgggagtaggtggctacgtctccgaactcacgaccgaaaagatca
agagcagcccgcatggatttgacttggtcagggccgagcctacatgtgcgaatgatgcc
catacttgagccacctaactttgttttagggcgactgccctgctgcgtaacatcgttgc
tgctgcgtaacatcgttgctgctccataacatcaaacatcgacccacggcgtaacgcgc
ttgctgcttggatgcccgaggcatagactgtacaaaaaaacagtcataacaagccatga
aaaccgccactgcgccgttaccaccgctgcgttcggtcaaggttctggaccagttgcgt
gagcgcatacgctacttgcattacagtttacgaaccgaacaggcttatgtcaactgggt
tcgtgccttcatccgtttccacggtgtgcgtcacccggcaaccttgggcagcagcgaag
tcgaggcatttctgtcctggctggcgaacgagcgcaaggtttcggtctccacgcatcgt
caggcattggcggccttgctgttcttctacggcaaggtgctgtgcacggatctgccctg
gcttcaggagatcggtagacctcggccgtcgcggcgcttgccggtggtgctgaccccgg
atgaagtggttcgcatcctcggttttctggaaggcgagcatcgtttgttcgcccaggac
tctagctatagttctagtggttggcctacgtacccgtagtggctatggcagggcttgcg
cttaatgcgccgctacagggcgcgtggggataccccctagagccccagctggttctttc
cgcctcagaagccatagagcccaccgcatccccagcatgcctgctattgtcttcccaat
cctccccttgctgtcctgccccaccccaccccccagaatagaatgacacctactcaga
caatgcgatgcaatttcctcattttattaggaaaggacagtgggagtggcaccttccag
ggtcaaggaaggcacgggggagggcaaacaacagatggctggcaactagaaggcacag
tcgaggctgatcagcgggtttaaacgggccctctagactcgagttaaagtcgacgcggg
gaggcggcccaaagggagatccgactcgtctgagggcgaaggcgaagacgcggaagagg
ccgcagagccggcagcaggccgcgggaaggaaggtccgctggattgagggccgaaggga
cgtagcagaaggacgtcccgcgcagaatccaggtggcaacacaggcgagcagccaagga
aaggacgatgatttccccgacaacaccacggaattgtcagtgcccaacagccgagcccc
tgtccagcagcgggcaaggcaggcggcgatgagttccgccgtggcaatagggaggggga
aagcgaaagtcccggaaaggagctgacaggtggtggcaatgccccaaccagtgggggtt
gcgtcagcaaacacagtgcacaccacgccacgttgcctgacaacgggccacaactcctc
ataaagagacagcaaccaggatttatacaaggaggagaaaatgaaagccatacgggaag
caatagcatgatacaaaggcattaaagcagcgtatccacatagcgtaaaaggagcaaca
tagttaagaataccagtcaatctttcacaaattttgtaatccagaggttgattgtcgac
ttaacgcgttgaattctcattaatggtgatggtgatgatgaccggtatgcatattcaga
tcctcttctgagatgagttttgttcgaagggccccttgtacagctcgtccatgccgag
agtgatcccggcggcggtcacgaactccagcaggaccatgtgatcgcgcttctcgttgg
ggtctttgctcagggcggactgggtgctcaggtagtggttgtcgggcagcagcacgggg
ccgtcgccgatgggggtgttctgctggtagtggtcggcgagctgcacgctgccgtcctc
gatgttgtggcggatcttgaagttcaccttgatgccgttcttctgcttgtcggccatga
tatagacgttgtggctgttgtagttgtactccagcttgtgccccaggatgttgccgtcc
tccttgaagtcgatgcccttcagctcgatgcggttcaccagggtgtcgccctcgaactt
cacctcggcgcgggtcttgtagttgccgtcgtccttgaagaagatggtgcgctcctgga
cgtagccttcgggcatggcggacttgaagaagtcgtgctgcttcatgtggtcggggtag

cggctgaagcactgcacgccgtaggtcagggtggtcacgagggtgggccagggcacggg
cagcttgccggtggtgcagatgaacttcagggtcagcttgccgtaggtggcatcgccct
cgccctcgccggacacgctgaacttgtggccgtttacgtcgccgtccagctcgaccagg
atgggcaccaccccggtgaacagctcctcgcccttgctcacgagggagccgccctaaga
gccgcccttgtacagctcgtccatgccgccggtggagtggcggccctcggcgcgttcgt
actgttccacgatggtgtagtcctcgttgtgggaggtgatgtccaacttgatgttgacg
ttgtaggcgccgggcagctgcacgggcttcttggccttgtaggtggtcttgacctcagc
gtcgtagtggccgccgtccttcagcttcagcctctgcttgatctcgcccttcagggcgc
cgtcctcggggtacatccgctcggaggaggcctcccagcccatggtcttcttctgcatt
acggggccgtcggaggggaagttggtgccgcgcagcttcaccttgtagatgaactcgcc
gtcctgcagggaggagtcctgggtcacggtcaccacgccgccgtcctcgaagttcatca
cgcgctcccacttgaagccctcggggaaggacagcttcaagtagtcggggatgtcggcg
gggtgcttcacgtaggccttggagccgtacatgaactgaggggacaggatgtcccaggc
gaagggcaggggccacccttggtcaccttcagcttggcggtctgggtgccctcgtagg
ggcggccctcgccctcgccctcgatctcgaactcgtggccgttcacggagccctccatg
tgcaccttgaagcgcatgaactccttgatgatggccatgttatcctcctcgcccttgct
caccatggtggcggcgctagccagcttgggtctccctatagtgagtcgtattaatttcg
ataagccagtaagcagtgggttctctagttagccagagagctctgcttatatagacctc
ccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacattttg
gaaagtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggaga
cttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaaccg
catcaccatggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtccc
ataaggtcatgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaata
gggggcgtacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgtaa
atagtccacccattgacgtcaatggaaagtccctattggcgttactatgggaacatacg
tcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttacc
gtaagttatgtaacgcggaactccatatatgggctatgaactaatgaccccgtaattga
ttactattaataactagtcaataatcaatgtcaacgcgtatatctggcccgtacatcgc
gaagcagcgcaaaacggatcctgcaggtatttgcggccgcggtccgtatactccggaat
attaatagatcatggagataattaaaatgataaccatctcgcaaataaataagtatttt
actgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcatac
cgtcccaccatcgggcgcgaactcctaaaaaaccgccacc atgaagtgccttttgtact
tagccttttattcattggggtgaattgcaagttcaccatagttttccacacaaccaa
aaaggaaactggaaaaatgttccttctaattaccattattgcccgtcaagctcagattt
aaattggcataatgacttaataggcacagccttacaagtcaaaatgcccaagagtcaca
aggctattcaagcagacggttggatgtgtcatgcttccaaatgggtcactacttgtgat
ttccgctggtatggaccgaagtatataacacattccatccgatccttcactccatctgt
agaacaatgcaaggaaagcattgaacaaacgaaacaaggaacttggctgaatccaggct
tccctcctcaaagttgtggatatgcaactgtgacggatgccgaagcagtgattgtccag
gtgactcctcaccatgtgctggttgatgaatacacaggagaatgggttgattcacagtt
catcaacggaaaatgcagcaattacatatgccccactgtccataactctacaacctggc
attctgactataaggtcaaagggctatgtgattctaacctcatttccatggacatcacc
ttcttctcagaggacggagagctatcatccctgggaaaggagggcacagggttcagaag
taactactttgcttatgaaactggaggcaaggcctgcaaaatgcaatactgcaagcatt
ggggagtcagactcccatcaggtgtctggttcgagatggctgataaggatctctttgct
gcagccagattccctgaatgcccagaagggtcaagtatctctgctccatctcagacctc
agtggatgtaagtctaattcaggacgttgagaggatcttggattattccctctgccaag
aaacctggagcaaaatcagagcgggtcttccaatctctccagtggatctcagctatctt

gctcctaaaaacccaggaaccggtcctgctttcaccataatcaatggtaccctaaaata
ctttgagaccagatacatcagagtcgatattgctgctccaatcctctcaagaatggtcg
gaatgatcagtggaactaccacagaaagggaactgtgggatgactgggcaccatatgaa
gacgtggaaattggacccaatggagttctgaggaccagttcaggatataagtttccttt
atacatgattggacatggtatgttggactccgatcttcatcttagctcaaaggctcagg
tgttcgaacatcctcacattcaagacgctgcttcgcaacttcctgatgatgagagttta
ttttttggtgatactgggctatccaaaaatccaatcgagcttgtagaaggttggttcag
tagttggaaaagctctattgcctctttttctttatcatagggttaatcattggactat
tcttggttctccgagttggtatccatctttgcattaaattaaagcacaccaagaaaaga
cagatttatacagacatagagatgaaccgacttggaaagtgataaggccaggccggcca
agcttgtcgagaagtactagaggatcataatcagccataccacatttgtagaggtttta
cttgctttaaaaaacctcccacacctccccctgaacctgaaacataaaatgaatgcaat
tgttgttgttaacttgtttattgcagcttataatggttacaaataaagcaatagcatca
caaatttcacaaataaagcatttttttcactgcattctagttgtggtttgtccaaactc
atcaatgtatcttatcatgtctggatctgatcactgcttgagcctaggagatccgaacc
agataagtgaaatctagttccaaactattttgtcatttttaatttcgtattagcttac
gacgctacacccagttcccatctattttgtcactcttccctaaataatccttaaaaact
ccatttccaccctcccagttcccaactattttgtccgcccacagcggggcatttttct
tcctgttatgttttaatcaaacatcctgccaactccatgtgacaaaccgtcatcttcg
gctacttt

pIDTSMART-4xU6MmPytR



Sequence color-coding key

Feature	Color
MmPytR	text
U6	text
KanR	text

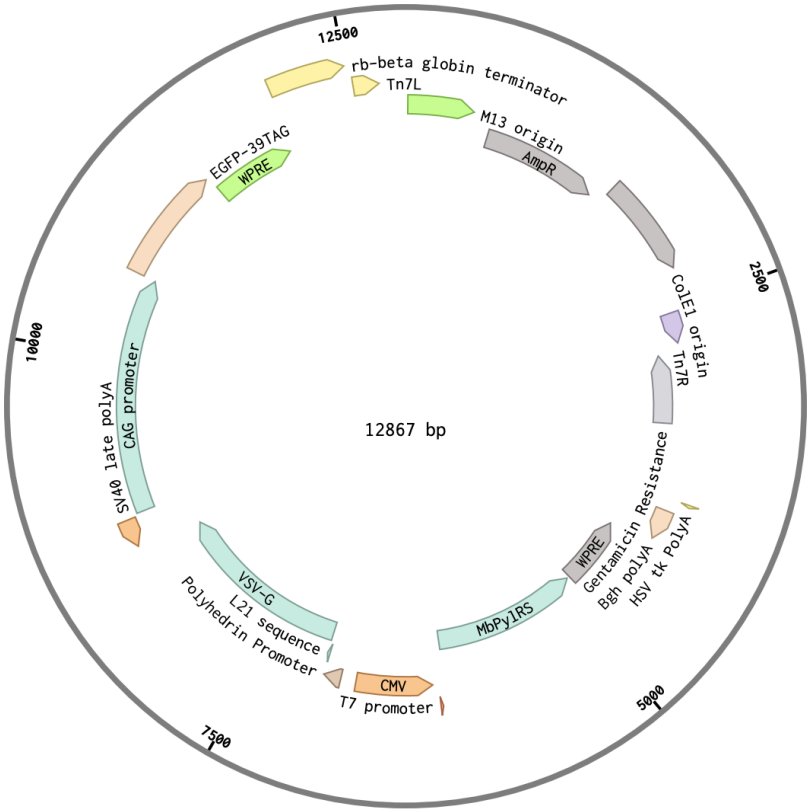
Sequence for HTS25 construct

cccgtgtaaaacgacggccagtttatctagtcagcttgattctagctgatcgtggaccg
gaaggtgagccagtgagttgattgcagtccagttacgctggagtctgaggctcgtcctg
aatgatatgcgaccgccggagggttgcgtttgagacgggcgacagatccagtcgcgctg
ctctcgtcgatccgctagcaaaaaacgggcagcccgggaatctaacccggctgaacgga
tttagagtccgttcgatctacatgatcagccgcccggtgtttcgtcctttccacaagat
atataaagccaagaaatcgaaatactttcaagttacggtaagcatatgatagtccattt
taaaacataattttaaaactgcaaactacccaagaaattattactttctacgtcacgta
ttttgtactaatatctttgtgtttacagtcaaattaattctaattatctctctaacagc
cttgtatcgtatatgcaaatatgaaggaatcatgggaaataggccctcttcctgcccga
ggtaccctagcaaaaaacgggcagcccgggaatctaacccggctgaacggatttagag
tccgttcgatctacatgatcagccgcccggtgtttcgtcctttccacaagatatataaa
gccaagaaatcgaaatactttcaagttacggtaagcatatgatagtccattttaaaaca
taattttaaaactgcaaactacccaagaaattattactttctacgtcacgtattttgta
ctaatatctttgtgtttacagtcaaattaattctaattatctctctaacagccttgtat
cgtatatgcaaatatgaaggaatcatgggaaataggccctcttcctgcccgaggtaccc
ctagcaaaaaacgggcagcccgggaatctaacccggctgaacggatttagagtccgttc
gatctacatgatcagccgcccggtgtttcgtcctttccacaagatatataaagccaaga

aatcgaaatactttcaagttacggtaagcatatgatagtccattttaaaacataatttt
aaaactgcaaactacccaagaaattattactttctacgtcacgtattttgtactaatat
ctttgtgtttacagtcaaattaattctaattatctctctaacagccttgtatcgtatat
gcaaatatgaaggaatcatgggaaataggccctcttcctgcccgaggtacccctagcaa
aaaacgggcagcccgggaatctaacccggctgaacggatttagagtccgttcgatctac
atgatcagccgcccggtgtttcgtcctttccacaagatatataaagccaagaaatcgaa
atactttcaagttacggtaagcatatgatagtccattttaaaacataatttttaaaactg
caaactacccaagaaattattactttctacgtcacgtattttgtactaatatctttgtg
tttacagtcaaattaattctaattatctctctaacagccttgtatcgtatatgcaaata
tgaaggaatcatgggaaataggccctcttcctgcccgaggtacccctaggggtgcgagc
ggatcgagcagtgtcgatcactactggaccgcgagctgtgctgcgacccgtgatcttac
ggcattatacgtatgatcggtccacgatcagctagattatctagtcagcttgatgtcat
agctgtttcctgaggctcaatactgaccatttaaatcatacctgacctccatagcagaa
agtcaaaagcctccgaccggaggcttttgacttgatcggcacgtaagaggttccaactt
tcaccataatgaaataagatcactaccgggcgtatttttgagttatcgagattttcag
gagctaaggaagctaaaatgagccatattcaacgggaaacgtcttgcttgaagccgcga
ttaaattccaacatggatgctgatttatatgggtataaatgggctcgcgataatgtcgg
gcaatcaggtgcgacaatctatcgattgtatgggaagcccgatgcgccagagttgtttc
tgaaacatggcaaaggtagcgttgccaatgatgttacagatgagatggtcaggctaaac
tggctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatga
tgcatggttactcaccactgcgatcccagggaaaacagcattccaggtattagaagaat
atcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccggttgcat
tcgattcctgtttgtaattgtccttttaacggcgatcgcgtatttcgtctcgctcaggc
gcaatcacgaatgaataacggtttggttggtgcgagtgattttgatgacgagcgtaatg
gctggcctgttgaacaagtctggaaagaaatgcataaactcttgccattctcaccggat
tcagtcgtcactcatggtgatttctcacttgataaccttatttttgacgaggggaaatt
aataggttgtattgatgttggacgagtcggaatcgcagaccgataccaggatcttgcca
tcctatggaactgcctcggtgagttttctccttcattacagaaacggctttttcaaaaa
tatggtattgataatcctgatatgaataaattgcagtttcacttgatgctcgatgagtt
tttctaatgaggacctaaatgtaatcacctggctcaccttcgggtgggcctttctgcgt
tgctggcgttttccataggctccgccccctgacgagcatcacaaaaatcgatgctca
agtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctggaag
ctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttc
tcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtg
taggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctg
cgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccac
tggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagag
ttcttgaagtggtggcctaactacggctacactagaagaacagtatttggtatctgcgc
tctgctgaagccagttacctcggaaaaagagttggtagctcttgatccggcaaacaaac
caccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaag
gatctcaagaagatcctttgattttctaccgaagaaaggccca

The sequence for the wild type PytR is:
ggaaacctgatcatgtagatcgaacggactctaaatccgttcagccgggttagattccc
ggggtttccg

pAcBac-CMV-MbPylRS-WPRE-CAG-GFPtag39AvrIIdel-WPRE



Sequence color-coding key

Feature	Color
M13 origin	text
AmpR	text
ColE1 origin	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
WPRE	text
MbPylRS	text
T7 promoter	text
CMV promoter	text
Polyhedrin Promoter	text
L21 sequence	text
VSV-G	text
SV40 late polyA	text
CAG promoter	text
EGFP-39tag	text
WPRE	text

rb-beta globin terminator	text
Tn7L	text

Sequence

ctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggt
tacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttct
tcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctc
cctttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgattaggg
tgatggttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttgg
agtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatc
tcggtctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaa
tgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgtttacaattt
caggtggcactttcggggaaatgtgcgcggaacccctatttgtttatttttctaaata
cattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattg
aaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcgg
cattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaa
gatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatcct
tgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctat
gtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacac
tattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatgg
catgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggcca
acttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatg
ggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaa
cgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaa
ctggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggat
aaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataa
atctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggta
agccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacga
aatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagacca
agtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatct
aggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttc
cactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttttct
gcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgc
cggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagata
ccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagc
accgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgata
agtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcg
ggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaact
gagatacctacagcgtgagcattgagaaagcgccacgcttcccgaagggagaaaggcgg
acaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccaggg
ggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcg
atttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcct
ttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcag
ccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcctgatgcggt
atttctccttacgcatctgtgcggtatttcacaccgcagaccagccgcgtaacctggc
aaaatcggttacggttgagtaataaatggatgccctgcgtaagcgggtgtgggcggaca

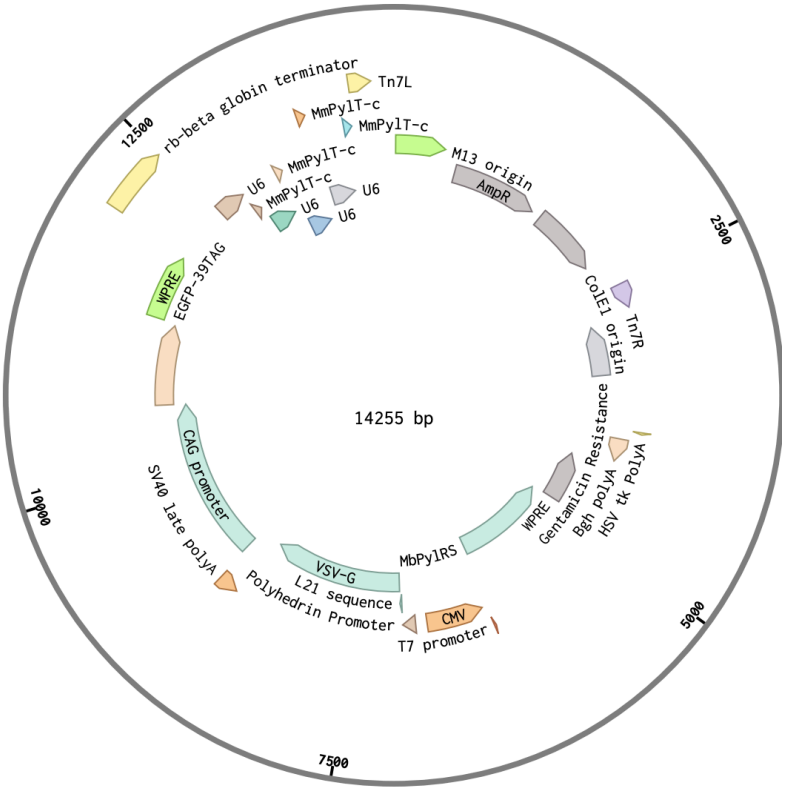
ataaagtcttaaactgaacaaaatagatctaaactatgacaataaagtcttaaactaga
cagaatagttgtaaactgaaatcagtccagttatgctgtgaaaaagcatactggacttt
tgttatggctaaagcaaactcttcattttctgaagtgcaaattgcccgtcgtattaaag
aggggcgtggccaagggcatggtaaagactatattcgcggcgttgtgacaatttaccga
acaactccgcggccgggaagccgatctcggcttgaacgaattgttaggtggcggtactt
gggtcgatatcaaagtgcatcacttcttcccgtatgcccaactttgtatagagagccac
tgcgggatcgtcaccgtaatctgcttgcacgtagatcacataagcaccaagcgcgttgg
cctcatgcttgaggagattgatgagcgcggtggcaatgccctgcctccggtgctcgccg
gagactgcgagatcatagatatagatctcactacgcggctgctcaaacctgggcagaac
gtaagccgcgagagcgccaacaaccgcttcttggtcgaaggcagcaagcgcgatgaatg
tcttactacggagcaagttcccgaggtaatcggagtccggctgatgttgggagtaggtg
gctacgtctccgaactcacgaccgaaaagatcaagagcagcccgcatggatttgacttg
gtcagggccgagcctacatgtgcgaatgatgcccatacttgagccacctaactttgttt
tagggcgactgccctgctgcgtaacatcgttgctgctgcgtaacatcgttgctgctcca
taacatcaaacatcgacccacggcgtaacgcgcttgctgcttggatgcccgaggcatag
actgtacaaaaaaacagtcataacaagccatgaaaaccgccactgcgccgttaccaccg
ctgcgttcggtcaaggttctggaccagttgcgtgagcgcatacgctacttgcattacag
tttacgaaccgaacaggcttatgtcaactgggttcgtgccttcatccgtttccacggtg
tgcgtcacccggcaaccttgggcagcagcgaagtcgaggcatttctgtcctggctggcg
aacgagcgcaaggtttcggtctccacgcatcgtcaggcattggcggccttgctgttctt
ctacggcaaggtgctgtgcacggatctgccctggcttcaggagatcggtagacctcggc
cgtcgcggcgcttgccggtggtgctgaccccggatgaagtggttcgcatcctcggtttt
ctggaaggcgagcatcgtttgttcgcccaggactctagctatagttctagtggttggct
acgtacccgtagtggctatggcagggcttgcgcttaatgcgccgctacagggcgcgtgg
ggataccccctagagccccagctggttctttccgcctcagaagccatagagcccaccgc
atccccagcatgcctgctattgtcttcccaatcctccccttgctgtcctgccccaccc
cacccccagaatagaatgacacctactcagacaatgcgatgcaatttcctcattttat
taggaaaggacagtgggagtggcaccttccagggtcaaggaaggcacgggggaggggca
aacaacagatggctggcaactagaaggcacagtcgaggctgatcagcgggtttaaacgg
gccctctagactcgagttaaagtcgacgcggggaggcggcccaaagggagatccgactc
gtctgagggcgaaggcgaagacgcggaagaggccgcagagccggcagcaggccgcggga
aggaaggtccgctggattgagggccgaagggacgtagcagaaggacgtcccgcgcagaa
tccaggtggcaacacaggcgagcagccaaggaaaggacgatgatttccccgacaacacc
acggaattgtcagtgcccaacagccgagcccctgtccagcagcgggcaaggcaggcggc
gatgagttccgccgtggcaatagggaggggaaagcgaaagtcccggaaaggagctgac
aggtggtggcaatgccccaaccagtgggggttgcgtcagcaaacacagtgcacaccacg
ccacgttgcctgacaacgggccacaactcctcataaagagacagcaaccaggatttata
caaggaggagaaaatgaaagccatacgggaagcaatagcatgatacaaaggcattaaag
cagcgtatccacatagcgtaaaaggagcaacatagttaagaataccagtcaatctttca
caaattttgtaatccagaggttgattgtcgacttaacgcgttgaattcttacagattgg
ttgaaatcccattatagtaagattcggaccttgatgccctcttaatgttttaaagccg
tgcataaccttgagcaagcgttcaagaccaaaacctgcacctatccatggtttgtcaat
accccattctctatcaagagaaactggcccgacgactgccgaagaaagctccaggtccc
cgtgcattatatcaagagtatccccaaagaccatacaggaatctcctacgatttcgaag
tcgatttccagatagtccagaaactctttgatgagagcttcaagattttcccgagtaca
tcccgaacccatctgacagaagttcaccatagtaaattcttccaggtgctctttgccgt
cagactctttccggtaacaaggtccgacttcgaaaatttttattgggcctggtaaaatc
ctatcgagtttcgcagatagttgtaaagagtcggggcaagcattggcctcaagcagag

attttatccacccggaagatctgttttgaaagttcagtatcattattaatacccattc
tctccacgtattccgccggaataaggataggagactttatctccagaaaaccccggtct
acgaaaaatttcgtaatatcacgttcgagtttaccgaggtagtcttctctatcattggt
atagagccgctgaaaatcgtttttttcttcttgtcacaagttcaggctcaagttccctga
aaggctttgccatatttagagaaattttatcctctggacttaagagagcctcaacccta
tcaagctggcttcttgtaagtgaaggagcaggagccgatgcgggaacagacgaatttgg
agttgattttgcaggcgaaggtacagatctggatgtgttcgttgatgcctttgcagaaa
cagaatttccagaggcttcggagcccttgaaactgatttcggcatagcttttttgacc
tttggagcagaaactaccctaactttcacactgtttttgctttcggttgatcttgtgag
aaaattattgatatcctcgtccgaaaccctacatcgtttgcaggttttctgtacttat
gatgtctgaatgctctggctgttctacaactcctggaattattcacaacaagatggtct
ccacacgccatttcaatgtatattttacttcttgagacctcatggtgcttgattttgtg
gagcgtgccagtcctggacatccagagcccggtcgcagatattaaaacatctaatggtt
ttttatccatggtggcgctagccagcttgggtctccctatagtgagtcgtattaatttc
gataagccagtaagcagtgggttctctagttagccagagagctctgcttatatagacct
cccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacatttt
ggaaagtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggag
acttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaacc
gcatcaccatggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcc
cataaggtcatgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaat
aggggggcgtacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgta
aatagtccacccattgacgtcaatggaaagtccctattggcgttactatgggaacatac
gtcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttac
cgtaagttatgtaacgcggaactccatatatgggctatgaactaatgaccccgtaattg
attactattaataactagtcaataatcaatgtcaacgcgtatatctggcccgtacatcg
cgaagcagcgcaaaacggatcctgcaggtatttgcggccgcggtccgtatactccggaa
tattaatagatcatggagataattaaaatgataaccatctcgcaaataaataagtattt
tactgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcata
ccgtcccaccatcgggcgcgaactcctaaaaaaccgccaccatgaagtgccttttgtac
ttagccttttttattcattggggtgaattgcaagttcaccatagttttccacacaacca
aaaaggaaactggaaaaatgttccttctaattaccattattgcccgtcaagctcagatt
taaattggcataatgacttaataggcacagccttacaagtcaaaatgcccaagagtcac
aaggctattcaagcagacggttggatgtgtcatgcttccaaatgggtcactacttgtga
tttccgctggtatggaccgaagtatataacacattccatccgatccttcactccatctg
tagaacaatgcaaggaaagcattgaacaaacgaaacaaggaacttggctgaatccaggc
ttccctcctcaaagttgtggatatgcaactgtgacggatgccgaagcagtgattgtcca
ggtgactcctcaccatgtgctggttgatgaatacacaggagaatgggttgattcacagt
tcatcaacggaaaatgcagcaattacatatgccccactgtccataactctacaacctgg
cattctgactataaggtcaaagggctatgtgattctaacctcatttccatggacatcac
cttcttctcagaggacggagagctatcatccctgggaaaggagggcacagggttcagaa
gtaactactttgcttatgaaactggaggcaaggcctgcaaaatgcaatactgcaagcat
tggggagtcagactcccatcaggtgtctggttcgagatggctgataaggatctctttgc
tgcagccagattccctgaatgcccagaagggtcaagtatctctgctccatctcagacct
cagtggatgtaagtctaattcaggacgttgagaggatcttggattattccctctgccaa
gaaacctggagcaaaatcagagcgggtcttccaatctctccagtggatctcagctatct
tgctcctaaaaacccaggaaccggtcctgctttcaccataatcaatggtaccctaaaat
actttgagaccagatacatcagagtcgatattgctgctccaatcctctcaagaatggtc
ggaatgatcagtggaactaccacagaaagggaactgtgggatgactgggcaccatatga

agacgtggaaattggacccaatggagttctgaggaccagttcaggatataagtttcctt
tatacatgattggacatggtatgttggactccgatcttcatcttagctcaaaggctcag
gtgttcgaacatcctcacattcaagacgctgcttcgcaacttcctgatgatgagagttt
atttttggtgatactgggctatccaaaaatccaatcgagcttgtagaaggttggttca
gtagttggaaaagctctattgcctctttttctttatcatagggttaatcattggacta
ttcttggttctccgagttggtatccatctttgcattaaattaaagcacaccaagaaaag
acagatttatacagacatagagatgaaccgacttggaaagtgataagtcgagaagtact
agaggatcataatcagccataccacatttgtagaggttttacttgctttaaaaaacctc
ccacacctccccctgaacctgaaacataaaatgaatgcaattgttgttgttaacttgtt
tattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaag
catttttttcactgcattctagttgtggtttgtccaaactcatcaatgtatcttatcat
gtctggatctgatcactgcttgagcctagttattaatagtaatcaattacggggtcatt
agttcatagcccatatatggagttccgcgttacataacttacggtaaatggcccgcctg
gctgaccgcccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagta
acgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaactgccca
cttggcagtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacg
gtaaatggcccgcctggcattatgcccagtacatgaccttatgggactttcctacttgg
cagtacatctacgtattagtcatcgctattaccatggtcgaggtgagccccacgttctg
cttcactctcccccatctccccccccctccccacccccaattttgtatttatttatttttt
aattattttgtgcagcgatgggggcgggggggggggggcgcgcgccaggcggggcggg
gcggggcgagggggcggggcggggcgaggcggagaggtgcggcggcagccaatcagagcg
gcgcgctccgaaagtttccttttatggcgaggcggcggcggcggcgcccctataaaaag
cgaagcgcgcggcgggcgggagtcgctgcgttgccttcgccccgtgccccgctccgcgc
cgcctcgcgccgcccgcccggctctgactgaccgcgttactcccacaggtgagcgggc
gggacggcccttctcctccgggctgtaattagcgcttggtttaatgacggctcgtttct
tttctgtggctgcgtgaaagccttaaagggctccgggagggccctttgtgcgggggga
gcggctcgggggggtgcgtgcgtgtgtgtgtgcgtggggagcgccgcgtgcggcccgcgc
tgcccggcggctgtgagcgctgcgggcgcggcgcggggctttgtgcgctccgcgtgtgc
gcgagggagcgcggccgggggcggtgccccgcggtgcgggggggctgcgagggaaca
aaggctgcgtgcggggtgtgtgcgtgggggggtgagcagggggtgtgggcgcggcggtc
gggctgtaacccccccctgcacccccctccccgagttgctgagcacggcccggcttcgg
gtgcggggctccgtgcggggcgtggcgcggggctcgccgtgccgggcgggggggtggcgg
caggtgggggtgccgggcgggggcgggggccgcctcgggccgggggagggctcgggggaggg
gcgcggcggcccccggagcgccggcggctgtcgaggcgcggcgagccgcagccattgcct
tttatggtaatcgtgcgagagggcgcagggacttcctttgtcccaaatctggcggagcc
gaaatctgggaggcgccgccgcaccccctctagcgggcgcgggcgaagcggtgcggcgc
cggcaggaaggaaatgggcggggagggccttcgtgcgtcgccgcgccgccgtcccctte
tccatctccagcctcggggctgccgcaggggacggctgccttcgggggggacggggca
gggcggggttcggcttctggcgtgtgaccggcggctctagagcctctgctaaccatgtt
catgccttcttctttttcctacagctcctgggcaacgtgctggttattgtgctgtctca
tcattttggcaaagaattggccaaggaggccaccatggtgagcaagggcgaggagctgt
tcaccggggtggtgcccatcctggtcgagctggacggcgacgtaaacggccacaagttc
agcgtgtccggcgagggcgagggcgatgccacttagggcaagctgaccctgaagttcat
ctgcaccaccggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacg
gcgtgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaagtcc
gccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacgacggcaacta
caagacccgcgccgaggtgaagttcgagggcgacaccctggtgaaccgcatcgagctga
agggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactac

aacagccacaacgtctatatcatggccgacaagcagaagaacggcatcaaggtgaactt
caagatccgccacaacatcgaggacggcagcgtgcagctcgccgaccactaccagcaga
acacccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagcacccag
tccgccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgt
gaccgccgccgggatcactctcggcatggacgagctgtacaaggggcccttcgaacaaa
aactcatctcagaagaggatctgaatatgcataccggtcatcatcaccatcaccattga
gaattcaacgcgttaagtcgacaatcaacctctggattacaaaatttgtgaaagattga
ctggtattcttaactatgttgctccttttacgctatgtggatacgctgctttaatgcct
ttgtatcatgctattgcttcccgtatggctttcattttctcctccttgtataaatcctg
gttgctgtctctttatgaggagttgtggcccgttgtcaggcaacgtggcgtggtgtgca
ctgtgtttgctgacgcaacccccactggttggggcattgccaccacctgtcagctcctt
tccgggactttcgctttcccctccctattgccacggcggaactcatcgccgcctgcct
tgcccgctgctggacaggggctcggctgttgggcactgacaattccgtggtgttgtcgg
ggaaatcatcgtcctttccttggctgctcgcctgtgttgccacctggattctgcgcggg
acgtccttctgctacgtcccttcggccctcaatccagcggaccttccttcccgcggcct
gctgccggctctgcggcctcttccgcgtcttcgccttcgccctcagacgagtcggatct
ccctttgggccgcctccccgcggcctggccggccgccagcacagtggtcgatcgaccaa
tgccctggctcacaaataccactgagatcttttcccctctgccaaaaattatggggaca
tcatgaagcccccttgagcatctgacttctggctaataaaggaaatttattttcattgca
atagtgtgttggaatttttttgtgtctctcactcggaaggacatatgggagggcaaatca
tttaaaacatcagaatgagtatttggtttagagtttggcaacatatgcccatatgctgg
ctgccatgaacaaaggttggctataaagaggtcatcagtatatgaaacagcccctgct
gtccattccttattccatagaaaagccttgacttgaggttagatttttttatattttg
ttttgtgttatttttttctttaacatccctaaaatttccttacatgttttactagcca
gatttttcctcctctcctgactactcccagtcatagctgtccctcttctcttatggaga
tccctcgacctgccctaggagatccgaaccagataagtgaaatctagttccaaactatt
ttgtcatttttaattttcgtattagcttacgacgctacacccagttcccatctattttg
tcactcttccctaaataatccttaaaaactccatttccaccctcccagttcccaacta
ttttgtccgcccacagcggggcatttttcttcctgttatgttttaatcaaacatcctg
ccaactccatgtgacaaaccgtcatcttcggctacttttttctctgtcacagaatgaaaa
ttttttctgtcatctcttcgttattaatgtttgtaattgactgaatatcaacgcttattt
gcagc

pAcBac-CMV-MbPylRS-CAG-GFPtagAvrIIdel-4xPytR



Sequence color-coding key

Feature	Color
M13 origin	text
AmpR	text
ColE1 origin	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
WPRE	text
MbPylRS	text
T7 promoter	text
UbiC promoter	text
Polyhedrin Promoter	text
L21 sequence	text
VSV-G	text
SV40 late polyA	text
CAG promoter	text
GFP39tag	text
WPRE	text

rb-beta globin terminator	text
U6	text
MmPytR	text
Tn7L	text

Sequence for HTS25 construct

ctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggt
tacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttct
tcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctc
cctttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgattaggg
tgatggttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttgg
agtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatc
tcggtctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaa
tgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgtttacaattt
caggtggcactttcggggaaatgtgcgcggaacccctatttgtttatttttctaaata
cattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattg
aaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcgg
cattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaa
gatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatcct
tgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctat
gtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacac
tattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatgg
catgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggcca
acttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatg
ggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaa
cgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaa
ctggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggat
aaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataa
atctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggta
agccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacga
aatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagacca
agtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatct
aggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttc
cactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttct
gcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgc
cggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcagata
ccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagc
accgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgata
agtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcg
ggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaact
gagatacctacagcgtgagcattgagaaagcgccacgcttcccgaagggagaaaggcgg
acaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccaggg
ggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcg
atttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcct
ttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcag
ccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcctgatgcggt

attttctccttacgcatctgtgcggtatttcacaccgcagaccagccgcgtaacctggc
aaaatcggttacggttgagtaataaatggatgccctgcgtaagcgggtgtgggcggaca
ataaagtcttaaactgaacaaaatagatctaaactatgacaataaagtcttaaactaga
cagaatagttgtaaactgaaatcagtccagttatgctgtgaaaaagcatactggacttt
tgttatggctaaagcaaactcttcattttctgaagtgcaaattgcccgtcgtattaaag
aggggcgtggccaagggcatggtaaagactatattcgcggcgttgtgacaatttaccga
acaactccgcggccgggaagccgatctcggcttgaacgaattgttaggtggcggtactt
gggtcgatatcaaagtgcatcacttcttcccgtatgcccaactttgtatagagagccac
tgcgggatcgtcaccgtaatctgcttgcacgtagatcacataagcaccaagcgcgttgg
cctcatgcttgaggagattgatgagcgcggtggcaatgccctgcctccggtgctcgccg
gagactgcgagatcatagatatagatctcactacgcggctgctcaaacctgggcagaac
gtaagccgcgagagcgccaacaaccgcttcttggtcgaaggcagcaagcgcgatgaatg
tcttactacggagcaagttcccgaggtaatcggagtccggctgatgttgggagtaggtg
gctacgtctccgaactcacgaccgaaaagatcaagagcagcccgcatggatttgacttg
gtcagggccgagcctacatgtgcgaatgatgcccatacttgagccacctaactttgttt
tagggcgactgccctgctgcgtaacatcgttgctgctgcgtaacatcgttgctgctccca
taacatcaaacatcgacccacggcgtaacgcgcttgctgcttggatgcccgaggcatag
actgtacaaaaaaacagtcataacaagccatgaaaaccgccactgcgccgttaccaccg
ctgcgttcggtcaaggttctggaccagttgcgtgagcgcatacgctacttgcattacag
tttacgaaccgaacaggcttatgtcaactgggttcgtgccttcatccgtttccacggtg
tgcgtcacccggcaaccttgggcagcagcgaagtcgaggcatttctgtcctggctggcg
aacgagcgcaaggtttcggtctccacgcatcgtcaggcattggcggccttgctgttctt
ctacggcaaggtgctgtgcacggatctgccctggcttcaggagatcggtagacctcggc
cgtcgcggcgcttgccggtggtgctgaccccggatgaagtggttcgcatcctcggtttt
ctggaaggcgagcatcgtttgttcgcccaggactctagctatagttctagtggttggct
acgtacccgtagtggctatggcagggcttgcgcttaatgcgccgctacagggcgcgtgg
ggataccccctagagccccagctggttctttccgcctcagaagccatagagcccaccgc
atccccagcatgcctgctattgtcttcccaatcctccccttgctgtcctgccccaccc
cacccccagaatagaatgacacctactcagacaatgcgatgcaatttcctcattttat
taggaaaggacagtgggagtggcaccttccagggtcaaggaaggcacgggggaggggca
aacaacagatggctggcaactagaaggcacagtcgaggctgatcagcgggtttaaacgg
gccctctagactcgagttaaagtcgacgcggggaggcggcccaaagggagatccgactc
gtctgagggcgaaggcgaagacgcggaagaggccgcagagccggcagcaggccgcggga
aggaaggtccgctggattgagggccgaagggacgtagcagaaggacgtcccgcgcagaa
tccaggtggcaacacaggcgagcagccaaggaaaggacgatgatttccccgacaacacc
acggaattgtcagtgcccaacagccgagcccctgtccagcagcgggcaaggcaggcggc
gatgagttccgccgtggcaatagggaggggggaaagcgaaagtcccggaaaggagctgac
aggtggtggcaatgccccaaccagtgggggttgcgtcagcaaacacagtgcacaccacg
ccacgttgcctgacaacgggccacaactcctcataaagagacagcaaccaggatttata
caaggaggagaaaatgaaagccatacgggaagcaatagcatgatacaaaggcattaaag
cagcgtatccacatagcgtaaaaggagcaacatagttaagaataccagtcaatctttca
caaattttgtaatccagaggttgattgtcgacttaacgcgttgaattcttacagattgg
ttgaaatcccattatagtaagattcggaccttgatgccctcttaatgttttaaagccg
tgcataaccttgagcaagcgttcaagaccaaaacctgcacctatccatggtttgtcaat
accccattctctatcaagagaaactggcccgacgactgccgaagaaagctccaggtccc
cgtgcattatatcaagagtatccccaaagaccatacaggaatctcctacgatttcgaag
tcgatttccagatagtccagaaactctttgatgagagcttcaagattttcccgagtaca
tcccgaacccatctgacagaagttcaccatagtaaattcttccaggtgctctttgccgt

cagactctttccggtaacaaggtccgacttcgaaaatttttattgggcctggtaaaatc
ctatcgagttttcgcagatagttgtaaagagtcggggcaagcattggcctcaagcagag
atttttatccacccggaagatctgttttgaaagttcagtatcattattaatacccattc
tctccacgtattccgccggaataaggataggagactttatctccagaaaaccccggtct
acgaaaaatttcgtaatatcacgttcgagtttaccgaggtagtcttctctatcattggt
atagagccgctgaaaatcgtttttttcttcttgtcacaagttcaggctcaagttccctga
aaggctttgccatatttagagaaattttatcctctggacttaagagagcctcaacccta
tcaagctggcttcttgtaagtgaaggagcaggagccgatgcgggaacagacgaatttgg
agttgattttgcaggcgaaggtacagatctggatgtgttcgttgatgcctttgcagaaa
cagaattttccagaggcttcggagcccttgaaactgatttcggcatagcttttttgacc
tttggagcagaaactaccctaactttcacactgtttttgctttcggttgatcttgtgag
aaaattattgatatcctcgtccgaaaccctacatcgtttgcaggttttctgtacttat
gatgtctgaatgctctggctgttctacaactcctggaattattcacaacaagatggtct
ccacacgccatttcaatgtatattttacttcttgagacctcatggtgcttgattttgtg
gagcgtgccagtcctggacatccagagcccggtcgcagatattaaaacatctaatggtt
ttttatccatggtggcgctagccagcttgggtctccctatagtgagtcgtattaatttc
gataagccagtaagcagtgggttctctagttagccagagagctctgcttatatagacct
cccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacatttt
ggaaagtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggag
acttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaacc
gcatcaccatggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcc
cataaggtcatgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaat
aggggggcgtacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgta
aatagtccacccattgacgtcaatggaaagtccctattggcgttactatgggaacatac
gtcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttac
cgtaagttatgtaacgcggaactccatatatgggctatgaactaatgaccccgtaattg
attactattaataactagtcaataatcaatgtcaacgcgtatatctggcccgtacatcg
cgaagcagcgcaaaacggatcctgcaggtatttgcggccgcggtccgtatactccggaa
tattaatagatcatggagataattaaaatgataaccatctcgcaaataaataagtattt
tactgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcata
ccgtcccaccatcgggcgcgaactcctaaaaaaccgccaccatgaagtgcctttgtac
ttagccttttttattcattggggtgaattgcaagttcaccatagttttccacacaacca
aaaaggaaactggaaaaatgttccttctaattaccattattgcccgtcaagctcagatt
taaattggcataatgacttaataggcacagccttacaagtcaaaatgcccaagagtcac
aaggctattcaagcagacggttggatgtgtcatgcttccaaatgggtcactacttgtga
tttccgctggtatggaccgaagtatataacacattccatccgatccttcactccatctg
tagaacaatgcaaggaaagcattgaacaaacgaaacaaggaacttggctgaatccaggc
ttccctcctcaaagttgtggatatgcaactgtgacggatgccgaagcagtgattgtcca
ggtgactcctcaccatgtgctggttgatgaatacacaggagaatgggttgattcacagt
tcatcaacggaaaatgcagcaattacatatgccccactgtccataactctacaacctgg
cattctgactataaggtcaaagggctatgtgattctaacctcatttccatggacatcac
cttcttctcagaggacggagagctatcatccctgggaaaggagggcacagggttcagaa
gtaactactttgcttatgaaactggaggcaaggcctgcaaaatgcaatactgcaagcat
tggggagtcagactcccatcaggtgtctggttcgagatggctgataaggatctctttgc
tgcagccagattccctgaatgcccagaagggtcaagtatctctgctccatctcagacct
cagtggatgtaagtctaattcaggacgttgagaggatcttggattattccctctgccaa
gaaacctggagcaaaatcagagcgggtcttccaatctctccagtggatctcagctatct
tgctcctaaaaacccaggaaccggtcctgctttcaccataatcaatggtaccctaaaat

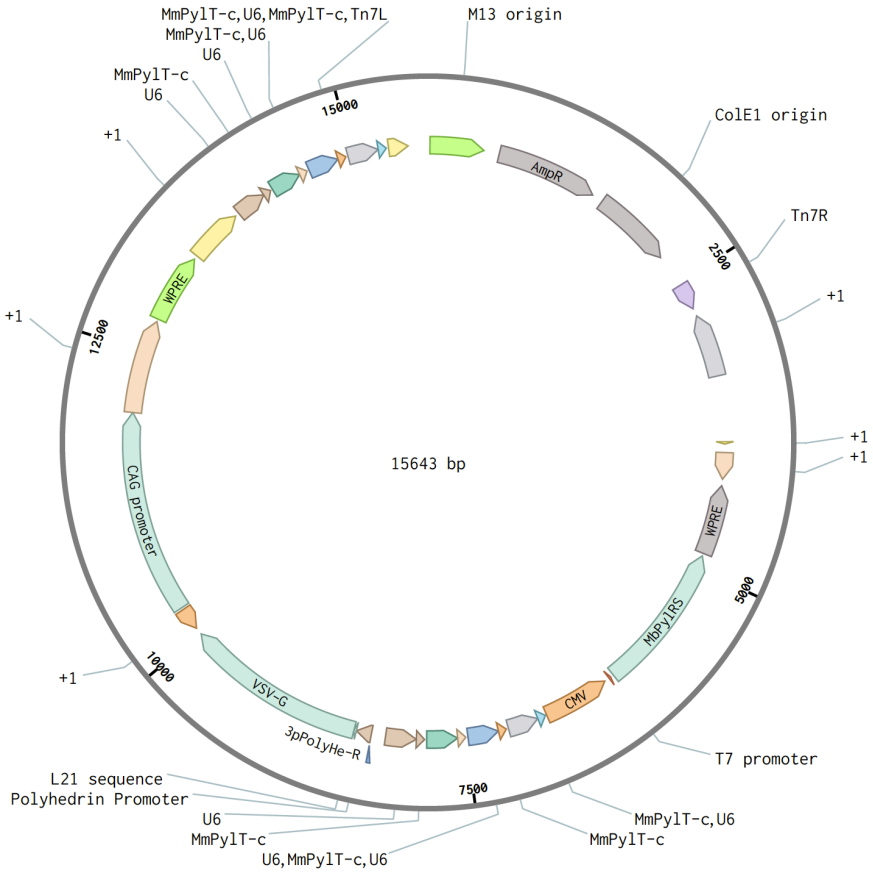
actttgagaccagatacatcagagtcgatattgctgctccaatcctctcaagaatggtc
ggaatgatcagtggaactaccacagaaagggaactgtgggatgactgggcaccatatga
agacgtggaaattggacccaatggagttctgaggaccagttcaggatataagtttcctt
tatacatgattggacatggtatgttggactccgatcttcatcttagctcaaaggctcag
gtgttcgaacatcctcacattcaagacgctgcttcgcaacttcctgatgatgagagttt
atttttggtgatactgggctatccaaaaatccaatcgagcttgtagaaggttggttca
gtagttggaaaagctctattgcctctttttttctttatcatagggttaatcattggacta
ttcttggttctccgagttggtatccatctttgcattaaattaaagcacaccaagaaaag
acagatttatacagacatagagatgaaccgacttggaaagtgataagtcgagaagtact
agaggatcataatcagccataccacatttgtagaggttttacttgctttaaaaaacctc
ccacacctccccctgaacctgaaacataaaatgaatgcaattgttgttgttaacttgtt
tattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaag
catttttttcactgcattctagttgtggtttgtccaaactcatcaatgtatcttatcat
gtctggatctgatcactgcttgagcctagttattaatagtaatcaattacggggtcatt
agttcatagcccatatatggagttccgcgttacataacttacggtaaatggcccgcctg
gctgaccgcccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagta
acgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaactgccca
cttggcagtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacg
gtaaatggcccgcctggcattatgcccagtacatgaccttatgggactttcctacttgg
cagtacatctacgtattagtcatcgctattaccatggtcgaggtgagccccacgttctg
cttcactctcccccatctccccccccctccccaccccccaattttgtatttatttatttttt
aattattttgtgcagcgatgggggcgggggggggggggcgcgcgccaggcggggcggg
gcggggcgagggggcggggcggggcgaggcggagaggtgcggcggcagccaatcagagcg
gcgcgctccgaaagtttccttttatggcgaggcggcggcggcggcgccctataaaaaag
cgaagcgcgcggcgggcgggagtcgctgcgttgccttcgccccgtgccccgctccgcgc
cgcctcgcgccgcccgcccccggctctgactgaccgcgttactcccacaggtgagcgggc
gggacggcccttctcctccgggctgtaattagcgcttggtttaatgacggctcgtttct
tttctgtggctgcgtgaaagccttaaagggctccgggagggcccttgtgcgggggggga
gcggctcgggggggtgcgtgcgtgtgtgtgtgcgtggggagcgccgcgtgcggcccgcgc
tgcccggcggctgtgagcgctgcgggcgcggcgcggggctttgtgcgctccgcgtgtgc
gcgaggggagcgcggccgggggcggtgccccgcggtgcgggggggctgcgaggggaaca
aaggctgcgtgcggggtgtgtgcgtgggggggtgagcagggggtgtgggcgcggcggtc
gggctgtaaccccccctgcacccccctccccgagttgctgagcacggcccggcttcgg
gtgcggggctccgtgcggggcgtggcgcggggctcgccgtgccgggcgggggggtggcgg
caggtgggggtgccgggcgggggcgggccgcctcgggccgggagggctcgggggaggg
gcgcggcggcccggagcgccggcggctgtcgaggcgcggcgagccgcagccattgcct
tttatggtaatcgtgcgagagggcgcagggacttcctttgtcccaaatctggcggagcc
gaaatctgggaggcgccgccgcaccccctctagcgggcgcgggcgaagcggtgcggcgc
cggcaggaaggaaatgggcggggagggccttcgtgcgtcgccgcgccgccgtccccttc
tccatctccagcctcggggctgccgcaggggacggctgccttcgggggggacggggca
gggcggggttcggcttctggcgtgtgaccggcggctctagagcctctgctaaccatgtt
catgccttcttctttttcctacagctcctgggcaacgtgctggttattgtgctgtctca
tcattttggcaaagaattggccaaggaggccaccatggtgagcaagggcgaggagctgt
tcaccggggtggtgcccatcctggtcgagctggacggcgacgtaaacggccacaagttc
agcgtgtccggcgagggcgagggcgatgccacttagggcaagctgaccctgaagttcat
ctgcaccaccggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacg
gcgtgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaagtcc
gccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacgacggcaacta

caagacccgcgccgaggtgaagttcgagggcgacaccctggtgaaccgcatcgagctga
agggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactac
aacagccacaacgtctatatcatggccgacaagcagaagaacggcatcaaggtgaactt
caagatccgccacaacatcgaggacggcagcgtgcagctcgccgaccactaccagcaga
acacccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagcacccag
tccgccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgt
gaccgccgccgggatcactctcggcatggacgagctgtacaaggggcccttcgaacaaa
aactcatctcagaagaggatctgaatatgcataccggtcatcatcaccatcaccattga
gaattcaacgcgttaagtcgacaatcaacctctggattacaaaatttgtgaaagattga
ctggtattcttaactatgttgctcctttacgctatgtggatacgctgctttaatgcct
ttgtatcatgctattgcttcccgtatggctttcattttctcctccttgtataaatcctg
gttgctgtctctttatgaggagttgtggcccgttgtcaggcaacgtggcgtggtgtgca
ctgtgtttgctgacgcaacccccactggttggggcattgccaccacctgtcagctcctt
tccgggactttcgctttcccctccctattgccacggcggaactcatcgccgcctgcct
tgcccgctgctggacaggggctcggctgttgggcactgacaattccgtggtgttgtcgg
ggaaatcatcgtcctttccttggctgctcgcctgtgttgccacctggattctgcgcggg
acgtccttctgctacgtcccttcggccctcaatccagcggaccttccttcccgcggcct
gctgccggctctgcggcctcttccgcgtcttcgccttcgccctcagacgagtcggatct
ccctttgggccgcctccccgcggcctggccggccgccagcacagtggtcgatcgaccaa
tgccctggctcacaaataccactgagatcttttcccctctgccaaaaattatggggaca
tcatgaagcccccttgagcatctgacttctggctaataaaggaaatttatttcattgca
atagtgtgttggaatttttgtgtctctcactcggaaggacatatgggagggcaaatca
tttaaaacatcagaatgagtatttggtttagagtttggcaacatatgcccatatgctgg
ctgccatgaacaaaggttggctataaagaggtcatcagtatatgaaacagccccctgct
gtccattccttattccatagaaaagccttgacttgaggttagatttttttatattttg
ttttgtgttatttttttctttaacatccctaaaatttccttacatgttttactagcca
gattttcctcctctcctgactactcccagtcatagctgtccctcttctcttatggaga
tccctcgacctgccctaggggtacctcgggcaggaagagggcctatttcccatgattcc
ttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttgactg
taaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtag
tttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaa
gtatttcgatttcttggctttatatatcttgtggaaaggacgaaacaccgggcggctga
tcatgtagatcgaacggactctaaatccgttcagccgggttagattcccgggctgcccg
ttttttgctaggggtacctcgggcaggaagagggcctatttcccatgattccttcatat
ttgcatatacgatacaaggctgttagagagataattagaattaatttgactgtaaacac
aaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcag
ttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttc
gatttcttggctttatatatcttgtggaaaggacgaaacaccgggcggctgatcatgta
gatcgaacggactctaaatccgttcagccgggttagattcccgggctgcccgttttttg
ctaggggtacctcgggcaggaagagggcctatttcccatgattccttcatatttgcata
tacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagata
ttagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaa
attatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttct
tggctttatatatcttgtggaaaggacgaaacaccgggcggctgatcatgtagatcgaa
cggactctaaatccgttcagccgggttagattcccgggctgcccgttttttgctagggg
tacctcgggcaggaagagggcctatttcccatgattccttcatatttgcatatacgata
caaggctgttagagagataattagaattaatttgactgtaaacacaaagatattagtac
aaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgt

tttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggcttt
atatatcttgtggaaaggacgaaacaccgggcggctgatcatgtagatcgaacggactc
taaatccgttcagccgggttagattcccgggctgcccgtttttgctaggagatccgaa
ccagataagtgaaatctagttccaaactattttgtcatttttaattttcgtattagctt
acgacgctacacccagttcccatctattttgtcactcttccctaaataatccttaaaaa
ctccatttccacccctcccagttcccaactattttgtccgcccacagcggggcatttt
cttcctgttatgttttttaatcaaacatcctgccaactccatgtgacaaaccgtcatctt
cggctacttttttctctgtcacagaatgaaaatttttctgtcatctcttcgttattaatg
tttgtaattgactgaatatcaacgcttatttgcagc

The sequence for the wild type PytR is:
ggaaacctgatcatgtagatcgaacggactctaaatccgttcagccgggttagattccc
ggggtttccg

pAcBac-CMV-MbPylRS-CAG-GFPtagAvrIIdel-8xPytR



Sequence color-coding key

Feature	Color
M13 origin	text
AmpR	text
ColE1 origin	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
WPRE	text
MbPylRS	text
T7 promoter	text
UbiC promoter	text
Polyhedrin Promoter	text
L21 sequence	text
VSV-G	text
SV40 late polyA	text
CAG promoter	text

GFP39tag	text
WPRE	text
rb-beta globin terminator	text
U6	text
MmPytR	text
Tn7L	text

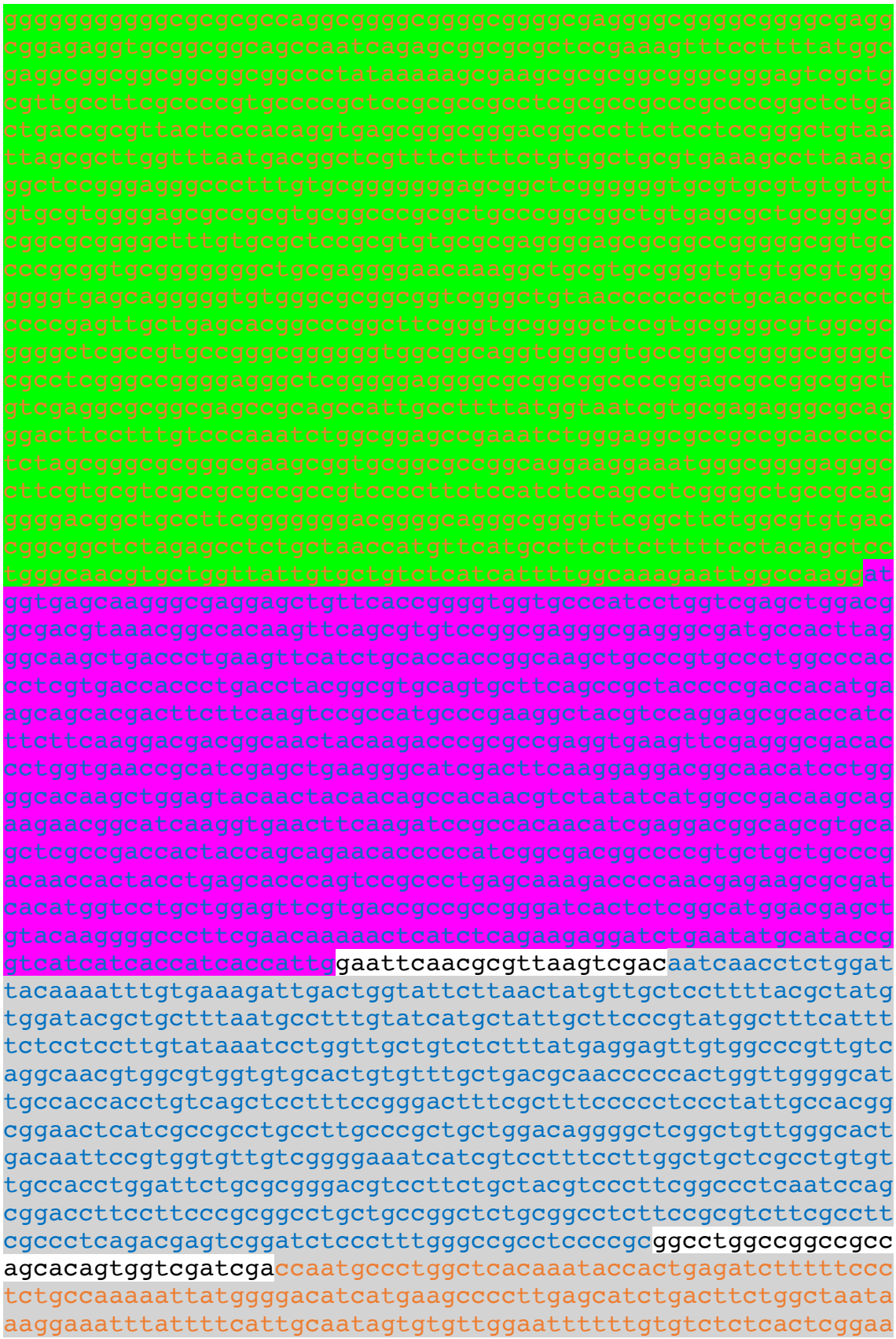
Sequence for HTS25 construct

ctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggt
tacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttct
tcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctc
cctttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgattaggg
tgatggttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttgg
agtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatc
tcggtctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaa
tgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgtttacaattt
caggtggcactttcggggaaatgtgcgcggaacccctatttgtttatttttctaaata
cattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattg
aaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcgg
catttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaa
gatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatcct
tgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctat
gtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacac
tattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatgg
catgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggcca
acttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatg
ggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaa
cgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaa
ctggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggat
aaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataa
atctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggta
agccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacga
aatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagacca
agtttactcatatatactttagattgatttaaaacttcattttaatttaaaaggatct
aggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttc
cactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttct
gcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgc
cggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcagata
ccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagc
accgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgata
agtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcg
ggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaact
gagatacctacagcgtgagcattgagaaagcgccacgcttcccgaagggagaaaggcgg
acaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccaggg
ggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcg
atttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcct

ttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcag
ccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcctgatgcggt
attttctccttacgcatctgtgcggtatttcacaccgcagaccagccgcgtaacctggc
aaaatcggttacggttgagtaataaatggatgccctgcgtaagcgggtgtgggcggaca
ataaagtcttaaactgaacaaaatagatctaaactatgacaataaagtcttaaactaga
cagaatagttgtaaactgaaatcagtccagttatgctgtgaaaaagcatactggacttt
tgttatggctaaagcaaactcttcattttctgaagtgcaaattgcccgtcgtattaaag
aggggcgtggccaagggcatggtaaagactatattcgcggcgttgtgacaatttaccga
acaactccgcggccgggaagccgatctcggcttgaacgaattgttaggtggcggtactt
gggtcgatatcaaagtgcatcacttcttcccgtatgcccaactttgtatagagagccac
tgcgggatcgtcaccgtaatctgcttgcacgtagatcacataagcaccaagcgcgttgg
cctcatgcttgaggagattgatgagcgcggtggcaatgccctgcctccggtgctcgccg
gagactgcgagatcatagatatagatctcactacgcggctgctcaaacctgggcagaac
gtaagccgcgagagcgccaacaaccgcttcttggtcgaaggcagcaagcgcgatgaatg
tcttactacggagcaagttcccgaggtaatcggagtccggctgatgttgggagtaggtg
gctacgtctccgaactcacgaccgaaaagatcaagagcagcccgcatggatttgacttg
gtcagggccgagcctacatgtgcgaatgatgcccatacttgagccacctaactttgttt
tagggcgactgccctgctgcgtaacatcgttgctgctgcgtaacatcgttgctgctcca
taacatcaaacatcgacccacggcgtaacgcgcttgctgcttggatgcccgaggcatag
actgtacaaaaaaacagtcataacaagccatgaaaaccgccactgcgccgttaccaccg
ctgcgttcggtcaaggttctggaccagttgcgtgagcgcatacgctacttgcattacag
tttacgaaccgaacaggcttatgtcaactgggttcgtgccttcatccgtttccacggtg
tgcgtcacccggcaaccttgggcagcagcgaagtcgaggcatttctgtcctggctggcg
aacgagcgcaaggtttcggtctccacgcatcgtcaggcattggcggccttgctgttctt
ctacggcaaggtgctgtgcacggatctgccctggcttcaggagatcggtagacctcggc
cgtcgcggcgcttgccggtggtgctgaccccggatgaagtggttcgcatcctcggtttt
ctggaaggcgagcatcgtttgttcgcccaggactctagctatagttctagtggttggct
acgtacccgtagtggctatggcagggcttgcgcttaatgcgccgctacagggcgcgtgg
ggataccccctagagccccagctggttctttccgcctcagaagccatagagcccaccgc
atccccagcatgcctgctattgtcttcccaatcctccccttgctgtcctgccccaccc
cacccccagaatagaatgacacctactcagacaatgcgatgcaatttcctcattttat
taggaaaggacagtgggagtggcaccttccagggtcaaggaaggcacgggggaggggca
aacaacagatggctggcaactagaaggcacagtcgaggctgatcagcgggtttaaacgg
gccctctagactcgagttaaagtcgacgcgggggaggcggcccaaagggagatccgactc
gtctgagggcgaaggcgaagacgcggaagaggccgcagagccggcagcaggccgcggga
aggaaggtccgctggattgagggccgaagggacgtagcagaaggacgtcccgcgcagaa
tccaggtggcaacacaggcgagcagccaaggaaaggacgatgatttccccgacaacacc
acggaattgtcagtgcccaacagccgagcccctgtccagcagcgggcaaggcaggcggc
gatgagttccgccgtggcaatagggaggggggaaagcgaaagtcccggaaaggagctgac
aggtggtggcaatgccccaaccagtgggggttgcgtcagcaaacacagtgcacaccacg
ccacgttgcctgacaacgggccacaactcctcataaagagacagcaaccaggatttata
caaggaggagaaaatgaaagccatacgggaagcaatagcatgatacaaaggcattaaag
cagcgtatccacatagcgtaaaaggagcaacatagttaagaataccagtcaatctttca
caaattttgtaatccagaggttgattgtcgacttaacgcgttgaattcttacagattgg
ttgaaatcccattatagtaagattcggaccttgatgccctcttaatgttttaaagccg
tgcataaccttgagcaagcgttcaagaccaaaacctgcacctatccatggtttgtcaat
accccattctctatcaagagaaactggcccgacgactgccgaagaaagctccaggtccc

cgtgcattatatcaagagtatccccaaagaccatacaggaatctcctacgatttcgaag
tcgatttccagatagtccagaaactctttgatgagagcttcaagattttcccgagtaca
tcccgaacccatctgacagaagttcaccatagtaaattcttccaggtgctctttgccgt
cagactctttccggtaacaaggtccgacttcgaaaatttttattgggcctggtaaaatc
ctatcgagttttcgcagatagttgtaaagagtcggggcaagcattggcctcaagcagag
atttttatccacccggaagatctgttttgaaagttcagtatcattattaatacccattc
tctccacgtattccgccggaataaggataggagactttatctccagaaaaccccggtct
acgaaaaatttcgtaatatcacgttcgagtttaccgaggtagtcttctctatcattggt
atagagccgctgaaaatcgtttttttcttcttgtcacaagttcaggctcaagttccctga
aaggctttgccatatttagagaaattttatcctctggacttaagagagcctcaacccta
tcaagctggcttcttgtaagtgaaggagcaggagccgatgcgggaacagacgaatttgg
agttgattttgcaggcgaaggtacagatctggatgtgttcgttgatgcctttgcagaaa
cagaattttccagaggcttcggagcccttgaaactgatttcggcatagcttttttgacc
tttggagcagaaactaccctaactttcacactgttttgctttcggttgatcttgtgag
aaaattattgatatcctcgtccgaaaccctacatcgtttgcaggttttctgtacttat
gatgtctgaatgctctggctgttctacaactcctggaattattcacaacaagatggtct
ccacacgccatttcaatgtatattttacttcttgagacctcatggtgcttgattttgtg
gagcgtgccagtcctggacatccagagcccggtcgcagatattaaaacatctaatggtt
ttttatccatggtggcgctagccagcttgggtctccctatagtgagtcgtattaatttc
gataagccagtaagcagtgggttctctagttagccagagagctctgcttatatagacct
cccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacatttt
ggaaagtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggag
acttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaacc
gcatcaccatggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcc
cataaggtcatgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaat
aggggggcgtacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgta
aatagtccacccattgacgtcaatggaaagtccctattggcgttactatgggaacatac
gtcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttac
cgtaagttatgtaacgcggaactccatatatgggctatgaactaatgaccccgtaattg
attactattaataactagcaaaaaacgggcagcccgggaatctaacccggctgaacgga
tttagagtccgttcgatctacatgatcagccgcccggtgtttcgtcctttccacaagat
atataaagccaagaaatcgaaatactttcaagttacggtaagcatatgatagtccattt
taaaacataatttaaaactgcaaactacccaagaaattattactttctacgtcacgta
ttttgtactaatatctttgtgtttacagtcaaattaattctaattatctctctaacagc
cttgtatcgtatatgcaaatatgaaggaatcatgggaaataggccctcttcctgcccga
ggtacccctagcaaaaaacgggcagcccgggaatctaacccggctgaacggatttagag
tccgttcgatctacatgatcagccgcccggtgtttcgtcctttccacaagatatataaa
gccaagaaatcgaaatactttcaagttacggtaagcatatgatagtccattttaaaaca
taattttaaaactgcaaactacccaagaaattattactttctacgtcacgtattttgta
ctaatatctttgtgtttacagtcaaattaattctaattatctctctaacagccttgtat
cgtatatgcaaatatgaaggaatcatgggaaataggccctcttcctgcccgaggtaccc
ctagcaaaaaacgggcagcccgggaatctaacccggctgaacggatttagagtccgttc
gatctacatgatcagccgcccggtgtttcgtcctttccacaagatatataaagccaaga
aatcgaaatactttcaagttacggtaagcatatgatagtccattttaaaacataatttt
aaaactgcaaactacccaagaaattattactttctacgtcacgtattttgtactaatat
ctttgtgtttacagtcaaattaattctaattatctctctaacagccttgtatcgtatat
gcaaatatgaaggaatcatgggaaataggccctcttcctgcccgaggtaccctagcaa
aaaacgggcagcccgggaatctaacccggctgaacggatttagagtccgttcgatctac

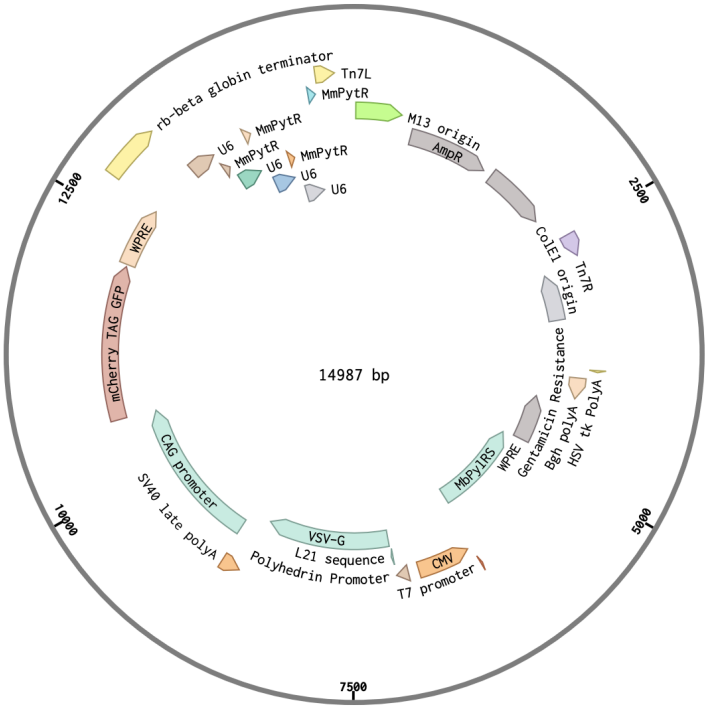
atgatcagccgcccggtgtttcgtcctttccacaagatatataaagccaagaaatcgaa
atactttcaagttacggtaagcatatgatagtccatttaaaacataatttaaaactg
caaactacccaagaaattattactttctacgtcacgtattttgtactaatatctttgtg
tttacagtcaaattaattctaattatctctctaacagccttgtatcgtatatgcaaata
tgaaggaatcatgggaaataggccctcttcctgcccgaggtacccctagtcaataatca
atgtcaacgcgtatatctggcccgtacatcgcgaagcagcgcaaaacggatcctgcagg
tatttgcggccgcggtccgtatactccggaatattaatagatcatggagataattaaaa
tgataaccatctcgcaaataaataagtattttactgttttcgtaacagttttgtaataa
aaaaacctataaatattccggattattcataccgtcccaccatcgggcgcgaactccta
aaaaaccgccaccatgaagtgccttttgtacttagccttttattcattggggtgaatt
gcaagttcaccatagtttttccacacaaccaaaaaggaaactggaaaaatgttccttct
aattaccattattgcccgtcaagctcagatttaaattggcataatgacttaataggcac
agccttacaagtcaaaatgcccaagagtcacaaggctattcaagcagacggttggatgt
gtcatgcttccaaatgggtcactacttgtgatttccgctggtatggaccgaagtatata
acacattccatccgatccttcactccatctgtagaacaatgcaaggaaagcattgaaca
aacgaaacaaggaacttggctgaatccaggcttccctcctcaaagttgtggatatgcaa
ctgtgacggatgccgaagcagtgattgtccaggtgactcctcaccatgtgctggttgat
gaatacacaggagaatgggttgattcacagttcatcaacggaaaatgcagcaattacat
atgccccactgtccataactctacaacctggcattctgactataaggtcaaagggctat
gtgattctaacctcatttccatggacatcaccttcttctcagaggacggagagctatca
tccctgggaaaggagggcacagggttcagaagtaactactttgcttatgaaactggagg
caaggcctgcaaaatgcaatactgcaagcattggggagtcagactcccatcaggtgtct
ggttcgagatggctgataaggatctctttgctgcagccagattccctgaatgcccagaa
gggtcaagtatctctgctccatctcagacctcagtggatgtaagtctaattcaggacgt
tgagaggatcttggattattccctctgccaagaaacctggagcaaaatcagagcgggtc
ttccaatctctccagtggatctcagctatcttgctcctaaaaacccaggaaccggtcct
gctttcaccataatcaatggtaccctaaaatactttgagaccagatacatcagagtcga
tattgctgctccaatcctctcaagaatggtcggaatgatcagtggaactaccacagaaa
gggaactgtgggatgactgggcaccatatgaagacgtggaaattggacccaatggagtt
ctgaggaccagttcaggatataagtttcctttatacatgattggacatggtatgttgga
ctccgatcttcatcttagctcaaaggctcaggtgttcgaacatcctcacattcaagacg
ctgcttcgcaacttcctgatgatgagagtttatttttggtgatactgggctatccaaa
aatccaatcgagcttgtagaaggttggttcagtagttggaaaagctctattgcctctt
tttctttatcatagggttaatcattggactattcttggttctccgagttggtatccatc
tttgcattaaattaaagcacaccaagaaaagacagatttatacagacatagagatgaac
cgacttggaaagtgataagtcgagaagtactagaggatcataatcagccataccacatt
tgtagaggttttacttgctttaaaaaacctcccacacctccccctgaacctgaaacata
aaatgaatgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataa
agcaatagcatcacaaatttcacaaataaagcatttttttcactgcattctagttgtgg
tttgtccaaactcatcaatgtatcttatcatgtctggatctgatcactgcttgagccta
gttattaatagtaatcaattacgggggtcattagttcatagcccatatatggagttccgc
gttacataacttacggtaaatggcccgcctggctgaccgcccaacgacccccgcccatt
gacgtcaataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtc
aatgggtggagtatttacggtaaactgcccacttggcagtacatcaagtgtatcatatg
ccaagtacgccccctattgacgtcaatgacggtaaatggcccgcctggcattatgccca
gtacatgaccttatgggactttcctacttggcagtacatctacgtattagtcatcgcta
ttaccatggtcgaggtgagccccacgttctgcttcactctcccatctccccccctcc
cccccccaatttgtatttatttatttttaattattttgtgcagcgatgggggcgggg


ggggggggggcgcgcgccaggcggggcggggcggggcgaggggcggggcggggcgagg
cggagaggtgcggcggcagccaatcagagcggcgcgctccgaaagtttccttttatggc
gaggcggcggcggcggcggccctataaaaagcgaagcgcgcggcgggcgggagtcgctg
cgttgccttcgccccgtgccccgctccgcgccgcctcgcgccgcccgccccggctctga
ctgaccgcgttactcccacaggtgagcgggcgggacggcccttctcctccgggctgtaa
ttagcgcttggtttaatgacggctcgtttctttttctgtggctgcgtgaaagccttaaag
ggctccgggagggccctttgtgcgggggggagcggctcgggggggtgcgtgcgtgtgtgt
gtgcgtgggggagcgccgcgtgcggcccgcgctgcccgggcggctgtgagcgctgcgggcg
cggcgcgggggctttgtgcgctccgcgtgtgcgcgaggggagcgcggccgggggcggtgc
cccgcggtgcgggggggctgcgagggggaacaaaggctgcgtgcgggggtgtgtgcgtggg
ggggtgagcaggggggtgtgggcgcggcggtcgggctgtaacccccccctgcacccccct
ccccgagttgctgagcacggcccggcttcgggtgcgggggctccgtgcgggggcgtggcgc
ggggctcgccgtgccgggcgggggggtggcggcaggtgggggtgccgggcgggggcgggggc
cgcctcgggccgggggagggctcgggggagggggcgcggcggcccccggagcgccggcggct
gtcgaggcgcggcgagccgcagccattgccttttatggtaatcgtgcgagagggcgcag
ggacttcctttgtcccaaatctggcggagccgaaatctgggaggcgccgccgcacccccc
tctagcgggcgcgggcgaagcggtgcggcgccggcaggaaggaaatgggcggggagggc
cttcgtgcgtcgccgcgccgccgtcccctttctccatctccagcctcggggctgccgcag
gggacggctgccttcgggggggacggggcagggcggggttcggcttctggcgtgtgac
cggcggctctagagcctctgctaaccatgttcatgccttcttctttttcctacagctcc
tgggcaacgtgctggttattgtgctgtctcatcattttggcaaagaattggccaaggat
ggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacg
gcgacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccacttag
ggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctggcccac
cctcgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgaccacatga
agcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcaccatc
ttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacac
cctggtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctgg
ggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaagcag
aagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgca
gctcgccgaccactaccagcagaacaccccatcggcgacggccccgtgctgctgcccg
acaaccactacctgagcacccagtccgccctgagcaaagaccccaacgagaagcgcgat
cacatggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacgagct
gtacaagggcccttcgaacaaaaactcatctcagaagaggatctgaatatgcataccg
gtcatcatcaccatcaccattgaattcaacgcgttaagtcgacaatcaacctctggat
tacaaaatttgtgaaagattgactggtattcttaactatgttgctccttttacgctatg
tggatacgctgctttaatgcctttgtatcatgctattgcttcccgtatggctttcattt
tctcctccttgtataaatcctggttgctgtctctttatgaggagttgtggcccgttgtc
aggcaacgtggcgtggtgtgcactgtgtttgctgacgcaaccccactggttggggcat
tgccaccacctgtcagctcctttccgggactttcgctttccccctccctattgccacgg
cggaactcatcgccgcctgccttgcccgctgctggacaggggctcggctgttgggcact
gacaattccgtggtgttgtcggggaaatcatcgtcctttccttggctgctcgcctgtgt
tgccacctggattctgcgcgggacgtccttctgctacgtcccttcggccctcaatccag
cggaccttccttcccgcggcctgctgccggctctgcggcctcttccgcgtcttcgcctt
cgccctcagacgagtcggatctccctttgggccgcctccccgcggcctggccggccgcc
agcacagtggtcgatcgaccaatgccctggctcacaaataccactgagatcttttcccc
tctgccaaaaattatggggacatcatgaagcccttgagcatctgacttctggctaata
aaggaaatttatttcattgcaatagtgtgttggaatttttgtgtctctcactcggaa

ggacatatgggagggcaaatcatttaaaacatcagaatgagtatttggtttagagtttg
gcaacatatgcccatatgctggctgccatgaacaaaggttggctataaagaggtcatca
gtatatgaaacagccccctgctgtccattccttattccatagaaaagccttgacttgag
gttagatttttttatattttgttttgtgttatttttttctttaacatccctaaaattt
tccttacatgttttactagccagatttttcctcctctcctgactactcccagtcatagc
tgtccctcttctcttatggagatccctcgacctgccctaggggtacctcgggcaggaag
agggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttagagag
ataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtag
aaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatc
atatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaa
ggacgaaacacccgggcggctgatcatgtagatcgaacggactctaaatccgttcagccg
ggttagattcccgggctgcccgttttttgctaggggtacctcgggcaggaagagggcct
atttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataatta
gaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaa
taatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgct
taccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaa
acacccgggcggctgatcatgtagatcgaacggactctaaatccgttcagccgggttaga
ttcccgggctgcccgttttttgctaggggtacctcgggcaggaagagggcctatttccc
atgattccttcatatttgcatatacgatacaaggctgttagagagataattagaattaa
tttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttc
ttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgta
acttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaacacccgg
gcggctgatcatgtagatcgaacggactctaaatccgttcagccgggttagattcccgg
gctgcccgttttttgctaggggtacctcgggcaggaagagggcctatttcccatgattc
cttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttgact
gtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggta
gtttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaa
agtatttcgatttcttggctttatatatcttgtggaaaggacgaacacccgggcggctg
atcatgtagatcgaacggactctaaatccgttcagccgggttagattcccgggctgccc
gttttttgctaggagatccgaaccagataagtgaaatctagttccaaactattttgtca
ttttttaattttcgtattagcttacgacgctacacccagttcccatctattttgtcactc
ttccctaaataatccttaaaaactccatttccacccctcccagttcccaactattttgt
ccgcccacagcggggcatttttcttcctgttatgttttaatcaaacatcctgccaact
ccatgtgacaaaccgtcatcttcggctacttttttctctgtcacagaatgaaaatttttc
tgtcatctcttcgttattaatgtttgtaattgactgaatatcaacgcttatttgcagc

The sequence for the wild type PytR is:
ggaaacctgatcatgtagatcgaacggactctaaatccgttcagccgggttagattccc
ggggtttccg

pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-4xHTS25



Sequence color-coding key

Feature	Color
M13 origin	text
AmpR	text
ColE1 origin	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
WPRE	text
MbPylRS	text
T7 promoter	text
CMV promoter	text
Polyhedrin Promoter	text
L21 sequence	text
VSV-G	text
SV40 late polyA	text
CAG promoter	text
mCherry TAG GFP	text
WPRE	text
rb-beta globin terminator	text
U6	text

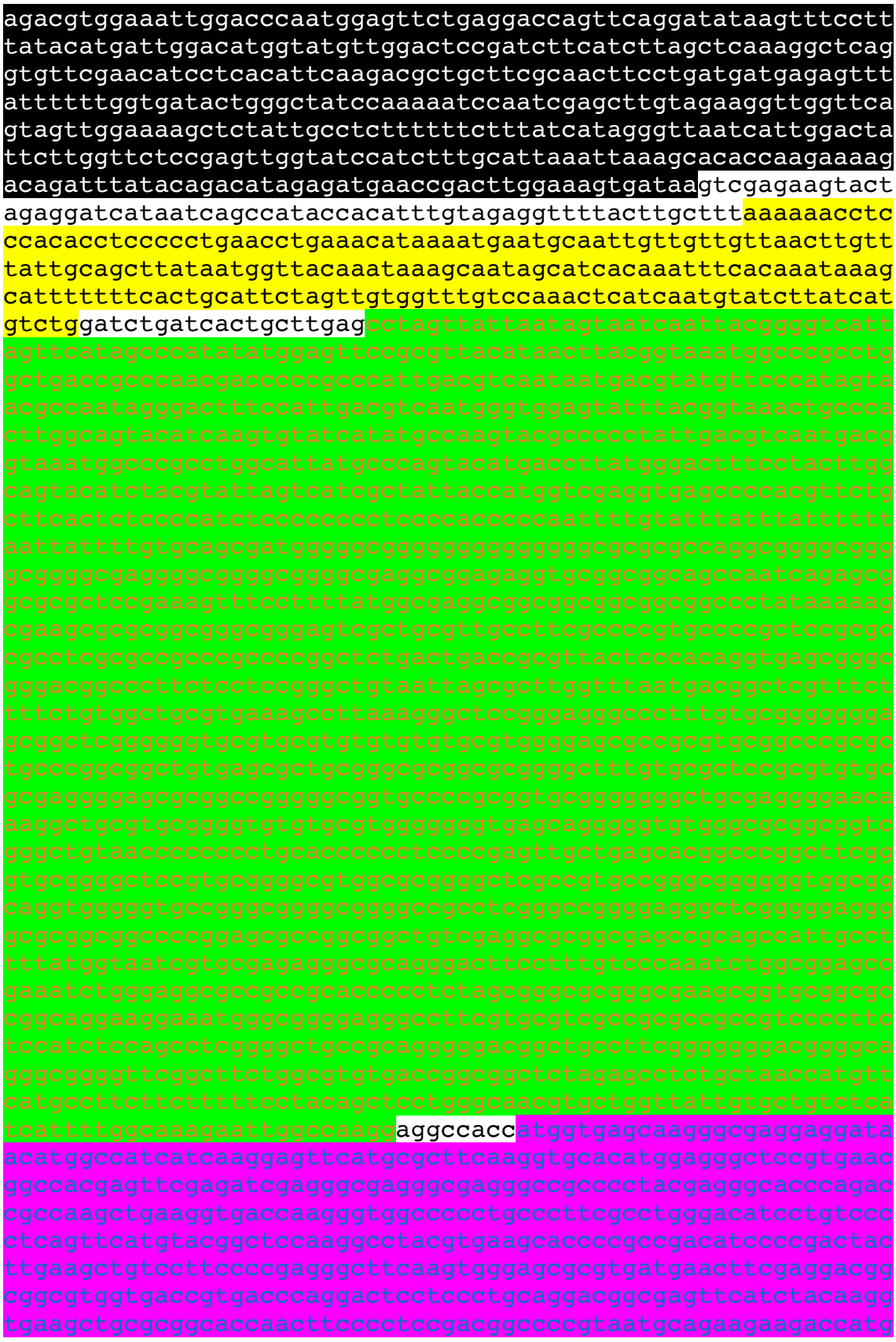
MmPytR	text
Tn7L	text

Sequence

ctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggt
tacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttct
tcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctc
cctttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgattaggg
tgatggttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttgg
agtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatc
tcggtctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaa
tgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgtttacaattt
caggtggcactttcggggaaatgtgcgcggaacccctatttgtttatttttctaaata
cattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattg
aaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcgg
cattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaa
gatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatcct
tgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctat
gtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacac
tattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatgg
catgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggcca
acttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatg
ggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaa
cgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaa
ctggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggat
aaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataa
atctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggta
agccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacga
aatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagacca
agtttactcatatatactttagattgatttaaaacttcattttaatttaaaaggatct
aggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttc
cactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttttct
gcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgc
cggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagata
ccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagc
accgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgata
agtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcg
ggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaact
gagatacctacagcgtgagcattgagaaagcgccacgcttcccgaagggagaaaggcgg
acaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccaggg
ggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcg
atttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcct
ttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcag
ccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcctgatgcggt
atttctccttacgcatctgtgcggtatttcacaccgcagaccagccgcgtaacctggc
aaaatcggttacggttgagtaataaatggatgccctgcgtaagcgggtgtgggcggaca

ataaagtcttaaactgaacaaaatagatctaaactatgacaataaagtcttaaactaga
cagaatagttgtaaactgaaatcagtccagttatgctgtgaaaaagcatactggacttt
tgttatggctaaagcaaactcttcattttctgaagtgcaaattgcccgtcgtattaaag
aggggcgtggccaagggcatggtaaagactatattcgcggcgttgtgacaatttaccga
acaactccgcggccgggaagccgatctcggcttgaacgaattgttaggtggcggtactt
gggtcgatatcaaagtgcatcacttcttcccgtatgcccaactttgtatagagagccac
tgcgggatcgtcaccgtaatctgcttgcacgtagatcacataagcaccaagcgcgttgg
cctcatgcttgaggagattgatgagcgcggtggcaatgccctgcctccggtgctcgccg
gagactgcgagatcatagatatagatctcactacgcggctgctcaaacctgggcagaac
gtaagccgcgagagcgccaacaaccgcttcttggtcgaaggcagcaagcgcgatgaatg
tcttactacggagcaagttcccgaggtaatcggagtccggctgatgttgggagtaggtg
gctacgtctccgaactcacgaccgaaaagatcaagagcagcccgcatggatttgacttg
gtcagggccgagcctacatgtgcgaatgatgcccatacttgagccacctaactttgttt
tagggcgactgccctgctgcgtaacatcgttgctgctgcgtaacatcgttgctgctcca
taacatcaaacatcgacccacggcgtaacgcgcttgctgcttggatgcccgaggcatag
actgtacaaaaaaacagtcataacaagccatgaaaaccgccactgcgccgttaccaccg
ctgcgttcggtcaaggttctggaccagttgcgtgagcgcatacgctacttgcattacag
tttacgaaccgaacaggcttatgtcaactgggttcgtgccttcatccgtttccacggtg
tgcgtcacccggcaaccttgggcagcagcgaagtcgaggcatttctgtcctggctggcg
aacgagcgcaaggtttcggtctccacgcatcgtcaggcattggcggccttgctgttctt
ctacggcaaggtgctgtgcacggatctgccctggcttcaggagatcggtagacctcggc
cgtcgcggcgcttgccggtggtgctgaccccggatgaagtggttcgcatcctcggtttt
ctggaaggcgagcatcgtttgttcgcccaggactctagctatagttctagtggttggct
acgtacccgtagtggctatggcagggcttgcgcttaatgcgccgctacagggcgcgtgg
ggataccccctagagccccagctggttctttccgcctcagaagccatagagcccaccgc
atccccagcatgcctgctattgtcttcccaatcctccccttgctgtcctgccccaccc
cacccccagaatagaatgacacctactcagacaatgcgatgcaatttcctcattttat
taggaaaggacagtgggagtggcaccttccagggtcaaggaaggcacgggggaggggca
aacaacagatggctggcaactagaaggcacagtcgaggctgatcagcgggtttaaacgg
gccctctagactcgagttaaagtcgacgcggggaggcggcccaaagggagatccgactc
gtctgagggcgaaggcgaagacgcggaagaggccgcagagccggcagcaggccgcggga
aggaaggtccgctggattgagggccgaagggacgtagcagaaggacgtcccgcgcagaa
tccaggtggcaacacaggcgagcagccaaggaaaggacgatgatttccccgacaacacc
acggaattgtcagtgcccaacagccgagcccctgtccagcagcgggcaaggcaggcggc
gatgagttccgccgtggcaatagggaggggaaagcgaaagtcccggaaaggagctgac
aggtggtggcaatgccccaaccagtgggggttgcgtcagcaaacacagtgcacaccacg
ccacgttgcctgacaacgggccacaactcctcataaagagacagcaaccaggatttata
caaggaggagaaaatgaaagccatacgggaagcaatagcatgatacaaaggcattaaag
cagcgtatccacatagcgtaaaaggagcaacatagttaagaataccagtcaatctttca
caaattttgtaatccagaggttgattgtcgacttaacgcgttgaattcttacagattgg
ttgaaatcccattatagtaagattcggaccttgatgccctcttaatgttttaaagccg
tgcataaccttgagcaagcgttcaagaccaaaacctgcacctatccatggtttgtcaat
accccattctctatcaagagaaactggcccgacgactgccgaagaaagctccaggtccc
cgtgcattatatcaagagtatccccaaagaccatacaggaatctcctacgatttcgaag
tcgatttccagatagtccagaaactctttgatgagagcttcaagattttcccgagtaca
tcccgaacccatctgacagaagttcaccatagtaaattcttccaggtgctctttgccgt
cagactctttccggtaacaaggtccgacttcgaaaatttttattgggcctggtaaaatc
ctatcgagtttcgcagatagttgtaaagagtcggggcaagcattggcctcaagcagag

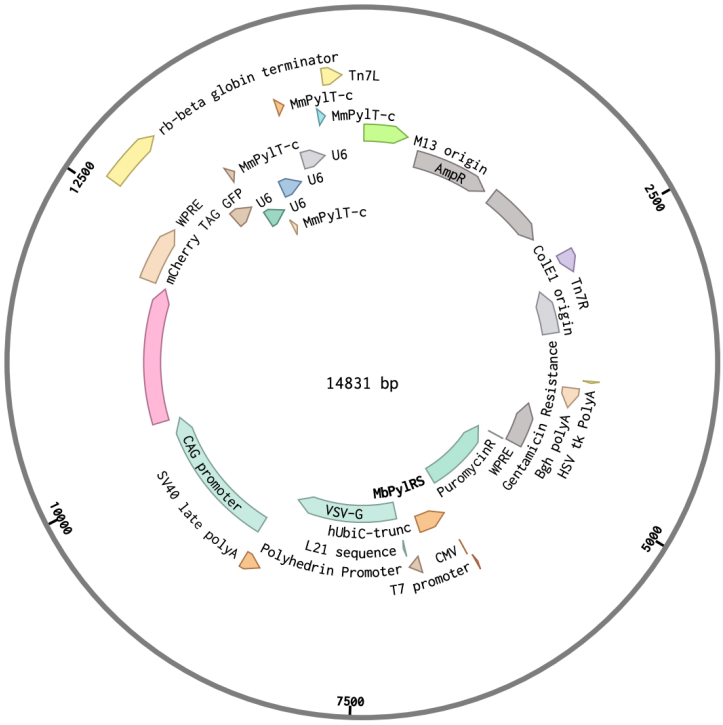
attttatccacccggaagatctgttttgaaagttcagtatcattattaatacccattc
tctccacgtattccgccggaataaggataggagactttatctccagaaaaccccggtct
acgaaaaatttcgtaatatcacgttcgagtttaccgaggtagtcttctctatcattggt
atagagccgctgaaaatcgtttttttcttcttgtcacaagttcaggctcaagttccctga
aaggctttgccatatttagagaaattttatcctctggacttaagagagcctcaacccta
tcaagctggcttcttgtaagtgaaggagcaggagccgatgcgggaacagacgaatttgg
agttgattttgcaggcgaaggtacagatctggatgtgttcgttgatgcctttgcagaaa
cagaatttccagaggcttcggagcccttgaaactgatttcggcatagcttttttgacc
tttggagcagaaactaccctaactttcacactgtttttgctttcggttgatcttgtgag
aaaattattgatatcctcgtccgaaaccctacatcgtttgcaggttttctgtacttat
gatgtctgaatgctctggctgttctacaactcctggaattattcacaacaagatggtct
ccacacgccatttcaatgtatattttacttcttgagacctcatggtgcttgattttgtg
gagcgtgccagtcctggacatccagagcccggtcgcagatattaaaacatctaatggtt
ttttatccatggtggcgctagccagcttgggtctccctatagtgagtcgtattaatttc
gataagccagtaagcagtgggttctctagttagccagagagctctgcttatatagacct
cccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacatttt
ggaaagtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggag
acttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaacc
gcatcaccatggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcc
cataaggtcatgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaat
aggggggcgtacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgta
aatagtccacccattgacgtcaatggaaagtccctattggcgttactatgggaacatac
gtcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttac
cgtaagttatgtaacgcggaactccatatatgggctatgaactaatgaccccgtaattg
attactattaataactagtcaataatcaatgtcaacgcgtatatctggcccgtacatcg
cgaagcagcgcaaaacggatcctgcaggtatttgcggccgcggtccgtatactccggaa
tattaatagatcatggagataattaaaatgataaccatctcgcaaataaataagtattt
tactgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcata
ccgtcccaccatcgggcgcgaactcctaaaaaaccgccaccatgaagtgccttttgtac
ttagccttttttattcattggggtgaattgcaagttcaccatagttttccacacaacca
aaaaggaaactggaaaaatgttccttctaattaccattattgcccgtcaagctcagatt
taaattggcataatgacttaataggcacagccttacaagtcaaaatgcccaagagtcac
aaggctattcaagcagacggttggatgtgtcatgcttccaaatgggtcactacttgtga
tttccgctggtatggaccgaagtatataacacattccatccgatccttcactccatctg
tagaacaatgcaaggaaagcattgaacaaacgaaacaaggaacttggctgaatccaggc
ttccctcctcaaagttgtggatatgcaactgtgacggatgccgaagcagtgattgtcca
ggtgactcctcaccatgtgctggttgatgaatacacaggagaatgggttgattcacagt
tcatcaacggaaaatgcagcaattacatatgccccactgtccataactctacaacctgg
cattctgactataaggtcaaagggctatgtgattctaacctcatttccatggacatcac
cttcttctcagaggacggagagctatcatccctgggaaaggagggcacagggttcagaa
gtaactactttgcttatgaaactggaggcaaggcctgcaaaatgcaatactgcaagcat
tggggagtcagactcccatcaggtgtctggttcgagatggctgataaggatctctttgc
tgcagccagattccctgaatgcccagaagggtcaagtatctctgctccatctcagacct
cagtggatgtaagtctaattcaggacgttgagaggatcttggattattccctctgccaa
gaaacctggagcaaaatcagagcgggtcttccaatctctccagtggatctcagctatct
tgctcctaaaaacccaggaaccggtcctgctttcaccataatcaatggtaccctaaaat
actttgagaccagatacatcagagtcgatattgctgctccaatcctctcaagaatggtc
ggaatgatcagtggaactaccacagaaagggaactgtgggatgactgggcaccatatga



ggctgggaggcctcctccgagcggatgtaccccgaggacggcgccctgaagggcgagat
caagcagaggctgaagctgaaggacggcggccactacgacgctgaggtcaagaccacct
acaaggccaagaagcccgtgcagctgcccggcgcctacaacgtcaacatcaagttggac
atcacctcccacaacgaggactacaccatcgtggaacagtacgaacgcgccgagggccg
ccactccaccggcggcatggacgagctgtacaagggcggctcttagggcggctccctcg
tgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacggc
gacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacgg
caagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctggcccaccc
tcgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgaccacatgaag
cagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcaccatctt
cttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccc
tggtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctgggg
cacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaagcagaa
gaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagc
tcgccgaccactaccagcagaacaccccatcggcgacggccccgtgctgctgcccgac
aaccactacctgagcacccagtccgccctgagcaaagaccccaacgagaagcgcgatca
catggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacgagctgt
acaaggggccttcgaacaaaaactcatctcagaagaggatctgaatatgcataccggt
catcatcaccatcaccattaatgagaattcaacgcgttaagtcgacaatcaacctctgg
attacaaaatttgtgaaagattgactggtattcttaactatgttgctccttttacgcta
tgtggatacgctgctttaatgcctttgtatcatgctattgcttcccgtatggctttcat
tttctcctccttgtataaatcctggttgctgtctctttatgaggagttgtggcccgttg
tcaggcaacgtggcgtggtgtgcactgtgtttgctgacgcaacccccactggttggggc
attgccaccacctgtcagctcctttccgggactttcgctttcccctccctattgccac
ggcggaactcatcgccgcctgccttgcccgctgctggacaggggctcggctgttgggca
ctgacaattccgtggtgttgtcggggaaatcatcgtcctttccttggctgctcgcctgt
gttgccacctggattctgcgcgggacgtccttctgctacgtcccttcggccctcaatcc
agcggaccttccttcccgcggcctgctgccggctctgcggcctcttccgcgtcttcgcc
ttcgccctcagacgagtcggatctcccttgggccgcctccccgcggcctggccggccg
ccagcacagtggtcgatcgaccaatgccctggctcacaaataccactgagatcttttc
cctctgccaaaaattatggggacatcatgaagcccttgagcatctgacttctggctaa
taaaggaaatttatttcattgcaatagtgtgttggaatttttgtgtctctcactcgg
aaggacatatgggagggcaaatcatttaaaacatcagaatgagtatttggtttagagtt
tggcaacatatgcccatatgctggctgccatgaacaaaggttggctataaagaggtcat
cagtatatgaaacagccccctgctgtccattccttattccatagaaaagccttgacttg
aggttagatttttttatattttgttttgtgttatttttttctttaacatccctaaaat
tttccttacatgttttactagccagatttttcctcctctcctgactactcccagtcata
gctgtccctcttctcttatggagatccctcgacctgccctagggtacctcgggcagga
agagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttagag
agataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgt
agaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatggacta
tcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtgga
aaggacgaaacaccgggcggctgatcatgtagatcgaacggactctaaatccgttcagc
cgggttagattcccgggctgcccgttttttgctaggggtacctcgggcaggaagagggc
ctatttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataat
tagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagt
aataatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatg
cttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacg

aaacaccgggcggctgatcatgtagatcgaacggactctaaatccgttcagccgggtta
gattcccgggctgcccgtttttgctaggggtacctcgggcaggaagagggcctatttc
ccatgattccttcatatttgcatatacgatacaaggctgttagagagataattagaatt
aatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatt
tcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgcttaccg
taacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacacc
gggcggctgatcatgtagatcgaacggactctaaatccgttcagccgggttagattccc
gggctgcccgtttttgctaggggtacctcgggcaggaagagggcctatttcccatgat
tccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttga
ctgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttggg
tagtttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttg
aaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacaccgggcggc
tgatcatgtagatcgaacggactctaaatccgttcagccgggttagattcccgggctgc
ccgtttttgctaggagatccgaaccagataagtgaaatctagttccaaactattttgt
cattttaattttcgtattagcttacgacgctacacccagttcccatctattttgtcac
tcttccctaaataatccttaaaaactccatttccacccctcccagttcccaactatttt
gtccgcccacagcggggcatttttcttcctgttatgttttaatcaaacatcctgccaa
ctccatgtgacaaaccgtcatcttcggctacttttttctctgtcacagaatgaaaatttt
tctgtcatctcttcgttattaatgtttgtaattgactgaatatcaacgcttatttgcag
c

pAcBac-UbiC-MbPylRS-CAG-mCherry-TAG-GFP-4xHTS25



Sequence color-coding key

Feature	Color
M13 origin	text
AmpR	text
ColE1 origin	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
WPRE	text
MbPylRS	text
T7 promoter	text
UbiC promoter	text
Polyhedrin Promoter	text
L21 sequence	text
VSV-G	text
SV40 late polyA	text
CAG promoter	text
mCherry TAG GFP	text
WPRE	text
rb-beta globin terminator	text
U6	text

MmPytR	text
Tn7L	text

Sequence

ctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggt
tacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttct
tcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctc
cctttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgattaggg
tgatggttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttgg
agtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatc
tcggtctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaa
tgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgtttacaattt
caggtggcactttcggggaaatgtgcgcggaacccctatttgtttatttttctaaata
cattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattg
aaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcgg
cattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaa
gatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatcct
tgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctat
gtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacac
tattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatgg
catgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggcca
acttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatg
ggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaa
cgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaa
ctggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggat
aaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataa
atctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggta
agccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacga
aatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagacca
agtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatct
aggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttc
cactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttttct
gcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgc
cggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagata
ccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagc
accgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgata
agtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcg
ggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaact
gagatacctacagcgtgagcattgagaaagcgccacgcttcccgaagggagaaaggcgg
acaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccaggg
ggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcg
atttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcct
ttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcag
ccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcctgatgcggt
atttctccttacgcatctgtgcggtatttcacaccgcagaccagccgcgtaacctggc
aaaatcggttacggttgagtaataaatggatgccctgcgtaagcgggtgtgggcggaca

ataaagtcttaaactgaacaaaatagatctaaactatgacaataaagtcttaaactaga
cagaatagttgtaaactgaaatcagtccagttatgctgtgaaaaagcatactggacttt
tgttatggctaaagcaaactcttcattttctgaagtgcaaattgcccgtcgtattaaag
aggggcgtggccaagggcatggtaaagactatattcgcggcgttgtgacaatttaccga
acaactccgcggccgggaagccgatctcggcttgaacgaattgttaggtggcggtactt
gggtcgatatcaaagtgcatcacttcttcccgtatgcccaactttgtatagagagccac
tgcgggatcgtcaccgtaatctgcttgcacgtagatcacataagcaccaagcgcgttgg
cctcatgcttgaggagattgatgagcgcggtggcaatgccctgcctccggtgctcgccg
gagactgcgagatcatagatatagatctcactacgcggctgctcaaacctgggcagaac
gtaagccgcgagagcgccaacaaccgcttcttggtcgaaggcagcaagcgcgatgaatg
tcttactacggagcaagttcccgaggtaatcggagtccggctgatgttgggagtaggtg
gctacgtctccgaactcacgaccgaaaagatcaagagcagcccgcatggatttgacttg
gtcagggccgagcctacatgtgcgaatgatgcccatacttgagccacctaactttgttt
tagggcgactgccctgctgcgtaacatcgttgctgctgcgtaacatcgttgctgctcca
taacatcaaacatcgacccacggcgtaacgcgcttgctgcttggatgcccgaggcatag
actgtacaaaaaaacagtcataacaagccatgaaaaccgccactgcgccgttaccaccg
ctgcgttcggtcaaggttctggaccagttgcgtgagcgcatacgctacttgcattacag
tttacgaaccgaacaggcttatgtcaactgggttcgtgccttcatccgtttccacggtg
tgcgtcacccggcaaccttgggcagcagcgaagtcgaggcatttctgtcctggctggcg
aacgagcgcaaggtttcggtctccacgcatcgtcaggcattggcggccttgctgttctt
ctacggcaaggtgctgtgcacggatctgccctggcttcaggagatcggtagacctcggc
cgtcgcggcgcttgccggtggtgctgaccccggatgaagtggttcgcatcctcggtttt
ctggaaggcgagcatcgtttgttcgcccaggactctagctatagttctagtggttggct
acgtacccgtagtggctatggcagggcttgcgcttaatgcgccgctacagggcgcgtgg
ggataccccctagagccccagctggttctttccgcctcagaagccatagagcccaccgc
atccccagcatgcctgctattgtcttcccaatcctccccttgctgtcctgccccaccc
cacccccagaatagaatgacacctactcagacaatgcgatgcaatttcctcattttat
taggaaaggacagtgggagtggcaccttccagggtcaaggaaggcacgggggaggggca
aacaacagatggctggcaactagaaggcacagtcgaggctgatcagcgggtttaaacgg
gccctctagactcgagttaaagtcgacgcggggaggcggcccaaagggagatccgactc
gtctgagggcgaaggcgaagacgcggaagaggccgcagagccggcagcaggccgcggga
aggaaggtccgctggattgagggccgaagggacgtagcagaaggacgtcccgcgcagaa
tccaggtggcaacacaggcgagcagccaaggaaaggacgatgatttccccgacaacacc
acggaattgtcagtgcccaacagccgagcccctgtccagcagcgggcaaggcaggcggc
gatgagttccgccgtggcaatagggaggggaaagcgaaagtcccggaaaggagctgac
aggtggtggcaatgccccaaccagtgggggttgcgtcagcaaacacagtgcacaccacg
ccacgttgcctgacaacgggccacaactcctcataaagagacagcaaccaggatttata
caaggaggagaaaatgaaagccatacgggaagcaatagcatgatacaaaggcattaaag
cagcgtatccacatagcgtaaaaggagcaacatagttaagaataccagtcaatctttca
caaattttgtaatccagaggttgattgtcgacttaacgcgttgaattcttacagattgg
ttgaaatcccattatagtaagattcggaccttgatgccctcttaatgttttaaagccg
tgcataaccttgagcaagcgttcaagaccaaaacctgcacctatccatggtttgtcaat
accccattctctatcaagagaaactggcccgacgactgccgaagaaagctccaggtccc
cgtgcattatatcaagagtatccccaaagaccatacaggaatctcctacgatttcgaag
tcgatttccagatagtccagaaactctttgatgagagcttcaagattttcccgagtaca
tcccgaacccatctgacagaagttcaccatagtaaattcttccaggtgctctttgccgt
cagactctttccggtaacaaggtccgacttcgaaaatttttattgggcctggtaaaatc
ctatcgagtttcgcagatagttgtaaagagtcggggcaagcattggcctcaagcagag

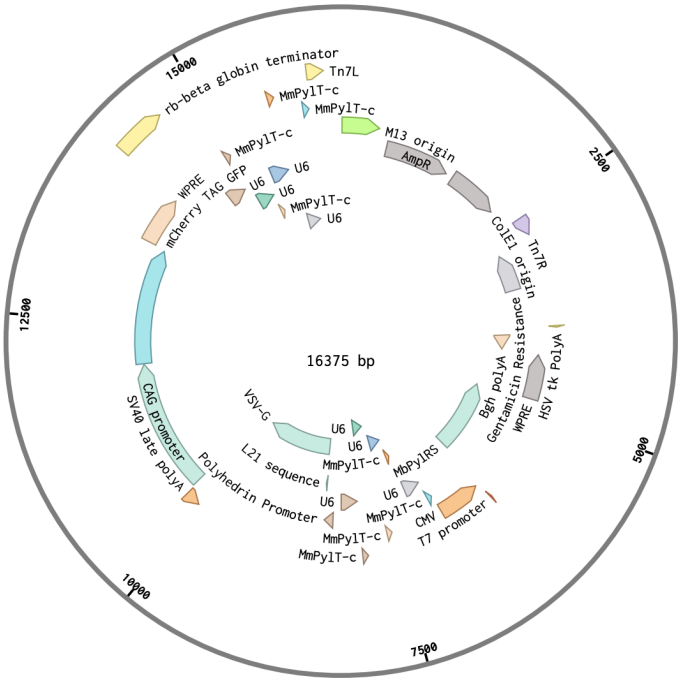
attttatccacccggaagatctgttttgaaagttcagtatcattattaatacccattc
tctccacgtattccgccggaataaggataggagactttatctccagaaaaccccggtct
acgaaaaatttcgtaatatcacgttcgagtttaccgaggtagtcttctctatcattggt
atagagccgctgaaaatcgtttttttcttcttgtcacaagttcaggctcaagttccctga
aaggctttgccatatttagagaaattttatcctctggacttaagagagcctcaacccta
tcaagctggcttcttgtaagtgaaggagcaggagccgatgcgggaacagacgaatttgg
agttgattttgcaggcgaaggtacagatctggatgtgttcgttgatgcctttgcagaaa
cagaatttccagaggcttcggagcccttgaaactgatttcggcatagcttttttgacc
tttggagcagaaactaccctaactttcacactgtttttgctttcggttgatcttgtgag
aaaattattgatatcctcgtccgaaaccctacatcgtttgcaggttttctgtacttat
gatgtctgaatgctctggctgttctacaactcctggaattattcacaacaagatggtct
ccacacgccatttcaatgtatattttacttcttgagacctcatggtgcttgattttgtg
gagcgtgccagtcctggacatccagagcccggtcgcagatattaaaacatctaatggtt
ttttatccatggtggcgctagccagcttgggtctccctatagtgagtcgtattaatttc
gataagccagtaagcagtgggttctctagttagccagagc ctagaccaagtgacgatc
acagcgatccacaaacaagaaccgcgacccaaatcccggctgcgacggaactagctgtg
ccacacccggcgcgtccttatataatcatcggcgttcaccgccccacggagatccctcc
gcagaatcgccgagaagggactacttttcctcgcctgttccgctctctggaaagaaaac
cagtgccctagagtcacccaagtcccgtcctaaaatgtccttctgctgatactggggtt
ctaaggccgagtcttatgagcagcgggccgctgtcctgagcgtccgggcggaaggatca
ggacgctcgctgcgcccttcgtctgacgtggcagcgctcgccgtgaggaggggggcgcc
cgcgggaggcgccaaaacccggcgcggaggccactagtcaataatcaatgtcaacgcgt
atatctggcccgtacatcgcgaagcagcgcaaaacggatcctgcaggtatttgcggccg
cggtccgtatactccggaatattaatagatcatggagataattaaaatgataaccatct
cgcaaataaataagtattttactgttttcgtaacagtttgtaataaaaaaacctataa
atattccggattattcataccgtcccaccatcgggcgcgaactcctaaaaaaccgccac
catgaagtgcctttgtacttagccttttattcattggggtgaattgcaagttcacca
tagttttccacacaaccaaaaaggaaactggaaaaatgttccttctaattaccattat
tgcccgtcaagctcagatttaaattggcataatgacttaataggcacagccttacaagt
caaaatgcccaagagtcacaaggctattcaagcagacggttggatgtgtcatgcttcca
aatgggtcactacttgtgatttccgctggtatggaccgaagtatataacacattccatc
cgatccttcactccatctgtagaacaatgcaaggaaagcattgaacaaacgaaacaagg
aacttggctgaatccaggcttccctcctcaaagttgtggatatgcaactgtgacggatg
ccgaagcagtgattgtccaggtgactcctcaccatgtgctggttgatgaatacacagga
gaatgggttgattcacagttcatcaacggaaaatgcagcaattacatatgccccactgt
ccataactctacaacctggcattctgactataaggtcaaagggctatgtgattctaacc
tcatttccatggacatcaccttcttctcagaggacggagagctatcatccctgggaaag
gagggcacagggttcagaagtaactactttgcttatgaaactggaggcaaggcctgcaa
aatgcaatactgcaagcattggggagtcagactcccatcaggtgtctggttcgagatgg
ctgataaggatctctttgctgcagccagattccctgaatgcccagaagggtcaagtatc
tctgctccatctcagacctcagtggatgtaagtctaattcaggacgttgagaggatctt
ggattattccctctgccaagaaacctggagcaaaatcagagcgggtcttccaatctctc
cagtggatctcagctatcttgctcctaaaaacccaggaaccggtcctgctttcaccata
atcaatggtaccctaaaatactttgagaccagatacatcagagtcgatattgctgctcc
aatcctctcaagaatggtcggaatgatcagtggaactaccacagaaagggaactgtggg
atgactgggcaccatatgaagacgtggaaattggacccaatggagttctgaggaccagt
tcaggatataagtttcctttatacatgattggacatggtatgttggactccgatcttca
tcttagctcaaaggctcaggtgttcgaacatcctcacattcaagacgctgcttcgcaac

ttcctgatgatgagagtttatttttggtgatactgggctatccaaaaatccaatcgag
cttgtagaaggttggttcagtagttggaaaagctctattgcctctttttctttatcat
agggttaatcattggactattcttggttctccgagttggtatccatctttgcattaaat
taaagcacaccaagaaaagacagatttatacagacatagagatgaaccgacttggaaag
tgataagtcgagaagtactagaggatcataatcagccataccacatttgtagaggtttt
acttgctttaaaaaacctcccacacctccccctgaacctgaaacataaaatgaatgcaa
ttgttgttgttaacttgtttattgcagcttataatggttacaaataaagcaatagcatc
acaaatttcacaaataaagcatttttttcactgcattctagttgtggtttgtccaaact
catcaatgtatcttatcatgtctggatctgatcactgcttgagcctagttattaatagt
aatcaattacggggtcattagttcatagcccatatatggagttccgcgttacataactt
acggtaaatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataat
gacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagt
atttacggtaaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccc
cctattgacgtcaatgacggtaaatggcccgcctggcattatgcccagtacatgacctt
atgggactttcctacttggcagtacatctacgtattagtcatcgctattaccatggtcg
aggtgagccccacgttctgcttcactctccccatctccccccccctccccacccccaatt
ttgtatttatttatttttaattattttgtgcagcgatgggggcgggggggggggggggc
gcgcgccaggcggggcggggcggggcgaggggcggggcggggcgaggcggagaggtgcg
gcggcagccaatcagagcggcgcgctccgaaagtttccttttatggcgaggcggcggcg
gcggcggccctataaaaagcgaagcgcgcggcgggcgggagtcgctgcgttgccttcgc
cccgtgccccgctccgcgccgcctcgcgccgcccgcccccggctctgactgaccgcgtta
ctcccacaggtgagcgggcgggacggcccttctcctccgggctgtaattagcgcttggt
ttaatgacggctcgtttcttttctgtggctgcgtgaaagccttaaagggctccgggagg
gccctttgtgcgggggggagcggctcgggggtgcgtgcgtgtgtgtgtgcgtgggggag
cgccgcgtgcggcccgcgctgcccggcggctgtgagcgctgcgggcgcggcgcggggct
ttgtgcgctccgcgtgtgcgcgaggggagcgcggccgggggcggtgccccgcggtgcgg
gggggctgcgaggggaacaaaggctgcgtgcgggtgtgtgcgtgggggggtgagcagg
gggtgtgggcgcggcggtcgggctgtaacccccccctgcacccccctcccccgagttgct
gagcacggcccggcttcgggtgcgggctccgtgcggggcgtggcgcggggctcgccgt
gccgggcgggggtggcggcaggtgggggtgccgggcgggggcggggccgcctcgggccg
gggagggctcgggggaggggcgcggcggcccccggagcgccggcggctgtcgaggcgcgg
cgagccgcagccattgccttttatggtaatcgtgcgagagggcgcagggacttcctttg
tcccaaatctggcggagccgaaatctgggaggcgccgccgcaccccctctagcgggcgc
gggcgaagcggtgcggcgccggcaggaaggaaatgggcggggagggccttcgtgcgtcg
ccgcgccgccgtcccctttctccatctccagcctcggggctgccgcaggggacggctgc
cttcgggggggacggggcagggcgggggttcggcttctggcgtgtgaccggcggctctag
agcctctgctaaccatgttcatgccttcttcttttcctacagctcctgggcaacgtgc
tggttattgtgctgtctcatcattttggcaaagaattggccaaggaggccaccatggtg
agcaagggcgaggaggataacatggccatcatcaaggagttcatgcgcttcaaggtgca
catggagggctccgtgaacggccacgagttcgagatcgagggcgagggcgagggccgcc
cctacgagggcacccagaccgccaagctgaaggtgaccaagggtggcccctgcccttc
gcctgggacatcctgtcccctcagttcatgtacggctccaaggcctacgtgaagcaccc
cgccgacatccccgactacttgaagctgtccttccccgagggcttcaagtgggagcgcg
tgatgaacttcgaggacggcggcgtggtgaccgtgacccaggactcctccctgcaggac
ggcgagttcatctacaaggtgaagctgcgcggcaccaacttcccctccgacggccccgt
aatgcagaagaagaccatgggctgggaggcctcctccgagcggatgtaccccgaggacg
gcgccctgaagggcgagatcaagcagaggctgaagctgaaggacggcggccactacgac
gctgaggtcaagaccacctacaaggccaagaagcccgtgcagctgcccggcgcctacaa

cgtcaacatcaagttggacatcacctcccacaacgaggactacaccatcgtggaacagt
acgaacgcgccgagggccgccactccaccggcggcatggacgagctgtacaagggcggc
tcttagggcggctccctcgtgagcaagggcgaggagctgttcaccggggtggtgcccat
cctggtcgagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgagggcg
agggcgatgccacctacggcaagctgaccctgaagttcatctgcaccaccggcaagctg
cccgtgccctggcccaccctcgtgaccaccctgacctacggcgtgcagtgcttcagccg
ctaccccgaccacatgaagcagcacgacttcttcaagtccgccatgcccgaaggctacg
tccaggagcgcaccatcttcttcaaggacgacggcaactacaagacccgcgccgaggtg
aagttcgagggcgacaccctggtgaaccgcatcgagctgaagggcatcgacttcaagga
ggacggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtctata
tcatggccgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatc
gaggacggcagcgtgcagctcgccgaccactaccagcagaacaccccatcggcgacgg
ccccgtgctgctgcccgacaaccactacctgagcacccagtccgccctgagcaaagacc
ccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccgccgggatcact
ctcggcatggacgagctgtacaaggggcccttcgaacaaaaactcatctcagaagagga
tctgaatatgcataccggtcatcatcaccatcaccattaatgagaattcaacgcgttaa
gtcgacaatcaacctctggattacaaaatttgtgaaagattgactggtattcttaacta
tgttgctccttttacgctatgtggatacgctgctttaatgcctttgtatcatgctattg
cttcccgtatggctttcattttctcctccttgtataaatcctggttgctgtctctttat
gaggagttgtggcccgttgtcaggcaacgtggcgtggtgtgcactgtgtttgctgacgc
aaccccccactggttggggcattgccaccacctgtcagctcctttccgggactttcgctt
tccccctccctattgccacggcggaactcatcgccgcctgccttgcccgctgctggaca
ggggctcggctgttgggcactgacaattccgtggtgttgtcggggaaatcatcgtcctt
tccttggctgctcgcctgtgttgccacctggattctgcgcgggacgtccttctgctacg
tcccttcggccctcaatccagcggaccttccttcccgcggcctgctgccggctctgcgg
cctcttccgcgtcttcgccttcgccctcagacgagtcggatctccctttgggccgcctc
cccgcggcctggccggccgccagcacagtggtcgatcgaccaatgccctggctcacaaa
taccactgagatcttttcccctctgccaaaaattatggggacatcatgaagccccttga
gcatctgacttctggctaataaaggaaatttattttcattgcaatagtgtgttggaatt
ttttgtgtctctcactcggaaggacatatgggagggcaaatcatttaaaacatcagaat
gagtatttggtttagagtttggcaacatatgcccatatgctggctgccatgaacaaagg
ttggctataaagaggtcatcagtatatgaaacagccccctgctgtccattccttattcc
atagaaaagccttgacttgaggttagatttttttatattttgttttgtgttatttttt
tctttaacatccctaaaatttccttacatgttttactagccagattttcctcctctc
ctgactactcccagtcatagctgtccctcttctcttatggagatccctcgacctgccct
aggggtacctcgggcaggaagagggcctatttcccatgattccttcatatttgcatata
cgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatatt
agtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaat
tatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttg
gctttatatatcttgtggaaaggacgaaacaccgggcggctgatcatgtagatcgaacg
gactctaaatccgttcagccgggttagattcccgggctgcccgtttttgctaggggta
cctcgggcaggaagagggcctatttcccatgattccttcatatttgcatatacgataca
aggctgttagagagataattagaattaatttgactgtaaacacaaagatattagtacaa
aatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttt
taaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttat
atatcttgtggaaaggacgaaacaccgggcggctgatcatgtagatcgaacggactcta
aatccgttcagccgggttagattcccgggctgcccgtttttgctaggggtacctcggg
caggaagagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgt

tagagagataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgt
gacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatg
gactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatctt
gtggaaaggacgaaacaccgggcggctgatcatgtagatcgaacggactctaaatccgt
tcagccgggttagattcccgggctgcccgtttttgctaggggtacctcgggcaggaag
agggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttagagag
ataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtag
aaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatc
atatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaa
ggacgaaacaccgggcggctgatcatgtagatcgaacggactctaaatccgttcagccg
ggttagattcccgggctgcccgtttttgctaggagatccgaaccagataagtgaaatc
tagttccaaactattttgtcatttttaattttcgtattagcttacgacgctacacccag
ttcccatctattttgtcactcttccctaaataatccttaaaaactccatttccacccct
cccagttcccaactattttgtccgcccacagcggggcatttttcttcctgttatgtttt
taatcaaacatcctgccaactccatgtgacaaaccgtcatcttcggctacttttttctct
gtcacagaatgaaaatttttctgtcatctcttcgttattaatgtttgtaattgactgaa
tatcaacgcttatttgcagc

pAcBac-CMV-MbPylRS-CAG-mCherry-TAG-GFP-8xPytR



Sequence color-coding key

Feature	Color
M13 origin	text
AmpR	text
ColE1 origin	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
WPRE	text
MbPylRS	text
T7 promoter	text
CMV promoter	text
Polyhedrin Promoter	text
L21 sequence	text
VSV-G	text
SV40 late polyA	text
CAG promoter	text
mCherry TAG GFP	text
WPRE	text
rb-beta globin terminator	text
U6	text
MmPytR	text

Tn7L	text

Sequence of HTS25 variant

ctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggt
tacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttct
tcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctc
cctttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgattaggg
tgatggttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttgg
agtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatc
tcggtctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaa
tgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgtttacaattt
caggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaata
cattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattg
aaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttttgcgg
cattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaa
gatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatcct
tgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctat
gtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacac
tattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatgg
catgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggcca
acttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatg
ggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaa
cgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaa
ctggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggat
aaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataa
atctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggta
agccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacga
aatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagacca
agtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatct
aggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttc
cactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttttct
gcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgc
cggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagata
ccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagc
accgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgata
agtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcg
ggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaact
gagatacctacagcgtgagcattgagaaagcgccacgcttcccgaagggagaaaggcgg
acaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccaggg
ggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcg
atttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcct
ttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcag
ccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcctgatgcggt
attttctccttacgcatctgtgcggtatttcacaccgcagaccagccgcgtaacctggc
aaaatcggttacggttgagtaataaatggatgccctgcgtaagcgggtgtgggcggaca
ataaagtcttaaactgaacaaaatagatctaaactatgacaataaagtcttaaactaga

cagaatagttgtaaactgaaatcagtccagttatgctgtgaaaaagcatactggacttt
tgttatggctaaagcaaactcttcattttctgaagtgcaaattgcccgtcgtattaaag
aggggcgtggccaagggcatggtaaagactatattcgcggcgttgtgacaatttaccga
acaactccgcggccgggaagccgatctcggcttgaacgaattgttaggtggcggtactt
gggtcgatatcaaagtgcatcacttcttcccgtatgcccaactttgtatagagagccac
tgcgggatcgtcaccgtaatctgcttgcacgtagatcacataagcaccaagcgcgttgg
cctcatgcttgaggagattgatgagcgcggtggcaatgccctgcctccggtgctcgccg
gagactgcgagatcatagatatagatctcactacgcggctgctcaaacctgggcagaac
gtaagccgcgagagcgccaacaaccgcttcttggtcgaaggcagcaagcgcgatgaatg
tcttactacggagcaagttcccgaggtaatcggagtccggctgatgttgggagtaggtg
gctacgtctccgaactcacgaccgaaaagatcaagagcagcccgcatggatttgacttg
gtcagggccgagcctacatgtgcgaatgatgcccatacttgagccacctaactttgttt
tagggcgactgccctgctgcgtaacatcgttgctgctgcgtaacatcgttgctgctcca
taacatcaaacatcgacccacggcgtaacgcgcttgctgcttggatgcccgaggcatag
actgtacaaaaaaacagtcataacaagccatgaaaaccgccactgcgccgttaccaccg
ctgcgttcggtcaaggttctggaccagttgcgtgagcgcatacgctacttgcattacag
tttacgaaccgaacaggcttatgtcaactgggttcgtgccttcatccgtttccacggtg
tgcgtcacccggcaaccttgggcagcagcgaagtcgaggcatttctgtcctggctggcg
aacgagcgcaaggtttcggtctccacgcatcgtcaggcattggcggccttgctgttctt
ctacggcaaggtgctgtgcacggatctgccctggcttcaggagatcggtagacctcggc
cgtcgcggcgcttgccggtggtgctgaccccggatgaagtggttcgcatcctcggtttt
ctggaaggcgagcatcgtttgttcgcccaggactctagctatagttctagtggttggct
acgtacccgtagtggctatggcagggcttgcgcttaatgcgccgctacagggcgcgtgg
ggataccccctagagccccagctggttctttccgcctcagaagccatagagcccaccgc
atccccagcatgcctgctattgtcttcccaatcctccccttgctgtcctgccccaccc
cacccccagaatagaatgacacctactcagacaatgcgatgcaatttcctcattttat
taggaaaggacagtgggagtggcaccttccagggtcaaggaaggcacgggggaggggca
aacaacagatggctggcaactagaaggcacagtcgaggctgatcagcgggtttaaacgg
gccctctagactcgagttaaagtcgacgcggggaggcggcccaaagggagatccgactc
gtctgagggcgaaggcgaagacgcggaagaggccgcagagccggcagcaggccgcggga
aggaaggtccgctggattgagggccgaagggacgtagcagaaggacgtcccgcgcagaa
tccaggtggcaacacaggcgagcagccaaggaaaggacgatgatttccccgacaacacc
acggaattgtcagtgcccaacagccgagcccctgtccagcagcgggcaaggcaggcggc
gatgagttccgccgtggcaatagggaggggaaagcgaaagtcccggaaaggagctgac
aggtggtggcaatgccccaaccagtgggggttgcgtcagcaaacacagtgcacaccacg
ccacgttgcctgacaacgggccacaactcctcataaagagacagcaaccaggatttata
caaggaggagaaaatgaaagccatacgggaagcaatagcatgatacaaaggcattaaag
cagcgtatccacatagcgtaaaaggagcaacatagttaagaataccagtcaatctttca
caaattttgtaatccagaggttgattgtcgacttaacgcgttgaattcttacagattgg
ttgaaatcccattatagtaagattcggaccttgatgccctcttaatgttttaaagccg
tgcataaccttgagcaagcgttcaagaccaaaacctgcacctatccatggtttgtcaat
accccattctctatcaagagaaactggcccgacgactgccgaagaaagctccaggtccc
cgtgcattatatcaagagtatccccaaagaccatacaggaatctcctacgatttcgaag
tcgatttccagatagtccagaaactctttgatgagagcttcaagattttcccgagtaca
tcccgaacccatctgacagaagttcaccatagtaaattcttccaggtgctctttgccgt
cagactctttccggtaacaaggtccgacttcgaaaatttttattgggcctggtaaaatc
ctatcgagttttcgcagatagttgtaaagagtcggggcaagcattggcctcaagcagag
atttttatccacccggaagatctgttttgaaagttcagtatcattattaatacccattc

tctccacgtattccgccggaataaggataggagactttatctccagaaaaccccggtct
acgaaaaatttcgtaatatcacgttcgagtttaccgaggtagtcttctctatcattggt
atagagccgctgaaaatcgtttttttcttcttgtcacaagttcaggctcaagttccctga
aaggctttgccatatttagagaaattttatcctctggacttaagagagcctcaacccta
tcaagctggcttcttgtaagtgaaggagcaggagccgatgcgggaacagacgaatttgg
agttgattttgcaggcgaaggtacagatctggatgtgttcgttgatgcctttgcagaaa
cagaattttccagaggcttcggagcccttgaaactgatttcggcatagcttttttgacc
tttggagcagaaactaccctaactttcacactgttttgctttcggttgatcttgtgag
aaaattattgatatcctcgtccgaaaccctacatcgtttgcaggtttttctgtacttat
gatgtctgaatgctctggctgttctacaactcctggaattattcacaacaagatggtct
ccacacgccatttcaatgtatattttacttcttgagacctcatggtgcttgattttgtg
gagcgtgccagtcctggacatccagagcccggtcgcagatattaaaacatctaatggtt
ttttatccatggtggcgctagccagcttgggtctccctatagtgagtcgtattaatttc
gataagccagtaagcagtgggttctctagttagccagagagctctgcttatatagacct
cccaccgtacacgcctaccgcccatttgcgtcaatggggcggagttgttacgacatttt
ggaaagtcccgttgattttggtgccaaaacaaactcccattgacgtcaatggggtggag
acttggaaatccccgtgagtcaaaccgctatccacgcccattgatgtactgccaaaacc
gcatcaccatggtaatagcgatgactaatacgtagatgtactgccaagtaggaaagtcc
cataaggtcatgtactgggcataatgccaggcgggccatttaccgtcattgacgtcaat
aggggggcgtacttggcatatgatacacttgatgtactgccaagtgggcagtttaccgta
aatagtccacccattgacgtcaatggaaagtccctattggcgttactatgggaacatac
gtcattattgacgtcaatgggcgggggtcgttgggcggtcagccaggcgggccatttac
cgtaagttatgtaacgcggaactccatatatgggctatgaactaatgaccccgtaattg
attactattaataactagcaaaaaacgggcagcccgggaatctaacccggctgaacgga
tttagagtccgttcgatctacatgatcagccgcccggtgtttcgtcctttccacaagat
atataaagccaagaaatcgaaatactttcaagttacggtaagcatatgatagtccattt
taaaacataatttaaaactgcaaactacccaagaaattattactttctacgtcacgta
ttttgtactaatatctttgtgtttacagtcaaattaattctaattatctctctaacagc
cttgtatcgtatatgcaaatatgaaggaatcatgggaaataggccctcttcctgcccga
ggtacccctagcaaaaaacgggcagcccgggaatctaacccggctgaacggatttagag
tccgttcgatctacatgatcagccgcccggtgtttcgtcctttccacaagatatataaa
gccaagaaatcgaaatactttcaagttacggtaagcatatgatagtccattttaaaaca
taattttaaaactgcaaactacccaagaaattattactttctacgtcacgtattttgta
ctaatatctttgtgtttacagtcaaattaattctaattatctctctaacagccttgtat
cgtatatgcaaatatgaaggaatcatgggaaataggccctcttcctgcccgaggtaccc
ctagcaaaaaacgggcagcccgggaatctaacccggctgaacggatttagagtccgttc
gatctacatgatcagccgcccggtgtttcgtcctttccacaagatatataaagccaaga
aatcgaaatactttcaagttacggtaagcatatgatagtccattttaaaacataatttt
aaaactgcaaactacccaagaaattattactttctacgtcacgtattttgtactaatat
ctttgtgtttacagtcaaattaattctaattatctctctaacagccttgtatcgtatat
gcaaatatgaaggaatcatgggaaataggccctcttcctgcccgaggtaccctagcaa
aaaacgggcagcccgggaatctaacccggctgaacggatttagagtccgttcgatctac
atgatcagccgcccggtgtttcgtcctttccacaagatatataaagccaagaaatcgaa
atactttcaagttacggtaagcatatgatagtccattttaaaacataatttaaaactg
caaactacccaagaaattattactttctacgtcacgtattttgtactaatatctttgtg
tttacagtcaaattaattctaattatctctctaacagccttgtatcgtatatgcaaata
tgaaggaatcatgggaaataggccctcttcctgcccgaggtaccctagtcaataatca
atgtcaacgcgtatatctggcccgtacatcgcgaagcagcgcaaaacggatcctgcagg

tatttgcggccgcggtccgtatactccggaatattaatagatcatggagataattaaaa
tgataaccatctcgcaaataaataagtattttactgttttcgtaacagtttgtaataa
aaaaacctataaatattccggattattcataccgtcccaccatcgggcgcgaactccta
aaaaaccgccaccatgaagtgcctttgtacttagccttttattcattggggtgaatt
gcaagttcaccatagtttttccacacaaccaaaaaggaaactggaaaaatgttccttct
aattaccattattgcccgtcaagctcagatttaaattggcataatgacttaataggcac
agccttacaagtcaaaatgcccaagagtcacaaggctattcaagcagacggttggatgt
gtcatgcttccaaatgggtcactacttgtgatttccgctggtatggaccgaagtatata
acacattccatccgatccttcactccatctgtagaacaatgcaaggaaagcattgaaca
aacgaaacaaggaacttggctgaatccaggcttccctcctcaaagttgtggatatgcaa
ctgtgacggatgccgaagcagtgattgtccaggtgactcctcaccatgtgctggttgat
gaatacacaggagaatgggttgattcacagttcatcaacggaaaatgcagcaattacat
atgccccactgtccataactctacaacctggcattctgactataaggtcaaagggctat
gtgattctaacctcatttccatggacatcaccttcttctcagaggacggagagctatca
tccctgggaaaggagggcacagggttcagaagtaactactttgcttatgaaactggagg
caaggcctgcaaaatgcaatactgcaagcattggggagtcagactcccatcaggtgtct
ggttcgagatggctgataaggatctctttgctgcagccagattccctgaatgcccagaa
gggtcaagtatctctgctccatctcagacctcagtggatgtaagtctaattcaggacgt
tgagaggatcttggattattccctctgccaagaaacctggagcaaaatcagagcgggtc
ttccaatctctccagtggatctcagctatcttgctcctaaaaacccaggaaccggtcct
gctttcaccataatcaatggtaccctaaaatactttgagaccagatacatcagagtcga
tattgctgctccaatcctctcaagaatggtcggaatgatcagtggaactaccacagaaa
gggaactgtgggatgactgggcaccatatgaagacgtggaaattggacccaatggagtt
ctgaggaccagttcaggatataagtttcctttatacatgattggacatggtatgttgga
ctccgatcttcatcttagctcaaaggctcaggtgttcgaacatcctcacattcaagacg
ctgcttcgcaacttcctgatgatgagagtttatttttggtgatactgggctatccaaa
aatccaatcgagcttgtagaaggttggttcagtagttggaaaagctctattgcctctt
tttctttatcatagggttaatcattggactattcttggttctccgagttggtatccatc
tttgcattaaattaaagcacaccaagaaaagacagatttatacagacatagagatgaac
cgacttggaaagtgataagtcgagaagtactagaggatcataatcagccataccacatt
tgtagaggttttacttgctttaaaaaacctcccacacctccccctgaacctgaaacata
aaatgaatgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataa
agcaatagcatcacaaatttcacaaataaagcatttttttcactgcattctagttgtgg
tttgtccaaactcatcaatgtatcttatcatgtctggatctgatcactgcttgagccta
gttattaatagtaatcaattacggggtcattagttcatagcccatatatggagttccgc
gttacataacttacggtaaatggcccgcctggctgaccgcccaacgacccccgcccatt
gacgtcaataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtc
aatgggtggagtatttacggtaaactgcccacttggcagtacatcaagtgtatcatatg
ccaagtacgccccctattgacgtcaatgacggtaaatggcccgcctggcattatgccca
gtacatgaccttatgggactttcctacttggcagtacatctacgtattagtcatcgcta
ttaccatggtcgaggtgagccccacgttctgcttcactctccccatctccccccctcc
ccacccccaattttgtatttatttatttttttaattattttgtgcagcgatgggggcggg
ggggggggggcgcgcgccaggcggggcggggcggggcgaggggcggggcggggcgagg
cggagaggtgcggcggcagccaatcagagcggcgcgctccgaaagtttccttttatggc
gaggcggcggcggcggcggccctataaaaagcgaagcgcgcggcgggcgggagtcgctg
cgttgccttcgccccgtgccccgctccgcgccgcctcgcgccgcccgcccggctctga
ctgaccgcgttactcccacaggtgagcgggcgggacggcccttctcctccgggctgtaa
ttagcgcttggtttaatgacggctcgtttctttttctgtggctgcgtgaaagccttaaag

ggctccgggagggcccttgtgcgggggggagcggctcgggggtgcgtgcgtgtgtgt
gtgcgtgggagcgccgcgtgcggcccgcgctgcccggcggctgtgagcgctgcgggcg
cggcgcggggctttgtgcgctccgcgtgtgcgcgaggggagcgcggccgggggcggtgc
cccgcggtgcgggggggctgcgagggaacaaaggctgcgtgcggggtgtgtgcgtggg
ggggtgagcaggggtgtgggcgcggcggtcgggctgtaacccccccctgcacccccct
ccccgagttgctgagcacggcccggcttcgggtgcgggctccgtgcgggcgtggcgc
ggggctcgccgtgccgggcgggggggtggcggcaggtgggggtgccgggcgggcgggcc
cgcctcgggccgggggagggctcgggggagggggcgcggcggcccccggagcgccggcggct
gtcgaggcgcggcgagccgcagccattgccttttatggtaatcgtgcgagagggcgcag
ggacttcctttgtcccaaatctggcggagccgaaatctgggaggcgccgccgcacccccc
tctagcgggcgcgggcgaagcggtgcggcgccggcaggaaggaaatgggcggggagggc
cttcgtgcgtcgccgcgccgccgtcccctttctccatctccagcctcggggctgccgcag
ggggacggctgccttcgggggggacggggcagggcggggttcggcttctggcgtgtgac
cggcggctctagagcctctgctaaccatgttcatgccttcttctttttcctacagctcc
tgggcaacgtgctggttattgtgctgtctcatcattttggcaaagaattggccaaggag
gccaccatggtgagcaagggcgaggaggataacatggccatcatcaaggagttcatgcg
cttcaaggtgcacatggagggctccgtgaacggccacgagttcgagatcgagggcgagg
gcgagggccgcccctacgagggcacccagaccgccaagctgaaggtgaccaagggtggc
cccctgcccttcgcctgggacatcctgtcccctcagttcatgtacggctccaaggccta
cgtgaagcaccccgccgacatccccgactacttgaagctgtccttccccgagggcttca
agtgggagcgcgtgatgaacttcgaggacggcggcgtggtgaccgtgacccaggactcc
tccctgcaggacggcgagttcatctacaaggtgaagctgcgcggcaccaacttcccctc
cgacggccccgtaatgcagaagaagaccatgggctgggaggcctcctccgagcggatgt
accccgaggacggcgccctgaagggcgagatcaagcagaggctgaagctgaaggacggc
ggccactacgacgctgaggtcaagaccacctacaaggccaagaagcccgtgcagctgcc
cggcgcctacaacgtcaacatcaagttggacatcacctcccacaacgaggactacacca
tcgtggaacagtacgaacgcgccgagggccgccactccaccggcggcatggacgagctg
tacaagggcggctcttagggcggctccctcgtgagcaagggcgaggagctgttcaccgg
ggtggtgcccatcctggtcgagctggacggcgacgtaaacggccacaagttcagcgtgt
ccggcgagggcgagggcgatgccacctacggcaagctgaccctgaagttcatctgcacc
accggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcgtgca
gtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaagtccgccatgc
ccgaaggctacgtccaggagcgcaccatcttcttcaaggacgacggcaactacaagacc
cgcgccgaggtgaagttcgagggcgacaccctggtgaaccgcatcgagctgaagggcat
cgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagcc
acaacgtctatatcatggccgacaagcagaagaacggcatcaaggtgaacttcaagatc
cgccacaacatcgaggacggcagcgtgcagctcgccgaccactaccagcagaacacccc
catcggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtccgccc
tgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgcc
gccgggatcactctcggcatggacgagctgtacaagggccccttcgaacaaaaactcat
ctcagaagaggatctgaatatgcataccggtcatcatcaccatcaccattaatgagaat
tcaacgcgttaagtcgacaatcaacctctggattacaaaatttgtgaaagattgactgg
tattcttaactatgttgctccttttacgctatgtggatacgctgctttaatgcctttgt
atcatgctattgcttcccgtatggctttcattttctcctccttgtataaatcctggttg
ctgtctctttatgaggagttgtggcccgttgtcaggcaacgtggcgtggtgtgcactgt
gtttgctgacgcaacccccactggttggggcattgccaccacctgtcagctcctttccg
ggactttcgctttccccctccctattgccacggcggaactcatcgccgcctgccttgcc
cgctgctggacaggggctcggctgttgggcactgacaattccgtggtgttgtcggggaa

atcatcgtcctttccttggctgctcgcctgtgttgccacctggattctgcgcgggacgt
ccttctgctacgtcccttcggccctcaatccagcggaccttccttcccgcggcctgctg
ccggctctgcggcctcttccgcgtcttcgccttcgccctcagacgagtcggatctccct
ttgggccgcctccccgcggcctggccggccgccagcacagtggtcgatcgaccaatgcc
ctggctcacaaataccactgagatcttttcccctctgccaaaaattatggggacatcat
gaagccccttgagcatctgacttctggctaataaaggaaatttattttcattgcaatag
tgtgttggaatttttgtgtctctcactcggaaggacatatgggagggcaaatcattta
aaacatcagaatgagtatttggtttagagtttggcaacatatgcccatatgctggctgc
catgaacaaaggttggctataaagaggtcatcagtatatgaaacagcccctgctgtcc
attccttattccatagaaaagccttgacttgaggttagatttttttatattttgtttt
gtgttatttttttctttaacatccctaaaattttccttacatgttttactagccagatt
tttcctcctctcctgactactcccagtcatagctgtccctcttctcttatggagatccc
tcgacctgccctaggggtacctcgggcaggaagagggcctatttcccatgattccttca
tatttgcatatacgatacaaggctgttagagagataattagaattaatttgactgtaaa
cacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttg
cagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtat
ttcgatttcttggctttatatatcttgtggaaaggacgaaacaccgggcggctgatcat
gtagatcgaacggactctaaatccgttcagccgggttagattcccgggctgcccgtttt
ttgctaggggtacctcgggcaggaagagggcctatttcccatgattccttcatatttgc
atatacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaag
atattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagtttt
aaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatt
tcttggctttatatatcttgtggaaaggacgaaacaccgggcggctgatcatgtagatc
gaacggactctaaatccgttcagccgggttagattcccgggctgcccgtttttgctag
gggtacctcgggcaggaagagggcctatttcccatgattccttcatatttgcatatacg
atacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatattag
tacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaatta
tgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggc
tttatatatcttgtggaaaggacgaaacaccgggcggctgatcatgtagatcgaacgga
ctctaaatccgttcagccgggttagattcccgggctgcccgtttttgctaggggtacc
tcgggcaggaagagggcctatttcccatgattccttcatatttgcatatacgatacaag
gctgttagagagataattagaattaatttgactgtaaacacaaagatattagtacaaaa
tacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgtttta
aaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatat
atcttgtggaaaggacgaaacaccgggcggctgatcatgtagatcgaacggactctaaa
tccgttcagccgggttagattcccgggctgcccgtttttgctaggagatccgaaccag
ataagtgaaatctagttccaaactattttgtcatttttaattttcgtattagcttacga
cgctacacccagttcccatctattttgtcactcttccctaaataatccttaaaaactcc
atttccacccctcccagttcccaactattttgtccgcccacagcggggcatttttcttc
ctgttatgttttaatcaaacatcctgccaactccatgtgacaaaccgtcatcttcggc
tacttttttctctgtcacagaatgaaaatttttctgtcatctcttcgttattaatgtttg
taattgactgaatatcaacgcttatttgcagc

The sequence for the wild type PytR is:
ggaaacctgatcatgtagatcgaacggactctaaatccgttcagccgggttagattccc
ggggtttccg

The sequence for the NES-MbPylRS is:

atggcctgcccagtgcccttcagttgcctccgctggagagactgaccctcgacgataa
aaaaccattagatgttttaatatctgcgaccgggctctggatgtccaggactggcacgc
tccacaaaatcaagcaccatgaggtctcaagaagtaaaatatacattgaaatggcgtgt
ggagaccatcttgttgtgaataattccaggagttgtagaacagccagagcattcagaca
tcataagtacagaaaaacctgcaaacgatgtagggtttcggacgaggatatcaataatt
ttctcacaagatcaaccgaaagcaaaaacagtgtgaaagttagggtagtttctgctcca
aaggtcaaaaaagctatgccgaaatcagtttcaagggctccgaagcctctggaaaattc
tgtttctgcaaaggcatcaacgaacacatccagatctgtaccttcgcctgcaaaatcaa
ctccaaattcgtctgttcccgcatcggctcctgctccttcacttacaagaagccagctt
gatagggttgaggctctcttaagtccagaggataaaatttctctaaatatggcaaagcc
tttcagggaacttgagcctgaacttgtgacaagaagaaaaaacgattttcagcggctct
ataccaatgatagagaagactacctcggtaaactcgaacgtgatattacgaaatttttc
gtagaccggggttttctggagataaagtctcctatccttattccggcggaatacgtgga
gagaatgggtattaataatgatactgaactttcaaaacagatcttccgggtggataaaa
atctctgcttgaggccaatgcttgccccgactctttacaactatctgcgaaaactcgat
aggattttaccaggcccaataaaaattttcgaagtcggaccttgttaccggaaagagtc
tgacggcaaagagcacctggaagaatttactatggtgaacttctgtcagatgggttcgg
gatgtactcgggaaaatcttgaagctctcatcaaagagtttctggactatctggaaatc
gacttcgaaatcgtaggagattcctgtatggtctttggggatactcttgatataatgca
cggggacctggagctttcttcggcagtcgtcgggccagtttctcttgatagagaatggg
gtattgacaaaccatggataggtgcaggttttggtcttgaacgcttgctcaaggttatg
cacggctttaaaaacattaagagggcatcaaggtccgaatcttactataatgggatttc
aaccaatctgtaa

The sequence for NBK-RS is:
atggataaaaaaccattagatgttttaatatctgcgaccgggctctggatgtccaggac
tggcacgctccacaaaatcaagcaccatgaggtctcaagaagtaaaatatacattgaaa
tggcgtgtggagaccatcttgttgtgaataattccaggagttgtagaacagccagagca
ttcagacatcataagtacagaaaaacctgcaaacgatgtagggtttcggacgaggatat
caataatttctcacaagatcaaccgaaagcaaaaacagtgtgaaagttagggtagttt
ctgctccaaaggtcaaaaaagctatgccgaaatcagtttcaagggctccgaagcctctg
gaaaattctgtttctgcaaaggcatcaacgaacacatccagatctgtaccttcgcctgc
aaaatcaactccaaattcgtctgttcccgcatcggctcctgctccttcacttacaagaa
gccagcttgatagggttgaggctctcttaagtccagaggataaaatttctctaaatatg
gcaaagcctttcagggaacttgagcctgaacttgtgacaagaagaaaaaacgatttca
gcggctctataccaatgatagagaagactacctcggtaaactcgaacgtgatattacga
aatttttcgtagaccggggttttctggagataaagtctcctatccttattccggcggaa
tacgtggagagatttggtattaataatgatactgaactttcaaaacagatcttccgggt
ggataaaaatctctgcttgaggccaatgcttagcccgactctttgcaactatatgcgaa
aactcgataggattttaccaggcccaataaaaattttcgaagtcggaccttgttaccgg
aaagagtctgacggcaaagagcacctggaagaatttactatggtgaacttctgtcagat
gggttcgggatgtactcgggaaaatcttgaagctctcatcaaagagtttctggactatc
tggaaatcgacttcgaaatcgtaggagattcctgtatggtctttggggatactcttgat
ataatgcacggggacctggagctttcttcggcagtcgtcgggccagtttctcttgatag
agaatggggtattgacaaaccatggataggtgcaggttttggtcttgaacgcttgctca
aggttatgcacggctttaaaaacattaagagggcatcaaggtccgaatcttactataat
gggatttcaaccaatctgtaa

The sequence for CouK-RS is:
atggataaaaaaccattagatgttttaatatctgcgaccgggctctggatgtccaggac
tggcacgctccacaaaatcaagcaccatgaggtctcaagaagtaaaatatacattgaaa
tggcgtgtggagaccatcttgttgtgaataattccaggagttgtagaacagccagagca
ttcagacatcataagtacagaaaaacctgcaaacgatgtagggtttcggacgaggatat
caataattttctcacaagatcaaccgaaagcaaaaacagtgtgaaagttagggtagttt
ctgctccaaaggtcaaaaaagctatgccgaaatcagtttcaagggctccgaagcctctg
gaaaattctgtttctgcaaaggcatcaacgaacacatccagatctgtaccttcgcctgc
aaaatcaactccaaattcgtctgttcccgcatcggctcctgctccttcacttacaagaa
gccagcttgatagggttgaggctctcttaagtccagaggataaaatttctctaaatatg
gcaaagcctttcagggaacttgagcctgaacttgtgacaagaagaaaaaacgattttca
gcggctctataccaatgatagagaagactacctcggtaaactcgaacgtgatattacga
aatttttcgtagaccggggttttctggagataaagtctcctatccttattccggcggaa
tacgtggagagaatgggtattaataatgatactgaactttcaaaacagatcttccgggt
ggataaaaatctctgcttgaggccaatgcttgccccgactcttgccaactatatgcgaa
aactcgataggattttaccaggcccaataaaaattttcgaagtcggaccttgttaccgg
aaagagtctgacggcaaagagcacctggaagaatttactatggtgaacttctgtcagat
gggttcgggatgtactcgggaaaatcttgaagctctcatcaaagagtttctggactatc
tggaaatcgacttcgaaatcgtaggagattcctgtatggtctttggggatactcttgat
ataatgcacggggacctggagctttcttcggcagtcgtcgggccagtttctcttgatag
agaatggggtattgacaaaccatggataggtgcaggttttggtcttgaacgcttgctca
aggttatgcacggctttaaaaacattaagagggcatcaaggtccgaatcttactataat
gggatttcaaccaatctgtaa

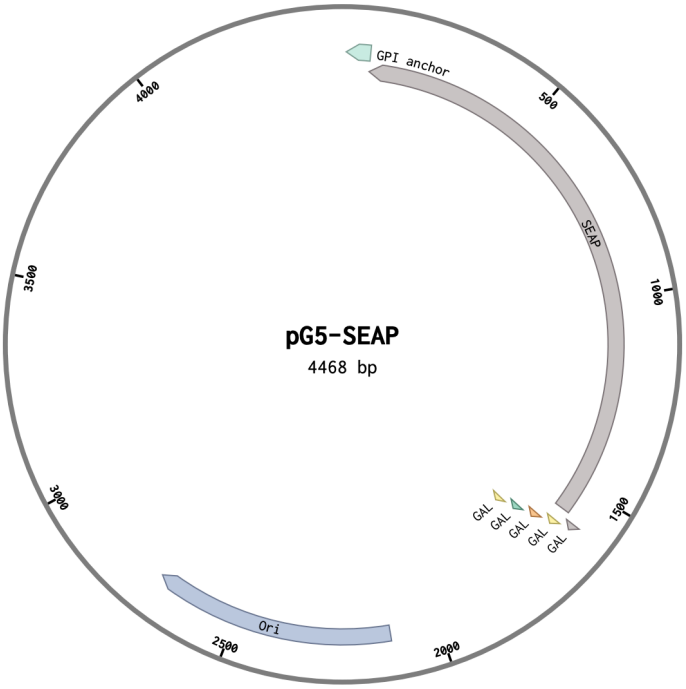
The sequence for AcK-RS is:
atggataaaaaaccattagatgttttaatatctgcgaccgggctctggatgtccaggac
tggcacgctccacaaaatcaagcaccatgaggtctcaagaagtaaaatatacattgaaa
tggcgtgtggagaccatcttgttgtgaataattccaggagttgtagaacagccagagca
ttcagacatcataagtacagaaaaacctgcaaacgatgtagggtttcgggtgaggatat
caataattttctcacaagatcaaccgaaagcaaaaacagtgtgaaagttagggtagttt
ctgctccaaaggtcaaaaaagctatgccgaaatcagtttcaagggctccgaagcctctg
gaaaattctgtttctgcaaaggcatcaacgaacacatccagatctgtaccttcgcctgc
aaaatcaactccaaattcgtctgttcccgcatcggctcctgctccttcacttacaagaa
gccagcttgatagggttgaggctctcttaagtccagaggataaaatttctctaaatatg
gcaaagcctttcagggaacttgagcctgaacttgtgacaagaagaaaaaacgattttca
gcggctctataccaatgatagagaagactacctcggtaaactcgaacgtgatattacga
aatttttcgtagaccggggttttctggagataaagtctcctatccttattccggcggaa
tacgtggagagaatgggtattaataatgatactgaactttcaaaacagatcttccgggt
ggataaaaatctctgcttgaggccaatggtggccccgactatttttaactatgcgcgaa
aactcgataggattttaccaggcccaataaaaattttcgaagtcggaccttgttaccgg
aaagagtctgacggcaaagagcacctggaagaatttactatggtgaacttctttcagat
gggttcgggatgtactcgggaaaatcttgaagctctcatcaaagagtttctggactatc
tggaaatcgacttcgaaatcgtaggagattcctgtatggtctttggggatactcttgat
ataatgcacggggacctggagctttcttcggcagtcgtcgggccagtttctcttgatag
agaatggggtattgacaaaccatggataggtgcaggttttggtcttgaacgcttgctca
aggttatgcacggctttaaaaacattaagagggcatcaaggtccgaatcttactataat
gggatttcaaccaatctgtaa

The sequence for BiotinK-RS is:
atggataaaaaaccattagatgttttaatatctgcgaccgggctctggatgtccaggac
tggcacgctccacaaaatcaagcaccatgaggtctcaagaagtaaaatatacattgaaa
tggcgtgtggagaccatcttgttgtgaataattccaggagttgtagaacagccagagca
ttcagacatcataagtacagaaaaacctgcaaacgatgtagggtttcggacgaggatat
caataattttctcacaagatcaaccgaaagcaaaaacagtgtgaaagttagggtagttt
ctgctccaaaggtcaaaaaagctatgccgaaatcagtttcaagggctccgaagcctctg
gaaaattctgtttctgcaaaggcatcaacgaacacatccagatctgtaccttcgcctgc
aaaatcaactccaaattcgtctgttcccgcatcggctcctgctccttcacttacaagaa
gccagcttgatagggttgaggctctcttaagtccagaggataaaatttctctaaatatg
gcaaagcctttcagggaacttgagcctgaacttgtgacaagaagaaaaaacgattttca
gcggctctataccaatgatagagaagactacctcggtaaactcgaacgtgatattacga
aatttttcgtagaccggggttttctggagataaagtctcctatccttattccggcggaa
tacgtggagagaatgggtattaataatgatactgaactttcaaaacagatcttccgggt
ggataaaaatctctgcttgaggccaatgcttgccccgactcttgcgaactatgtgcgaa
aactcgataggattttaccaggcccaataaaaattttcgaagtcggaccttgttaccgg
aaagagtctgacggcaaagagcacctggaagaatttactatggtgaacttcgtgcagta
tggttcgggatgtactcgggaaaatcttgaagctctcatcaaagagtttctggactatc
tggaaatcgacttcgaaatcgtaggagattcctgtatggtctttggggatactcttgat
ataatgcacggggacctggagctttcttcggcagtcgtcgggccacgttctcttgatag
agaatggggtattgacaaaccatggataggtgcaggttttggtcttgaacgcttgctca
aggttatgcacggctttaaaaacattaagagggcatcaaggtccgaatcttactataat
gggatttcaaccaatctgtaa

The sequence for DiazK-RS is:
atggataaaaaaccattagatgttttaatatctgcgaccgggctctggatgtccaggac
tggcacgctccacaaaatcaagcaccatgaggtctcaagaagtaaaatatacattgaaa
tggcgtgtggagaccatcttgttgtgaataattccaggagttgtagaacagccagagca
ttcagacatcataagtacagaaaaacctgcaaacgatgtagggtttcggacgaggatat
caataattttctcacaagatcaaccgaaagcaaaaacagtgtgaaagttagggtagttt
ctgctccaaaggtcaaaaaagctatgccgaaatcagtttcaagggctccgaagcctctg
gaaaattctgtttctgcaaaggcatcaacgaacacatccagatctgtaccttcgcctgc
aaaatcaactccaaattcgtctgttcccgcatcggctcctgctccttcacttacaagaa
gccagcttgatagggttgaggctctcttaagtccagaggataaaatttctctaaatatg
gcaaagcctttcagggaacttgagcctgaacttgtgacaagaagaaaaaacgattttca
gcggctctataccaatgatagagaagactacctcggtaaactcgaacgtgatattacga
aatttttcgtagaccggggttttctggagataaagtctcctatccttattccggcggaa
tacgtggagagaatgggtattaataatgatactgaactttcaaaacagatcttccgggt
ggataaaaatctctgcttgaggccaatgcttgccccgactctgtataactatatgcgaa
aactcgataggattttaccaggcccaataaaaattttcgaagtcggaccttgttaccgg
aaagagtctgacggcaaagagcacctggaagaatttactatggtgaacttcgctcagat
gggttcgggatgtactcgggaaaatcttgaagctctcatcaaagagtttctggactatc
tggaaatcgacttcgaaatcgtaggagattcctgtatggtctttggggatactcttgat
ataatgcacggggacctggagctttcttcggcagtcgtcgggccagtttctcttgatag
agaatggggtattgacaaaccatggataggtgcaggttttggtcttgaacgcttgctca
aggttatgcacggctttaaaaacattaagagggcatcaaggtccgaatcttactataat
gggatttcaaccaatctataa

Chapter 5 Plasmids

pG5-SEAP



Sequence color-coding key

Feature	Color
GPI anchor	text
SEAP	text
GAL	text
GAL	text
Ori	text

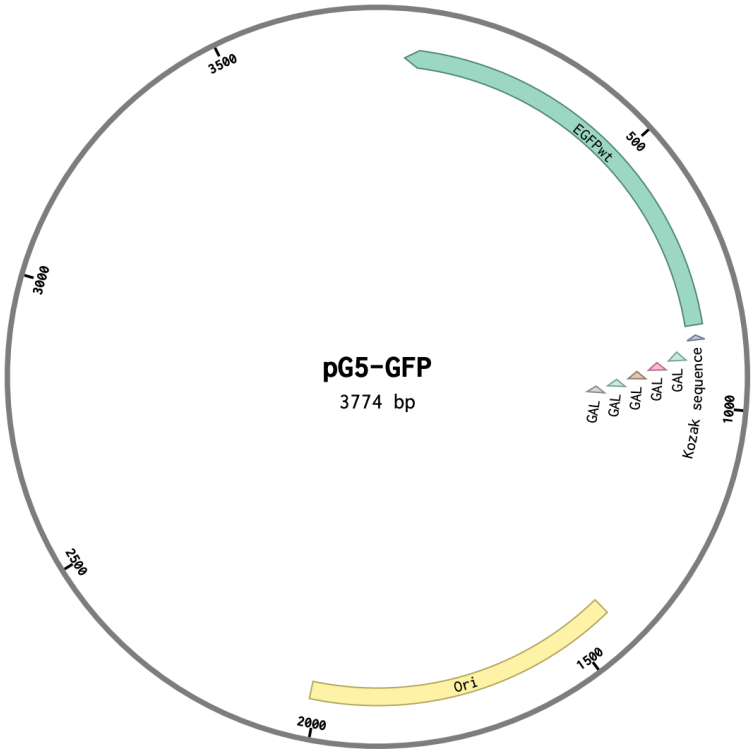
Sequence

caatgtatcttatcatgtctgctcgaagcggccggccgccccgactctagagtaacccg
ggtgcgcggcgtcggtggtgccggcgggggcgccaggtcgcaggcggtgtagggctcc
aggcaggcggcgaaggccatgacgtgcgctatgaaggtctgctcctgcacgccgtgaac
caggtgcgcctgcgggccgcgcgcgaacaccgccacgtcctcgcctgcgtgggtctctt
cgtccaggggcactgctgactgctgccgatactcggggctcccgctctcgctctcggta
acatccggccgggcgccgtccttgagcacatagcctggaccgtttccgtataggaggac
cgtgtaggccttcctgtcccgggccttgccaggggccagcccgaagatggagctcctc
gcaggggtagcctccgaaggagaagacgtgggagtggtcggcagtgacgaggctcagc
gtgtcctcctcgctggtgagctggcccgccctctcaatggcgtcgtcgaacatgatcgt
ctcagtcagtgcccggtaagccctgctttcatgatgaccatggtcgatgcgaccaccct

ccacgaagaggaagaagccgcgggggttcctgctcagcaggcgcagggcagcctctgtc
atctccatcagggaggggtccagtgtggagtctcggtggatctcgtatttcatgtctcc
aggctcaaagagacccatgagatgggtcacagacgggtccagggaagcctgcatgagct
cagtgcggttccacacataccgggcaccctggcgcttcgccagccattcctgcaccaga
ttcttcccgtccagcctggtcccaccttggctgtagtcatctgggtactcagggtctgg
ggttcccatgcgaaacatgtactttcggcctccacctaggatcacgtcaatgtccatgt
tggagatgagctgcgtagcgatgtcctggcacccctcctggcgggccgaggcaggcacg
tcggcgtccgagtaccagttgcggttcaccgtgtgggcgtaggtgccggctggcgaggc
gtgctgcactcgtgtggtggttaccactcccactgacttccctgctttcttggcccgat
tcatcacggagatgacctcgttgccgcgtgtcgtgttgcactggttaaagcgggcggct
gcactcaagccaatggtctggaagttgcccttgaccccgcacaggtaggccgtggctgt
ggctccactgtctggcacatgtttgtctacattgtatgtcttggacagagccacatatg
ggaagcggtccatggccaggggtatctcaggccccagtttgtccttcttctgccctttt
aggatcctggcagctgtcaccgtagacacccccatcccatcgcccaggaagatgatgag
gttcttggcggctgtctgtgcaggctgcagcttcttggcggcacccagggcctcggctg
cctcgcggttccagaagtccgggttctcctcctcaactgggatgatgcccagggagagc
tgtagcctcaggcccagcagcagcagcagcagcagcatgctagagtcgacggatcggag
tactgtcctccgagcggagtactgtcctcgagcggagtactgtcctccgagcggagtac
tgtcctccgagcggagtactgtcctccgagcggatccccgggaattagcttggcgtaat
catggtcatagctgtttcctgtgtgaaattgttatccgctcacaattccacacaacata
cgagccggaagcataaagtgtaaagcctggggtgcctaatgagtgagctaactcacatt
aattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcatt
aatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcc
tcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactc
aaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgag
caaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccat
aggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaa
cccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctc
ctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtg
gcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaa
gctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaact
atcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggt
aacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcc
taactacggctacactagaagaacagtatttggtatctgcgctctgctgaagccagtta
ccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggt
ggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcc
tttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattt
tggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagt
tttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaat
cagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactcc
ccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatg
ataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccgg
aagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaatt
gttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgcc
attgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccgg
ttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagct
ccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggtt
atggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgac

tggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctctt
gcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatc
attggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccag
ttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcg
tttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgaca
cggaaatgttgaatactcatactcttccttttcaatattattgaagcatttatcaggg
ttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagggg
ttccgcgcacatttccccgaaaagtgccacctgacgtctaagaaaccattattatcatg
acattaacctataaaaataggcgtatcacgaggccctttcgtctcgcgcgtttcggtga
tgacggtgaaaacctctgacacatgcagctcccggagacggtcacagcttgtctgtaag
cggatgccgggagcagacaagcccgtcagggcgcgtcagcgggtgttggcgggtgtcgg
ggctggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatgcggtg
tgaaataccgcacagatgcgtaaggagaaaataccgcatcaggcgccattcgccattca
ggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgctattacgccagctg
gcgaaaggggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtc
acgacgttgtaaaacgacggccagtgaattcgagctcggtaccttgttgttaacttgtt
tattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaag
catttttttcactgcattctagttgtggtttgtccaaactcat

pG5-GFP



Sequence color-coding key

Feature	Color
eGFP	text
Kozak sequence	text
GAL	text
GAL	text
Ori	text

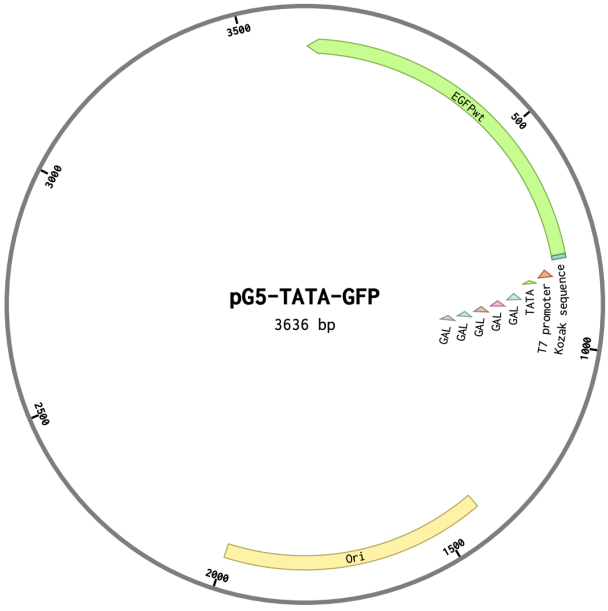
Sequence

caatgtatcttatcatgtctgctcgaagcggccggccgccccgactctagatcattaat
ggtgatggtgatgatgaccggtatgcatattcagatcctcttctgagatgagttttgt
tcgaagggccccttgtacagctcgtccatgccgagagtgatcccggcggcggtcacgaa
ctccagcaggaccatgtgatcgcgcttctcgttggggtctttgctcagggcggactggg
tgctcaggtagtggttgtcgggcagcagcacggggccgtcgccgatgggggtgttctgc
tggtagtggtcggcgagctgcacgctgccgtcctcgatgttgtggcggatcttgaagtt
caccttgatgccgttcttctgcttgtcggccatgatatagacgttgtggctgttgtagt
tgtactccagcttgtgccccaggatgttgccgtcctccttgaagtcgatgcccttcagc
tcgatgcggttcaccagggtgtcgccctcgaacttcacctcggcgcgggtcttgtagtt
gccgtcgtccttgaagaagatggtgcgctcctggacgtagccttcgggcatggcggact
tgaagaagtcgtgctgcttcatgtggtcggggtagcggctgaagcactgcacgccgtag
gtcagggtggtcacgagggtgggccagggcacgggcagcttgccggtggtgcagatgaa

cttcagggtcagcttgccgtaggtggcatcgccctcgccctcgccggacacgctgaact
tgtggccgtttacgtcgccgtccagctcgaccaggatgggcaccaccccggtgaacagc
tcctcgcccttgctcaccatggtggcggcgctagccagcggtggcggcgcatgctagag
tcgacggatcggagtactgtcctccgagcggagtactgtcctcgagcggagtactgtcc
tccgagcggagtactgtcctccgagcggagtactgtcctccgagcggatccccgggaat
tagcttggcgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaa
ttccacacaacatacgagccggaagcataaagtgtaaagcctggggtgcctaatgagtg
agctaactcacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtc
gtgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggc
gctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcg
gtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcagg
aaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgc
tggcgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagt
cagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctggaagctc
cctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcc
cttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtag
gtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgc
cttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactgg
cagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttc
ttgaagtggtggcctaactacggctacactagaagaacagtatttggtatctgcgctct
gctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaacca
ccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaagga
tctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactc
acgttaagggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaa
attaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagt
taccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccat
agttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggcc
ccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaata
aaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccat
ccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgc
gcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggct
tcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaa
aaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgt
tatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaaga
tgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcg
accgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactt
taaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccg
ctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttt
tactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagg
gaataagggcgacacggaaatgttgaatactcatactcttccttttcaatattattga
agcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaa
taaacaaataggggttccgcgcacatttccccgaaaagtgccacctgacgtctaagaaa
ccattattatcatgacattaacctataaaaataggcgtatcacgaggccctttcgtctc
gcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggagacggtcac
agcttgtctgtaagcggatgccgggagcagacaagcccgtcagggcgcgtcagcgggtg
ttggcgggtgtcggggctggcttaactatgcggcatcagagcagattgtactgagagtg
caccatatgcggtgtgaaataccgcacagatgcgtaaggagaaaataccgcatcaggcg
ccattcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgc

tattacgccagctggcgaaaggggatgtgctgcaaggcgattaagttgggtaacgcca
gggttttcccagtcacgacgttgtaaaacgacggccagtgaattcgagctcggtacctt
gttgttaacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaa
tttcacaaataaagcatttttttcactgcattctagttgtggtttgtccaaactcat

pG5-TATA-GFP



Sequence color-coding key

Feature	Color
eGFP	text
Kozak sequence	text
T7 promoter	text
TATA box	text
GAL	text
Ori	text

Sequence

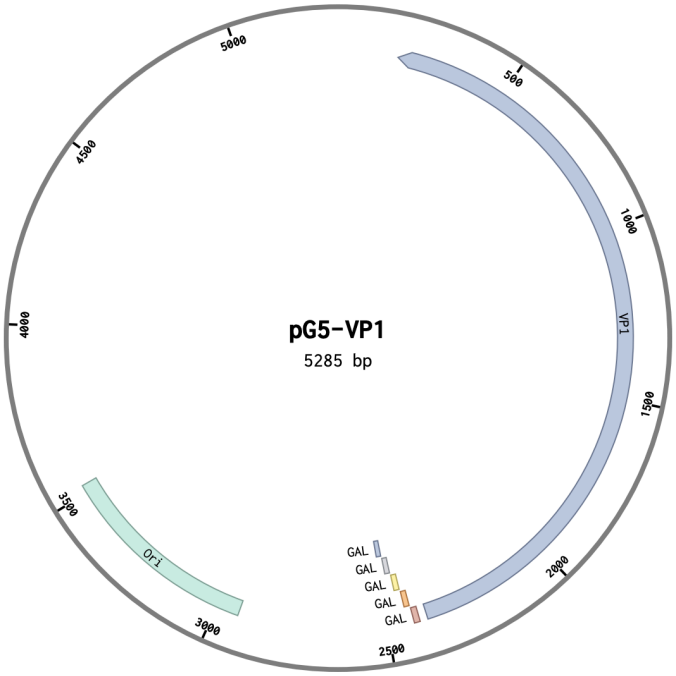
gtacctcaatggtgatggtgatgatgaccggtatgcatattcagatcctcttctgagat
gagtttttgttcgaagggccccttgtacagctcgtccatgccgagagtgatcccggcgg
cggtcacgaactccagcaggaccatgtgatcgcgcttctcgttggggtctttgctcagg
gcggactgggtgctcaggtagtggttgtcgggcagcagcacggggccgtcgccgatggg
ggtgttctgctggtagtggtcggcgagctgcacgctgccgtcctcgatgttgtggcgga
tcttgaagttcaccttgatgccgttcttctgcttgtcggccatgatatagacgttgtgg
ctgttgtagttgtactccagcttgtgccccaggatgttgccgtcctccttgaagtcgat
gcccttcagctcgatgcggttcaccagggtgtcgccctcgaacttcacctcggcgcggg
tcttgtagttgccgtcgtccttgaagaagatggtgcgctcctggacgtagccttcgggc
atggcggacttgaagaagtcgtgctgcttcatgtggtcggggtagcggctgaagcactg
cacgccgtaggtcagggtggtcacgagggtgggccagggcacgggcagcttgccggtgg
tgcagatgaacttcagggtcagcttgccgtaggtggcatcgccctcgccctcgccggac
acgctgaacttgtggccgtttacgtcgccgtccagctcgaccaggatgggcaccacccc
ggtgaacagctcctcgcccttgctcaccatggtggcggcgctagccagcttgggtctcc
ctatagtgagtcgtattaatttctttataaatgcatgctagagtcgacggatcggagt

actgtcctccgagcggagtactgtcctcgagcggagtactgtcctccgagcggagtact
gtcctccgagcggagtactgtcctccgagcggatccccgggaattagcttggcgtaatc
atggtcatagctgtttcctgtgtgaaattgttatccgctcacaattccacacaacatac
gagccggaagcataaagtgtaaagcctggggtgcctaatgagtgagctaactcacatta
attgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcatta
atgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcct
cgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactca
aaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagc
aaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccata
ggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaac
ccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcc
tgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtgg
cgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaag
ctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaacta
tcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggta
acaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcct
aactacggctacactagaagaacagtatttggtatctgcgctctgctgaagccagttac
cttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtg
gtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcct
ttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggatttt
ggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagtt
ttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatc
agtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccc
cgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatga
taccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccgga
agggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattg
ttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgcca
ttgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggt
tcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctc
cttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggtta
tggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgact
ggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttg
cccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatca
ttggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagt
tcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgt
ttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacac
ggaaatgttgaatactcatactcttccttttcaatattattgaagcatttatcagggt
tattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagggat
tccgcgcacatttccccgaaaagtgccacctgacgtctaagaaaccattattatcatga
cattaacctataaaaataggcgtatcacgaggccctttcgtctcgcgcgtttcggtgat
gacggtgaaaacctctgacacatgcagctcccggagacggtcacagcttgtctgtaagc
ggatgccgggagcagacaagcccgtcagggcgcgtcagcgggtgttggcgggtgtcggg
gctggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatgcggtgt
gaaataccgcacagatgcgtaaggagaaaataccgcatcaggcgccattcgccattcag
gctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgctattacgccagctgg
cgaaaggggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtca
cgacgttgtaaaacgacggccagtgaattcgagctcg

The sequence for the YBTATA minimal promoter is:
tctagagggtatataatgggggcca

The sequence for the minCMV minimal promoter is:
Gtaggcgtgtacggtgggaggtctatataagcagagctcgtttagtgaaccgtcagatc

pG5-VP1



Sequence color-coding key

Feature	Color
VP1	text
GAL	text
Ori	text

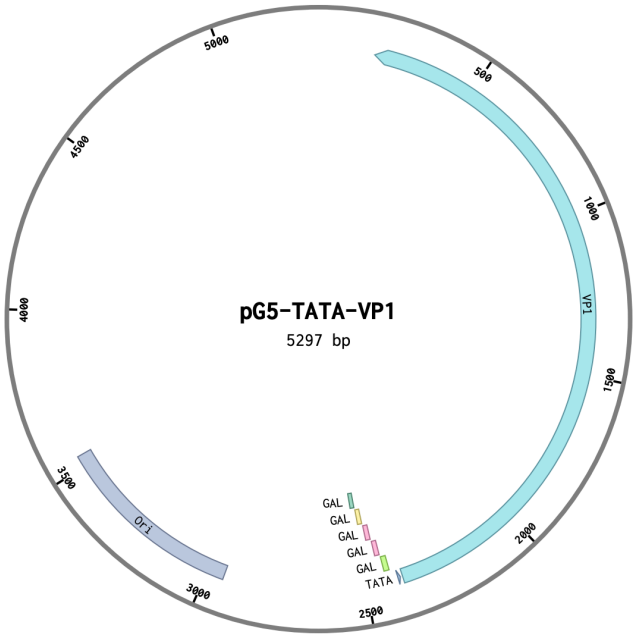
Sequence

caatgtatcttatcatgtctgctcgaagcggccggccgccccgactctagagctgtagt
taatgattaacccgccatgctacttatctacgtagccatggaaactagataagaaagaa
atacgcagagaccaaagttcaactgaaacgaattaaacggtttattgattaacaagcaa
ttacagattacgagtcaggtatctggtgccaatggggcgaggctctgaatacacgccat
tagtgtccacagtaaagtccacattaacagacttgttgtagttggaagtgtactgaatt
tcgggattccagcgtttgctgttttccttctgcagctcccactcgatctccacgctgac
ctgtcccgtggagtactgtgtgatgaaggaagcaaactttgccgcactgaaggtggtcg
aaggattcgcaggtaccggggtgttcttgatgagaatctgtggaggagggtgtttaagt
ccgaatccacccatgaggggagaggggtgaaaatgtccgtccgtgtgtggaatctttgc
ccagatgggcccctgaaggtacacatctctgtcctgccagaccatgcctggaagaacgc
cttgtgtgttgacatctgcggtagctgcttgtctgttgcctctctggaggttggtagat
acagaaccatactgctccgtagccacgggattggttgtcctgatttcctcttcgtctgt
aatcatgacctttttcaatgtccacatttgttttctctgagccttgcttcccaaagatga
gaaccccgctctgaggaaaaaactttcttcatcgtccttgtggcttgccatggccggg
cccggattcaccagagagtctctgccattgaggtggtacttggtagctccagtccacga
gtattcactgttgttgttatccgcagatgtctttgatactcgctgctggcggtaacagg
gtccaggaagccagttcctagactggtcccgaatgtcactcgctccggcctgagaaaac

tgaagccttgactgcgtggtggttccacttggagtgtttgttctgctcaagtaatacag
gtactggtcgatgagaggattcatgagacggtccagactctggctgtgagcgtagctgc
tgtggaaaggaacgtcctcaaaagtgtagctgaaggtaaagttgtttccggtacgcagc
atctgagaaggaaagtactccaggcagtaaaatgaagagcgtcctactgcctgactccc
gttgttcagggtgaggtatccatactgtggcaccatgaagacgtctgctgggaacggcg
ggaggcatccttgatgcgccgagccgaggacgtacgggagctggtactccgagtcagta
aacacctgaaccgtgctggtaaggttattggcaatcgtcgtcgtaccgtcattctgcgt
gacctctttgacttgaatgttaaagagcttgaagttgagtctcttgggtcggaatcccc
agttgttgttgatgagtctttgccagtcacgtggtgaaaagtggcagtggaatctgttg
aagtcaaaataccccaaggggtgctgtagccaaagtagtgattgtcgttcgaggctcc
tgattggctggaaatttgtttgtagaggtggttgttgtaggtgggcagggcccaggttc
gggtgctggtggtgatgactctgtcgccgagccatgtggaatcgcaatgccaatttccc
gaggaattacccactccgtcggcgccctcgttattgtctgcgagtggtgcgccactgcc
tgtagcgagggtattagttccgagaccagagggggctgctggtggctgtccgagaggct
gggggtcaggtactgagtctgcgtctccagtctgaccaaaattcaatctttttcttgca
ggctgctggcccgcctttccggttcccgaggaggagtctggctccacaggagagtgctc
taccggcctctttttcccggagcggtcttaacaggttcctcaaccaggcccagaggtt
caagaaccctctttttcgcctggaagactgctcgtccgaggttgcccccaaaagacgta
tcttctttaaggcgctcctgaaactccgcgtcggcgtggttgtacttgaggtacgggtt
gtctccgctgtcgagctgccggtcgtaggctttgtcgtgctcgagggccgcggcgtctg
cctcgttgaccggctctcccttgtcgagtccgttgaagggtccgaggtacttgtaccca
ggaagcacaagacccctgctgtcgtccttatgccgctctgcgggctttggtggtggtgg
gccaggtttgagcttccaccactgtcttattccttcagagagagtgtcctcgagccaat
ctggaagataaccatcggcagccatgcatgctagagtcgacggatcggagtactgtcct
ccgagcggagtactgtcctcgagcggagtactgtcctccgagcggagtactgtcctccg
agcggagtactgtcctccgagcggatccccgggaattagcttggcgtaatcatggtcat
agctgtttcctgtgtgaaattgttatccgctcacaattccacacaacatacgagccgga
agcataaagtgtaaagcctggggtgcctaatgagtgagctaactcacattaattgcgtt
gcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcg
gccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcact
gactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggt
aatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggcc
agcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgc
cccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacagg
actataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccga
ccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttct
catagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctg
tgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttg
agtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggatt
agcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacgg
ctacactagaagaacagtatttggtatctgcgctctgctgaagccagttaccttcggaa
aaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttt
gtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatctt
ttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatga
gattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatca
atctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggc
acctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgt
agataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcga

gacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccga
gcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccggg
aagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctaca
ggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacg
atcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtc
ctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagca
ctgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagta
ctcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgt
caatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaa
cgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgta
acccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggt
gagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgt
tgaatactcatactcttccttttcaatattattgaagcatttatcagggttattgtct
catgagcggatacatatttgaatgtatttagaaaaataaacaaataggggttccgcgca
catttccccgaaaagtgccacctgacgtctaagaaaccattattatcatgacattaacc
tataaaaataggcgtatcacgaggccctttcgtctcgcgcgtttcggtgatgacggtga
aaacctctgacacatgcagctcccggagacggtcacagcttgtctgtaagcggatgccg
ggagcagacaagcccgtcagggcgcgtcagcgggtgttggcgggtgtcggggctggctt
aactatgcggcatcagagcagattgtactgagagtgcaccatatgcggtgtgaaataacc
gcacagatgcgtaaggagaaaataccgcatcaggcgccattcgccattcaggctgcgca
actgttgggaagggcgatcggtgcgggcctcttcgctattacgccagctggcgaaaggg
ggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacgacgttg
taaaacgacggccagtgaattcgagctcggtaccttgttgttaacttgtttattgcagc
ttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttt
cactgcattctagttgtggtttgtccaaactcat

pG5-TATA-VP1



Sequence color-coding key

Feature	Color
VP1	text
TATA box	text
GAL	text
Ori	text

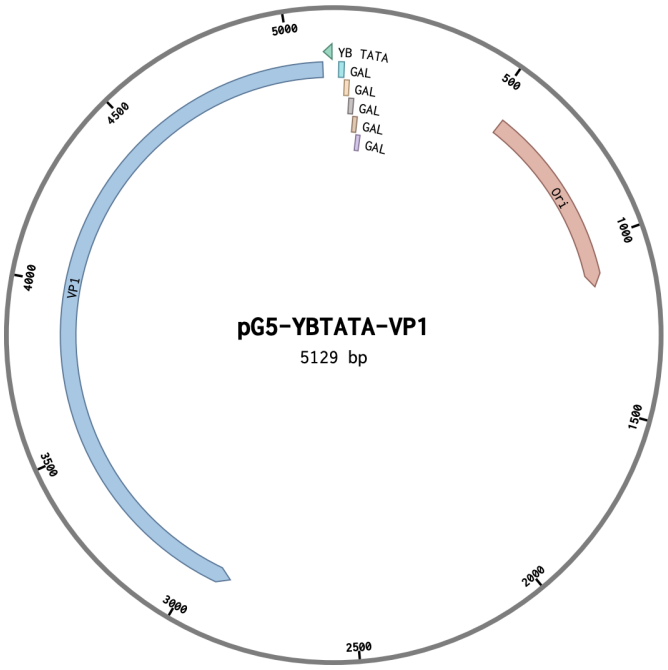
Sequence

caatgtatcttatcatgtctgctcgaagcggccggccgccccgactctagagctgtagt
taatgattaacccgccatgctacttatctacgtagccatggaaactagataagaaagaa
atacgcagagaccaaagttcaactgaaacgaattaaacggtttattgattaacaagcaa
ttacagattacgagtcaggtatctggtgccaatggggcgaggctctgaatacacgccat
tagtgtccacagtaaagtccacattaacagacttgttgtagttggaagtgtactgaatt
tcgggattccagcgtttgctgttttccttctgcagctcccactcgatctccacgctgac
ctgtcccgtggagtactgtgtgatgaaggaagcaaactttgccgcactgaaggtggtcg
aaggattcgcaggtaccggggtgttcttgatgagaatctgtggaggagggtgtttaagt
ccgaatccacccatgaggggagaggggtgaaaatgtccgtccgtgtgtggaatctttgc
ccagatgggcccctgaaggtacacatctctgtcctgccagaccatgcctggaagaacgc
cttgtgtgttgacatctgcggtagctgcttgtctgttgcctctctggaggttggtagat
acagaaccatactgctccgtagccacgggattggttgtcctgatttcctcttcgtctgt
aatcatgacctttttcaatgtccacatttgttttctctgagccttgcttcccaaagatga
gaaccccgctctgaggaaaaaactttcttcatcgtccttgtggcttgccatggccggg
cccggattcaccagagagtctctgccattgaggtggtacttggtagctccagtccacga
gtattcactgttgttgttatccgcagatgtctttgatactcgctgctggcggtaacagg
gtccaggaagccagttcctagactggtcccgaatgtcactcgctccggcctgagaaaac

tgaagccttgactgcgtggtggttccacttggagtgtttgttctgctcaagtaatacag
gtactggtcgatgagaggattcatgagacggtccagactctggctgtgagcgtagctgc
tgtggaaaggaacgtcctcaaaagtgtagctgaaggtaaagttgtttccggtacgcagc
atctgagaaggaaagtactccaggcagtaaaatgaagagcgtcctactgcctgactccc
gttgttcagggtgaggtatccatactgtggcaccatgaagacgtctgctgggaacggcg
ggaggcatccttgatgcgccgagccgaggacgtacgggagctggtactccgagtcagta
aacacctgaaccgtgctggtaaggttattggcaatcgtcgtcgtaccgtcattctgcgt
gacctctttgacttgaatgttaaagagcttgaagttgagtctcttgggtcggaatcccc
agttgttgttgatgagtctttgccagtcacgtggtgaaaagtggcagtggaatctgttg
aagtcaaaataccccaaggggtgctgtagccaaagtagtgattgtcgttcgaggctcc
tgattggctggaaatttgtttgtagaggtggttgttgtaggtgggcagggcccaggttc
gggtgctggtggtgatgactctgtcgccgagccatgtggaatcgcaatgccaatttccc
gaggaattacccactccgtcggcgccctcgttattgtctgcgagtggtgcgccactgcc
tgtagcgagggtattagttccgagaccagagggggctgctggtggctgtccgagaggct
gggggtcaggtactgagtctgcgtctccagtctgaccaaaattcaatctttttcttgca
ggctgctggcccgcctttccggttcccgaggaggagtctggctccacaggagagtgctc
taccggcctctttttcccggagcggtcttaacaggttcctcaaccaggcccagaggtt
caagaaccctctttttcgcctggaagactgctcgtccgaggttgcccccaaaagacgta
tcttctttaaggcgctcctgaaactccgcgtcggcgtggttgtacttgaggtacgggtt
gtctccgctgtcgagctgccggtcgtaggctttgtcgtgctcgagggccgcggcgtctg
cctcgttgaccggctctcccttgtcgagtccgttgaagggtccgaggtacttgtaccca
ggaagcacaagacccctgctgtcgtccttatgccgctctgcgggctttggtggtggtgg
gccaggtttgagcttccaccactgtcttattccttcagagagagtgtcctcgagccaat
ctggaagataaccatcggcagccatatttcttttatagcatgctagagtcgacggatcg
gagtactgtcctccgagcggagtactgtcctcgagcggagtactgtcctccgagcggag
tactgtcctccgagcggagtactgtcctccgagcggatccccgggaattagcttggcgt
aatcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaattccacacaac
atacgagccggaagcataaagtgtaaagcctggggtgcctaatgagtgagctaactcac
attaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgc
attaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgct
tcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctca
ctcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgt
gagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttc
cataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcg
aaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgct
ctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagc
gtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctc
caagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggta
actatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccact
ggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtg
gcctaactacggctacactagaagaacagtatttggtatctgcgctctgctgaagccag
ttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagc
ggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaaga
tcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaaggga
ttttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatga
agttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgctt
aatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgac
tccccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgca

atgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagc
cggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctatta
attgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgtt
gccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctc
cggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggtta
gctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatg
gttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgt
gactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgct
cttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctc
atcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatc
cagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcacca
gcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcg
acacggaaatgttgaatactcatactcttccttttcaatattattgaagcatttatca
gggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatag
gggttccgcgcacatttccccgaaaagtgccacctgacgtctaagaaaccattattatc
atgacattaacctataaaaataggcgtatcacgaggccctttcgtctcgcgcgtttcgg
tgatgacggtgaaaacctctgacacatgcagctcccggagacggtcacagcttgtctgt
aagcggatgccgggagcagacaagcccgtcagggcgcgtcagcgggtgttggcgggtgt
cggggctggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatgcg
gtgtgaaataccgcacagatgcgtaaggagaaaataccgcatcaggcgccattcgccat
tcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgctattacgccag
ctggcgaaaggggatgtgctgcaaggcgattaagttgggtaacgccagggttttccca
gtcacgacgttgtaaaacgacggccagtgaattcgagctcggtaccttgttgttaactt
gtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaata
aagcatttttttcactgcattctagttgtggtttgtccaaactcat

pG5-YBTATA-VP1



Sequence color-coding key

Feature	Color
GAL	text
Ori	text
VP1	text
YBTATA	text

Sequence

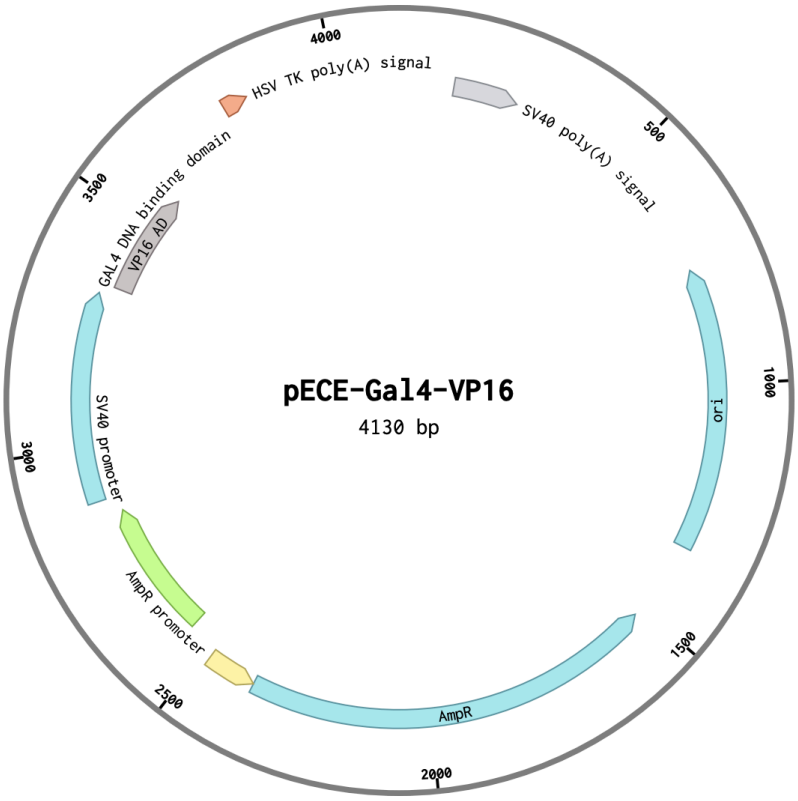
ctagagtcgacggatcggagtactgtcctccgagcggagtactgtcctcgagcggagta
ctgtcctccgagcggagtactgtcctccgagcggagtactgtcctccgagcggatcccc
gggaattagcttggcgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgc
tcacaattccacacaacatacgagccggaagcataaagtgtaaagcctggggtgcctaa
tgagtgagctaactcacattaattgcgttgcgctcactgcccgctttccagtcgggaaa
cctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgta
ttgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcgg
cgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataa
cgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccg
cgttgctggcgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgc
tcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctgg
aagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcct
ttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcg
gtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccg
ctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgc
cactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctaca

gagttcttgaagtggtggcctaactacggctacactagaagaacagtatttggtatctg
cgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaac
aaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaa
aaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacga
aaactcacgttaagggattttggtcatgagattatcaaaaaggatcttcacctagatcc
ttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtct
gacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttc
atccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccat
ctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatca
gcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgc
ctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaata
gtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggt
atggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgtt
gtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccg
cagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatcc
gtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtat
gcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagca
gaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatc
ttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagc
atcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaa
aaaagggaataagggcgacacggaaatgttgaatactcatactcttccttttcaatat
tattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtattta
gaaaaataaacaaatagggttccgcgcacatttccccgaaaagtgccacctgacgtct
aagaaaccattattatcatgacattaacctataaaaataggcgtatcacgaggcccttt
cgtctcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggagac
ggtcacagcttgtctgtaagcggatgccgggagcagacaagcccgtcagggcgcgtcag
cgggtgttggcgggtgtcggggctggcttaactatgcggcatcagagcagattgtactg
agagtgcaccatatgcggtgtgaaataccgcacagatgcgtaaggagaaaataccgcat
caggcgccattcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcct
cttcgctattacgccagctggcgaaaggggatgtgctgcaaggcgattaagttgggta
acgccagggttttcccagtcacgacgttgtaaaacgacggccagtgaattcgctgtagt
taatgattaacccgccatgctacttatctacgtagccatggaaactagataagaaagaa
atacgcagagaccaaagttcaactgaaacgaattaaacggtttattgattaacaagcaa
ttacagattacgagtcaggtatctggtgccaatggggcgaggctctgaatacacgccat
tagtgtccacagtaaagtccacattaacagacttgttgtagttggaagtgtactgaatt
tcgggattccagcgtttgctgttttccttctgcagctcccactcgatctccacgctgac
ctgtcccgtggagtactgtgtgatgaaggaagcaaactttgccgcactgaaggtggtcg
aaggattcgcaggtaccggggtgttcttgatgagaatctgtggaggagggtgtttaagt
ccgaatccacccatgaggggagaggggtgaaaatgtccgtccgtgtgtggaatctttgc
ccagatgggcccctgaaggtacacatctctgtcctgccagaccatgcctggaagaacgc
cttgtgtgttgacatctgcggtagctgcttgtctgttgcctctctggaggttggtagat
acagaaccatactgctccgtagccacgggattggttgtcctgatttcctcttcgtctgt
aatcatgacctttttcaatgtccacatttgttttctctgagccttgcttcccaaagatga
gaaccccgctctgaggaaaaaactttttcttcatcgtccttgtggcttgccatggccggg
cccggattcaccagagagtctctgccattgaggtggtacttggtagctccagtccacga
gtattcactgttgttgttatccgcagatgtctttgatactcgctgctggcggtaacagg
gtccaggaagccagttcctagactggtcccgaatgtcactcgctccggcctgagaaaac
tgaagccttgactgcgtggtggttccacttggagtgtttgttctgctcaagtaatacag

gtactggtcgatgagaggattcatgagacggtccagactctggctgtgagcgtagctgc
tgtggaaaggaacgtcctcaaaagtgtagctgaaggtaaagttgtttccggtacgcagc
atctgagaaggaaagtactccaggcagtaaaatgaagagcgtcctactgcctgactccc
gttgttcagggtgaggtatccatactgtggcaccatgaagacgtctgctgggaacggcg
ggaggcatccttgatgcgccgagccgaggacgtacgggagctggtactccgagtcagta
aacacctgaaccgtgctggtaaggttattggcaatcgtcgtcgtaccgtcattctgcgt
gacctctttgacttgaatgttaaagagcttgaagttgagtctcttgggtcggaatcccc
agttgttgttgatgagtctttgccagtcacgtggtgaaaagtggcagtggaatctgttg
aagtcaaaataccccaaggggtgctgtagccaaagtagtgattgtcgttcgaggctcc
tgattggctggaaatttgtttgtagaggtggttgttgtaggtgggcagggcccaggttc
gggtgctggtggtgatgactctgtcgccgagccatgtggaatcgcaatgccaatttccc
gaggaattacccactccgtcggcgccctcgttattgtctgcgagtggtgcgccactgcc
tgtagcgagggtattagttccgagaccagaggggctgctggtggctgtccgagaggct
gggggtcaggtactgagtctgcgtctccagtctgaccaaaattcaatctttttcttgca
ggctgctggcccgcctttccggttcccgaggaggagtctggctccacaggagagtgctc
taccggcctctttttttcccggagcggtcttaacaggttcctcaaccaggcccagaggtt
caagaaccctctttttcgcctggaagactgctcgtccgaggttgcccccaaaagacgta
tcttctttaaggcgctcctgaaactccgcgtcggcgtggttgtacttgaggtacgggtt
gtctccgctgtcgagctgccggtcgtaggctttgtcgtgctcgagggccgcggcgtctg
cctcgttgaccggctctcccttgtcgagtccgttgaagggtccgaggtacttgtaccca
ggaagcacaagacccctgctgtcgtccttatgccgctctgcgggctttggtggtggtgg
gccaggtttgagcttccaccactgtcttattccttcagagagagtgtcctcgagccaat
ctggaagataaccatcggcagccattggcccccattatataccctctagagcatg

The sequence for the minCMV minimal promoter is:
gatctgacggttcactaaacgagctctgcttatatagacctcccaccgtacacgcctac

pECE-Gal4-VP16



Sequence color-coding key

Feature	Color
SV40 poly(A) signal	text
Ori	text
AmpR	text
AmpR promoter	text
SV40 promoter	text
GAL4 DNA binding domain	text
VP16 AD	text
HSV TK poly(A) signal	text

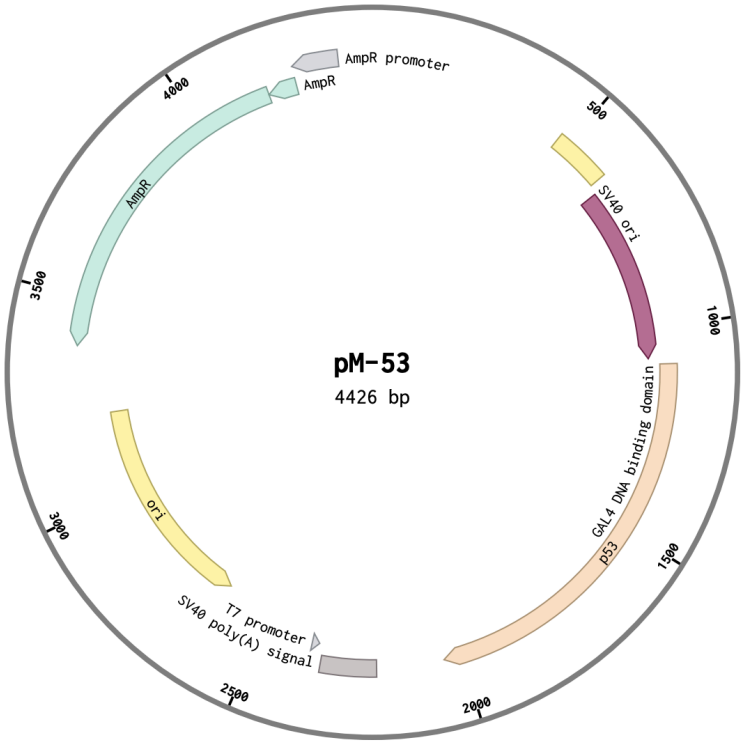
Sequence

taagtaatgatcataatcagccatatcacatctgtagaggttttacttgctttaaaaaa
cctcccacacctccccctgaacctgaaacataaaatgaatgcaattgttgttgttaact
tgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaat
aaagcatttttttcactgcattctagttgtggtttgtccaaactcatcaatgtatctta
tcatgtctggatcctctacgccggacgcatcgtggccggcatcaccggcgccacaggtg
cggttgctggcgcctatatcgccgacatcaccgatggggaagatcgggctcgccacttc
gggctcatgagcgcttgtttcggcgtgggtatggtggcaggccccgtggccgggggact
gttgggcgccatctccttgcatgcaccattccttgcggcggcggtgctcaacggcctca
acctactactgggctgcttcctaatgcaggagtcgcataagggagagcgtgcggtgtga

aataccgcacagatgcgtaaggagaaaataccgcatcaggcgctcttccgcttcctcgc
tcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaag
gcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaa
aggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggc
tccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccg
acaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgt
tccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgc
tttctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctg
ggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcg
tcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaaca
ggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaac
tacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttacctt
cggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtt
tttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttg
atcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggt
catgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagtttta
aatcaatctaaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcag
tgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccg
tcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgata
ccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaag
ggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgtt
gccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccatt
gctgcaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttc
ccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctcct
tcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatg
gcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactgg
tgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcc
cggcgtcaacacgggataataccgcgccacatagcagaactttaaaagtgctcatcatt
ggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttc
gatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgttt
ctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacgg
aaatgttgaatactcatactcttccttttttcaatattattgaagcatttatcagggtta
ttgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagggggttc
cgcgcacatttccccgaaaagtgccacctgcttctgaggcggaaagaaccagctgtgga
atgtgtgtcagttaggtgtggaaagtccccaggctccccagcaggcagaagtatgcaa
agcatgcatctcaattagtcagcaaccaggtgtggaaagtccccaggctccccagcagg
cagaagtatgcaaagcatgcatctcaattagtcagcaaccatagtcccgccccctaactc
cgcccatcccgcccctaactccgcccagttccgcccattctccgccccatggctgacta
attttttttatttatgcagaggccgaggccgcctcggcctctgagctattccagaagta
gtgaagaggcttttttggaggagatctaagcttgaagcaagcctcctgaaagatgaagc
tactgtcttctatcgaacaagcatgcgatatttgccgacttaaaaagctcaagtgctcc
aaagaaaaaccgaagtgcgccaagtgtctgaagaacaactgggagtgtcgctactctcc
caaaaccaaaaggtctccgctgactagggcacatctgacagaagtggaatcaaggctag
aaagactggaacagctatttctactgattttcctcgagaagaccttgacatgattttg
aaaatggattctttacaggatataaaagcattgttaacaggattatttgtacaagataa
tgtgaataaagatgccgtcacagatagattggcttcagtggagactgatatgcctctaa
cattgagacagcatagaataagtgcgacatcatcatcggaagagagtagtaacaaaggt
caaagacagttgactgtatcgccggaattcccggggatctgggccccccgaccgatgt

cagcctgggggacgagctccacttagacggcgaggacgtggcgatggcgcatgccgacg
cgctagacgatttcgatctggacatgttgggggacggggattccccgggtccgggattt
acccccccacgactccgccccctacggcgctctggatatggccgacttcgagtttgagca
gatgtttaccgatgcccttggaattgacgagtacggtgggtaggggggcgcgaccggacc
cgcatcccccgtctgggttttcccctcccgtccccggttcgtatccacaataaacacga
gcacatacattacaaaaacctgcggttgtcgtctgattatttggtggatccgggggaga
tgggggaggctaactgaaacacggaaggagacaataccggaaggaacccgcgctatgac
ggcaataaaaagacagaataaaacgcacgggtgttgggtcgtttgttcataaacgcggg
gttcggtcccagggctggcactctgtcgataccccaccgagaccccattggggccaata
cgcccgcgtttcttccttttccccacccccacccccaagttcgggtgaaggcccagggc
tcgcagccaacgtcggggcggcaggccctgccatagccactggccccgtgggttaggga
cggggtcccccatggggaatggtttatggttcgtgggggttattattttgggcgttgcg
tggggtcaggtccacgacccaagcttgtcgacggtaccccggggaattcgagctctaga

pM-53



Sequence color-coding key

Feature	Color
SV40 ori	text
GAL4 DNA binding domain	text
p53	text
SV40 poly(A) signal	text
T7 promoter	text
Ori	text
AmpR	text
AmpR promoter	text

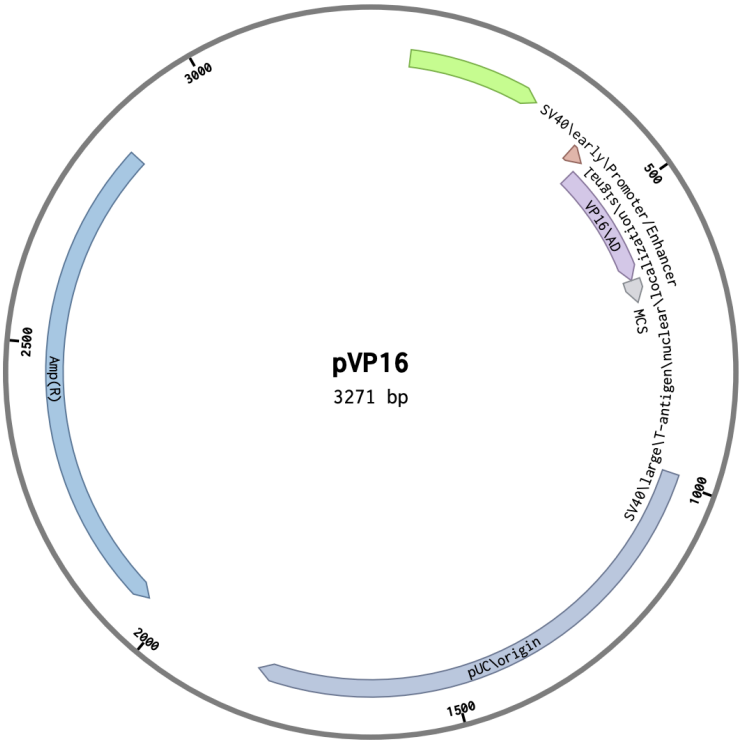
Sequence

tcacgaggcctttcgtctcgcgcgtttcggtgatgacggtgaaaacctctgacacatg
cagctcccggagacggtcacagcttgtctgtaagcggatgccgggagcagacaagcccg
tcagggcgcgtcagcgggtgttggcgggtgtcggggctggcttaactatgcggcatcag
agcagattgtactgagagtgcaccatatggacatattgtcgttagaacgcggctacaat
taatacataaccttatgtatcatacacatacgatttaggtgacactatagaactcgatg
tggaatgtgtgtcagttagggtgtggaaagtccccaggctccccagcaggcagaagtat
gcaaagcatgcatctcaattagtcagcaaggaaagtccccaggctccccagcaggcaga
agtatgcaaagcatgcatctcaattagtcagcaaccatagtcccgccctaactccgcc
catcccgccctaactccgcccagttccgcccattctccgccccatggctgactaattt
ttttatttatgcagaggccgaggccgcctcggcctctgagctattccagaagtagtga

agaggctttttttggaggagatctaagctagcctcctgaaagatgaagctactgtcttct
atcgaacaagcatgcgatatttgccgacttaaaaagctcaagtgctccaaagaaaaacc
gaagtgcgccaagtgtctgaagaacaactgggagtgtcgctactctcccaaaaccaaaa
ggtctccgctgactagggcacatctgacagaagtggaatcaaggctagaaagactggaa
cagctatttctactgatttttcctcgagaagaccttgacatgattttgaaaatggattc
tttacaggatataaaagcattgttaacaggattatttgtacaagataatgtgaataaag
atgccgtcacagatagattggcttcagtggagactgatatgcctctaacattgagacag
catagaataagtgcgacatcatcatcggaagagagtagtaacaaaggtcaaagacagtt
gactgtatcgccggaattccctgtcaccgagacccctgggccagtggcccctgccccag
ccactccatggcccctgtcatcttttgtcccttctcaaaaaacttaccagggcaactat
ggcttccacctgggcttcctgcagtctgggacagccaagtctgttatgtgcacgtactc
tcctcccctcaataagctattctgccagctggcgaagacgtgccctgtgcagttgtggg
tcagcgccacacctccagctgggagccgtgtccgcgccatggccatctacaagaagtca
cagcacatgacggaggtcgtgagacgctgcccccaccatgagcgctgctccgatggtga
tggcctggctcctccccagcatcttatccgggtggaaggaaatttgtatcccgagtatc
tggaagacaggcagacttttcgccacagcgtggtggtaccttatgagccacccgaggcc
ggctctgagtataccaccatccactacaagtacatgtgtaatagctcctgcatgggggg
catgaaccgccgacctatccttaccatcatcacactggaagactccagtgggaaccttc
tgggacgggacagctttgaggttcgtgtttgtgcctgccctgggagagaccgccgtaca
gaagaagaaaatttccgcaaaaaggaagtcctttgccctgaactgcccccagggagcgc
aaagagagcgctgcccacctgcacaagcgcctctcccccgcaaaagaaaaaaccacttg
atggagagtatttcaccctcaagatccgcgggcgtaaacgcttcgagatgttccgggag
ctgaatgaggccttagagttaaaggatgcccatgctacagaggagtctggagacagcag
ggctcactccagctacctgaagaccaagaagggccagtctacttcccgccataaaaaaa
caatggtcaagaaagtggggcctgactcagactgactgcctctgcatcccggatccgtc
gacgcgtctgcagaagcttctagataagtaatgatcataatcagccatatcacatctgt
agaggttttacttgctttaaaaaacctcccacacctccccctgaacctgaaacataaaa
tgaatgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataaagc
aatagcatcacaaatttcacaaataaagcatttttttcactgcattctagttgtggttt
gtccaaactcatcaatgtatcttatcatgtctggatctgccggtctccctatagtgagt
cgtattaatttcgataagccaggttaacctgcattaatgaatcggccaacgcgcgggga
gaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcg
gtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccac
agaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccagga
accgtaaaaaggccgcgttgctggcgtttttccataggctccgccccctgacgagcat
cacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagatacca
ggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccg
gatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgt
aggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccc
cgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaa
gacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtat
gtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaac
agtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagct
cttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcag
attacgcgcagaaaaaaggatctcaagaagatcctttgatcttttctacggggtctga
cgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaaaaagga
tcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatat
gagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgat

ctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatac
gggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccg
gctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcc
tgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagta
gttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtca
cgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttac
atgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtca
gaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctctt
actgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcatt
ctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataata
ccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcga
aaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacc
caactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaa
ggcaaaatgccgcaaaaagggaataagggcgacacggaaatgttgaatactcatactc
ttccttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacat
atttgaatgtatttagaaaaataaacaaatagggttccgcgcacatttccccgaaaag
tgccacctgacgtctaagaaaccattattatcatgacattaacctataaaaataggcgt
a

pVP16



Sequence color-coding key

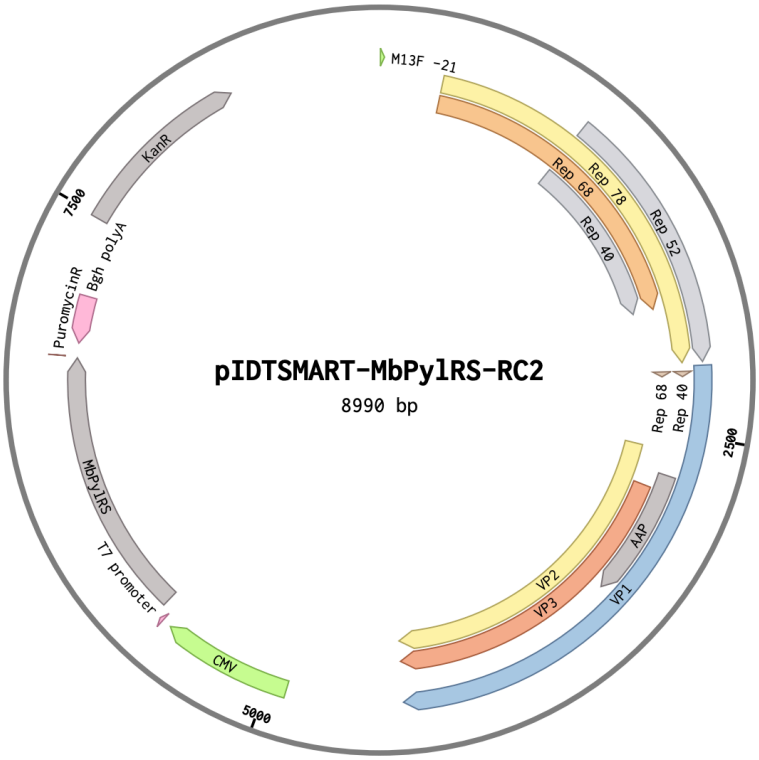
Feature	Color
SV40 early Promoter/Enhancer	text
SV40 large T-antigen nuclear localization signal	text
VP16-AD	text
Multiple cloning site	text
pUC Ori	text
AmpR	text

Sequence

tatgtatcatacacatacgatttaggtgacactatagaactcgatgtggaatgtgtgtc
agttaggtgtggaaagtccccaggctccccagcaggcagaagtatgcaaagcatgcat
ctcaattagtcagcaaggaaagtccccaggctccccagcaggcagaagtatgcaaagca
tgcatctcaattagtcagcaaccatagtcccgcccctaactccgcccatcccgccccta
actccgcccagttccgcccattctccgccccatggctgactaatttttttatttatgc
agaggccgaggccgcctcggcctctgagctattccagaagtagtgaagaggcttttttg
gaggagatctaagctagcgccgccaccatgggccctaaaaagaagcgtaaagtcgcccc
cccgaccgatgtcagcctgggggacgagctccacttagacggcgaggacgtggcgatgg
cgcatgccgacgcgctagacgatttcgatctggacatgttgggggacggggattccccg
ggtccgggatttacccccacgactccgccccctacggcgctctggatatggccgactt
cgagtttgagcagatgtttaccgatgcccttggaattgacgagtacggtggggaattcc
cggggatccgtcgacgcgtctgcagaagcttctagataagtaatgatcataatcagcca

tatcacatctgtagaggttttacttgctttaaaaaacctcccacacctccccctgaacc
tgaaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttataatggt
tacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcactgcattc
tagttgtggtttgtccaaactcatcaatgtatcttatcatgtctggatctgccggtctc
cctatagtgagtcgtattaatttcgataagccaggttaacctgcattaatgaatcggcc
aacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgac
tcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaat
acggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagc
aaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgcccc
cctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggact
ataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccc
tgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcat
agctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgt
gcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagt
ccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagc
agagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggcta
cactagaagaacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaa
gagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtt
tgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatctttc
tacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagat
tatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatc
taaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacc
tatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtaga
taactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagac
ccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcg
cagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaag
ctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggc
atcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatc
aaggcgagttacatgatccccccatgttgtgcaaaaaagcggttagctccttcggtcctc
cgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactg
cataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactc
aaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaa
tacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgt
tcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacc
cactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgag
caaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttga
atactcatactcttccttttcaatattattgaagcatttatcagggttattgtctcat
gagcggatacatatttgaatgtatttagaaaaataaacaaatagggttccgcgcacat
ttccccgaaaagtgccacctgacgtctaagaaaccattattatcatgacattaacctat
aaaaataggcgtatcacgaggccctttcgtctcgcgcgtttcggtgatgacggtgaaaa
cctctgacacatgcagctcccggagacggtcacagcttgtctgtaagcggatgccggga
gcagacaagcccgtcagggcgcgtcagcgggtgttggcgggtgtcggggctggcttaac
tatgcggcatcagagcagattgtactgagagtgcaccatatggacatattgtcgttaga
acgcggctacaattaatacataacct

pIDT-MbPylRS-RC2



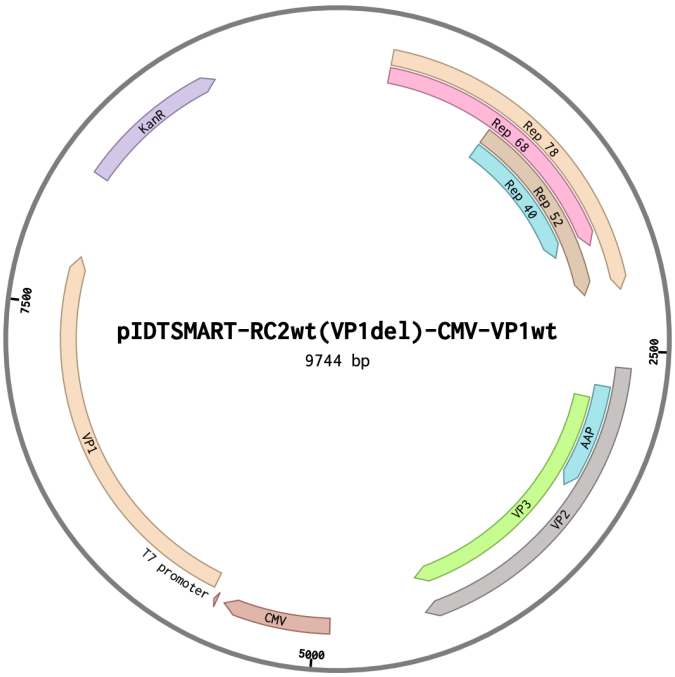
The sequence for the wild type RC2 sequence can be found in the Chapter 2 Plasmids section of this appendix.

The sequence for the VP1del *Cap* gene is:

CTCgctgccgatggttatcttccagattggctcgaggacactctctctgaaggaataag
acagtggtggaagctcaaacctggcccaccaccaccaaagcccgcagagcggcataagg
acgacagcaggggtcttgtgcttcctgggtacaagtacctcggacccttcaacggactc
gacaagggagagccggtcaacgaggcagacgccgcggccctcgagcacgacaaagccta
cgaccggcagctcgacagcggagacaacccgtacctcaagtacaaccacgccgacgcgg
agtttcaggagcgccttaaagaagatacgtcttttgggggcaacctcggacgagcagtc
ttccaggcgaaaaagagggttcttgaacctctgggcctggttgaggaacctgttaagac
ggctccgggaaaaaagaggccggtagagcactctcctgtggagccagactcctcctcgg
gaaccggaaaggcgggccagcagcctgcaagaaaaagattgaattttggtcagactgga
gacgcagactcagtacctgacccccagcctctcggacagccaccagcagccccctctgg
tctgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataacgagg
gcgccgacggagtgggtaattcctcgggaaattggcattgcgattccacatggatgggc
gacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaaccacct
ctacaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttggctaca
gcacccctgggggtattttgacttcaacagattccactgccacttttcaccacgtgac
tggcaaagactcatcaacaacaactggggattccgacccaagagactcaacttcaagct
ctttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgccaata
accttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacgtcctc

ggctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtgccaca
gtatggatacctcaccctgaacaacgggagtcaggcagtaggacgctcttcattttact
gcctggagtactttccttctcagatgctgcgtaccggaaacaactttaccttcagctac
acttttgaggacgttcctttccacagcagctacgctcacagccagagtctggaccgtct
catgaatcctctcatcgaccagtacctgtattacttgagcagaacaaacactccaagtg
gaaccaccacgcagtcaaggcttcagtttctcaggccggagcgagtgacattcgggac
cagtctaggaactggcttcctggaccctgttaccgccagcagcgagtatcaaagacatc
tgcggataacaacaacagtgaatactcgtggactggagctaccaagtaccacctcaatg
gcagagactctctggtgaatccgggcccggccatggcaagccacaaggacgatgaagaa
aagtttttcctcagagcggggttctcatctttgggaagcaaggctcagagaaaacaaa
tgtggacattgaaaaggtcatgattacagacgaagaggaaatcaggacaaccaatcccg
tggctacggagcagtatggttctgtatctaccaacctccagagaggcaacagacaagca
gctaccgcagatgtcaacacacaaggcgttcttccaggcatggtctggcaggacagaga
tgtgtaccttcaggggcccatctgggcaaagattccacacacggacggacattttcacc
cctctcccctcatgggtggattcggacttaaacaccctcctccacagattctcatcaag
aacaccccggtacctgcgaatccttcgaccaccttcagtgcggcaaagtttgcttcctt
catcacacagtactccacgggacaggtcagcgtggagatcgagtgggagctgcagaagg
aaaacagcaaacgctggaatcccgaaattcagtacacttccaactacaacaagtctgtt
aatgtggactttactgtggacactaatggcgtgtattcagagcctcgccccattggcac
cagatacctgactcgtaatctgtaa

pIDTSMART-RC2-VP1del-CMV-VP1wt



Sequence color-coding key

Feature	Color
Rep	text
Cap (VP1del)	text
CMV promoter	text
T7 promoter	text
VP1 wild type	text
KanR	text

Sequence

cccgtgtaaaacgacggccagtttatctagtcagcttgattctagctgatcgtggaccg
gaaggtgagccagtgagttgattgcagtccagttacgctggagtctgaggctcgtcctg
aatgatatgcgaccgccggagggttgcgtttgagacgggcgacagatccagtcgcgctg
ctctcgtcgatccgctagggcggccgctctagaactagtggatcccccggaagatcaga
agttcctattccgaagttcctattctctagaaagtataggaacttctgatctgcgcagc
cgccatgccgggggttttacgagattgtgattaaggtccccagcgaccttgacgagcatc
tgcccggcatttctgacagctttgtgaactgggtggccgagaaggaatgggagttgccg
ccagattctgacatggatctgaatctgattgagcaggcacccctgaccgtggccgagaa
gctgcagcgcgactttctgacggaatggcgccgtgtgagtaaggccccggaggcccttt
tctttgtgcaatttgagaagggagagagctacttccacatgcacgtgctcgtggaaacc
accggggtgaaatccatggttttgggacgtttcctgagtcagattcgcgaaaaactgat
tcagagaatttaccgcgggatcgagccgactttgccaaactggttcgcggtcacaaaga
ccagaaatggcgccggaggcgggaacaaggtggtggatgagtgctacatccccaattac
ttgctccccaaaacccagcctgagctccagtgggcgtggactaatatggaacagtattt

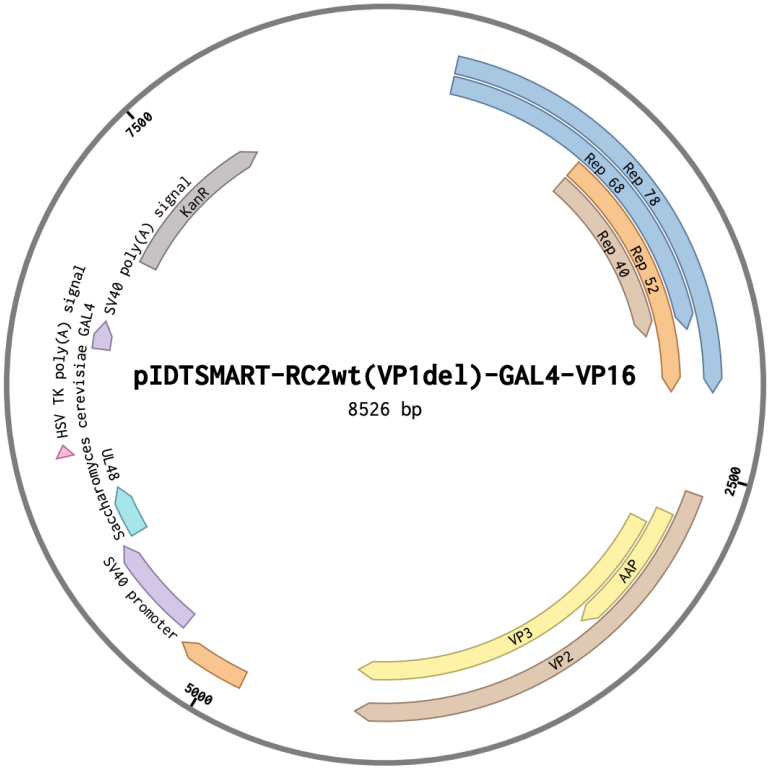
aagcgcctgtttgaatctcacggagcgtaaacggttggtggcgcagcatctgacgcacg
tgtcgcagacgcaggagcagaacaaagagaatcagaatcccaattctgatgcgccggtg
atcagatcaaaaacttcagccaggtacatggagctggtcgggtggctcgtggacaaggg
gattacctcggagaagcagtggatccaggaggaccaggcctcatacatctccttcaatg
cggcctccaactcgcggtcccaaatcaaggctgccttggacaatgcgggaaagattatg
agcctgactaaaaccgcccccgactacctggtgggccagcagcccgtggaggacatttc
cagcaatcggatttataaaatttggaactaaacgggtacgatccccaatatgcggctt
ccgtctttctgggatgggccacgaaaaagttcggcaagaggaacaccatctggctgttt
gggcctgcaactaccgggaagaccaacatcgcggaggccatagcccacactgtgccctt
ctacgggtgcgtaaactggaccaatgagaactttcccttcaacgactgtgtcgacaaga
tggtgatctggtgggaggaggggaagatgaccgccaaggtcgtggagtcggccaaagcc
attctcggaggaagcaaggtgcgcgtggaccagaaatgcaagtcctcggcccagataga
cccgactcccgtgatcgtcacctccaacaccaacatgtgcgccgtgattgacgggaact
caacgaccttcgaacaccagcagccgttgcaagaccggatgttcaaatttgaactcacc
cgccgtctggatcatgactttgggaaggtcaccaagcaggaagtcaaagacttttccg
gtgggcaaaggatcacgtggttgaggtggagcatgaattctacgtcaaaaagggtggag
ccaagaaaagacccgcccccagtgacgcagatataagtgagcccaaacgggtgcgcgag
tcagttgcgcagccatcgacgtcagacgcggaagcttcgatcaactacgcagacaggta
ccaaaacaaatgttctcgtcacgtgggcatgaatctgatgctgtttccctgcagacaat
gcgagagaatgaatcagaattcaaatatctgcttcactcacggacagaaagactgttta
gagtgctttcccgtgtcagaatctcaacccgtttctgtcgtcaaaaaggcgtatcagaa
actgtgctacattcatcatatcatgggaaaggtgccagacgcttgcactgcctgcgatc
tggtcaatgtggatttggatgactgcatctttgaacaataaatgatttaaatcaggtct
cgctgccgatggttatcttccagattggctcgaggacactctctctgaaggaataagac
agtggtggaagctcaaacctggcccaccaccaccaaagcccgcagagcggcataaggac
gacagcaggggtcttgtgcttcctgggtacaagtacctcggacccttcaacggactcga
caagggagagccggtcaacgaggcagacgccgcggccctcgagcacgacaaagcctacg
accggcagctcgacagcggagacaacccgtacctcaagtacaaccacgccgacgcggag
tttcaggagcgccttaaagaagatacgtcttttgggggcaacctcggacgagcagtctt
ccaggcgaaaaagagggttcttgaacctctgggcctggttgaggaacctgttaagacgg
ctccgggaaaaaagaggccggtagagcactctcctgtggagccagactcctcctcggga
accggaaaggcgggccagcagcctgcaagaaaaagattgaattttggtcagactggaga
cgcagactcagtacctgacccccagcctctcggacagccaccagcagccccctctggtc
tgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataacgagggc
gccgacggagtgggtaattcctcgggaaattggcattgcgattccacatggatgggcga
cagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaaccacctct
acaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttggctacagc
acccttgggggtattttgacttcaacagattccactgccacttttcaccacgtgactg
gcaaagactcatcaacaacaactggggattccgacccaagagactcaacttcaagctct
ttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgccaataac
cttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacgtcctcgg
ctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtgccacagt
atggatacctcaccctgaacaacgggagtcaggcagtaggacgctcttcattttactgc
ctggagtactttccttctcagatgctgcgtaccggaaacaactttaccttcagctacac
ttttgaggacgttcctttccacagcagctacgctcacagccagagtctggaccgtctca
tgaatcctctcatcgaccagtacctgtattacttgagcagaacaaacactccaagtgga
accaccacgcagtcaaggcttcagttttctcaggccggagcgagtgacattcgggacca
gtctaggaactggcttcctggaccctgttaccgccagcagcgagtatcaaagacatctg

cggataacaacaacagtgaatactcgtggactggagctaccaagtaccacctcaatggc
agagactctctggtgaatccgggcccggccatggcaagccacaaggacgatgaagaaaa
gtttttttcctcagagcggggttctcatctttgggaagcaaggctcagagaaaacaaatg
tggacattgaaaaggtcatgattacagacgaagaggaaatcaggacaaccaatcccgtg
gctacggagcagtatggttctgtatctaccaacctccagagaggcaacagacaagcagc
taccgcagatgtcaacacacaaggcgttcttccaggcatggtctggcaggacagagatg
tgtaccttcaggggcccatctgggcaaagattccacacacggacggacattttcacccc
tctcccctcatgggtggattcggacttaaacaccctcctccacagattctcatcaagaa
caccccggtacctgcgaatccttcgaccaccttcagtgcggcaaagtttgcttccttca
tcacacagtactccacgggacaggtcagcgtggagatcgagtgggagctgcagaaggaa
aacagcaaacgctggaatcccgaaattcagtacacttccaactacaacaagtctgttaa
tgtggactttactgtggacactaatggcgtgtattcagagcctcgccccattggcacca
gatacctgactcgtaatctgtaattgcttgttaatcaataaaccgtttaattcgtttca
gttgaactttggtctctgcgtatttctttcttatctagtttccatggctacgtagataa
gtagcatggcgggttaatcattaactacagcccgggcgtttaaacagcgggcggagggg
tggagtcgtgacgtgaattacgtcatagggttagggaggtcctgtattagaggtcacgt
gagtgttttgcgacattttgcgacaccatgtggtctcgctgggggggggggcccgagtg
agcacgcagggtctccattttgaagcgggaggtttgaacgagcgctggcgcgctcactg
gccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgcct
tgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcc
cttcccatgcatcggccgcaaatacctgcaggatccgttttgcgctgcttcgcgatgta
cgggccagatatacgcgttgacattgattattgactagttattaatagtaatcaattac
ggggtcattagttcatagcccatatatggagttccgcgttacataacttacggtaaatg
gcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtatgtt
cccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacggta
aactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgacg
tcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt
cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttg
gcagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacc
ccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgt
cgtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtcta
tataagcagagctctctggctaactagagaacccactgcttactggcttatcgaaatta
atacgactcactatagggagacccaagctggctagcatggctgccgatggttatcttcc
agattggctcgaggacactctctctgaaggaataagacagtggtggaagctcaaacctg
gcccaccaccaccaaagcccgcagagcggcataaggacgacagcaggggtcttgtgctt
cctgggtacaagtacctcggacccttcaacggactcgacaagggagagccggtcaacga
ggcagacgccgcggccctcgagcacgacaaagcctacgaccggcagctcgacagcggag
acaacccgtacctcaagtacaaccacgccgacgcggagtttcaggagcgccttaaagaa
gatacgtcttttgggggcaacctcggacgagcagtcttccaggcgaaaaagagggttct
tgaacctctgggcctggttgaggaacctgttaagaccgctccgggaaaaaagaggccgg
tagagcactctcctgtggagccagactcctcctcgggaaccggaaaggcgggccagcag
cctgcaagaaaaagattgaattttggtcagactggagacgcagactcagtacctgaccc
ccagcctctcggacagccaccagcagccccctctggtctcggaactaataccctcgcta
caggcagtggcgcaccactcgcagacaataacgagggcgccgacggagtgggtaattcc
tcgggaaattggcattgcgattccacatggctcggcgacagagtcatcaccaccagcac
ccgaacctgggccctgcccacctacaacaaccacctctacaaacaaatttccagccaat
caggagcctcgaacgacaatcactactttggctacagcacccttgggggtattttgac
ttcaacagattccactgccacttttcaccacgtgactggcaaagactcatcaacaacaa

ctggggattccgacccaagagactcaacttcaagctctttaacattcaagtcaaagagg
tcacgcagaatgacggtacgacgacgattgccaataaccttaccagcacggttcaggtg
tttactgactcggagtaccagctcccgtacgtcctcggctcggcgcatcaaggatgcct
cccgccgttcccagcagacgtcttcatggtgccacagtatggatacctcaccctgaaca
acgggagtcaggcagtaggacgctcttcattttactgcctggagtactttccttctcag
atgctgcgtaccggaaacaactttaccttcagctacacttttgaggacgttcctttcca
cagcagctacgctcacagccagagtctggaccgtctcatgaatcctctcatcgaccagt
acctgtattacttgagcagaacaaacactccaagtggaaccaccacgcagtcaaggctt
cagttttctcaggccggagcgagtgacattcgggaccagtctaggaactggcttcctgg
accctgttaccgccagcagcgagtatcaaagacatctgcggataacaacaacagtgaat
actcgtggactggagctaccaagtaccacctcaatggcagagactctctggtgaatccg
ggcccggccatggcaagccacaaggacgatgaagaaaagtttttttcctcagagcggggt
tctcatctttgggaagcaaggctcagagaaaacaaatgtggacattgaaaaggtcatga
ttacagacgaagaggaaatcaggacaaccaatcccgtggctacggagcagtatggttct
gtatctaccaacctccagagaggcaacagacaagcagctaccgcagatgtcaacacaca
aggcgttcttccaggcatggtctggcaggacagagatgtgtaccttcaggggcccatct
gggcaaagattccacacacggacggacattttcacccctctcccctcatgggtggattc
ggacttaaacaccctcctccacagattctcatcaagaacaccccggtacctgcgaatcc
ttcgaccaccttcagtgcggcaaagtttgcttccttcatcacacagtactccacgggac
aggtcagcgtggagatcgagtgggagctgcagaaggaaaacagcaaacgctggaatccc
gaaattcagtacacttccaactacaacaagtctgttaatgtggactttactgtggacac
taatggcgtgtattcagagcctcgccccattggcaccagatacctgactcgtaatctgt
aattgcttgttaatcaataaaccgtttaattcgtttcagttgaactttggtctctgcgt
atttctttcttatctagtttccatggctacgtagataagtagcatggcgggttaatcat
taactacagccctaggggtgcgagcggatcgagcagtgtcgatcactactggaccgcga
gctgtgctgcgacccgtgatcttacggcattatacgtatgatcggtccacgatcagcta
gattatctagtcagcttgatgtcatagctgtttcctgaggctcaatactgaccatttaa
atcatacctgacctccatagcagaaagtcaaaagcctccgaccggaggcttttgacttg
atcggcacgtaagaggttccaactttcaccataatgaaataagatcactaccgggcgta
ttttttgagttatcgagattttcaggagctaaggaagctaaaatgagccatattcaacg
ggaaacgtcttgcttgaagccgcgattaaattccaacatggatgctgatttatatgggt
ataaatgggctcgcgataatgtcgggcaatcaggtgcgacaatctatcgattgtatggg
aagcccgatgcgccagagttgtttctgaaacatggcaaaggtagcgttgccaatgatgt
tacagatgagatggtcaggctaaactggctgacggaatttatgcctcttccgaccatca
agcattttatccgtactcctgatgatgcatggttactcaccactgcgatcccagggaaa
acagcattccaggtattagaagaatatcctgattcaggtgaaaatattgttgatgcgct
ggcagtgttcctgcgccggttgcattcgattcctgtttgtaattgtccttttaacggcg
atcgcgtatttcgtctcgctcaggcgcaatcacgaatgaataacggtttggttggtgcg
agtgattttgatgacgagcgtaatggctggcctgttgaacaagtctggaaagaaatgca
taaactcttgccattctcaccggattcagtcgtcactcatggtgatttctcacttgata
accttatttttgacgaggggaaattaataggttgtattgatgttggacgagtcggaatc
gcagaccgataccaggatcttgccatcctatggaactgcctcggtgagtttctccttc
attacagaaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgc
agtttcacttgatgctcgatgagtttttctaatgaggacctaaatgtaatcacctggct
caccttcgggtgggcctttctgcgttgctggcgttttccataggctccgccccctga
cgagcatcacaaaaatcgatgctcaagtcagaggtggcgaaacccgacaggactataaa
gataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccg
cttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctc

acgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacg
aaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaac
ccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagc
gaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacacta
gaagaacagtatttggtatctgcgctctgctgaagccagttacctcggaaaaagagttg
gtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaag
cagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgattttctaccgaag
aaaggccca

pIDTSMART-RC2-VP1del-GAL4-VP16



Sequence color-coding key

Feature	Color
Rep	text
Cap (VP1del)	text
SV40 promoter	text
Gal4	text
UL48 (VP16)	text
HSV TK poly(A) signal	text
SV40 poly(A) signal	text
KanR	text

Sequence

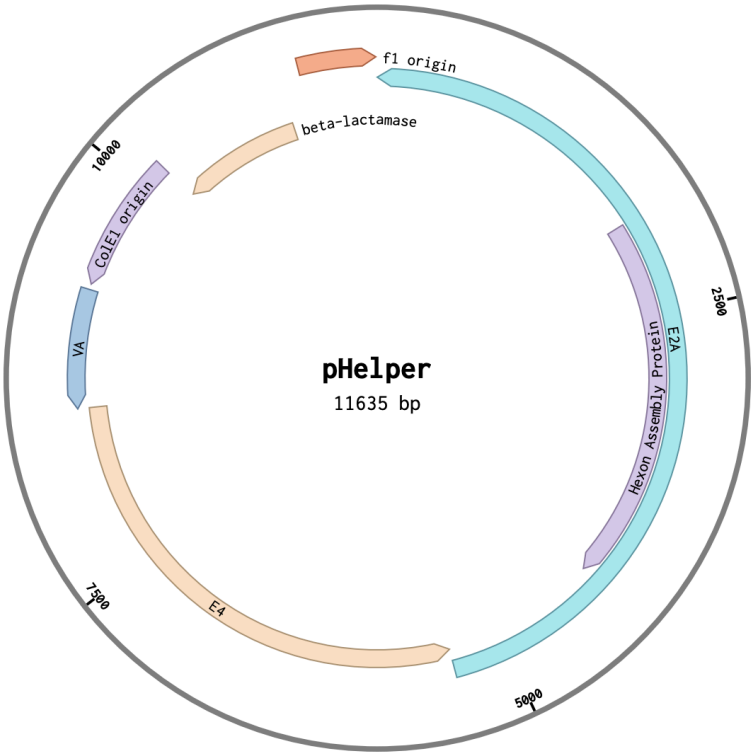
cccgtgtaaaacgacggccagtttatctagtcagcttgattctagctgatcgtggaccg
gaaggtgagccagtgagttgattgcagtccagttacgctggagtctgaggctcgtcctg
aatgatatgcgaccgccggagggttgcgtttgagacgggcgacagatccagtcgcgctg
ctctcgtcgatccgctagggcggccgctctagaactagtggatcccccggaagatcaga
agttcctattccgaagttcctattctctagaaagtataggaacttctgatctgcgcagc
cgccatgccggggttttacgagattgtgattaaggtccccagcgaccttgacgagcatc
tgcccggcatttctgacagctttgtgaactgggtggccgagaaggaatgggagttgccg
ccagattctgacatggatctgaatctgattgagcaggcacccctgaccgtggccgagaa
gctgcagcgcgactttctgacggaatggcgccgtgtgagtaaggccccggaggcccttt

tctttgtgcaatttgagaagggagagagctacttccacatgcacgtgctcgtggaaacc
accggggtgaaatccatggttttgggacgtttcctgagtcagattcgcgaaaaactgat
tcagagaatttaccgcgggatcgagccgactttgccaaactggttcgcggtcacaaaga
ccagaaatggcgccggaggcgggaacaaggtggtggatgagtgctacatccccaattac
ttgctccccaaaacccagcctgagctccagtgggcgtggactaatatggaacagtattt
aagcgcctgtttgaatctcacggagcgtaaacggttggtggcgcagcatctgacgcacg
tgtcgcagacgcaggagcagaacaaagagaatcagaatcccaattctgatgcgccggtg
atcagatcaaaaacttcagccaggtacatggagctggtcgggtggctcgtggacaaggg
gattacctcggagaagcagtggatccaggaggaccaggcctcatacatctccttcaatg
cggcctccaactcgcggtcccaaatcaaggctgccttggacaatgcgggaaagattatg
agcctgactaaaaccgcccccgactacctggtgggccagcagcccgtggaggacatttc
cagcaatcggatttataaaattttggaactaaacgggtacgatccccaatatgcggctt
ccgtctttctgggatgggccacgaaaaagttcggcaagaggaacaccatctggctgttt
gggcctgcaactaccgggaagaccaacatcgcggaggccatagcccacactgtgccctt
ctacgggtgcgtaaactggaccaatgagaactttcccttcaacgactgtgtcgacaaga
tggtgatctggtgggaggaggggaagatgaccgccaaggtcgtggagtcggccaaagcc
attctcggaggaagcaaggtgcgcgtggaccagaaatgcaagtcctcggcccagataga
cccgactcccgtgatcgtcacctccaacaccaacatgtgcgccgtgattgacgggaact
caacgaccttcgaacaccagcagccgttgcaagaccggatgttcaaatttgaactcacc
cgccgtctggatcatgactttgggaaggtcaccaagcaggaagtcaaagacttttccg
gtgggcaaaggatcacgtggttgaggtggagcatgaattctacgtcaaaaagggtggag
ccaagaaaagacccgcccccagtgacgcagatataagtgagcccaaacgggtgcgcgag
tcagttgcgcagccatcgacgtcagacgcggaagcttcgatcaactacgcagacaggta
ccaaaacaaatgttctcgtcacgtgggcatgaatctgatgctgtttccctgcagacaat
gcgagagaatgaatcagaattcaaatatctgcttcactcacggacagaaagactgttta
gagtgctttcccgtgtcagaatctcaacccgtttctgtcgtcaaaaaggcgtatcagaa
actgtgctacattcatcatatcatgggaaaggtgccagacgcttgcactgcctgcgatc
tggtcaatgtggatttggatgactgcatctttgaacaataaatgatttaaatcaggtct
cgctgccgatggttatcttccagattggctcgaggacactctctctgaaggaataagac
agtggtggaagctcaaacctggcccaccaccaccaaagcccgcagagcggcataaggac
gacagcaggggtcttgtgcttcctgggtacaagtacctcggacccttcaacggactcga
caagggagagccggtcaacgaggcagacgccgcggccctcgagcacgacaaagcctacg
accggcagctcgacagcggagacaacccgtacctcaagtacaaccacgccgacgcggag
tttcaggagcgccttaaagaagatacgtcttttgggggcaacctcggacgagcagtctt
ccaggcgaaaaagagggttcttgaacctctgggcctggttgaggaacctgttaagacgg
ctccgggaaaaaagaggccggtagagcactctcctgtggagccagactcctcctcggga
accggaaaggcgggccagcagcctgcaagaaaaagattgaattttggtcagactggaga
cgcagactcagtacctgacccccagcctctcggacagccaccagcagccccctctggtc
tgggaactaatacgatggctacaggcagtggcgcaccaatggcagacaataacgagggc
gccgacggagtgggtaattcctcgggaaattggcattgcgattccacatggatgggcga
cagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaaccacctct
acaaacaaatttccagccaatcaggagcctcgaacgacaatcactactttggctacagc
acccttgggggtattttgacttcaacagattccactgccacttttcaccacgtgactg
gcaaagactcatcaacaacaactggggattccgacccaagagactcaacttcaagctct
ttaacattcaagtcaaagaggtcacgcagaatgacggtacgacgacgattgccaataac
cttaccagcacggttcaggtgtttactgactcggagtaccagctcccgtacgtcctcgg
ctcggcgcatcaaggatgcctcccgccgttcccagcagacgtcttcatggtgccacagt
atggatacctcaccctgaacaacgggagtcaggcagtaggacgctcttcattttactgc

ctggagtactttccttctcagatgctgcgtaccggaaacaactttaccttcagctacac
ttttgaggacgttcctttccacagcagctacgctcacagccagagtctggaccgtctca
tgaatcctctcatcgaccagtacctgtattacttgagcagaacaaacactccaagtgga
accaccacgcagtcaaggcttcagttttctcaggccggagcgagtgacattcgggacca
gtctaggaactggcttcctggaccctgttaccgccagcagcgagtatcaaagacatctg
cggataacaacaacagtgaatactcgtggactggagctaccaagtaccacctcaatggc
agagactctctggtgaatccgggcccggccatggcaagccacaaggacgatgaagaaaa
gtttttcctcagagcggggttctcatctttgggaagcaaggctcagagaaaacaaatg
tggacattgaaaaggtcatgattacagacgaagaggaaatcaggacaaccaatcccgtg
gctacggagcagtatggttctgtatctaccaacctccagagaggcaacagacaagcagc
taccgcagatgtcaacacacaaggcgttcttccaggcatggtctggcaggacagagatg
tgtaccttcaggggcccatctgggcaaagattccacacacggacggacattttcacccc
tctcccctcatgggtggattcggacttaaacaccctcctccacagattctcatcaagaa
caccccggtacctgcgaatccttcgaccaccttcagtgcggcaaagtttgcttccttca
tcacacagtactccacgggacaggtcagcgtggagatcgagtgggagctgcagaaggaa
aacagcaaacgctggaatcccgaaattcagtacacttccaactacaacaagtctgttaa
tgtggactttactgtggacactaatggcgtgtattcagagcctcgccccattggcacca
gatacctgactcgtaatctgtaattgcttgttaatcaataaaccgtttaattcgtttca
gttgaactttggtctctgcgtatttctttcttatctagtttccatggctacgtagataa
gtagcatggcgggttaatcattaactacagcccgggcgtttaaacagcgggcggagggg
tggagtcgtgacgtgaattacgtcatagggttagggaggtcctgtattagaggtcacgt
gagtgttttgcgacattttgcgacaccatgtggtctcgctgggggggggggcccgagtg
agcacgcagggtctccattttgaagcgggaggtttgaacgagcgctggcgcgctcactg
gccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgcct
tgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcc
cttcccatgcatcggccgcaaatacctgcaggggtgtggaaagtccccaggctccccag
caggcagaagtatgcaaagcatgcatctcaattagtcagcaaccaggtgtggaaagtcc
ccaggctccccagcaggcagaagtatgcaaagcatgcatctcaattagtcagcaaccat
agtcccgccctaactccgcccatcccgcccctaactccgcccagttccgcccattctc
cgccccatggctgactaatttttttttatttatgcagaggccgaggccgcctcggcctct
gagctattccagaagtagtgaagaggcttttttggaggagatctaagcttgaagcaagc
ctcctgaaagatgaagctactgtcttctatcgaacaagcatgcgatatttgccgactta
aaaagctcaagtgctccaaagaaaaaccgaagtgcgccaagtgtctgaagaacaactgg
gagtgtcgctactctcccaaaaccaaaaggtctccgctgactagggcacatctgacaga
agtggaatcaaggctagaaagactggaacagctatttctactgatttttcctcgagaag
accttgacatgattttgaaaatggattctttacaggatataaaagcattgttaacagga
ttatttgtacaagataatgtgaataaagatgccgtcacagatagattggcttcagtgga
gactgatatgcctctaacattgagacagcatagaataagtgcgacatcatcatcggaag
agagtagtaacaaaggtcaaagacagttgactgtatcgccggaattcccggggatctgg
gcccccccgaccgatgtcagcctgggggacgagctccacttagacggcgaggacgtggc
gatggcgcatgccgacgcgctagacgatttcgatctggacatgttgggggacggggatt
ccccgggtccgggatttacccccacgactccgccccctacggcgctctggatatggcc
gacttcgagtttgagcagatgtttaccgatgcccttggaattgacgagtacggtgggta
gggggcgcgaccggacccgcatcccccgtctgggttttcccctcccgtccccggttcgt
atccacaataaacacgagcacatacattacaaaaacctgcggttgtcgtctgattattt
ggtggatccgggggagatgggggaggctaactgaaacacggaaggagacaataccggaa
ggaacccgcgctatgacggcaataaaaagacagaataaaacgcacgggtgttgggtcgt
ttgttcataaacgcggggttcggtcccagggctggcactctgtcgataccccaccgaga

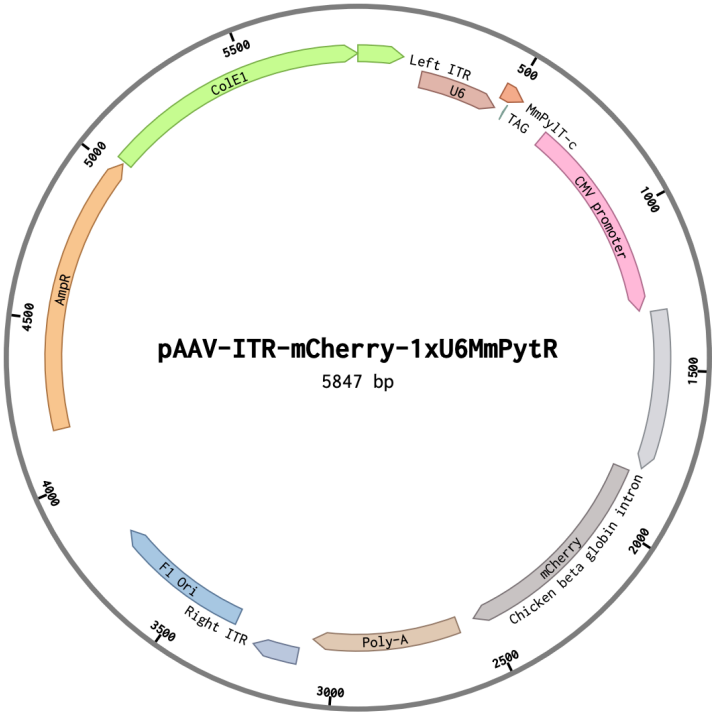
ccccattggggccaatacgcccgcgtttcttcctttccccaccccacccccaagttc
gggtgaaggcccagggctcgcagccaacgtcggggcggcaggccctgccatagccactg
gccccgtgggttagggacggggtcccccatggggaatggtttatggttcgtgggggtta
ttattttgggcgttgcgtggggtcaggtccacgacccaagcttgtcgacggtaccccgg
ggaattcgagctctagattaagtaatgatcataatcagccatatcacatctgtagaggt
tttacttgctttaaaaaacctcccacacctccccctgaacctgaaacataaaatgaatg
caattgttgttgttaacttgtttattgcagcttataatggttacaaataaagcaatagc
atcacaaatttcacaaataaagcatttttttcactgcattctagttgtggtttgtccaa
actcatcaatgtatcttatcatgtctggatcctaggggtgcgagcggatcgagcagtg
tcgatcactactggaccgcgagctgtgctgcgacccgtgatcttacggcattatacgta
tgatcggtccacgatcagctagattatctagtcagcttgatgtcatagctgtttcctga
ggctcaatactgaccatttaaatcatacctgacctccatagcagaaagtcaaaagcctc
cgaccggaggcttttgacttgatcggcacgtaagaggttccaactttcaccataatgaa
ataagatcactaccgggcgtatttttgagttatcgagattttcaggagctaaggaagc
taaaatgagccatattcaacgggaaacgtcttgcttgaagccgcgattaaattccaaca
tggatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatcaggtgcg
acaatctatcgattgtatgggaagcccgatgcgccagagttgtttctgaaacatggcaa
aggtagcgttgccaatgatgttacagatgagatggtcaggctaaactggctgacggaat
ttatgcctcttccgaccatcaagcattttatccgtactcctgatgatgcatggttactc
accactgcgatcccagggaaaacagcattccaggtattagaagaatatcctgattcagg
tgaaaatattgttgatgcgctggcagtgttcctgcgccggttgcattcgattcctgttt
gtaattgtccttttaacggcgatcgcgtatttcgtctcgctcaggcgcaatcacgaatg
aataacggtttggttggtgcgagtgattttgatgacgagcgtaatggctggcctgttga
acaagtctggaaagaaatgcataaactcttgccattctcaccggattcagtcgtcactc
atggtgatttctcacttgataaccttatttttgacgaggggaaattaataggttgtatt
gatgttggacgagtcggaatcgcagaccgataccaggatcttgccatcctatggaactg
cctcggtgagttttctccttcattacagaaacggcttttcaaaaatatggtattgata
atcctgatatgaataaattgcagtttcacttgatgctcgatgagtttttctaatgagga
cctaaatgtaatcacctggctcaccttcgggtgggcctttctgcgttgctggcgttttt
ccataggctccgccccctgacgagcatcacaaaaatcgatgctcaagtcagaggtggc
gaaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgc
tctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaag
cgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgct
ccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggt
aactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccac
tggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggt
ggcctaactacggctacactagaagaacagtatttggtatctgcgctctgctgaagcca
gttacctcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagc
ggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaggatctcaagaaga
tcctttgatttctaccgaagaaaggccca

pHelper



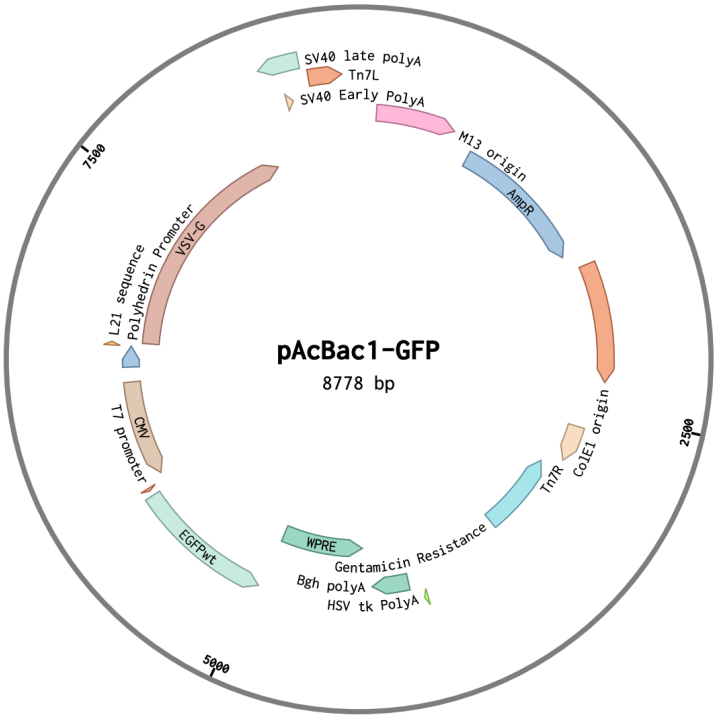
The sequence for this plasmid is the same as the pHelper sequence found in the Chapter 2 Plasmids section of this appendix.

pAAV-mCherry+1xMmPytR



The sequence for this plasmid is the same as the pAAV-mCherry+1xMmPytR sequence found in the Chapter 2 Plasmids section of this appendix.

pAcBac1-GFPwt



Sequence color-coding key

Feature	Color
M13 origin	text
AmpR	text
ColE1 origin	text
Tn7R	text
GentamicinR	text
HSV tk PolyA	text
Bgh polyA	text
WPRE	text
EGFPwt	text
T7 promoter	text
CMV promoter	text
Polyhedrin Promoter	text
L21 sequence	text
VSV-G	text
SV40 late polyA	text
Tn7L	text

Sequence

ttctctgtcacagaatgaaaatttttctgtcatctcttcgttattaatgtttgtaattg
actgaatatcaacgcttatttgcagcctgaatggcgaatgggacgcgccctgtagcggc

gcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgc
cctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttc
cccgtcaagctctaaatcgggggctcccttttagggttccgatttagtgctttacggcac
ctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgata
gacggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgttcc
aaactggaacaacactcaaccctatctcggtctattcttttgatttataagggattttg
ccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattt
taacaaaatattaacgtttacaatttcaggtggcacttttcggggaaatgtgcgcggaa
cccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataa
ccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccg
tgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaa
cgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaa
ctggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaat
gatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggc
aagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcacca
gtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccat
aaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaagg
agctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaa
ccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaat
ggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaac
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggccctt
ccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtat
cattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacgg
ggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactg
attaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaa
acttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgacca
aaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaa
ggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaacc
accgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaagg
taactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagtta
ggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgtt
accagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgat
agttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagc
ttggagcgaacgacctacaccgaactgagatacctacagcgtgagcattgagaaagcgc
cacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacag
gagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcggg
tttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcct
atggaaaaacgccagcaacgcggcctttttacggttcctggccttttgctggccttttg
ctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgccttt
gagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcga
ggaagcggaagagcgcctgatgcggtattttctccttacgcatctgtgcggtatttcac
accgcagaccagccgcgtaacctggcaaaatcggttacggttgagtaataaatggatgc
cctgcgtaagcgggtgtgggcggacaataaagtcttaaactgaacaaaatagatctaaa
ctatgacaataaagtcttaaactagacagaatagttgtaaactgaaatcagtccagtta
tgctgtgaaaaagcatactggacttttgttatggctaaagcaaactcttcattttctga
agtgcaaattgcccgtcgtattaaagaggggcgtggccaagggcatggtaaagactata
ttcgcggcgttgtgacaatttaccgaacaactccgcggccgggaagccgatctcggctt
gaacgaattgttaggtggcggtacttgggtcgatatcaaagtgcatcacttcttcccgt

atgcccaactttgtatagagagccactgcgggatcgtcaccgtaatctgcttgcacgta
gatcacataagcaccaagcgcgttggcctcatgcttgaggagattgatgagcgcggtgg
caatgccctgcctccggtgctcgccggagactgcgagatcatagatatagatctcacta
cgcggctgctcaaacctgggcagaacgtaagccgcgagagcgccaacaaccgcttcttg
gtcgaaggcagcaagcgcgatgaatgtcttactacggagcaagttcccgaggtaatcgg
agtccggctgatgttgggagtaggtggctacgtctccgaactcacgaccgaaaagatca
agagcagcccgcatggatttgacttggtcagggccgagcctacatgtgcgaatgatgcc
catacttgagccacctaactttgttttagggcgactgccctgctgcgtaacatcgttgc
tgctgcgtaacatcgttgctgctccataacatcaaacatcgacccacggcgtaacgcgc
ttgctgcttggatgcccgaggcatagactgtacaaaaaaacagtcataacaagccatga
aaaccgccactgcgccgttaccaccgctgcgttcggtcaaggttctggaccagttgcgt
gagcgcatacgctacttgcattacagtttacgaaccgaacaggcttatgtcaactgggt
tcgtgccttcatccgtttccacggtgtgcgtcacccggcaaccttgggcagcagcgaag
tcgaggcatttctgtcctggctggcgaacgagcgcaaggtttcggtctccacgcatcgt
caggcattggcggccttgctgttcttctacggcaaggtgctgtgcacggatctgccctg
gcttcaggagatcggtagacctcggccgtcgcggcgcttgccggtggtgctgaccccgg
atgaagtggttcgcatcctcggttttctggaaggcgagcatcgtttgttcgcccaggac
tctagctatagttctagtggttggcctacgtacccgtagtggctatggcagggcttgcg
cttaatgcgccgctacagggcgcgtggggataccccctagagccccagctggttctttc
cgcctcagaagccatagagcccaccgcatccccagcatgcctgctattgtcttcccaat
cctccccttgctgtcctgccccaccccaccccccagaatagaatgacacctactcaga
caatgcgatgcaatttcctcattttattaggaaaggacagtgggagtggcaccttccag
ggtcaaggaaggcacgggggagggcaaacaacagatggctggcaactagaaggcacag
tcgaggctgatcagcgggtttaaacgggccctctagactcgagttaaagtcgacgcggg
gaggcggcccaaagggagatccgactcgtctgagggcgaaggcgaagacgcggaagagg
ccgcagagccggcagcaggccgcgggaaggaaggtccgctggattgagggccgaaggga
cgtagcagaaggacgtcccgcgcagaatccaggtggcaacacaggcgagcagccaagga
aaggacgatgatttccccgacaacaccacggaattgtcagtgcccaacagccgagcccc
tgtccagcagcgggcaaggcaggcggcgatgagttccgccgtggcaatagggaggggga
aagcgaaagtcccggaaaggagctgacaggtggtggcaatgccccaaccagtgggggtt
gcgtcagcaaacacagtgcacaccacgccacgttgcctgacaacgggccacaactcctc
ataaagagacagcaaccaggatttatacaaggaggagaaaatgaaagccatacgggaag
caatagcatgatacaaaggcattaaagcagcgtatccacatagcgtaaaaggagcaaca
tagttaagaataccagtcaatctttcacaaattttgtaatccagaggttgattgtcgac
ttaacgcgttgaattctcattaatggtgatggtgatgatgaccggtatgcatattcaga
tcctcttctgagatgagttttgttcgaagggccccttgtacagctcgtccatgccgag
agtgatcccggcggcggtcacgaactccagcaggaccatgtgatcgcgcttctcgttgg
ggtctttgctcagggcggactgggtgctcaggtagtggttgtcgggcagcagcacgggg
ccgtcgccgatgggggtgttctgctggtagtggtcggcgagctgcacgctgccgtcctc
gatgttgtggcggatcttgaagttcaccttgatgccgttcttctgcttgtcggccatga
tatagacgttgtggctgttgtagttgtactccagcttgtgccccaggatgttgccgtcc
tccttgaagtcgatgcccttcagctcgatgcggttcaccagggtgtcgccctcgaactt
cacctcggcgcgggtcttgtagttgccgtcgtccttgaagaagatggtgcgctcctgga
cgtagccttcgggcatggcggacttgaagaagtcgtgctgcttcatgtggtcggggtag
cggctgaagcactgcacgccgtaggtcagggtggtcacgagggtgggccagggcacggg
cagcttgccggtggtgcagatgaacttcagggtcagcttgccgtaggtggcatcgccct
cgccctcgccggacacgctgaacttgtggccgtttacgtcgccgtccagctcgaccagg
atgggcaccaccccggtgaacagctcctcgcccttgctcaccatggtggcggcgctagc

cagcttgggtctccctatagtgagtcgtattaatttcgataagccagtaagcagtgggt
tctctagttagccagagagctctgcttatatagacctcccaccgtacacgcctaccgcc
catttgcgtcaatggggcggagttgttacgacattttggaaagtcccgttgattttggt
gccaaaacaaactcccattgacgtcaatggggtggagacttggaaatccccgtgagtca
aaccgctatccacgcccattgatgtactgccaaaaccgcatcaccatggtaatagcgat
gactaatacgtagatgtactgccaagtaggaaagtcccataaggtcatgtactgggcat
aatgccaggcgggccatttaccgtcattgacgtcaataggggggcgtacttggcatatga
tacacttgatgtactgccaagtgggcagtttaccgtaaatagtccacccattgacgtca
atggaaagtccctattggcgttactatgggaacatacgtcattattgacgtcaatgggc
gggggtcgttgggcggtcagccaggcgggccatttaccgtaagttatgtaacgcggaac
tccatatatgggctatgaactaatgaccccgtaattgattactattaataactagtcaa
taatcaatgtcaacgcgtatatctggcccgtacatcgcgaagcagcgcaaaacggatcc
tgcaggtatttgcggccgcggtccgtatactccggaatattaatagatcatggagataa
ttaaaatgataaccatctcgcaaataaataagtattttactgttttcgtaacagttttg
taataaaaaaacctataaatattccggattattcataccgtcccaccatcgggcgcgaa
ctcctaaaaaaccgccaccatgaagtgccttttgtacttagccttttttattcattgggg
tgaattgcaagttcaccatagttttccacacaaccaaaaaggaaactggaaaaatgtt
ccttctaattaccattattgcccgtcaagctcagatttaaattggcataatgacttaat
aggcacagccttacaagtcaaaatgcccaagagtcacaaggctattcaagcagacggtt
ggatgtgtcatgcttccaaatgggtcactacttgtgatttccgctggtatggaccgaag
tatataacacattccatccgatccttcactccatctgtagaacaatgcaaggaaagcat
tgaacaaacgaaacaaggaacttggctgaatccaggcttccctcctcaaagttgtggat
atgcaactgtgacggatgccgaagcagtgattgtccaggtgactcctcaccatgtgctg
gttgatgaatacacaggagaatgggttgattcacagttcatcaacggaaaatgcagcaa
ttacatatgccccactgtccataactctacaacctggcattctgactataaggtcaaag
ggctatgtgattctaacctcatttccatggacatcaccttcttctcagaggacggagag
ctatcatccctgggaaaggagggcacagggttcagaagtaactactttgcttatgaaac
tggaggcaaggcctgcaaaatgcaatactgcaagcattggggagtcagactcccatcag
gtgtctggttcgagatggctgataaggatctctttgctgcagccagattccctgaatgc
ccagaagggtcaagtatctctgctccatctcagacctcagtggatgtaagtctaattca
ggacgttgagaggatcttggattattccctctgccaagaaacctggagcaaaatcagag
cgggtcttccaatctctccagtggatctcagctatcttgctcctaaaaacccaggaacc
ggtcctgctttcaccataatcaatggtaccctaaaatactttgagaccagatacatcag
agtcgatattgctgctccaatcctctcaagaatggtcggaatgatcagtggaactacca
cagaaagggaactgtgggatgactgggcaccatatgaagacgtggaaattggacccaat
ggagttctgaggaccagttcaggatataagtttcctttatacatgattggacatggtat
gttggactccgatcttcatcttagctcaaaggctcaggtgttcgaacatcctcacattc
aagacgctgcttcgcaacttcctgatgatgagagtttatttttggtgatactgggcta
tccaaaaatccaatcgagcttgtagaaggttggttcagtagttggaaaagctctattgc
ctcttttttctttatcatagggttaatcattggactattcttggttctccgagttggta
tccatctttgcattaaattaaagcacaccaagaaaagacagatttatacagacatagag
atgaaccgacttggaaagtgataaggccaggccggccaagcttgtcgagaagtactaga
ggatcataatcagccataccacatttgtagaggttttacttgctttaaaaaacctccca
cacctccccctgaacctgaaacataaaatgaatgcaattgttgttgttaacttgtttat
tgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcat
tttttcactgcattctagttgtggtttgtccaaactcatcaatgtatcttatcatgtc
tggatctgatcactgcttgagcctaggagatccgaaccagataagtgaaatctagttcc
aaactatttttgtcatttttaatttttcgtattagcttacgacgctacacccagttcccat

ctattttgtcactcttccctaaataatccttaaaaactccatttccaccccctcccagtt
cccaactattttgtccgcccacagcggggcatttttcttcctgttatgttttaatcaa
acatcctgccaactccatgtgacaaaccgtcatcttcggctacttt

Appendix II

Oligonucleotide primers

Chapter 2 primers

Primer name	Oligonucleotide sequence (5' → 3')
Consensus-splice-donor-site-sequence-F	CGCGGAAGCTTCGATCAACTACGCAGACAGGTAAGTAAACAAATGTTCTCGTCACGTGGG
NheI-p40-F	ATTATTAGCTAGCAGTTGCGCAGCCATCGACGTCAGACCGGAAGCTTCGATCAACTACG
Cap-SalI-R	ATTATTAGTCGACTTACAGATTACGAGTCAGGTATCTGGCAATGG
SbfI-Rep-F	AATTAATTCCTGCAGGCTCTAGAACTAGTGGATCCCCCGGAAGATC
NotI-Rep-R	AATTAATTGCGGCCGCTCAGAGAGAGTGTCCTCGAGCCAATCTGG

Chapter 3 primers

Primer name	Oligonucleotide sequence (5' → 3')
AAPstop60-F	GGATGGGCGACAGAGTGATCACCACCAGCACC
AAPstop60-R	GGTGCTGGTGGTGATCACTCTGTCGCCCATCC
RC2int-HindIII-F	CGTCAGACGCGGAAGCTTCGATCAAC
RC2-SbfI-R	GCAAATACCTGCAGGATCCGTTTTGCGCTG
NheI-Kozak-AAP-WT-F	ATTATTGCTAGCGCCGCCACCATGGCCTAGCTGGAGACGCAGACTCAGTACCTGACC
XhoI-AAP-R	ATTATTCTCGAGTCAGGGTGAGGTATCCATACTGTGGCACC
AAP-L12TAG-F	CCTGACCCCCAGCTAGTCGGACAGCC
AAP-L12TAG-R	GGTGGCTGTCCGACTAGCTGGGGGTCAGG
AAP-T78TAG-F	GGGCCCTGCCCACCTACATAGACCACCTCTACAAACAAATTTCC
AAP-T78TAG-R	GTTTGTAGAGGTGGTCTATGTAGGTGGGCAGGGCCCAGGTTCGGG
AAP-T124TAG-F	CCACGTGACTGGCAAAGACTAGTCAACAACAACTGGGG
AAP-T124TAG-R	CCAGTTGTTGTTGACTAGTCTTTGCCAGTCACGTGGTG
AAP-T177TAG-F	GTACCAGCTCCCGTTAGTCCTCGGCTCGGCGCA
AAP-T177TAG-R	CCGAGCCGAGGACTAACGGGAGCTGGTACTCCG
AAP-T97TAG-F	CGAACGACAATCACTTAGTTGGCTACAGCACCCCTTGG
AAP-T97TAG-R	GGGGTGCTGTAGCCAACTAAGTGATTGTCGTTCGAGG
AAP-T110TAG-F	GGGTATTTTGACTTCATAGGATTCCACTGCCACTTTTC
AAP-T110TAG-R	GTGGCAGTGGAATCCTATGAAGTCAAAATACCCC
AAP-T78TGA-F	GGGCCCTGCCCACCTACATGAACCACCTCTACAAACAAATTTCC
AAP-T78TGA-R	GTTTGTAGAGGTGGTTCATGTAGGTGGGCAGGGCCCAGGTTCGGG
AAP-T97TGA-F	CGAACGACAATCACTTGATTGGTAACAGCACCCCTTGG
AAP-T97TGA-R	GGGGTGCTGTAGCCAATCAAGTGATTGTCGTTCGAGG
AAP-T110TGA-R	GTGGCAGTGGAATCTCATGAAGTCAAAATACCCC
AAP-T110TGA-F	GGGTATTTTGACTTAATGAGATTCCACTGCCACTTTTC
AAP-T124TGA-F	CCACGTGACTGGCAAAGACTGATCAACAACAACTGGGG
AAP-T124TGA-R	CCAGTTGTTGTTGATCAGTCTTTGCCAGTCACGTGGTG
AvrII-CMV-WRS-F	ATTATTCCTAGGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCC
SpeI-WRS-R	AATAATACTAGTCCATAGAGCCCACCGCATCCCCAGC

Chapter 4 primers

Primer name	Oligonucleotide sequence (5' → 3')
MTH-mCherry-SbfI-F	ATTATTCCTGCAGGACCATGGCTCGAGATCCCGGGTGATC
MTH-mCherry-NotI-R	AATAATGCGGCCAGACATGATAAGATACATTGATGAGTTTG GACAAACC
NheI-linker-EGFP-F	TAATAATGCTAGCAGGGCGGCTCCGTGAGCAAGGGCGAGGA GCTGTTCACCGGG
GFP-EcoRI-6xHis-R	ATTATTGAATTCCTCATTAATGGTGATGGTGATGATGACCG G
NheI-Kozak-mCherry-F	ATTATTGCTAGCGCCGCCACCATGGTGAGCAAGGGCGAGGA GG
mCherry-linker-AvrII-R	TAATAATCCTAGGAGCCGCCCTTGTACAGCTCGTCCATGCC GCCGGTGGAGTGG
Gibson-mCherry-F	GTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTGGCCAA GGAGGCCACCATGGTGAGC
Gibson-WPRE-R	GTCGATCGACCACTGTGCTGGCGGCCGGCCAGGCCGCGGGG AGG
Gibson-UbiC-F	CCAGTAAGCAGTGGGTTCTCTAGTTAGCCAGAGAGCTCTAG ACCAAGTGACGATCACAGC
Gibson-Ubic-R	GTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAG TGGCCTCCGCGCCGG
NheI-Lemke-NES-MbPylRS-F	ATTATTGCTAGCGCCACCATGGCCTGCCCAGTGCCCCTTCA GTTGCCTCCGCTGGAGAGACTGACCCTCGACGATAAAAAAC CATTAGATGTTTTAATATCTGCGACCGG
MbPylRS-WPRE-XhoI-R	AATAATCTCGAGTTAAAGTCGACGCGGGG
L274A-F	GCTTGCCCCGACTCTTGCCAACTATGTGCGAAAACTCGATA GGATTTTACCAGGCCC
L274A-mut-R	CCATGGTTTGTCAATACCCCATTCTCTATCAAGAGACCGTG GCCCGACGACTGC
V370R-F	CTTCGGCAGTCGTCGGGCCACGGTCTCTTGATAGAGAATGG GGTATTGACAAACCATG
V730R-mut-R	CCATGGTTTGTCAATACCCCATTCTCTATCAAGAGACCGTG GCCCGACGACTGC
C313V-M315Y-F	GAGCACCTGGAAGAATTTACTATGGTGAACTTCGTGCAGTA CGGTTCGGGATGTACTCGGGAAAATCTTG
C313V-M315Y-mut-R	CAAGATTTTCCCGAGTACATCCCGAACCGTACTGCACGAAG TTCACCATAGTAAATTC
MbPylRS-Nterm-F	ATTATTGCTAGCGCCACCAATGGATAAAAAACCATTAG
MbPylRS-Cterm-R	AATAATGAATTCTTACAGATTGGTTGAAATCCCATTATAGT AAGATTCGGACC

Chapter 5 primers

Primer name	Oligonucleotide sequence (5' → 3')
XbaI-eGFP-R	AATAATTCTAGATCATTAATGGTGATGGTGATGGTGATGATGACCGG
SphI-Kozak-GFP-F	ATTATTGCATGCGCCGCCACCGCTGGCTAGCGCCGCCACCATGGTGAGCAAGGGCG
TATA-GFP-F	ATTATTGCATGCATTTATAAAAGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGG
YB-TATA-GFP-F	ATTATTGCATGCATTTCTAGAGGGTATATAATGGGGGCCAAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGC
min-CMV-GFP-F	ATTATTGCATGCATTGTAGGCGTGTACGGTGGGAGGTGTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCAAATTAATACGACTCACTATAGGGACCAAGCTGG
KpnI-GFP-R	AATAATGGTACCTCAATGGTGATGGTGATGATGACCGGTATGC
SphI-VP1-only-F	ATTATTGCATGCATGGCTGCCGATGGTTATCTTCCAGATTG
XbaI-VP1-only-R	AATAATTCTAGATGTAGTTAATGATTAACCCGCCATGCTACTTATCTACG
SphI-TATA-VP1-only-F	ATTATTGCATGCTATAAAAGAAATATGGCTGCCGATGGTTATCTTCCAGATTG
SphI-YBTATA-VP1-only-F	TATTGCATGCTCTAGAGGGTATATAATGGGGGCCAATGGCTGCCGATGGTTATCTTCCAG
SphI-minCMV-VP1-only-F	ATTATTGCATGCGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCATGGCTGCCGATGGTTATCTTCCAGATT
SbfI-Gal4-VP16-F	ATTATTCCTGCAGGGGTGTGGAAAGTCCCCAGGCTCCCC
AvrII-Gal4-VP16-R	AATAATCCTAGGGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACC