2013

Orthogonal Expression of Metabolic Pathways

George Howard McArthur IV
Virginia Commonwealth University

Orthogonal Expression of Metabolic Pathways

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy in Engineering

at

VIRGINIA COMMONWEALTH UNIVERSITY

by
George H. McArthur IV
B.S., University of Virginia, 2008
M.S., Virginia Commonwealth University, 2009
Whitaker Fellow, Imperial College London, 2011-12

Directed by Stephen S. Fong, PhD
Associate Professor, Chemical and Life Science Engineering
Fellow, Center for the Study of Biological Complexity

Virginia Commonwealth University
Richmond, Virginia, USA
April 2013

This dissertation is dedicated to my lovely wife

CHRISTINA NOELLE

who holds my hand in life and
who has already made many of my childhood dreams come true

and to my fantastic parents

ROSIE and GEORGE

who first encouraged me to dream and
who gave me the tools to realize my dreams

Acknowledgments

As with life, working toward and earning a PhD is not something one can truly do on his own. I have been shaped and will continue to be shaped by those around me – and I am incredibly thankful to have been blessed by the family, friends, colleagues and mentors who have helped make me into the person I am and who have urged me to accomplish my goals. This is especially true of my wife and my parents and I am deeply thankful for them.

The work described in this document was enabled by many people and funding sources. I am grateful for the plasmid DNA that I have received as gifts from members of the synthetic biology community: pSB1C3 from the Registry of Standard Biological Parts, pJ401-RFP from Travis Bayer, pAC-LYC (and strain EcHW2f) from Harris Wang and the synthesis of pGM-O-RFP was paid for by Jean Peccoud. Oligonucleotides used to build out the promoter libraries were paid for by a NASA grant to Stephen Fong.

I am thankful for receiving the Whitaker International Fellowship and the time I spent in Travis Bayer's lab. My year at Imperial College London's Centre for Synthetic Biology and Innovation helped me mature into a fearless experimentalist, willing to fail and to learn from my failures. During that time I learned from countless individuals in the lab, but am most thankful for the mentorship of Eric Davidson, a postdoc in the Bayer Lab.

My graduate research (here and outside the scope of this dissertation) has been challenged and refined by my students, my instructors and my peers. I have enjoyed mentoring students of all ages (high school through PhD), including advising several iGEM teams, and I am thankful for the relationships made and the lessons learned through these experiences. My committee members have inspired me in a myriad of ways to become a better student, researcher and leader. All of my lab mates have encouraged me even through the most discouraging times and I have enjoyed a constructive, friendly atmostphere in lab throughout my graduate work. In particular, I want to thank three individuals. Mike, thanks for pushing and pushing and pushing me to accept who I am and urging me to find my own path. Adam, thanks for keeping me engaged daily and for your tremendous help at the bench. Chris, thanks for helping me get my hands dirty at the beginning of this journey and for sticking with me ever since.

I enrolled at VCU in 2008 to work with Stephen Fong believing that it would be a great experience. It ended up being even better than I had imagined. The intersection of synthetic biology, systems biology and metabolic engineering is a fascinating area in which to carry out research and it has been a blast thinking through problems with Steve. Perhaps even more important, however, is that Steve has helped me become a better person by simply being a great example of a wonderful human being. Steve, thank you for investing in me, for cultivating good qualities in me and for being patient with me.

Dissertation Committee

Accepted by

Stephen S. Fong, PhD
Associate Professor, Department of Chemical and Life Science Engineering
Fellow, Center for the Study of Biological Complexity

Accepted by

B. Frank Gupton, PhD
Chair, Department of Chemical and Life Science Engineering
Research Professor, Department of Chemistry

Accepted by

Allison A. Johnson, PhD
Assistant Director, Center for the Study of Biological Complexity

Accepted by

James M. McCollum, PhD
Senior Software Engineer, Google Inc.

Accepted by

Michael H. Peters, PhD
Professor, Department of Chemical and Life Science Engineering
Gross and Developmental Anatomy Faculty, VCU School of Medicine
Associate Member, VCU Massey Cancer Center
Senior Fellow, Center for the Study of Biological Complexity

Table of Contents

List of Abbreviations

CTD – carboxyl terminal domain

CV – coefficient of variation

DMAPP – dimethylallyl diphosphate

DNA – deoxyribonucleic acid

dsDNA – double stranded DNA

DXP – 1-deoxy-D-xylulose-5-phosphate

EMSA – electrophoretic mobility shift assay

IPP – isopentenyl diphosphate

IPTG – isopropyl β-D-1-thiogalactopyranoside

LSD – lysergic acid diethylamide

MCS – multiple cloning site

mRNA – messenger ribonucleic acid

NTP – nucleoside triphosphate

PCR – polymerase chain reaction

PDB – protein data bank

RNA – ribonucleic acid

RNAP – RNA polymerase

SELEX – systematic evolution of ligands by exponential enrichment

ssDNA – single stranded DNA

TSS – transcription start site

UV – ultraviolet

ORTHOGONAL EXPRESSION OF METABOLIC PATHWAYS

By

George H. McArthur IV, PhD

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

Virginia Commonwealth University, 2013

Advisor: Stephen S. Fong, PhD
Associate Professor, Department of Chemical and Life Science Engineering
Fellow, Center for the Study of Biological Complexity

Microbial metabolism can be tailored to meet human specifications, but the degree to which these living systems can be repurposed is still unknown. Artificial biological control strategies are being developed with the goal of enabling the predictable implementation of novel biological functions (e.g., engineered metabolism). This dissertation project contributes genetic tools useful for modulating gene expression levels (extending promoters with UP elements) and isolating transcription and translation of engineered DNA from the endogenous cellular network (expression by orthogonal cellular machinery), which have been demonstrated in *Escherichia coli* for the production of lycopene, a 40-carbon tetraterpene carotenoid with antioxidant activity and a number of other desirable properties.

Chapter I – Introduction

"Noah began to be a man of the soil, and he planted a vineyard. He drank of the wine and became drunk…" – Genesis 9:20-21

Chemicals from microbes – the oldest biotechnology in the world

According to the Biblical narrative, Noah, having survived the great deluge nearly 4,000 years ago, planted a vineyard and made wine. It is unlikely that Noah knew that the sugar in grape juice was chemically transformed to alcohols and other byproducts in microscopic living cells by way of enzyme-facilitated reactions. Indeed, harnessing natural fermentation by yeasts such as *Saccharomyces cerevisiae* to produce food products (e.g., beer, bread, yogurt, cheese) is considered to be the oldest biotechnology in human history, with evidence indicating that the Sumerians may have been brewing beer in what is now Iraq some 8,000 years ago – or even earlier.

The vinification of grape juice is just the start of this story (though, happily, still part of the story). In 1673, the Dutch scientist Antony van Leeuwenhoek fathered the field of microbiology by constructing an improved microscope to visualize and describe single cells, including yeast, for the first time in history. Louis Pasteur, a French chemist and microbiologist, presented convincing evidence in 1861 that ethanol fermentation is inextricably linked to yeast growth. Pasteur also discovered that some bacteria are capable of fermenting butanol and other alcohols. In 1916 at the University of Manchester, Chaim Weizmann isolated *Clostridium acetobutylicum*, a bacterium capable of producing acetone, butanol and ethanol at high yields in a 3:6:1 ratio (a process known as ABE fermentation).

During World War I (1914-1918), the Allies used Weizmann's ABE process – at an industrial scale – to convert maize to acetone, which was used to manufacture various cordite propellants. This marked the beginning of industrial fermentation (although the ABE fermentation is no longer economical due to cheaper petroleum-based processes). Interestingly, biotechnology after World War I aimed to develop sustainable chemical production processes by fermenting agricultural waste into a variety of products. Sir Alexander Fleming publicized the discovery of penicillin-type antibiotics in 1928 – and many other natural products were discovered and produced at an industrial scale in the 20th century that have been medically or industrially useful.

A huge step change occurred after major discoveries were made in the field of molecular biology. Watson and Crick (working with the data of Frankin and others) determined the structure of DNA in 1953. A few years later, Jacob and Monod postulated the existence of messenger RNA (mRNA). Immediately after, Crick and Brenner cracked the genetic code – and Crick articulated the "Central Dogma of Molecular Biology." In the early 1970s, work done by Berg and Cohen & Boyer (Stanford University) culminated in the advent of recombinant DNA technology (unfortunately also known by its misnomer, genetic engineering). Recombinant DNA technology, based on plasmid propagation and cutting and pasting fragments of DNA using restriction enzymes and ligases, marked the beginning of molecular cloning and also the conceptual birth of *synthetic* biology (i.e., building biological systems). The first commercial products arising from this technology, somatostatin and insulin (both human), were produced at an industrial scale by recombinant bacteria. In 1983, Kary Mullis developed the polymerase chain reaction (PCR) to amplify a piece of DNA *in vitro* across several orders of magnitude. These techniques

provided biotechnologists a fairly powerful method of writing DNA. Overproduction of individual proteins is a much easier task than trying to coordinate the expression of multiple enzymes and, as a result, metabolic engineering efforts were put aside in favor of making valuable therapeutic proteins throughout the last two decades of the 20th century.

Economic, environmental and geopolitical concerns are now (again) driving research efforts to replace fossil fuel-based chemical manufacturing with renewable, bio-based processes that are cheaper, greener and able to be carried out entirely domestically[1]. The key components of these processes will be microorganisms that have been engineered to efficiently carry out a desired metabolic function, converting inexpensive carbon substrates (e.g., glucose, CO_2, lignocellulosic biomass) to valuable molecular products. In 2003, the successful engineering of *Escherichia coli* for the production of the monomer 1,3-propanediol by Genencor and DuPont marked an important milestone for metabolic engineering[2]. According to DuPont, biologically produced 1,3-propanediol contributes to about 37% of the mass of Dupont's Sorona polymer fiber and is likely to become the first billion-dollar, non-pharmaceutical industrial biotechnology product.

Writing DNA – the true engineering of genetics

Prior to recombinant DNA technology, industrial microbiologists would induce genomic mutations via UV irradiation or by exposure to chemical mutagens and subsequently screen for an improved phenotype. This completely random approach was usurped by methods that allowed the heterologous expression of foreign DNA encoding enzymes of interest. Although recombinant DNA technology has produced many important and groundbreaking products, it has limited molecular cloning and synthetic biology efforts

because it is dependent on specific DNA sequences to work. The ever-decreasing cost of *de novo* DNA synthesis combined with facile, sequence-independent assembly methods have now opened the doors to a more interesting future for genetic engineering[3]. Instead of cobbling together stories from words and phrases clipped from magazines and newspapers, synthetic biologists now have their printing press, their word processor. The real challenge now is not figuring out *how* to write DNA, but rather determining *what* DNA should be written – that is, how to *design* genetically encoded biological function.

Programming cells to carry out desired metabolism

At its core, metabolic engineering depends on a variety of tools for the control of gene expression. This is because the turnover rate of molecules through any enzymatic pathway (termed the metabolic *flux*) is a function of the specific activity of the enzyme(s) in the pathway *and* the number of enzymes available to facilitate the reaction. For example, if the specific activity (i.e., the turnover rate of a single molecule) of an enzyme is very high, then only a few enzymes might be needed to maintain a physiologically relevant flux through that particular pathway. On the other hand, if the specific activity of an enzyme is very low, then many copies of that enzyme will likely be required to balance the flux through that particular pathway. Therefore, there are two general approaches to engineering metabolism in cells. The first is to engineer the actual enzyme itself thereby modifying its specific activity (and, possibly, its substrate specificity, product distribution, etc.). The second is to attack the *expression* of enzymes. Controlling and regulating *which*, *when*, *where* and *how much* of these enzymes are made is a powerful and straightforward way to affect metabolic flux[4]. Both approaches require engineers to write or encode instructions on DNA molecules for the cell to realize desired functions or behaviors.

Overview of the dissertation

My dissertation research began with a naïve question: can bacterial promoters be extended using UP elements to accomplish the fine-tuning of gene expression? Although this is an interesting biological question in its own right, this question was really motivated metabolic engineering applications. In particular, metabolic engineering is dependent on genetic tools that provide regulation of enzyme-encoding genes such that carbon, energy and electron fluxes are balanced. Therefore, an ability to finely tune expression levels by altering the rate of transcription initiation (through the modification of the promoter sequence) is extremely useful. Furthermore, UP elements should provide a facile method for engineering promoters that have constraints on their architecture

By driving the expression of a red fluorescent protein (*rpf*) using designed extended promoters (σ^{70} core) in *Escherichia coli*, I determined that UP elements are not only useful for modulating overall gene expression levels, but also appear to limit gene expression noise in some cases. Noise reduction in genetic networks is critical to the development of reliable or predictable gene expression strategies.

UP elements are known to contribute to the binding affinity between RNAP and the promoter region. This work also reveals that 1) a nearly dysfunctional core promoter can restored by the addition of an UP element and 2) a very strong core promoter that has been extended with an UP element will result in attenuated gene expression levels. To resolve these fascinating observations, I performed binding experiments to determine the equilibrium association constant (a relatively good measure of binding affinity) between many extended promoters and σ^{70}-saturated RNAP. The data suggest that RNAP:promoter

binding affinity can become so strong that the RNAP gets glued to the DNA such that it cannot escape the promoter region to proceed to elongate the transcript and, thus, gene expression is decreased.

It is not yet clear whether or not orthogonal (i.e., non-cross-reacting) processes will contribute to the success of metabolic engineering. A second question, therefore, is simply: do orthogonal gene expression systems offer any benefit to the metabolic engineering objective? I hypothesized that orthogonal gene expression offers a way to limit the metabolic burden placed on the cell by introduced pathways. To address this, I first performed a promoter engineering study using bacteriophage T7 RNAP and its cognate promoter. Surprisingly – because T7 RNAP does not have alpha subunits – extended T7 promoters were influenced by UP element sequences. Therefore, the promoter extension approach seems to be applicable outside of the scope of bacterial promoters alone.

To achieve a fully orthogonal gene expression, orthogonal translation must occur. A computationally derived orthogonal 16S rRNA (in particular, the anti-Shine-Dalgarno sequence) and Shine-Dalgarno pair was used to implement a translation process that was insulated from the native cellular machinery. This system was demonstrated to be functional by expressing *rfp* using only orthogonal machinery. Perhaps the most impactful finding in this dissertation is that orthogonal gene expression offers a straightforward way to improve the productivity of an engineering metabolic pathway. Lycopene biosynthesis (a three gene pathway from a common metabolite in *E coli*) under orthogonal expression, for example, was shown to occur at a rate similar to that of a specifically and highly engineered strain of *E. coli*.

These tools (UP elements and orthogonal gene expression) should be useful for future metabolic engineering efforts – or anything that requires regulating gene expression, for that matter.

Chapter 2 – Modulating gene expression in *E. coli* with extended promoters

"Great things are done by a series of small things brought together." – van Gogh

The process of gene expression is central to all known living systems. It begins with the initiation of transcription in which the sigma subunit of the protein complex RNA polymerase recognizes, binds to and unwinds DNA at a particular regulatory sequence known as the promoter (i.e., these sequences "promote" gene expression). After successful transcription initiation, the RNA polymerase (RNAP) moves along the DNA molecule to elongate/polymerize the RNA product. The process is terminated when the RNAP reaches the terminator sequence that follows the protein coding sequence of a gene.

Although post-transcription, translation and post-translation each offer a multitude of targets for tuning gene expression, the initiation of transcription is attractive because bacteria already use transcriptional regulation for the majority of their cellular decision-making. In addition, promoter engineering offers a straightforward approach to studying gene expression perturbations. That is, modifying the nucleic acid sequence of a particular promoter (which is a small DNA sequence, relative to other parts) will likely result in a change in the rate of transcription initiation (affecting either binding or unwinding or both) and, therefore, overall gene expression levels.

The intrinsic noise associated with gene expression limits the precision of molecular-level biological control. However, promoter engineering will certainly contribute to the development of the most precise, quantitative control of gene expression and perhaps limit extrinsic noise. A number of promoter libraries have been generated,

characterized, modeled and applied over the years[5-7]. These libraries have focused on the core promoter region. Here I describe extending core promoters to include upstream regions that contribute to gene expression.

Extending promoters with near-consensus UP elements

The σ^{70} subunit of RNAP is a protein responsible for initating transcription through interactions with the -10 and -35 promoter elements[8]. This particular transcription initiation factor interacts with the promoters of about 70% of the genes in *E. coli* because it drives expression of "housekeeping" genes that are essential for maintaining cellular activity. In a few cases, an additional interaction between the RNAP and promoter occurs. It is known that promoters of ribosomal operons, for example, share a curved, AT-rich region upstream of the -35 hexamer (-38 to -59), termed UP elements[9]. These UP elements can interact directly with the carboxyl terminal domains of the alpha subunits of RNAP (Figure 1, below)[10].

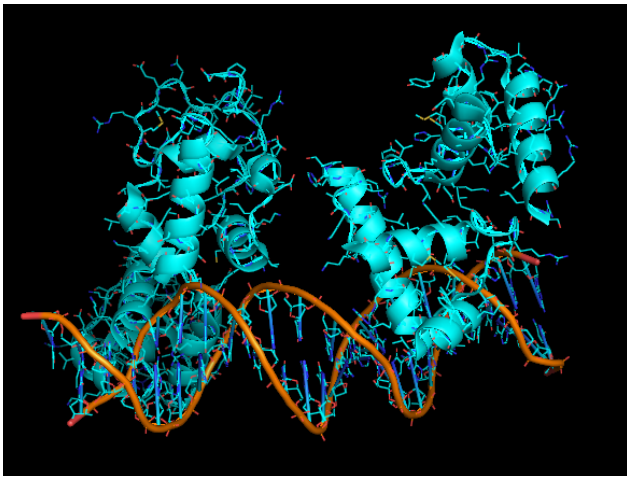

Figure 1 | The carboxyl terminal domains (CTDs) of the RNAP alpha subunits (colored teal) can bind an AT-rich DNA sequence upstream of the canonical promoter (the UP element[9]) with high specificifity, making it an attractive target for extending the repertoire of promoter engineering. Designed UP elements can be used to modulate the rate of transcription intitation and, therefore, overall gene expression rate. This image was generated using PDB ID 3N97 in PyMol.

The most studied naturally occurring UP element is associated with the rRNA promoter *rrnB* P1 in *E. coli*, which has been characterized *in vivo* using beta-galactosidase activity as a reporter for promoter activity[11]. In this foundational study, the *rrnB* P1 with its UP element was reported to increase overall activity of beta-galactosidase by at least 30-fold compared to the activity of the promoter alone. An additional study investigated different combinations of naturally occurring UP elements with various natural promoter sequences[12]. In this study, the overall activity of the reporter protein was increased 1.5 to 90 times and *in vitro* transcription was shown to increase without the presence of transcription factors – only the carboxyl terminal domains (CTDs) of the RNAP alpha subunits were required. Overall, studies on UP elements have largely focused on naturally occurring UP elements and found that the addition of an UP element to a core promoter increases gene expression[13-16].

Previous work with UP elements led to the derivation of a consensus sequence containing highly conserved A and T regions[17]. Estrem *et al*. generated the consensus sequence by using a modified SELEX procedure in which a combinatorial library of synthetic UP elements was produced and placed upstream of an *E. coli* rrnB P1 core promoter, again driving the expression of LacZ. From this genetic library, 31 functional UP elements were identified and characterized using β-galactosidase activity. Tabulated sequence information (Figure 2, bottom) was used to derive a consensus UP element sequence (Figure 2, top) where a single nucleotide was called if it made up more than 55% of the nucleotides at a given position. Likewise, if two nucleotides together made up more

than 95% of the nucleotides at any single position, then they were both selected (e.g. position -54).

N N A W W W W W T T T T T A A A A A A R N R N N

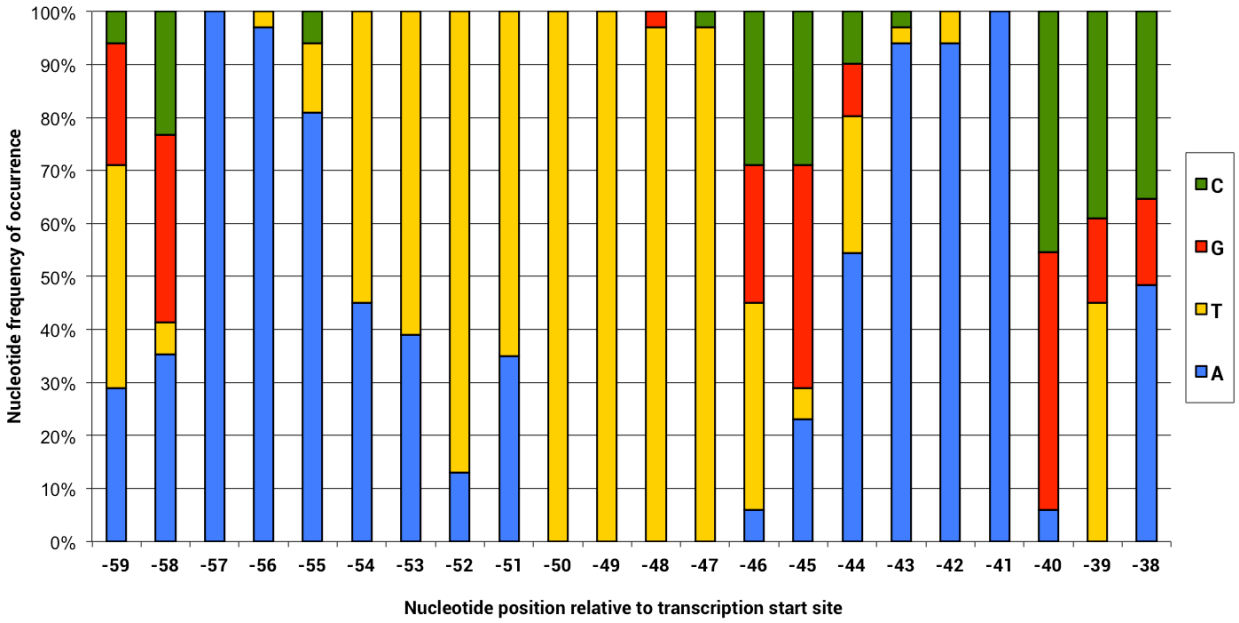

Figure 2 | An UP element consensus sequence (top) has been derived by the Gourse Lab by tabulating sequence information of synthetic (but not designed) UP elements[17]. Each nucleotide position is labeled with respect to the transcription start site (TSS) and the colors correspond to different nucleotides (C, G, T and A).

I hypothesized that UP elements could provide an intuitive and simple way to finely tune gene expression by altering the interactions between the RNAP and the promoter sequence. In particular, the UP element could be a useful *enhancer* of transcription such that it would always increase gene expression levels. Using the consensus sequence as a starting point, I designed six near-consensus UP element sequences by varying the nucleotides at the -58 and -59 positions (i.e., sequences with only 2 mutations away from the consensus). These positions were found to be highly variable (i.e., not conserved) and

mutations in this region should not sabotage UP element function but rather slightly affect its ability to bind RNAP α-CTDs, resulting in small changes in gene expression levels.

To test this idea, I constructed a series of plasmids starting from the high-copy number plasmid pSB1C3, which carries the coding sequence for a red fluorescent protein called mRFP1 (Figure 3)[18]. Upstream of the *rfp* there is a standard ribosome binding site (RBS) sequence used by Elowitz to construct the repressilator[19], one of the research projects considered by many to have kicked off the modern era of synthetic biology (building novel genetic circuits and networks from well-characterized parts). The influence of specific promoter sequences on gene expression can be measured by simply inserting designed core and extended promoters upstream of this reporter (the Elowitz RBS and mRFP1 translational unit) and measuring fluorescence using flow cytometry or fluorescence microscopy.

Each promoter was placed upstream of the same translational unit containing the Elowitz RBS and a sequence encoding the RFP variant and inserted into pSB1C3 to minimize variation in gene expression caused by differences in the rate of translation (initiation and elongation) or plasmid copy number. To maintain optimal spacing from the transcription start site (TSS), a two base-pair spacer (CT) was added to the 3' end of each designed extended promoters. This spacer is meant to maintain proper alignment with the RNAP complex and is based on the sequences of the natural *rrnB* P1 and *rrnD* P1 UP elements. After being manufactured by a DNA synthesis company, these synthetic promoters were cloned into pSB1C3 using a standard restriction/ligation scheme known as BioBrick assembly in which four restriction sites are used over and over to achieve

idempotent assembly[20]. BioBrick assembly makes use of standard restriction sites that are reusable as smaller DNA fragments are combined into larger constructs. The upstream part (promoter or UP element-promoter) was cut with EcoRI and SpeI, the downstream part (reporter device) with XbaI and PstI and the backbone expression plasmid (pSB1C3) with EcoRI and PstI such that a mixed-site, 8 base scar was formed between the upstream part and the reporter device. The pSB1C3 plasmid carries a chloramphenicol (Cam)-resistance gene for positive selection. Electrocompetent *E. coli* NEB10β cells were transformed with the resultant constructs by a standard electroporation transformation procedure.

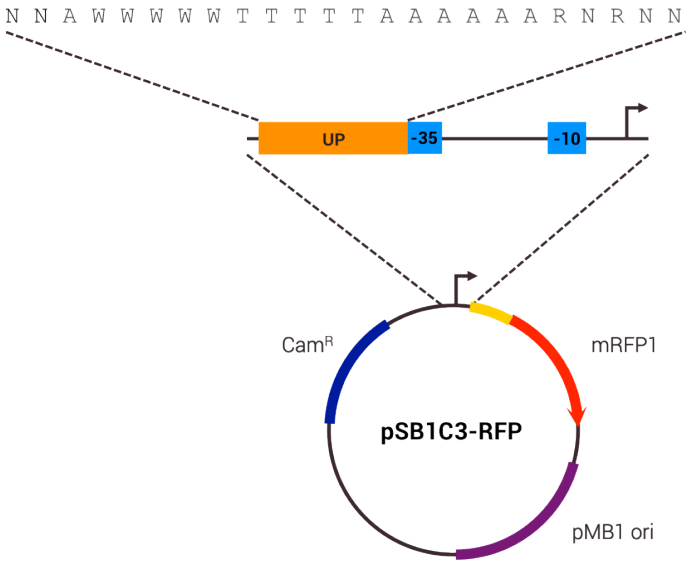

Figure 3 | A series of plasmids was constructed by inserting core and extended promoters (using a BioBricks MCS) upstream of *rfp* (red) preceded by a standard RBS sequence (yellow). *E. coli* were transformed with these plasmids and the resulting fluorescence levels were measured using flow cytometry as a means for characterizing promoter strength.

A near-consensus σ^{70} core promoter was selected from a promoter library generated by J. Chris Anderson, a collaborator at UC Berkeley. This library had been generated using error-prone PCR to introduce mutations in the -10 and -35 hexamer elements of the consensus σ^{70} core promoter (including an optimal 17 base-pair spacer between the -10 and -35 elements). Two mutations away from the consensus sequence

(one in each of the conserved regions) were sufficient to decrease transcriptional output of this chosen promoter such that it promoted gene expression at a rate similar to natural promoters (hereafter called "moderate core promoter").

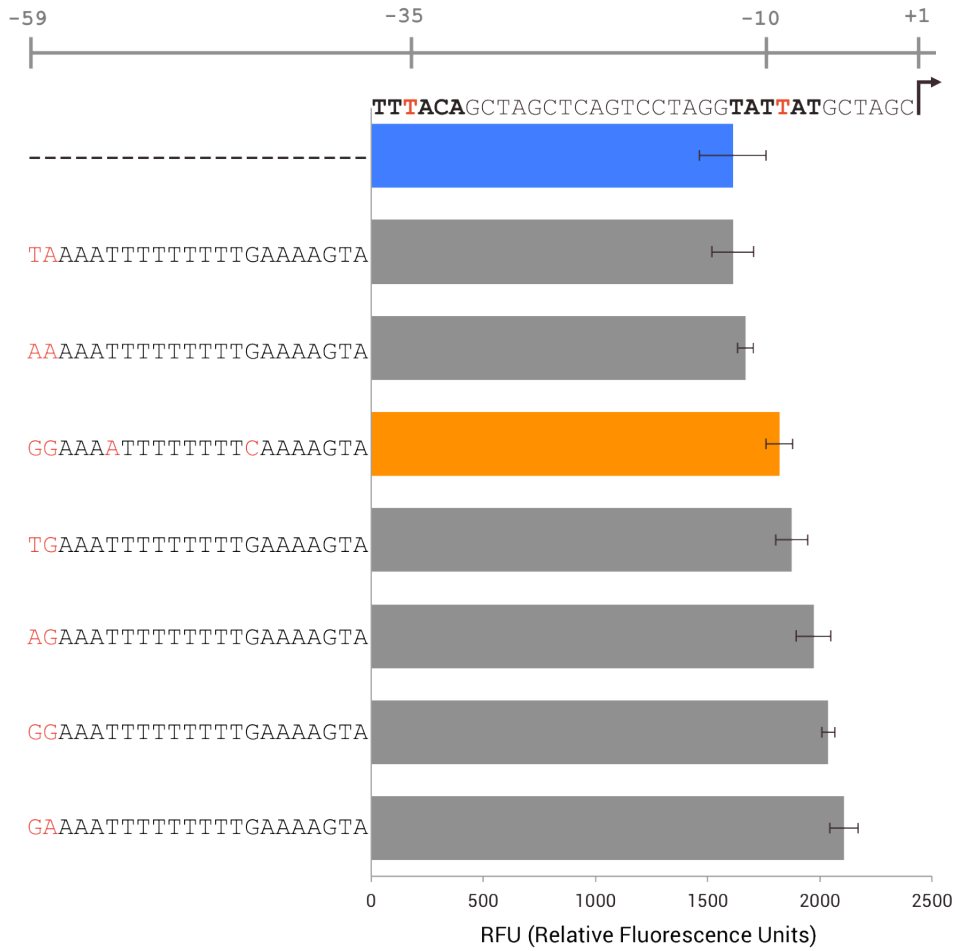

Figure 4 | A small library of extended promoters was designed by placing near-consensus UP element sequences (differing in the -58 and -59 positions, in red on 5' end) ahead of a moderate core promoter sequence (top of figure). In this case characterization by flow cytometry, used here to measure fluorescence that is reported in relative fluorescence units (RFUs), showed that these UP elements are able to finely increase gene expression levels, which is not easily accomplished using core promoter sequences alone.

As expected, the near-consensus UP elements finely increased gene expression from the level of the moderate core promoter alone (blue bar, Figure 4). I included the strongest known UP element (orange bar, Figure 4) from the study by Estrem *et al*. for comparison[17]. These initial constructs indicated that extending promoters with UP elements could offer a

powerful method for fine-tuning gene expression. However, genetic context has a great influence on the performance or behavior of parts such as promoters and RBSs[21]. Therefore, I decided to examine how these same UP elements influence gene expression levels when combined with weaker and stronger core promoters. I had hypothesized that a strong promoter (i.e., a promoter that drives gene expression to high levels) could be made even stronger by increasing its binding affinity to RNAP through the addition of an UP element.

I selected two representative UP elements and placed them upstream of a very weak (nearly dysfunctional with 2 mutations in each hexamer element) promoter and upstream of the consensus core promoter from the same Anderson promoter library. These extended and core promoters were characterized by measuring fluorescence from red fluorescent protein (excitation wavelength = 584 nm, emission wavelength = 607 nm) on the same pSB1C3 plasmid. Unexpectedly, these near-consensus UP elements demonstrated an ability to 1) restore a nearly dysfunctional promoter (top group, Figure 5) and 2) *decrease* overall gene expression when placed upstream of a very strong core promoter, the consensus σ^{70} sequence (bottom group, Figure 5). Why would this happen?

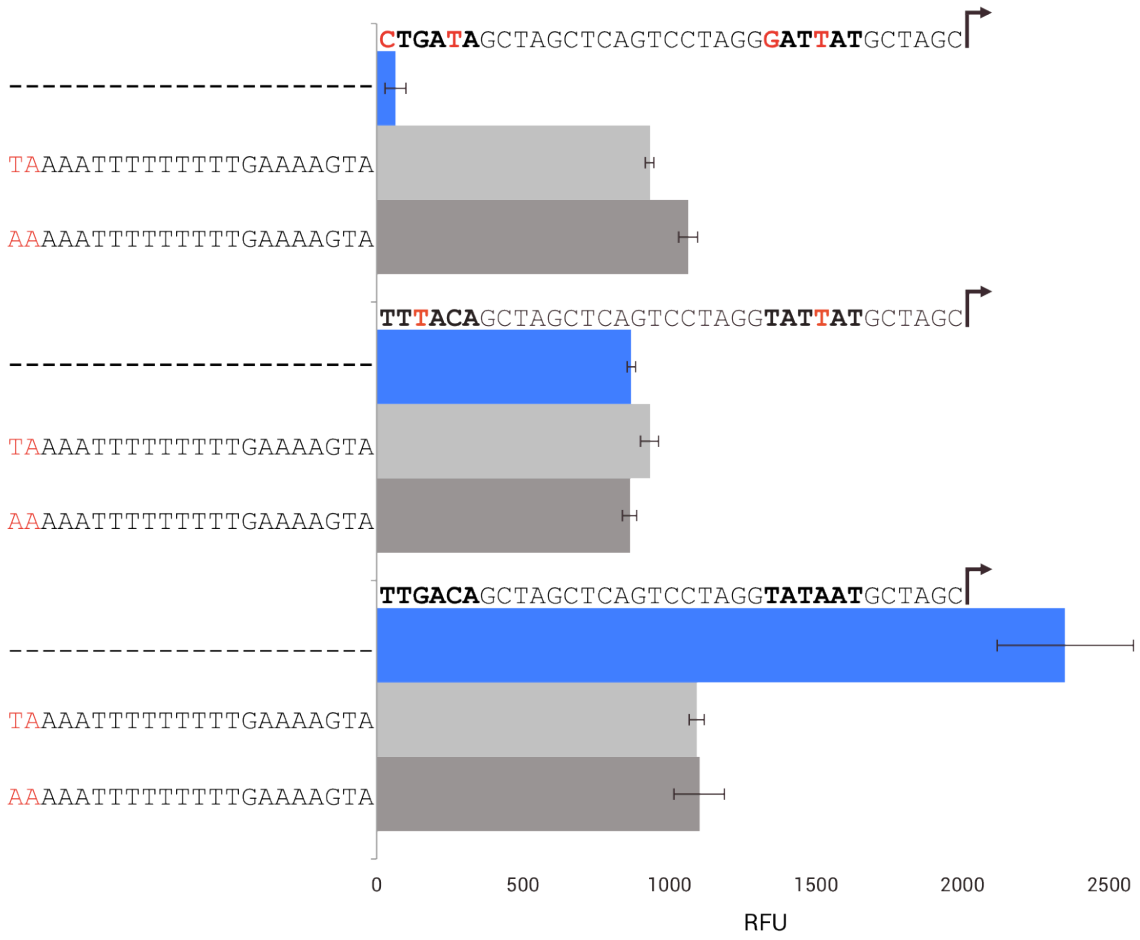

Figure 5 | A second library of extended promoters was constructed using a couple UP elements from the first promoter library and different core promoters: a weak, moderate and strong core promoter (from top to bottom, in blue). These data show that UP elements 1) have the ability to rescue a sabotaged core promoter by contributing binding affinity for the RNAP (top group) and that they 2) do not always increase gene expression levels (e.g., in the case of a very strong promoter gene expression is attenuated when it is extended with an UP element, bottom group).

The process of transcription initiation can be broken down into two major steps: 1) the recruitment of RNAP to the DNA and isomerization to an open complex at the promoter region and 2) the escape of RNAP from the promoter and transition to processive elongation. Either step can limit the overall rate of transcription, depending on the system. A potential drawback to a promoter sequence that has a high binding affinity to RNAP is that the RNAP may not be able to transition to transcription elongation because it is too tightly bound to the DNA (i.e., it will not be able to escape the promoter region). It is known that certain RNAP:promoter complexes experience *abortive transcription* in which many

short mRNAs are made before the RNAP can escape the promoter. This phenomenom may offer insights into characterizing extended promoters, as I describe later in this section (p. 27).

A strong UP element paired with a strong core promoter may result in a lower overall transcription rate than that of the core promoter alone even though RNAP may be recruited more frequently. To determine the role of binding affinity in these systems, I performed a series of electrophoretic mobility shift assays (EMSAs) to measure approximate values for K_d (the dissociation constant) at equilibrium, which can be used to calculate K_a (the association constant)[22]. Binding reactions (20 uL) contained purified *E. coli* RNAP saturated with σ^{70} at various concentrations (12.5 – 400 nM), PCR-generated, linear DNA fragments containing the promoter region (5 or 10 nM) and a simple binding buffer. Binding was allowed to come to equilibrium (30 minutes at 37°C) before heparin (200 ug/mL) was added as a competitor and incubated for an additional 10 minutes. The binding mixture was loaded on a 2.0% agarose-TAE gel and run for 2.5 hours at 6 V/cm to separate free DNA from bound DNA. Gels were visualized after being stained with SYBR Green and subsequently analyzed using ImageJ to quantify fluorescence intensity of each DNA band and to determine the fraction of bound DNA. A one-site total binding model in Prism was used to calculate dissociation constants.

Model for one-site, total binding:

$$Y = Y_{max}\,[RNAP]/(Kd+[RNAP]) + NS\,[RNAP]$$

Y is the fraction of DNA that is bound by RNAP.
Ymax is the maximum specific binding.
NS is the contribution of nonspecific binding (in inverse nM units).

Promoter architecture description	K_d (nM)	R^2	K_a (nM^{-1})
Dysfunctional core promoter (no UP)	311.4	0.924	0.00
UP element + dysfunctional promoter	15.53	0.981	0.06
UP element 2 + dysfunctional core promoter	13.73	0.997	0.07
Moderate core promoter (no UP)	11.66	0.987	0.09
UP element + moderate core promoter	10.21	0.987	0.10
Strong (consensus) core promoter (No UP)	2.97	0.990	0.34
UP element + strong (consensus) core promoter	0.67	0.996	1.48

Table 1 | An electrophoretic mobility shift assay (EMSA) was used to calculate dissociation/association constants for select extended promoters. Mixtures of DNA fragments containing the promoter region (10 nM) and RNAP (various concentrations, 12.5 – 400 nM) were allowed to come to equilibrium and separated using gel electrophoresis. The fraction of DNA bound increases with RNAP concentration and is used to calculate the dissocation constant k_d using a simple binding model in Prism (described above).

An example of a plot generated in Prism is given in Figure 6. In this case, the K_d of the consensus σ^{70} promoter was determined to be about 3 nM, which indicates that the RNAP:promoter affinity is high (natural, strong *E. coli* promoters typically have a K_d of 9-10 nM). Most of the promoter dissociation constants measured here were around 10-15 nM, although in the case of the extended consensus promoter (that is, the σ^{70} consensus promoter with an added near-consensus UP element), the K_d was found to be around 0.67 nM, indicating *very* high affinity. A more intuitive way to visualize this data, however, is to calculate the association constant and then plot measured fluorescence values against K_a (Figure 7) since the association constant is a good metric for binding affinity.

It appears that the addition of an UP element to the weak/dysfunctional core promoter rescues its binding functionality. Therefore, it can be assumed that the mutations in the -10 hexamer element do not significantly impact DNA unwinding (also known as promoter melting). Because UP elements contribute to RNAP *binding* – and not DNA *unwinding* – it is impossible for UP elements to rescue a completely sabotaged core

promoter[23]. It is also not possible for the UP element to initiate transcription alone; it requires at least a functional -10 hexamer to recruit the sigma factor and to facilitate DNA unwinding (namely an A_{-11} and a T_{-7})[24]. To verify this, I executed a series of EMSAs using the same DNA fragments but replaced the σ^{70}-saturated RNAP with the RNAP holoenzyme (no sigma factor). No binding was detected (data not shown).

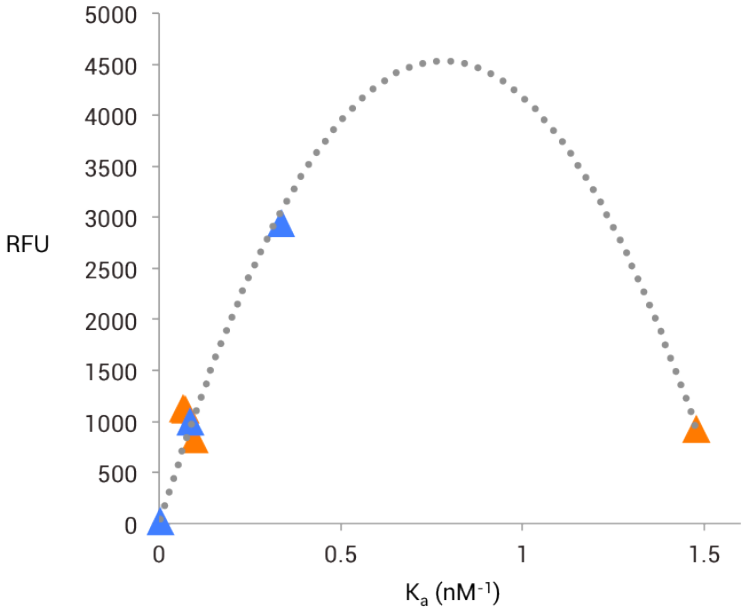

Figure 6 | EMSA-determined association constants ($1/K_d$) for each of the promoters characterized in Figure 6 (core promoters in blue, extended promoters in orange) were plotted against gene expression levels (as measured by fluorescence, in RFUs) to outline the general relationship between binding affinity and gene expression levels (i.e., transcription initiation rate). At an unknown K_a, gene expression begins to decrease with increased affinity, likely due to the RNAP being "glued down" to the DNA such that the RNAP does not escape the promoter region as often (represented by the speculative dashed, grey line). It should be possible to experimentally determine the maximum rate of transcription initiation (a function of both RNAP binding and DNA unwinding) by characterizing a large library of extended promoters.

These results indicate that UP elements are context-dependent, due to overall RNAP:promoter binding affinity, and may enhance or attenuate gene expression depending on the core promoter sequence. Furthermore, for the three core promoters tested, near-consensus UP elements appear to normalize gene expression to a particular level. Therefore, well-characterized UP elements should not only be useful for understanding

how natural UP elements function, but also for predictably fine-tuning gene expression of novel biological systems.

I have demonstrated that the overall binding affinity of RNAP:promoter complex can be broken apart into two contributing modules, the core promoter (which interacts with the sigma subunit) and the UP element (which interacts with the alpha subunit). The overall rate of transcription (i.e., the actual flux of RNAPs along the DNA or the synthesis rate of the mRNA) is tunable by both of these modules (because the binding affinity is tunable by both). Therefore, it should be possible to find the maximum rate of transcription *in vivo* by increasing the rate of RNAP recruitment *and* simultaneously increasing the rate of promoter escape. Balancing the flux of RNA polymerases in this manner will require a large combinatorial library of -10, -35 and UP elements.

Throughout the course of these measurements, a trend in the distribution of fluorescence intensity was noted across populations of cells harboring different designed promoters. In particular, I observed a heavy tail in the population distribution for measurements taken of cells expressing RFP by the strong, consensus core promoter. Interestingly, when near-consensus UP elements are added upstream of this core promoter, the distribution tightens up and becomes similar to the distribution seen when measuring RFP expression driven by the moderate core promoter or the rescued weak promoter (Figure 9). Time-lapse fluorescence microscopy (data not shown) also indicated that most of the fluorescence was being produced by a only a small fraction of the total population of cells. In fact, most cells were hardly producing RFP. The brightest fluorescing cells typically did not divide, but rather grew unusually long. It is possible that very high

expression of RFP may produce so much protein that it begins to interfere with the cellular division machinery.

A more-detailed analysis of the flow cytometry data revealed that the coefficient of variation (CV, defined as the ratio of the standard deviation to the mean), a reasonable measure of variability, is proportional to fluorescence (Figure 8). This is somewhat surprising because noise in gene expression typically increases as the process(es) become more stochastic (i.e., as the absolute numbers of species involved approach zero)[25]. In the case of the consensus core promoter, fluorescence is high which is likely proportional to the number of RFP mRNA molecules present in the cell. Therefore, it is reasonable to assume that the free ribosomes are saturated with RFP mRNA, which can plague the cell. High levels of expression (i.e., high rates of transcription) could impact cell-wide gene expression by tying up all of the free ribosomes[26]. This would not only affect the translation of critical structural proteins and enzymes, but also the translation of RNA polymerase – contributing even more noise to the system.

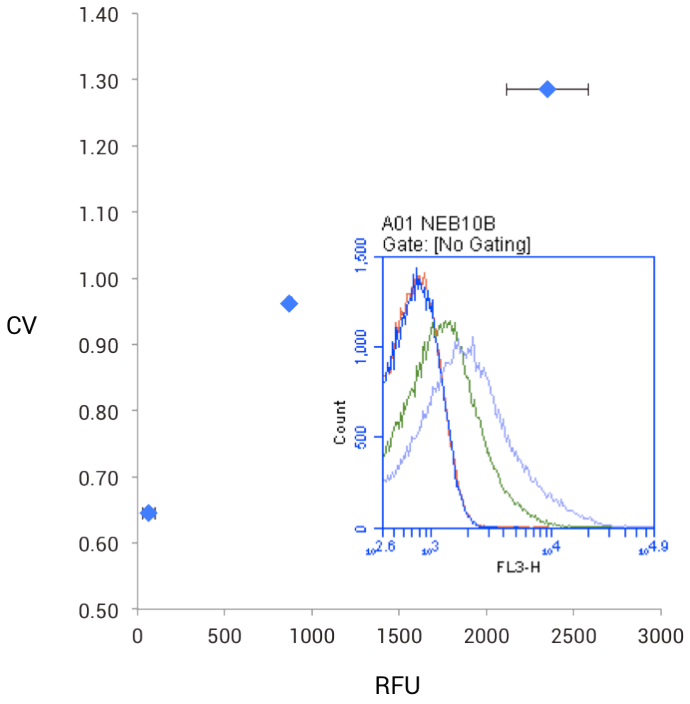

Figure 7 | The coefficient of variation (CV, defined as the ratio of the standard deviation to the mean) for each of the measured core promoters (blue dots) is plotted against overall gene expression (fluorescence). These data suggest that variability or "noise" in gene expression is high when the rate of transcription initiation is very high, possibly due to the saturation of free ribosomes with mRNAs. This contribution to extrinsic noise may be compounded by the decreased translation of RNAPs due to a decrease in available ribosomes. The *inset* shows raw fluorescence data collected by flow cytometry (FL3-H) for each of the core promoters and a negative control (negative control in dark blue, weak promoter in red, moderate promoter in green and strong promoter in violet). It is worth noting that the *x*-axis is scaled logarithmically.

UP elements can potentially limit noise in gene expression. In these experiments, it was observed that UP elements not only modulate gene expression levels, but also contribute to variability in gene expression across an isogenic, clonal population. In the next section I describe an UP element library that limits noise in gene expression even as it increases overall gene expression levels.

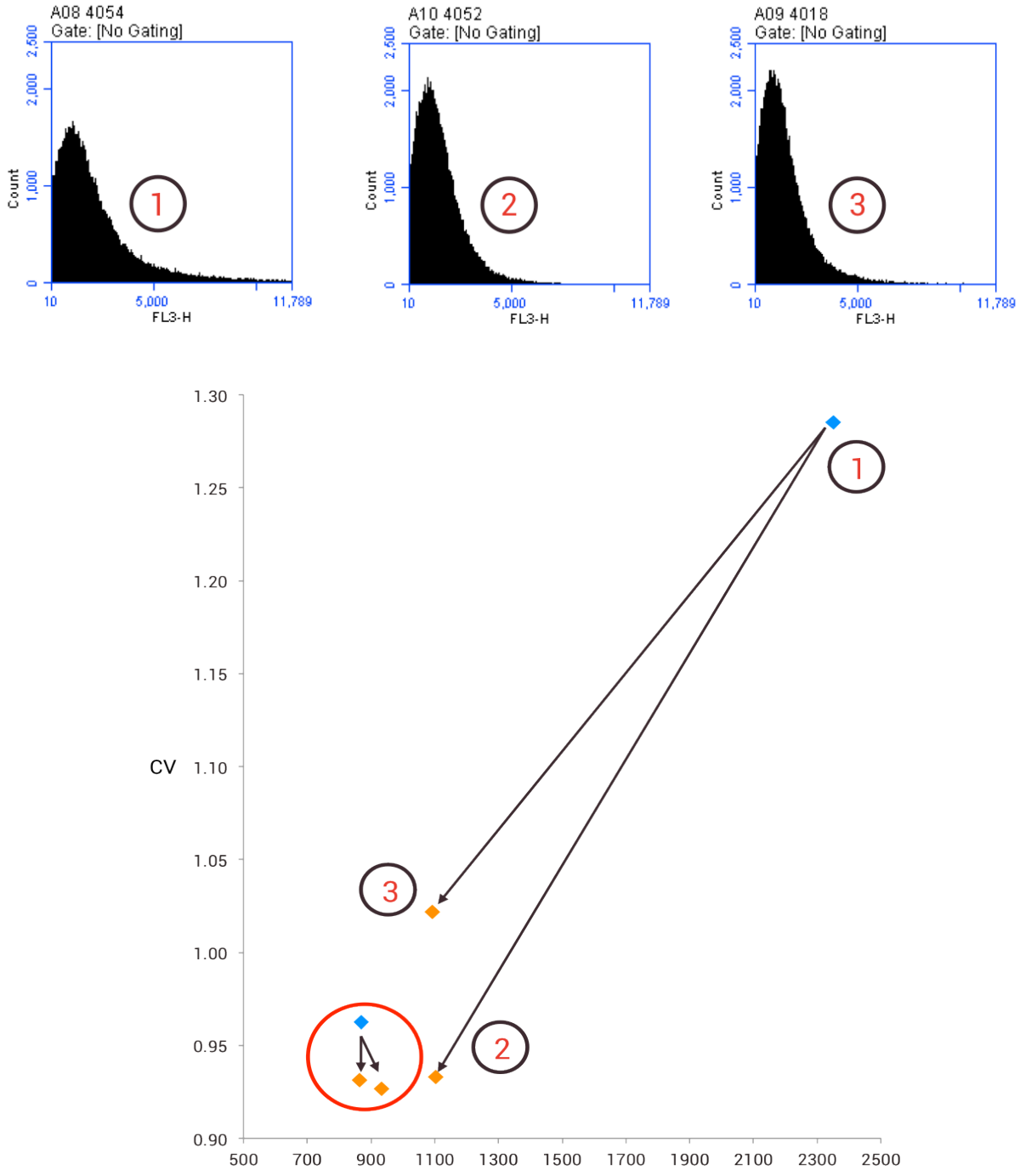

Figure 8 | The coefficient of variation (CV, defined as the ratio of the standard deviation to the mean) for select core promoters (blue dots) and extended promoters (orange dots) is plotted against overall gene expression (fluorescence). The addition of near-consensus UP elements to the strong consensus core promoter result in significantly reduced fluorescence (core promoter is labeled 1; extended promoters are 2 and 3). In this case, noise is also reduced but to varying degrees by each of the UP elements, although they are very similar in sequence. The three fluorescence districbutions at the top reveal heavy tails. It is clear that adding these UP elements to the strong core promoter tightens up the distribution. Interestingly, the addition of the same UP elements to the moderate core promoter does not significantly change fluorescence (as seen in Figure 5), but they do offer an advantage with respect to modulating noise in gene expression.

Characterizing a library of designed UP elements

Although near-consensus UP elements are useful for learning about UP element function, a more diverse set of sequences must be investigated to take full advantage of this modular genetic element. To design a more diverse set of UP element sequences, I aligned upstream sequences from three strong native *E. coli* promoters (Figure 10), *recA*, *rrnB*-P1 and *rrnD*-P1 and made nucleotide changes at positions that were moderately or highly conserved[27]. This small UP element library (including the upstream sequences of *recA*, *rrnB*-P1 and *rrnD*-P1) was designed and built to be upstream of the moderate core σ^{70} promoter.

Figure 9 | An alignment of sequences upstream of the core promoters (-38 to -59) for *recA*, *rrnB* and *rrnD* was used to generate a sequence logo (frequency plot) for a more natural UP element sequence (i.e., not near-consensus). This served as a starting point for designed a diverse library of UP elements.

A 5' UTR translational insulator was included downstream of the TSS and upstream of the RBS. The insulator, RiboJ, is a ribozyme that co-transcriptionally cleaves the 5' UTR, eliminating any variability between genetic systems caused by differences in 5' mRNA stability. A strong RBS (arbitrary translation initiation rate = 80,000) was designed using the RBS calculator[21]. Furthermore, a transcriptional insulator with 50% GC content was included upstream of the UP element in an effort to insulate the gene cassette from the genetic context of the plasmid.

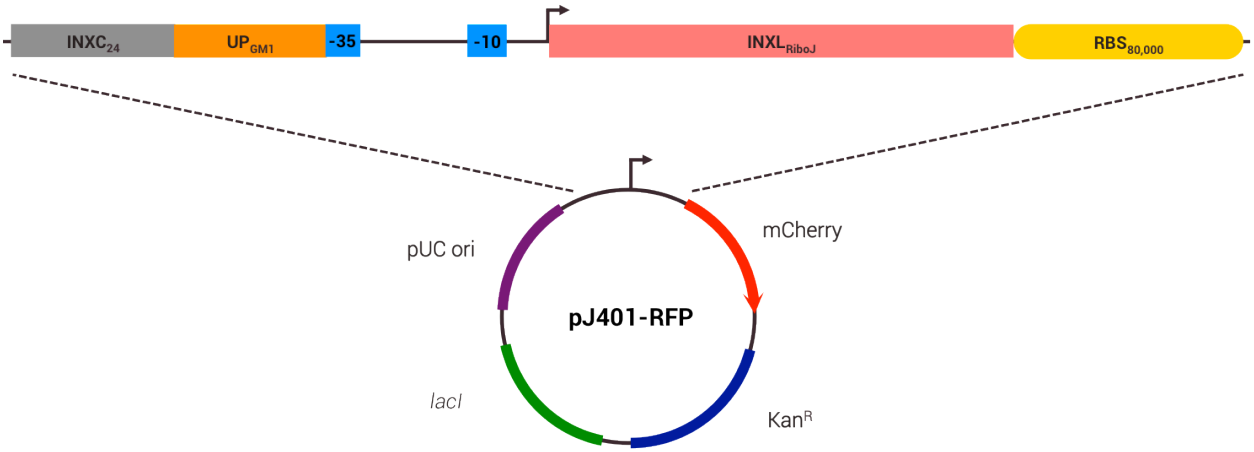

Figure 10 | The construction of an UP element library (and the construction of subsequent libraries) was carried out using isothermal DNA assembly in which six overlapping 60mer oligonucleotides and a linearized plasmid (pJ401-RFP) were assembled together in a one-pot reaction. To insulate the promoters from genetic context, a transcriptional insulator (a random sequence with 50% GC content) was placed upstream of the promoter and a post-transcriptional insulator (the RiboJ ribozyme) was placed immediately upstream of a designed, strong RBS.

The assembly of DNA into larger constructs has long been time-consuming and expensive, although the price is dropping steadily. To build out my designed UP element library, I had to turn to a cheaper, faster and more flexible DNA assembly method. The state-of-the-art assembly method when I started this project was (and still is, for the time being) isothermal DNA assembly, developed by Daniel Gibson and colleagues at the J. Craig Venter Institute[28-30]. The assembly method allows for the sequence-independent construction of DNA molecules from smaller pieces, both double stranded and single stranded. Isothermal DNA assembly uses three enzymes: a T5 exonuclease, a high-fidelity thermostable DNA polymerase (e.g., Phusion) and a thermostable ligase (e.g., *Taq* ligase) to chew back overlapping homologous regions, fill in gaps and seal the DNA backbone respectively. I adopted this assembly technique for my routine cloning due to its simplicity and because I could avoid any scarring resulting from restriction digestion and ligation.

Isothermal DNA assembly can be used to stitch together ssDNA oligonucleotides into larger DNA fragments, even recombining them with dsDNA such as linearized

plasmids. This was demonstrated by assembling the entire mouse mitochondrial genome from synthetic 60mer oligos[31]. I used this approach to assemble together members of an extended promoter library from six overlapping ssDNA oligos (detailed in Appendix B – DNA Sequences) and developed a small MATLAB program to automate the design of ssDNA oligos to use in this assembly scheme (Appendix A – Methods).

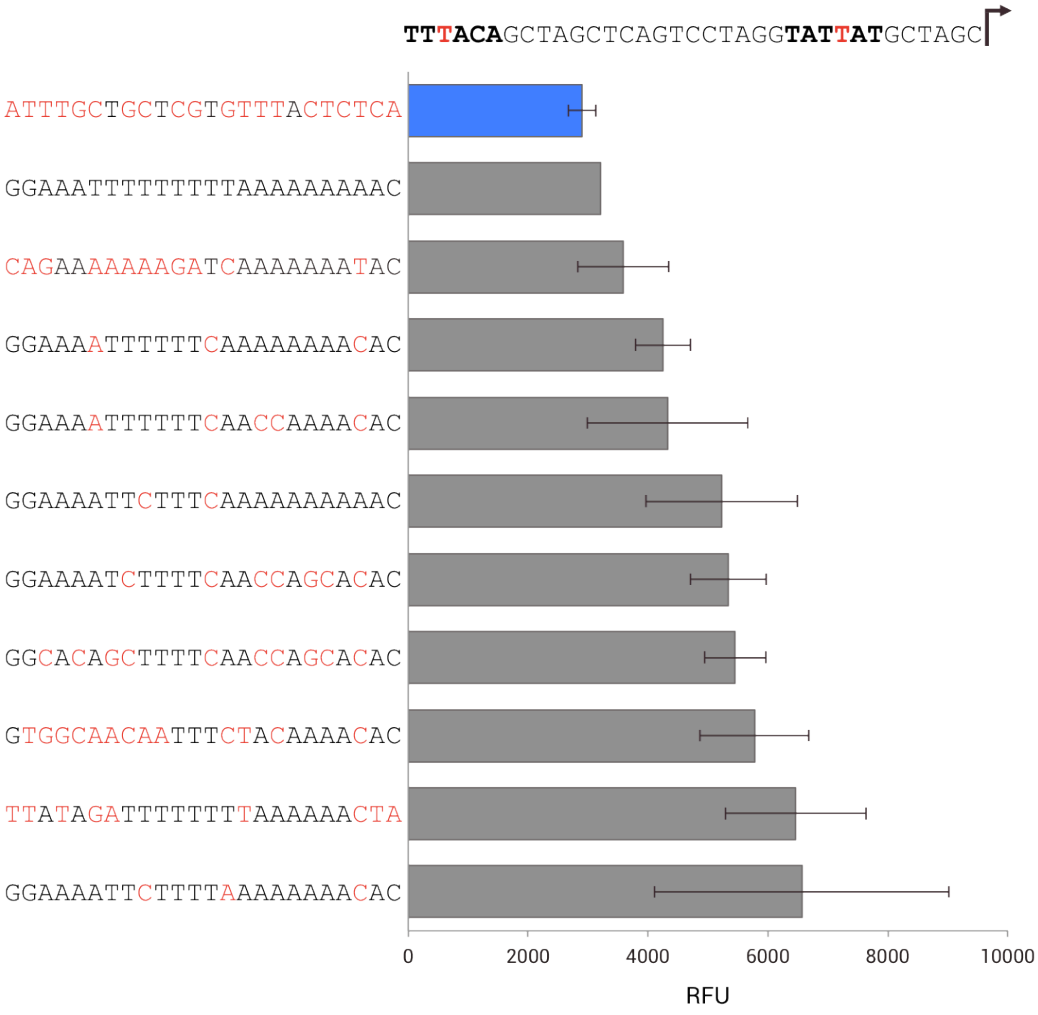

Figure 11 | A diverse UP element library was characterized using flow cytometry. These UP elements were placed upstream of the moderate core promoter (top) and are shown here to increase fluorescence by significant amounts.

Figure 12 shows that the extension of the moderate core promoter with any of these UP element library members increases fluorescence. Perhaps what is more interesting is

that although these UP elements increase overall gene expression levels, they simulataneously drive down the variability in gene expression (Figure 13). My hypothesis is that these sequences are near a "sweet spot" for transcription in which the two parts of transcription initiation are nearly balanced such that there is relatively high frequency of RNAP recruitment but limited abortive transcription. This of course can be tested by monitoring abortive transcription[32]. Future work will include an *in vitro* time-course transcription assay in which binding reactions are carried out in a manner similar to the EMSA described earlier. After the incubation period, NTP mix and heparin would be added to the mixture to start *in vitro* transcription and compete with the template DNA, respectively. Aliquots (5 uL) from the reaction mixture would be taken at 5 minute intervals and loaded onto a gel to be analyzed. Extended promoters containing UP elements are expected to show a shift for longer times than promoters lacking UP elements, which will dissociate sooner. In addition, promoter sequences that limit promoter escape by over-binding RNAP are expected to produce a large amount of abortive transcript relative to promoters that do not limit promoter escape (those would instead produce full length transcripts).

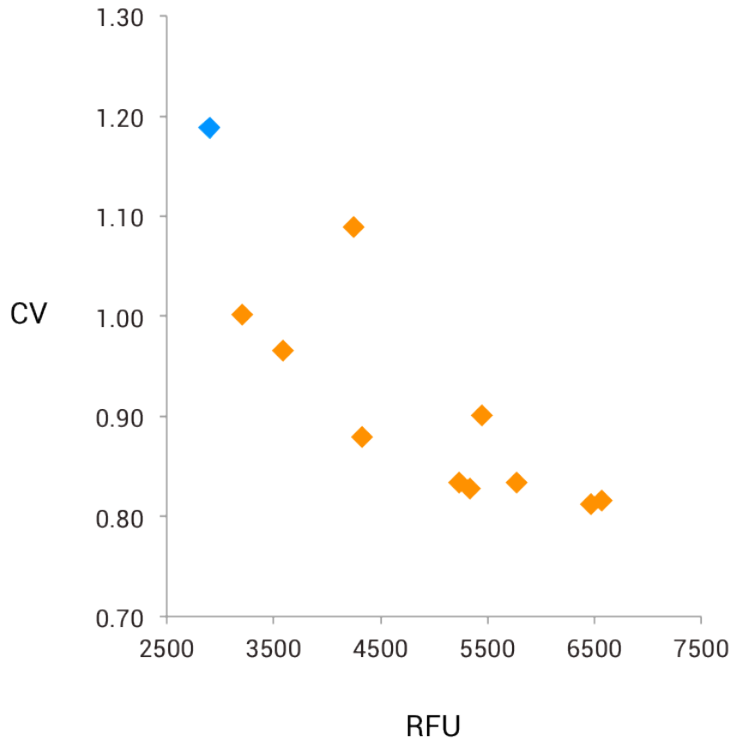

Figure 12 | The data generated from the promoter set in Figure 12 is plotted here as CV (a measure of noise) against fluorescence. A moderate core promoter (blue dot) is enhanced by the addition of UP elements (orange dots) by not only increasing gene expression but also decreasing noise in gene expression. These data suggest that under these conditions, RNAP:promoter binding (and not the transition to an open complex or transcription elongation) is the main contributor to transcriptional noise and can be adjusted by modifying the promoter sequence.

Regulated expression using lactose-inducible promoters

One argument against using UP elements to modulate gene expression levels is that it is simple enough to modify the -10 and -35 elements of the core promoter to modulate gene expression levels. However, in addition to offering a way to make *small* changes in the rate of transcription initiation (which is difficult to achieve by mutagenizing the core promoter) and a way to potentially decrease the noise associated with gene expression, UP elements provide a facile method for engineering promoters that have constraints on their architecture. For example, regulated promoters in bacteria often have an operator motif immediately upstream, downstream or embedded within the promoter sequence. In this

case, it is difficult to modulate the RNAP:promoter binding affinity by changing the nucleotide sequence of the promoter without altering the regulatory properties of the system. UP elements appear to be a very useful tool to get around these constraints.

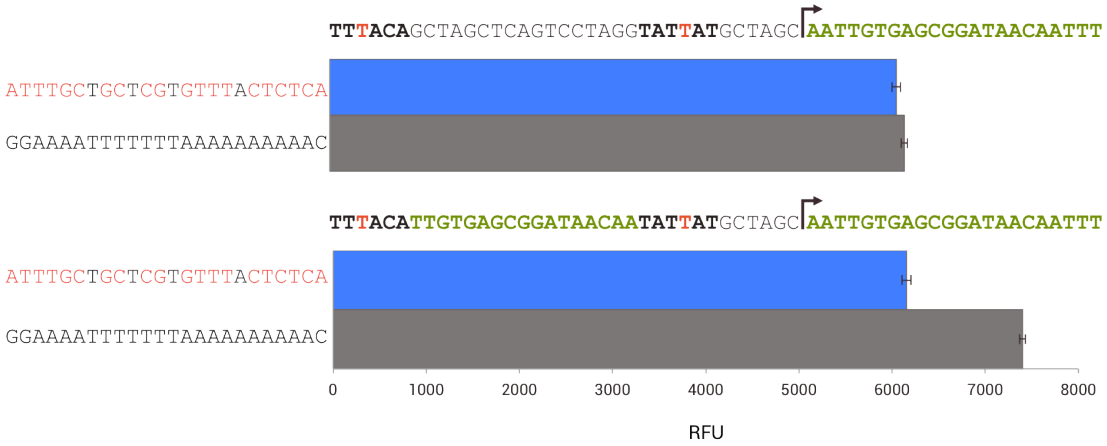

Figure 13 | A near-consensus UP element was placed upstream of two synthetic *lac* promoters, which can be induced by IPTG. After 6 hours at 1 mM IPTG, the extended promoter with two lacO operator sites showed an increased dynamic range than the core promoter with the same operator sites. This result indicates that UP elements are especially useful for modulating the binding of promoters that are constrained by their architecture (e.g., by operator sites).

To determine whether or not UP elements can modulate the dynamic range of gene expression of regulated promoters, I designed and characterized two ITPG-inducible promoters using the common lacO operator sequence. The fluorescence data in Figure 14 suggest that in certain cases UP elements can be used to extend promoters with constrained architectures in order to achieve modulated gene expression levels. In particular, when the regulatory protein binding is higher (two operator sites), then the addition of the UP element can help drive higher gene expression.

Chapter 3 – A platform for orthogonal gene expression

"An inefficient virus kills its host. A clever virus stays with it." – James Lovelock

Bacteriophage T7 RNAP and its cognate promoter have been used for biotechnological applications for decades. This is mainly due to T7 phage's high selectivity and processivity that results in a very high rate of transcription. However, because it is so processive, it can quickly consume cellular resources and result in severely decreased growth and possible cellular death[33]. Therefore, T7 promoters and/or T7 RNAP are commonly regulated when used to drive protein synthesis[34].

The goal of engineering microorganisms to perform desired tasks such as chemical biosynthesis is usually at odds with the cell's objective (to grow and to multiply). Therefore, metabolic engineers should start thinking like viruses – or perhaps use viral components to achieve their goals. For example, T7 RNAP carries out a transcription process that is independent or orthogonal to the endogenous cellular process. In other words, the T7 RNAP is non-cross-reacting with native *E. coli* promoters and the *E. coli* RNAP (of any sigma subunit) cannot recognize or act on T7 promoters. This property makes this an ideal transcription platform for metabolic engineering applications because it offers tunable and controllable transcription. There are likely many more viral genetic elements that would be useful for helping metabolic engineers repurpose the cell.

Characterizing native T7 genomic promoters

Bacteriophage T7 offers metabolic engineers 17 native promoters from its genome[35]. However, the transcriptional output of these promoters (i.e. their relative strength) has not been characterized. I designed an expression system nearly identical to

the one described in Chapter 2 for the extended σ^{70} promoter UP element library (swapping the σ^{70} promoters with the T7 genomic promoters). A consensus T7 promoter sequence was included as a reference. The fluorescence data are presented in Figure 15.

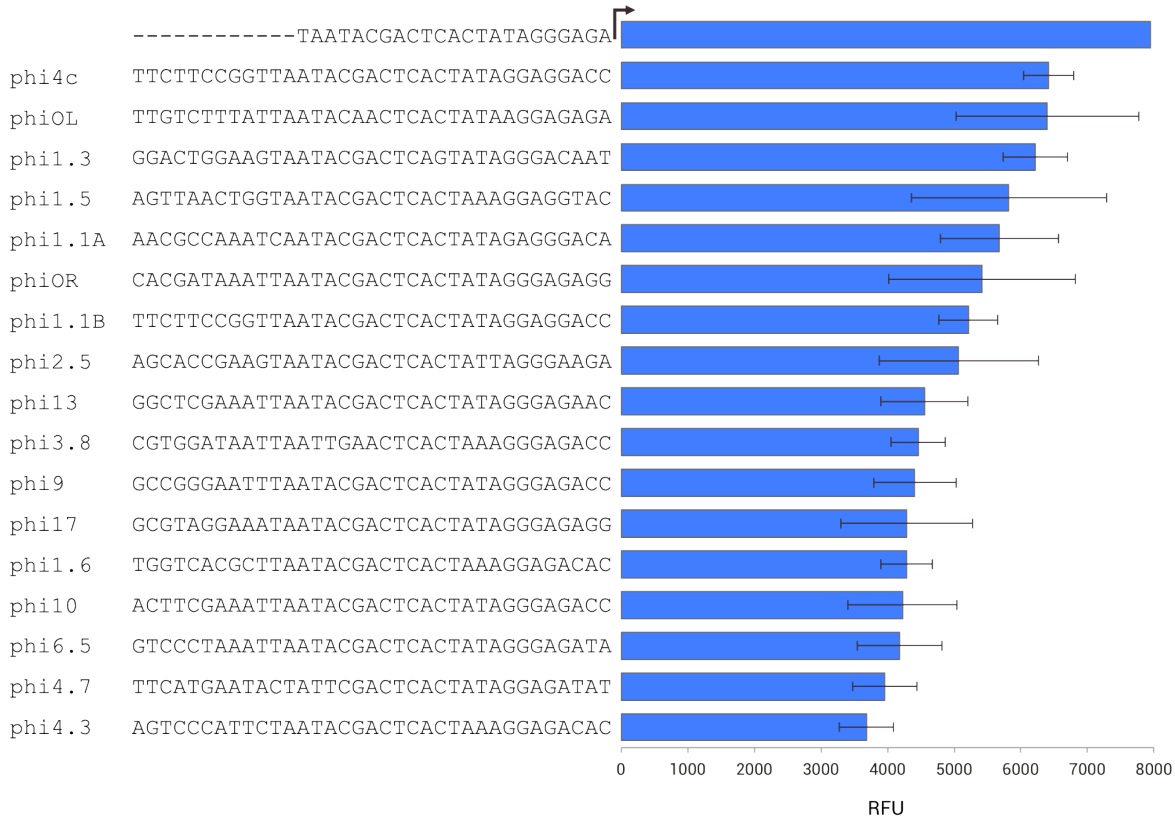

Figure 14 | The 17 genomic promoters of bacteriophage T7 were characterized alongside the consensus T7 core promoter using flow cytometry. The consensus sequence (top) drives gene expression higher than the native promoters.

Extending T7 promoters with designed UP elements

T7 RNAP:T7 promoter specificity is mediated by a specificity loop (residues 742-773) that interacts with -10 and -11. The -7 and -9 postitions are also highly conserved[35]. In fact, changing the nucleotide sequences of the core T7 promoter at any position may destroy its specificity and therefore render it non-orthogonal. A study in 2005 reported that an AT-rich region upstream of the core promoter (i.e., upstream of -17) contributes to

the strength of T7 Class III promoters whereas weaker Class II promoters have more GC content in this region[36]. The authors suggest that a T7 UP element may exist despite the fact that the T7 RNAP is a single catalytic protein and does not have a subunit similar to *E. coli*'s RNAP alpha subunit. I sought to extend bacteriophage T7 promoters in a manner similar to what I have described above for endogenous *E. coli* promoters. In addition to the existing UP element library for *E. coli* promoters, a multiple sequence alignment of the 17 native T7 promoters served as a guide for designing a variety of UP element sequences (Figure 16). I placed this library of UP elements upstream of the consensus T7 promoter and characterized these new transcriptional units using flow cytometry.

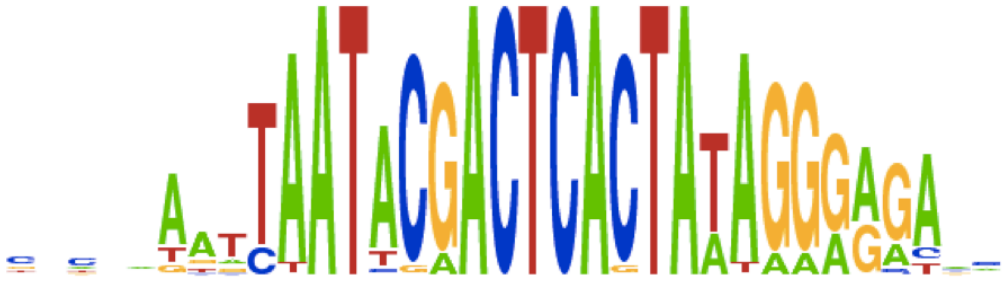

Figure 15 | A sequence logo for the T7 promoter was generated using a multiple sequence alignment of the 17 native T7 promoters, servicing as a guide for designing novel T7 UP elements (-27 through -18).

The data show that each new construct produces less fluorescence than the reference promoter (Figure 17). This is not entirely surprising since these UP elements are contributing to the overall binding of a promoter that already has a high affinity for T7 RNAP. AT-rich sequences such as these upstream of T7 promoters may not be binding in a highly specific manner (as the *E. coli* UP elements bind to the CTDs of the RNAP alpha subunits), but nonspecific interactions and DNA curvature may be helping the T7 RNAP bind to its promoter region[37].

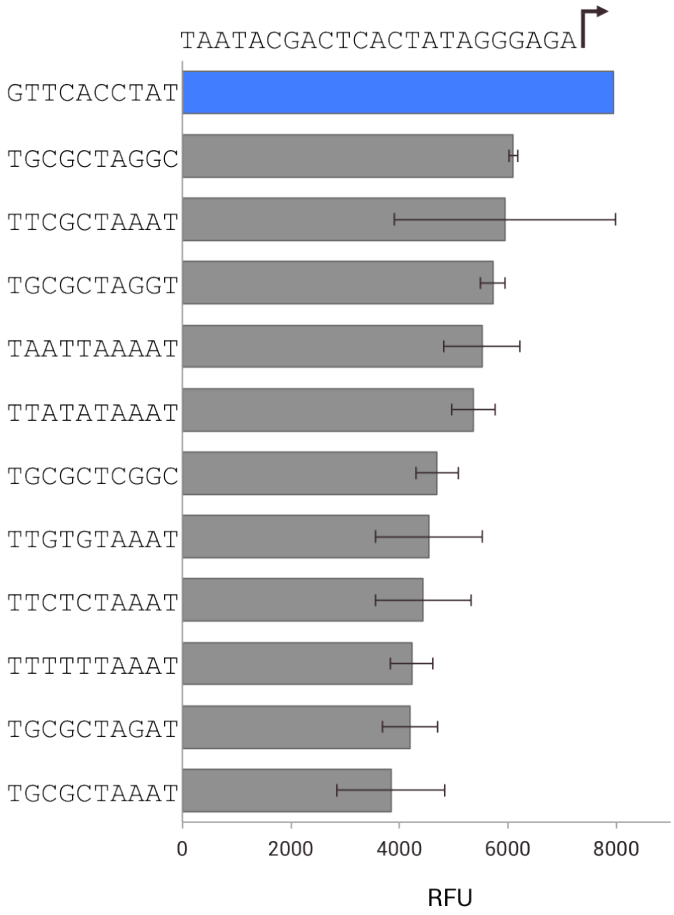

Figure 16 | T7 UP elements from a designed library were placed upstream of the consensus core promoter and measured using flow cytometry. Gene expression levels decrease upon addition of any UP element most likely due to the already high affinity the consensus core promoter has for T7 RNAP.

Characterization of an orthogonal gene expression system

Gene expression tuning has been the focus of a large body of work produced by the metabolic engineering community due to the need to balance energy and material (e.g., carbon) flux through the synthetic metabolic pathway and through the interface to the host metabolism. However, it is unreasonable to assume that there is no need to balance the flux of *information* through gene expression pathways (i.e., transcription and translation). That is, we should not assume that the concentration of available RNAPs and ribosomes in the cell are not rate-limiting[26]. Building from recent work that has demonstrated functional

transcription-translation processes that are orthogonal (i.e., parallel, but independent) to endogenous cellular processes (Figure 18)[38], I successfully express RFP using an orthogonal gene expression system (O-system) and describe (in the following chapter) the orthogonal expression of a three-step metabolic pathway.

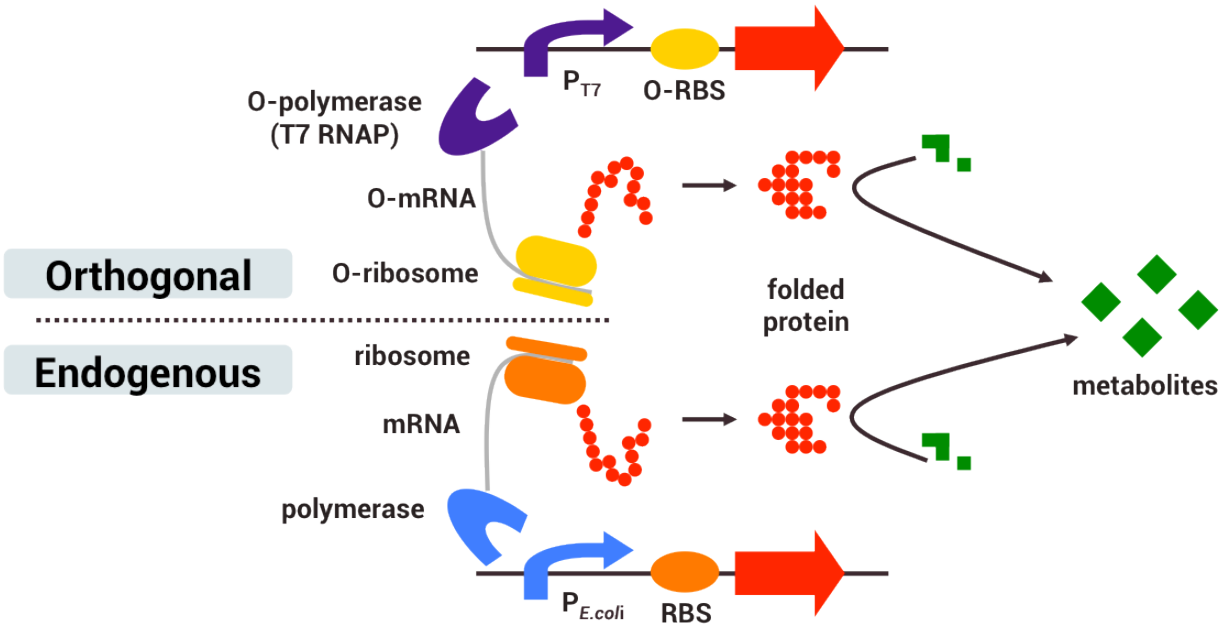

Figure 17 | The concept of an orthogonal gene expression is illustrated above. Expression of genes encoding enzymes can be carried out by native, endogenous cellular machinery (i.e., *E. coli* σ^{70} RNAP, *E. coli* ribosome). However, it can also be carried out by parallel but independent (i.e., orthogonal) processes in which orthogonal cellular machinery (e.g., T7 RNAP, O-ribosome) do not cross-react with native regulation. An orthogonal process may decrease the overall metabolic burden placed on the cell by the demands of the introduced pathway or system.

Orthogonal gene expression should provide a stable, steady supply of cellular machinery dedicated to expressing the gene(s) of interest. In the case of enzymes, this approach may dramatically increase productivity by decoupling the expression of the pathway genes from native cellular processes. To demonstrate orthogonal gene expression, I designed a plasmid (pGM-O-RFP) that houses two modules, the O-system self generator and the O-system reporter. The self generator module uses an initial T7 RNAP input (expressed from a chromosomally integrated gene) to bootstrap production of additional T7 RNAP and orthogonal 16S rRNA (for orthogonal ribosome assembly). Once a sufficient

pool is established, the orthogonal machinery should begin to express its intended target, an RFP reporter in this scenario (Figure 19).

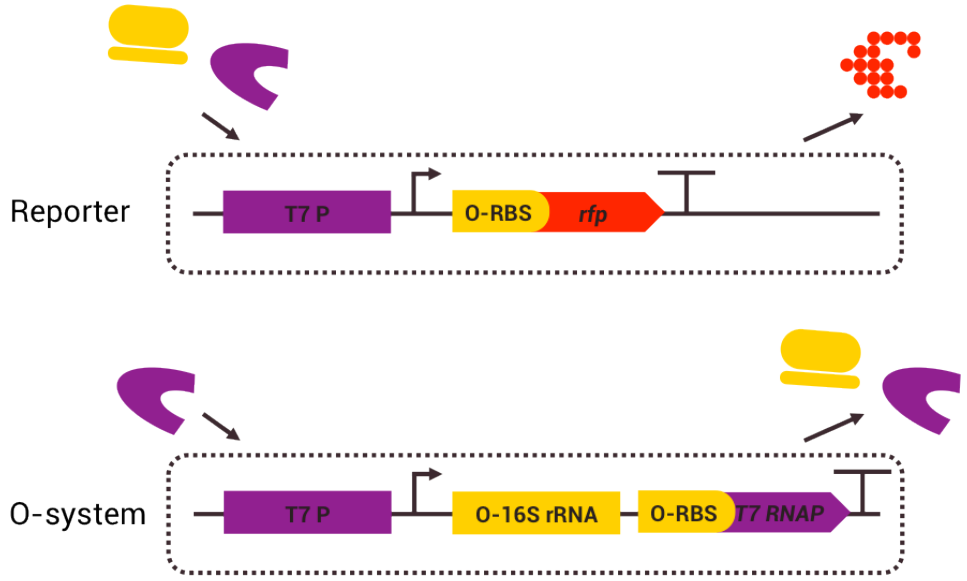

Figure 18 | The prototype O-system is autocatalytic, but requires an initial T7 RNAP pool to boot up. T7 RNAP is expressed from the chromosome of *E. coli* strain BLR(DE3) by inducing with IPTG (1 mM). The O-system is built from two components, T7 RNAP and a 16S rRNA sequence that includes an orthogonal anti-RBS. Once a sufficient pool of orthogonal components is established, the reporter (*rfp*) is orthogonally expressed.

As previously mentioned, the bacteriophage T7 promoter and RNAP pair have been used extensively for over-expressing recombinant proteins. T7 RNAP has a high affinity for the T7 promoter, which is responsible for the high gene expression levels (relative to endogenous expression levels). However, the specificity of the molecular interactions involved in T7 transcription makes this machinery orthogonal to the host, which can be exploited to achieve sophisticated tasks without concerns for directly burdening native transcription processes. Although phage-based orthogonal transcription machinery has been studied and used for decades, no such natural translation machinery has been elucidated. Orthogonal translation was first engineered and demonstrated twenty years ago, but has not been extensively adopted. Recently, an orthogonal gene expression system was demonstrated to work but not used for any application[38]. The orthogonal ribosomes

used in this work were experimentally generated via a high-throughout dual screen of a combinatorial library to identify orthogonal 16S rRNA (i.e., an anti-RBS sequence within the 16S rRNA that is orthogonal to native RBSs). In another study, researchers computationally generate a family of orthogonal 16S rRNA anti-RBS and RBS pairs[39]. A pair from this study was selected to be incorporated into my O-system (Figure 20).

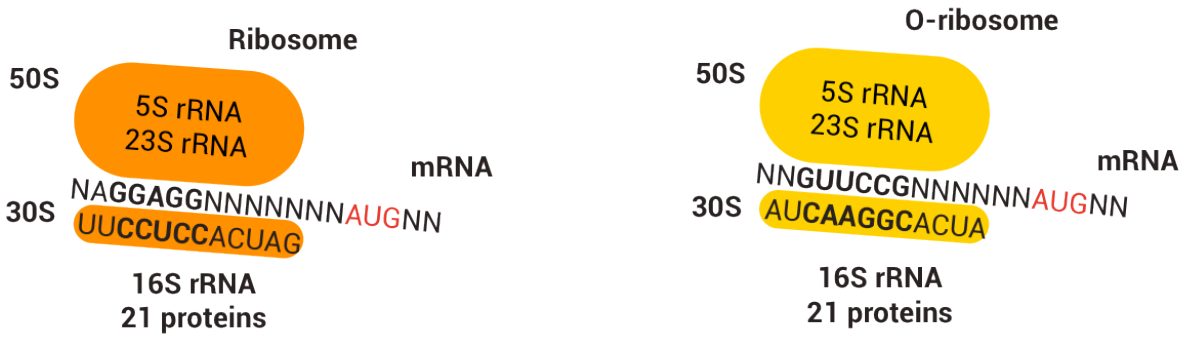

Figure 19 | The orthogonal ribosome (in yellow) differs from the native ribosome (in orange) only by the anti-RBS (bold) found within the 16S rRNA subunit, which recognizes a different RBS (bold). This particular orthogonal RBS and anti-RBS pair was generated computationally and reported by Rao and colleagues.

The O-system was turned on by inducing the expression of T7 RNAP from the chromosome of *E. coli* strain BLR(DE3) by adding 1.0 mM IPTG to a culture in the exponential growth phase (~OD=0.2). After 6 hours of additional growth, samples were analyzed using flow cytometry and fluorescence was measured at the appropriate wavelength. Surprisingly, the designed O-system worked as intended without modification (Figure 21). However, the system appears to be very leaky, with RFP being expressed from a culture that was not induced. This is likely due to basal expression of T7 RNAP expressed from the chromosome. Because T7 RNAP is incredibly processive, even a small amount can accumulate product.

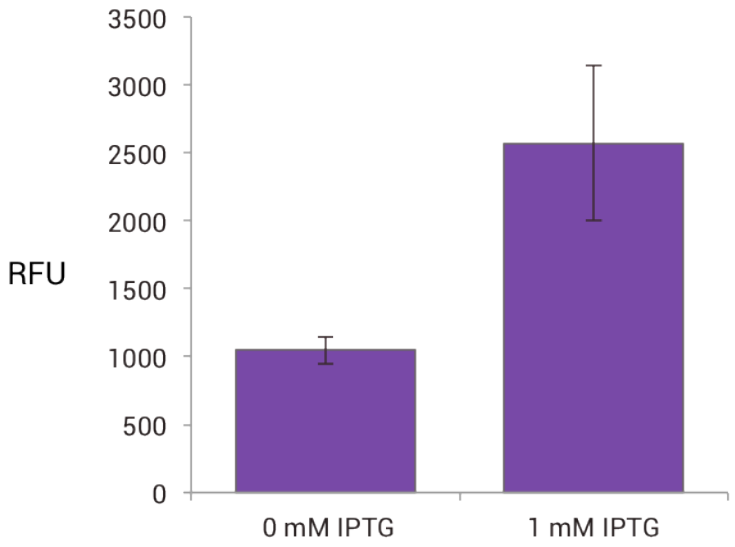

Figure 20 | The O-system was turned on by adding IPTG (1 mM) to the culture. After 6 hours, fluorescence was measured using flow cytometry. Although the O-system appears to be leaky (likely due to basal amounts of T7 RNAP expression from the chromosome), it appears to be functional.

Chapter 4 – Orthogonal expression of a lycopene biosynthetic pathway

"Imagination is more important than knowledge." – Einstein

There are countless chemicals produced by living systems all over the world, most of which have yet to be identified. Metabolic engineers will take advantage of the synthetic capabilities of biology and will eventually replace most of the synthetic chemistry that has dominated the chemical industry (due to petroleum feedstock). But what do we start with? A large number of natural products are already known. Many of these are high value per unit volume molecules and would seem to be economically sound targets.

One of these compounds, lycopene, is a common target for metabolic engineering projects due to its red pigmentation – is it easy to detect and measure[6,40-43]. In addition, lycopene (an effective antioxidant) and other carotenoids such as β-carotene have long been recognized as important, beneficial compounds to human and animal health. Indeed, an entire industry has been built around nutraceuticals (projected to be a $250 billion market by 2018) in which nutrients, dietary supplements and preventative medicines are sold as consumer products.

I sought to express in *E. coli* (natively non-carotenogenic) a lycopene biosynthesis pathway originally from the phytopathogenic bacterium *Erwinia herbicola* (also known as *Pantoea agglomerans*). The genes encoding the enzymes in this pathway are found in the *crt* gene cluster. Starting with farnesyl diphosphate (FPP) in *E. coli* (a product of the native DXP pathway), three exogenous genes must be expressed for the pathway to be realized (Figure 22). The chemicals transformations are facilitated by enzymes encoded by *crtEBI*, illustrated in Figure 22 on the right. The three-step pathway condenses an FPP and an IPP

into one GGPP which is subsequently polymerize into phytoene. Phyotene is the first commited step in the carotenoid pathway. This molecule can be desaturated to form lycopene in a complex, multi-step reaction facilitated by CrtI. The step beyond lycopene in this pathway produces β-carotene.

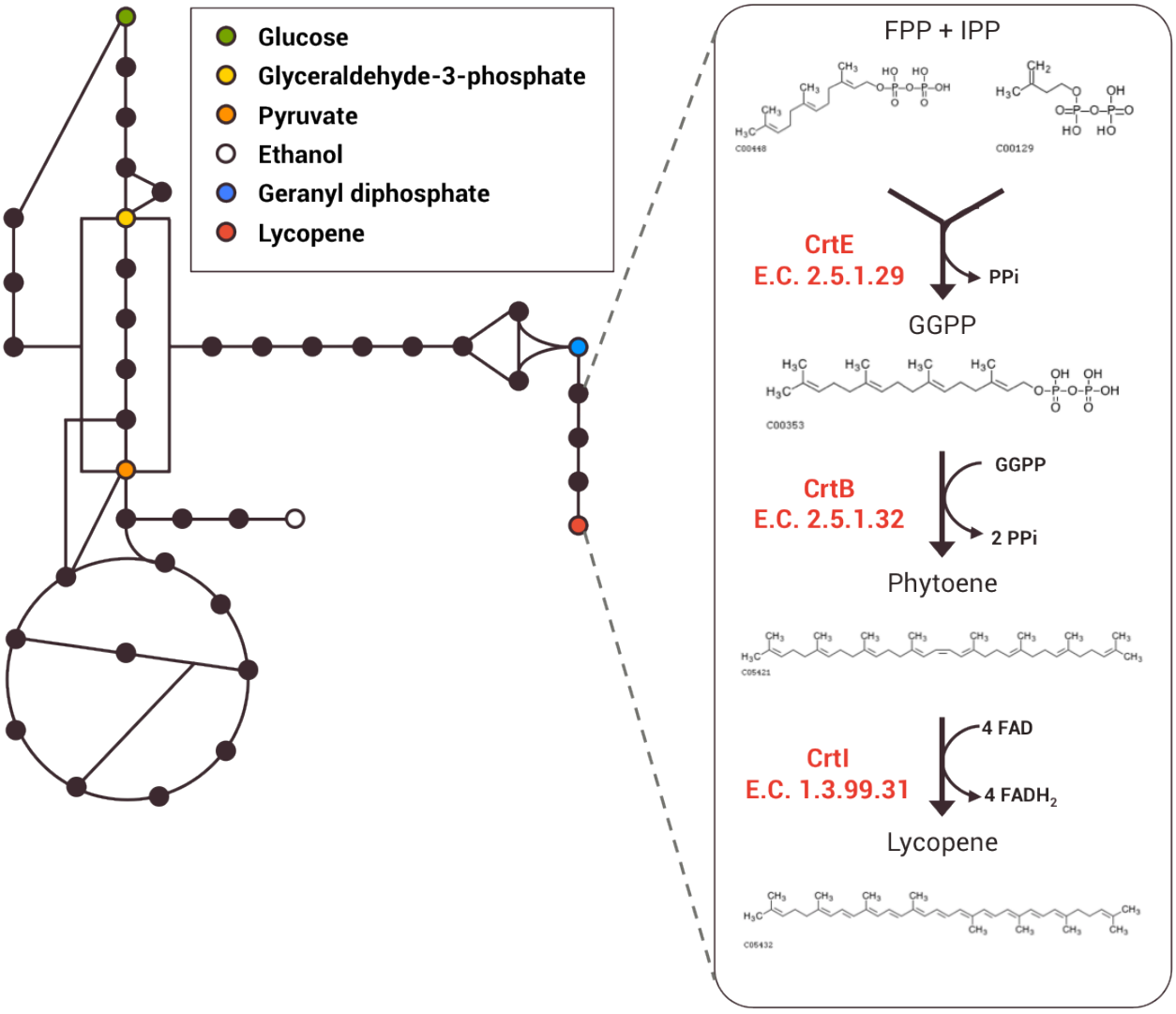

Figure 21 | A cartoon of central metabolism shows the relationship between the lycopene biosynthesis pathway and glycolysis and the TCA cycle (left). Lycopene is synthesized in three enzymatic steps from products of the DXP pathway (right). These enzymes, encoded by *crtEBI*, are not native to *E. coli* and must be added in order to achieve lycopene production.

The pAC-LYC plasmid has been used for decades by researches interested in lycopene biosynthesis[43,44]. I designed a series of plasmids based on the pJ401 plasmid previously used for promoter characterization (Figure 23). These plasmids were designed to express the gene cluster in the order of *crtEIB*. The first plasmid, pJ401-crtEIB, uses native transcription and translation to express the pathway genes. The second plasmid, pJ401-T7-crtEIB, uses orthogonal transcription and native translation to express the genes. The third plasmid, pJ401-O-crtEIB, uses the complete O-system to express the lycopene pathway genes.

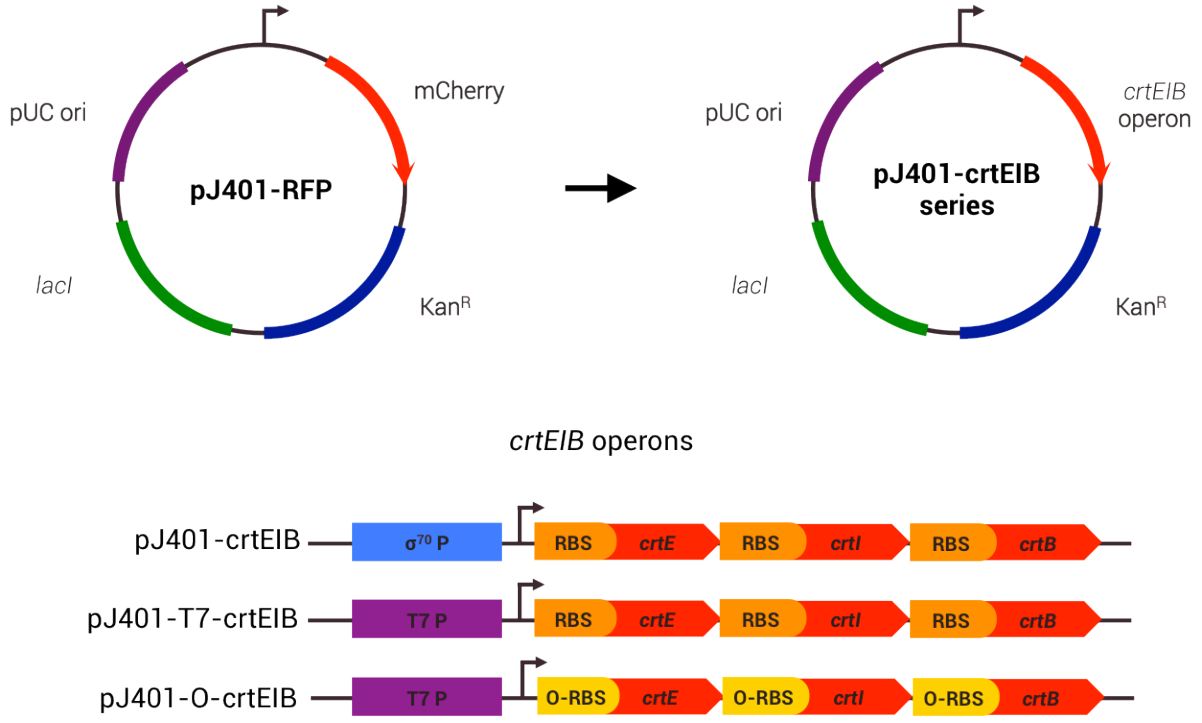

Figure 22 | A series of plasmids were constructed to introduce the lycopene pathway genes to *E. coli*. The *rfp* in pJ401-RFP (previously used for promoter characterization) was swapped out for a refactored operon, *crtEIB* using isothermal DNA assembly (top). This operon was expressed in three ways: 1) native transcription, native translation; 2) orthogonal transcription, native translation; and 3) orthogonal transcription, orthogonal translation (bottom). NOTE: Due to a DNA assembly design error, the O-16S rRNA sequence was not added to the plasmid requiring orthogonal translation.

These plasmids were assembled using isothermal DNA assembly from PCR-amplified fragments (from the pAC-LYC and pJ401 plasmids, see Appendix B for genbank files). DNA primers, PCR and assembly protocols were designed using the web-based J5 DeviceEditor software in less than 30 minutes[45]. The pieces assembled together as expected and were used to transform *E. coli* strain BLR(DE3). Transformants were selected and grown shaking overnight (~16 hours) at 37°C. These cultures were diluted by 100-fold in 10 mL cultures and grown shaking at room temperature for 36 hours before being characterized. Samples were collected and lycopene was extracted using a standard acetone extraction procedure[46]. Lycopene content was inferred by measuring the absorbance of the extraction supernatant at a wavelength of 475 nm. These values were then divided by the cell density of the population (measured by flow cytometry) and then multipled by an arbitrary factor to yield the data presented in Figure 24.

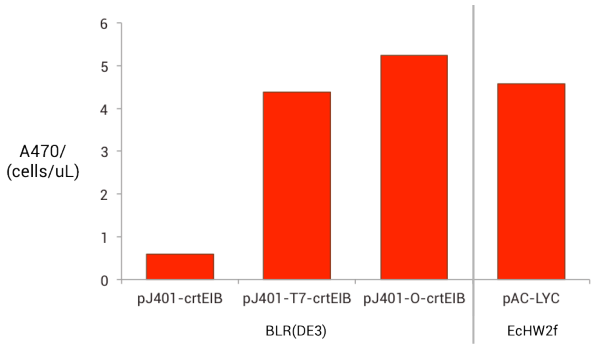

Figure 23 | Lycopene production was determined using a simple colormetric assay in which absorbance at 470 nm of extracted lycopene in acetone was measured using a spectrophotometer. These values were normalized to the density of the bacterial culture (i.e., the concentration of cells), which was measured using flow cytometry. The bars on the left of the gray line represent the constructed plasmids expressed in BLR(DE3). The all-native expression of the *crtEIB* operon did not result in high lycopene levels. Unfortunately, a design flaw for the assembly of the all-orthogonal expression of the operon resulted in the exclusion of the O-16S rRNA subunit, which is theoretically necessary for expression. Lycopene pathway genes were clearly expressed, however. This indicates that native ribosomes must have translated the orthogonal transcript and, therefore, the orthogonality of this RBS and anti-RBS pair is not as it was purported to be. Two of these constructs performed roughly as well as a previously reported optimized strain, EcHW2f (right of gray line). Future work includes combining the engineered plasmids presented in this work with the optimized strain, which may improve lycopene production to unprecedented levels.

Chapter 5 – Conclusions and perspective – next-generation metabolic engineering

"What I cannot build, I do not understand." – Richard Feynman

Despite large advances, the implementation of predictable, novel functions in cells remains an unmet goal of synthetic biology. To overcome this challenge, the synthetic biology community is working to generate and characterize large libraries of genetic parts with the aim of elucidating hidden biological design principles and eventually develop models for the *de novo* forward engineering of genetic systems[47-49]. For example, significant progress in this area has been demonstrated for ribosome binding sites. One of the developments that has facilitated this approach is large-scale *de novo* DNA synthesis, allowing the engineering of DNA sequences with single-nucleotide resolution. In addition, single-cell measurement technologies such as flow cytometry and time-lapse microscopy have enabled quantitative measurements of part performance and analysis of relevant system parameters[50].

While research in metabolic engineering has continued to advance with a growing number of genetic, analytical, and computational tools, the application of synthetic biology to metabolic engineering has the potential to create a paradigm shift. Rather than starting with the full complement of components in a wild-type organism and piece-wise modifying and streamlining its function, metabolic engineering can be attempted using a parts-based approach to design by carefully and rationally specifying the inclusion of each necessary component. Just as cars are routinely built from modular components in assembly lines, rapid, modularized construction of microbial chemical factories is on the horizon. For example, putative genes identified computationally in a metagenomics study need not be

isolated and cloned; the sequence information can be outsourced to synthesis companies and subsequently screened for function as Bayer *et al*. recently demonstrated[51]. This capability opens many opportunities for metabolic engineers. First of all, the convenience of this approach over traditional cloning allows for the systematic generation of genetic part variants such as promoter libraries. Secondly, it provides a practical way for eliminating restriction sites or undesirable RNA secondary structures and to codon-optimize genes for the expression in heterologous hosts. Thirdly and perhaps most importantly, it decouples molecular cloning from genetic design.

I have learned several lessons thoughout the course of this dissertation research. First of all, what I have built, I do not necessarily understand. It is apparent that for even *E. coli*, the best-characterized model organism, we know so little about biology. I was perplexed to learn that my high-copy number plasmids could actually vary from 50-300 copies per cell, even though they are routinely described as ~100 copies/cell plasmids. The difference between expressing an RFP reporter in a low-, medium- and high-copy number plasmid can be discerned by the naked eye (Figure 25). Yet the literature fails to address these basic issues.

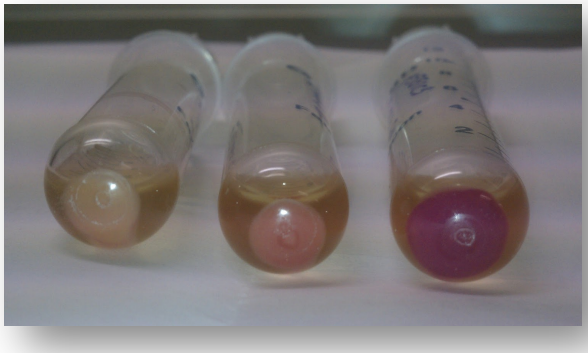

Figure 24 | The copy number of an expression plasmid has a huge impact on gene expression levels (and, unsurprisingly, noise in gene expression). Here, RFP is being made in the same strain of *E. coli* grown to the same OD in the same media using the same expression cassette but on different plasmids that have different replication origins. From left to right: pSB3C5 (low copy # of ~5, pSC101 ori), pSB3C4 (medium copy # of 10-12, p15A ori), pSB1C3 (high copy # of 100-300, pMB1 ori).

It turns out that environmental, cellular and genetic context are very important when trying to characterize or engineer a genetic system. For example, the expression of the consensus core promoter in NEB10β is very different from the expression of the same piece of DNA and in the same environmental conditions in BLR(DE3) (Figure 26). This demonstrates that cellular context (i.e., the strain background as it relates to the physiology of the cell) plays an important role in gene expression. Other considerations include culture medium, growth conditions such as aeration and temperature, antibiotic selection marker, plasmid size, plasmid-based or chromosomal expression and variations of codon usage in coding sequences, just to name a few.

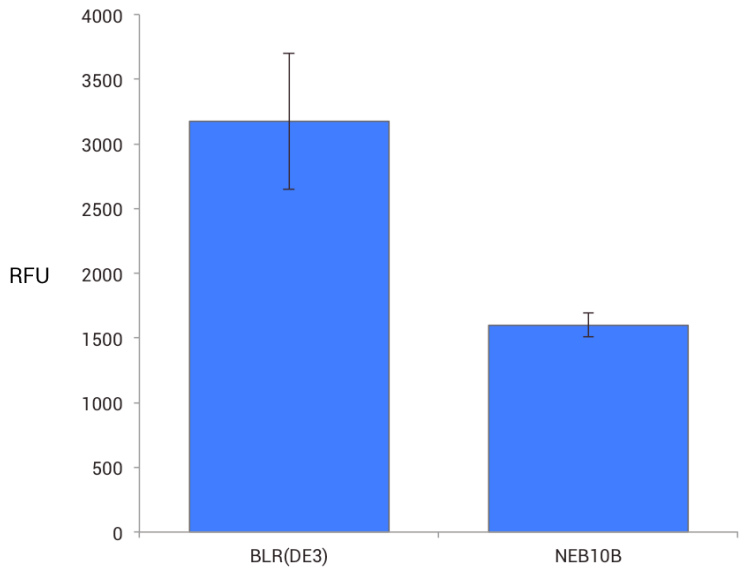

Figure 25 | The expression of *rfp* using the same plasmid in two different strains of *E. coli* results in very different protein production levels (even when grown in the same conditions). This demonstrates the importance of choosing the appropriate strain background (i.e., physiology, determined by its metabolic network) when attempting to express non-native gene products.

As new genetic systems are developed and characterization methods mature, it will join three other major areas of research that will converge to transform metabolic engineering into the platform discipline of the next industrial revolution. The three other areas are: genome sequencing and bioinformatics, systems and computational biology, and genome editing and synthetic genomics. It is critical that metabolic engineers combine knowledge of genome-scale metabolic modeling/simulation, methodologies for building and editing entire genomes, measurement techniques and analysis of omics data, and the predictable design of genetic systems from well-characterized components (Figure 27).

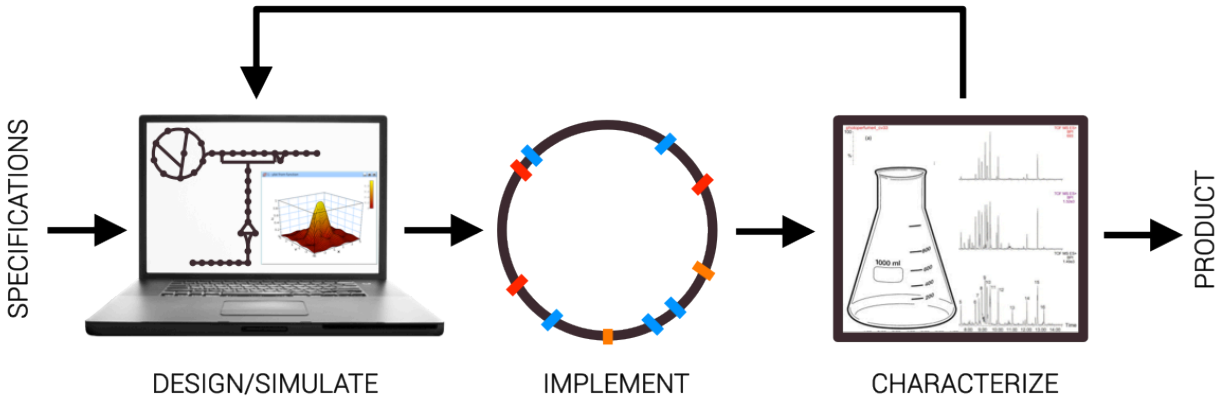

Figure 26 | Microbial strains will soon be designed to specificiations by synthetic biology companies and sold as products to customers such as chemical manufacturers. To achieve this, an engineering framework must be used to integrate model-guided design, implementation and characterization of synthetic microbial metabolism. An important step will be bridging the model-guided design and implementation steps. For example, once a genomic target has been identified for upregulation, what DNA sequence should be added or inserted to the chromosome to accomplish this? The genotype-phenotype relationship will slowly become illuminated as more and more genetic elements and gene networks are built and analzed.

The tools that I have described in this dissertation should be useful to the metabolic engineering community and beyond. I have demonstrated that UP elements can be used to extend core promoters to 1) finely modulate gene expression through altering binding affinity for RNAP, 2) reduce variation in gene expression of otherwise noisy promoters and 3) modify promoters that have constrained architectures. I have shown that UP elements can be used to affect transcription driven by T7 promoters. Finally, I have demonstrated a functional orthogonal gene expression system and have expressed a lycopene biosynthesis pathway using native and orthogonal cellular machinery.

References

1. Keasling, J. D. Manufacturing Molecules Through Metabolic Engineering. *Science* **330**, 1355–1358 (2010).
2. Nakamura, C. E. & Whited, G. M. Metabolic engineering for the microbial production of 1,3-propanediol. *Current Opinion in Biotechnology* **14**, 454–459 (2003).
3. Esvelt, K. M. & Wang, H. H. Genome-scale engineering for systems and synthetic biology. *Molecular Systems Biology* **9**, 1–17 (2013).
4. McArthur, G. H. & Fong, S. S. Toward Engineering Synthetic Microbial Metabolism. *Journal of Biomedicine and Biotechnology* **2010**, 1–11 (2010).
5. Cox, R. S., Surette, M. G. & Elowitz, M. B. Programming gene expression with combinatorial promoters. *Molecular Systems Biology* **3**, (2007).
6. Alper, H. H., Fischer, C. C., Nevoigt, E. E. & Stephanopoulos, G. G. Tuning genetic control through promoter engineering. *Proc Natl Acad Sci U S A* **102**, 12678–12683 (2005).
7. Hammer, K. K., Mijakovic, I. I. & Jensen, P. R. P. Synthetic promoter libraries - tuning of gene expression. *Trends in Biotechnology* **24**, 3–3 (2006).
8. Browning, D. F. & Busby, S. J. The regulation of bacterial transcription initiation. *Nat Rev Micro* **2**, 57–65 (2004).
9. Ross, W. *et al.* A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* **262**, 1407–1413 (1993).
10. Meng, W. *et al.* UP element-dependent transcription at the Escherichia coli rrnB P1 promoter: positional requirements and role of the RNA polymerase α subunit linker. *Nucleic Acids Research* **29**, 4166–4178 (2001).
11. Aiyar, S. E. S., Gourse, R. L. R. & Ross, W. W. Upstream A-tracts increase bacterial promoter activity through interactions with the RNA polymerase alpha subunit. *Proc Natl Acad Sci U S A* **95**, 14652–14657 (1998).
12. Ross, W., Aiyar, S. E., Salomon, J. & Gourse, R. L. Escherichia coli promoters with UP elements of different strengths: modular structure of bacterial promoters. *Journal of Bacteriology* **180**, 5375–5383 (1998).
13. Ross, W. W., Ernst, A. A. & Gourse, R. L. R. Fine structure of E. coli RNA polymerase-promoter interactions: alpha subunit binding to the UP element minor groove. *Genes & Development* **15**, 491–506 (2001).
14. Lloyd, G. S. G., Niu, W. W., Tebbutt, J. J., Ebright, R. H. R. & Busby, S. J. W. S. Requirement for two copies of RNA polymerase alpha subunit C-terminal domain for synergistic transcription activation at complex bacterial promoters. *Genes & Development* **16**, 2557–2565 (2002).
15. Gaal, T. *et al.* DNA-binding determinants of the alpha subunit of RNA polymerase: novel DNA-binding domain architecture. *Genes & Development* **10**, 16–26 (1996).
16. Gourse, R. L. R., Ross, W. W. & Gaal, T. T. UPs and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition. *Molecular Microbiology* **37**, 687–695 (2000).
17. Estrem, S. T., Gaal, T., Ross, W. & Gourse, R. L. Identification of an UP element

consensus sequence for bacterial promoters. *Proc Natl Acad Sci U S A* **95**, 9761–9766 (1998).
18. Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nature Biotechnology* **22**, 1567–1572 (2004).
19. Elowitz, M. B. M. & Leibler, S. S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).
20. Shetty, R. P., Endy, D. & Knight, T. F. Engineering BioBrick vectors from BioBrick parts. *J Biol Eng* **2**, 5–5 (2008).
21. Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nature Biotechnology* **27**, 946–950 (2009).
22. Fried, M. G. Measurement of protein-DNA interaction parameters by electrophoresis mobility shift assay. *Electrophoresis* **10**, 366–376 (1989).
23. Niedziela-Majka, A. A. & Heyduk, T. T. Escherichia coli RNA polymerase contacts outside the -10 promoter element are not essential for promoter melting. *The journal of biological chemistry* **280**, 38219–38227 (2005).
24. Schroeder, L. A., Choi, A.-J. & DeHaseth, P. L. The -11A of promoter DNA and two conserved amino acids in the melting region of sigma70 both directly affect the rate limiting step in formation of the stable RNA polymerase-promoter complex, but they do not necessarily interact. *Nucleic Acids Research* **35**, 4141–4153 (2007).
25. Swain, P. S. P., Elowitz, M. B. M. & Siggia, E. D. E. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci U S A* **99**, 12795–12800 (2002).
26. Vind, J. J., Sørensen, M. A. M., Rasmussen, M. D. M. & Pedersen, S. S. Synthesis of proteins in Escherichia coli is limited by the concentration of free ribosomes. Expression from reporter genes does not always reflect functional mRNA levels. *Journal of Molecular Biology* **231**, 678–688 (1993).
27. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Research* **14**, 1188–1190 (2004).
28. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth* **6**, 343–345 (2009).
29. Gibson, D. G. *et al.* Complete Chemical Synthesis, Assembly, and Cloning of a Mycoplasma genitalium Genome. *Science* **319**, 1215–1220 (2008).
30. Gibson, D. G. *Chapter 15 - Enzymatic Assembly of Overlapping DNA Fragments*. *Synthetic Biology Part B* **498**, 349–361 (Elsevier Inc., 2011).
31. Gibson, D. G. D., Smith, H. O. H., Hutchison, C. A. C., Venter, J. C. J. & Merryman, C. C. Chemical synthesis of the mouse mitochondrial genome. *Nat Meth* **7**, 901–903 (2010).
32. Borukhov, S. S. & Severinov, K. K. Role of the RNA polymerase sigma subunit in transcription initiation. *Res Microbiol* **153**, 557–562 (2002).
33. Davidson, E. A. E., van Blarcom, T. T., Levy, M. M. & Ellington, A. D. A. Emulsion based selection of t7 promoters of varying activity. *CORD Conference Proceedings* 433–443 (2010).
34. Lopez, P. J. P., Guillerez, J. J., Sousa, R. R. & Dreyfus, M. M. On the mechanism of inhibition of phage T7 RNA polymerase by lac repressor. *Journal of Molecular Biology*

276, 15–15 (1998).
35. Ikeda, R. A. R., Ligman, C. M. C. & Warshamana, S. S. T7 promoter contacts essential for promoter activity in vivo. *Nucleic Acids Research* **20**, 2517–2524 (1992).
36. Tang, G.-Q., Bandwar, R. P. & Patel, S. S. Extended upstream AT sequence increases T7 promoter strength. *The journal of biological chemistry* **280**, 40707–40713 (2005).
37. Nickerson, C. A. C. & Achberger, E. C. E. Role of curved DNA in binding of Escherichia coli RNA polymerase to promoters. *Journal of Bacteriology* **177**, 5756–5761 (1995).
38. An, W. & Chin, J. W. Synthesis of orthogonal transcription-translation networks. *Proc Natl Acad Sci U S A* **106**, 8477–8482 (2009).
39. Chubiz, L. M. & Rao, C. V. Computational design of orthogonal ribosomes. *Nucleic Acids Research* **36**, 4038–4046 (2008).
40. Alper, H., Miyaoku, K. & Stephanopoulos, G. Construction of lycopene-overproducing E. coli strains by combining systematic and combinatorial gene knockout targets. *Nature Biotechnology* **23**, 612–616 (2005).
41. Kang, M. J. *et al.* Identification of genes affecting lycopene accumulation inEscherichia coli using a shot-gun method. *Biotechnol. Bioeng.* **91**, 636–642 (2005).
42. Farmer, W. R. & Liao, J. C. Improving lycopene production in Escherichia coli by engineering metabolic control. *Nature Biotechnology* **18**, 533–537 (2000).
43. Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894–898 (2009).
44. Hundle, B. S. B., Beyer, P. P., Kleinig, H. H., Englert, G. G. & Hearst, J. E. J. Carotenoids of Erwinia herbicola and an Escherichia coli HB101 strain carrying the Erwinia herbicola carotenoid gene cluster. *Photochem Photobiol* **54**, 89–93 (1991).
45. Hillson, N. J., Rosengarten, R. D. & Keasling, J. D. j5 DNA Assembly Design Automation Software. *ACS Synth. Biol.* **1**, 14–21 (2012).
46. Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* 1–6 (2009).doi:10.1038/nature08187
47. Mutalik, V. K. *et al.* Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat Meth* 1–15 (2013).doi:10.1038/nmeth.2404
48. Mutalik, V. K. *et al.* Quantitative estimation of activity and quality for collections of functional genetic elements. *Nat Meth* 1–12 (2013).doi:10.1038/nmeth.2403
49. Endy, D. Foundations for engineering biology. *Nature* **438**, 449–453 (2005).
50. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).
51. Bayer, T. S. *et al.* Synthesis of Methyl Halides from Biomass Using Engineered Microbes. *J. Am. Chem. Soc.* **131**, 6508–6515 (2009).

Appendix A – Methods

Bacterial strains and growth conditions

Strains NEB10β and BLR(DE3) were used to characterize native and orthogonal gene expression, respectively. The genotype of NEB10β is: *araD139 Δ(ara-leu)7697 fhuA lacX74 galK (ϕ80 Δ(lacZ)M15) mcrA galU recA1 endA1 nupG rpsL* (Str^R)*Δ(mrr-hsdRMS-mcrBC)*. The genotype of BLR(DE3) is: *F- ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm*(DE3) Δ(srl-recA)306::Tn*10* (Tet^R). All experiments used *E. coli* cultured in Luria-Bertani (LB) media (Sigma-Aldrich) at 37°C and skaking at 250 rpm.

Plasmid assembly

Designed promoters were inserted into either plasmid pSB1C3 (partsregistry.org) or pJ401 (DNA2.0) upstream of a red fluorescence protein (*rfp*) coding sequence. These insertions were accomplished using a ssDNA isothermal DNA assembly protocol built upon recently published method. The following MATLAB program is used to expedite DNA assembly design by generating overlapping 60mer oligos from an input DNA sequence.

```
function [H]=oligos(DNA)
% function [H]=oligos(DNA)
%
% Author: George McArthur
% Adapted from a PCA oligo generator written by my student Jaeeung Kim
% Tool to split a DNA sequence into alternating 60-mers with 20bp overlaps
for ssDNA isothermal DNA assembly
% Note about the input DNA sequence: (DNA length - 20) must be evenly
divisible by 40.
%
% C is Coding strand oligos
% T is Template strand oligos
% H is Hybrid oligos (alternating oligos between T and C strands)
% All oligos are 60-mers with 20bp overlaps and listed 5' - 3'

LengthDNA=length(DNA);
NumRow=(LengthDNA-20)/40; % How many rows (oligos) will be in final output
if rem(LengthDNA-20,40)~=0
    error('notDivisible','Gene length cannot be evenly divided into
60mers!');
end

DNA=upper(DNA); % Makes sequence uppercase
SP=1;   % Starting point index
EP=60;  % End point index
oligosC=num2str(zeros(60,NumRow));

for k=1:NumRow
    oligosC((1+60*(k-1)):60+60*(k-1))=DNA(SP:EP);
    SP=SP+40;
    EP=EP+40;
end
```

```
oligosC=oligosC';
C=oligosC; % Coding sequence! (5' --> 3')

Size=size(oligosC);
Row=Size(1);
Col=Size(2);
oligosT=oligosC;

for k=1:(Row*Col)
    if oligosC(k) == 'A'
        oligosT(k)= 'T';
    elseif oligosC(k) == 'T'
        oligosT(k)= 'A';
    elseif oligosC(k) == 'G'
        oligosT(k) = 'C';
    elseif oligosC(k) == 'C'
        oligosT(k)= 'G';
    end
end

T=oligosT; % Template sequence! (3' --> 5')
rT=seqreverse(T); % Template sequence (5' --> 3')

NumOdd=[1:2:NumRow];
NumEven=[2:2:NumRow];
% start making H (alternating between Template and Coding sequences)
for k=1:length(NumEven)
    H(NumEven(k),:)=C(NumEven(k),:);
end

for k=1:length(NumOdd)
    H(NumOdd(k),:)=rT(NumOdd(k),:);
end

end
```

Promoter characterization using flow cytometry

E. coli were transformed with these new plasmids using electroporation and selected with the appropriate antibiotic (25 μg chloramphenicol/mL LB or 30 μg kanamycin/mL). Overnight cultures (5 mL) were inoculated from single colonies (triplicates) and grown overnight (~16 hours) at 37°C with shaking at 250 rpm. Cells were then diluted in 10 mL of fresh medium in an Erlenmeyer flask (100X dilution or an OD ~0.05). After 6 hours of growth at 37°C with shaking, 1 mL of culture sample was pelleted and subsequently resuspended in 1.0 mL of phosphate buffer solution. Samples were analyzed by measuring fluorescence from the RFP reporter (excitation wavelength = 584 nm, emission wavelength = 607 nm) using an Accuri C6 flow cytometer (BD Biosciences). Collected fluorescence data were gated based on cell size (determined using a negative control).

Electrophoretic mobility shift assay (EMSA) to determine binding affinity

Binding reactions (20 uL) contained purified E. coli RNAP saturated with σ70 (Epicentre) at various concentrations (12.5 - 400 nM), linearized DNA containing the promoter region (5 or 10 nM) and a simple binding buffer. Sequences of the primers used to PCR-amplify the DNA fragments (~1200 bp) from pSB1C3 are provided in the Supporting Information. The binding buffer (5X) is 750 mM KCl, 0.5 mM DTT, 0.5 mM EDTA and 50 mM Tris at pH 7.4. Reactions were carried out for 30 minutes at 37°C before heparin (200 μg/mL) was added as a competitor and incubated for an additional 10 minutes. Half of each reaction mixture (10 μL) was mixed with 2 μL of EMSA gel-loading solution (Invitrogen) and loaded onto a pre-running, 2.0% agarose gel (1X TAE) for 2.5 hours at 6 V/cm to separate free DNA from bound DNA. Each gel was stained with SYBR Green (Invitrogen) for 20 minutes, washed with water and visualized using a UV transilluminator (BioRad). Gel images were analyzed using ImageJ to quantify fluorescence intensity of each band and determine the fraction of bound DNA. A one-site total binding model in Prism was used to calculate dissociation constants, which is described in the main text.

Lycopene assay

E. coli were electrotransformed with a series of plasmids containing the *crtEBI* genes under the expression of a variety of native and/or orthogonal components and inserted in pJ401 (replacing the *rfp*). 5 mL tube cultures were inoculated with individual colonies picked from selective media and subsequently grown overnight at 37°C with skaking at 250 rpm for ~16 hours. 10 mL cultures in Erlenmeyer flasks were inoculated (100X dilution) and grown for 24 hours at 37°C with skaking at 250 rpm. At this point these cultures were removed from the incubator and grown for an additional 48 hours at room temperature with shaking at 250 rpm. To measure the lycopene content of the various engineered strains, a 1 mL culture sample was centrifuged at 16,000*g* for 1 minute to remove the broth and the pellet was resuspended in 1 mL of water. To extract the lycopene, the sample was centrifuged again and then resuspended in 200 uL of acetone. The mixture was incubated in the dark for 15 minutes at 55°C with vortexting every three minutes. Samples were centrifuged again to remove cellular debris and the relative lycopene content in the supernatant was determined by measuring the absorbance at 470 nm.

Appendix B – DNA sequences

List of synthetic oligonucleotides used to build libraries

Name	Sequence
GM1	TTCCGCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACATAAATA
GM2	GCCCATTATAACACCTGTATTTATTTATGTCAAGTTTTTTTTTTAAAAAAATTTTC
GM3	GACATAAATAAATACAGGTGTTATAATGGGCCAAGAAAAGACGGAACAA
GM4	ATATGTGTCCTCTTTGGTTGTTCCGTCTTTTCTTGGCCCAT
GM5	ATGCATTTAGAATATTCCGCATTTGTATGTTCACCTATGG
GM6	ACTTGGAAACCATATATGTGTCCTCTTTGGTTGTTCC
GM7	AGAGGACACATATATGGTTTCCAAGTGCGAGGA
GM8	ATACAAATGCGGAATATTCTAAATGCATAATAAATACTGATAACATCT
GM9	CCATAGGTGAACATACAAATGCGGAATATTCTAAATGCAT
GM10	GGAACAACCAAAGAGGACACATATATGGTTTCCAAGT
GM11	GGGCGACACAAAATTTATTCTAAATGCATAATAAATACTGATAACATC
GM12	ATGGTTTCCAAGTGCGAGGAGGATAAC
GM13	GTGAGTCGTATTAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAA
GM14	CACCTATTAATACGACTCACTATAGGGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGA
GM15	CAAAATTATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCC
GM16	AGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGAGGAGGA
GM17	CACCTATTAATACGACTCACTATWGGRNNNAGCTGTCACCGGATGTGCTTTCCGGTCTGA
GM18	GTTGTATTAATAAAGACAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM19	TTTGTCTTTATTAATACAACTCACTATAAGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
GM20	TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC
GM21	CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA
GM22	GTCGTATTGATTTGGCGTTATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM23	TAACGCCAAATCAATACGACTCACTATAGAGGGACAAGCTGTCACCGGATGTGCTTTCCG
GM24	GTCGTATTAACCGGAAGAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM25	TTTCTTCCGGTTAATACGACTCACTATAGGAGGACCAGCTGTCACCGGATGTGCTTTCCG
GM26	GTCGTATTACTTCCAGTCCATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM27	TGGACTGGAAGTAATACGACTCAGTATAGGGACAATAGCTGTCACCGGATGTGCTTTCCG
GM28	GTCGTATTACCAGTTAACTATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM29	TAGTTAACTGGTAATACGACTCACTAAAGGAGGTACAGCTGTCACCGGATGTGCTTTCCG
GM30	GTCGTATTAAGCGTGACCAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM31	TTGGTCACGCTTAATACGACTCACTAAAGGAGACACAGCTGTCACCGGATGTGCTTTCCG
GM32	GTCGTATTACTTCGGTGCTATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM33	TAGCACCGAAGTAATACGACTCACTATTAGGGAAGAAGCTGTCACCGGATGTGCTTTCCG
GM34	GTTCAATTAATTATCCACGATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM35	TCGTGGATAATTAATTGAACTCACTAAAGGGAGACCAGCTGTCACCGGATGTGCTTTCCG
GM36	GTCGGATTGTCTCAGTCGGATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM37	TCCGACTGAGACAATCCGACTCACTAAAGAGAGAGAAGCTGTCACCGGATGTGCTTTCCG
GM38	GTCGTATTAGAATGGGACTATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM39	TAGTCCCATTCTAATACGACTCACTAAAGGAGACACAGCTGTCACCGGATGTGCTTTCCG
GM40	GTCGAATAGTATTCATGAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM41	TTTCATGAATACTATTCGACTCACTATAGGAGATATAGCTGTCACCGGATGTGCTTTCCG
GM42	GTCGTATTAATTTAGGGACATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM43	TGTCCCTAAATTAATACGACTCACTATAGGGAGATAAGCTGTCACCGGATGTGCTTTCCG
GM44	GTCGTATTAAATTCCCGGCATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM45	TGCCGGGAATTTAATACGACTCACTATAGGGAGACCAGCTGTCACCGGATGTGCTTTCCG

GM46 GTCGTATTAATTTCGAAGTATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM47 TACTTCGAAATTAATACGACTCACTATAGGGAGACCAGCTGTCACCGGATGTGCTTTCCG
GM48 GTCGTATTAATTTCGAGCCATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM49 TGGCTCGAAATTAATACGACTCACTATAGGGAGAACAGCTGTCACCGGATGTGCTTTCCG
GM50 GTCGTATTATTTCCTACGCATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM51 TGCGTAGGAAATAATACGACTCACTATAGGGAGAGGAGCTGTCACCGGATGTGCTTTCCG
GM52 GTCGTATTAATTTATCGTGATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM53 TCACGATAAATTAATACGACTCACTATAGGGAGAGGAGCTGTCACCGGATGTGCTTTCCG
GM54 GTCGTATTAGCCGAGCGCAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM55 TTGCGCTCGGCTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
GM56 GTCGTATTAGCCTAGCGCAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM57 TTGCGCTAGGCTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
GM58 GTCGTATTAACCTAGCGCAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM59 TTGCGCTAGGTTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
GM60 GTCGTATTAATCTAGCGCAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM61 TTGCGCTAGATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
GM62 GTCGTATTAATTTAGCGCAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM63 TTGCGCTAAATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
GM64 GTCGTATTAATTTAGCGAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM65 TTTCGCTAAATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
GM66 GTCGTATTAATTTAGAGAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM67 TTTCTCTAAATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
GM68 GTCGTATTAATTTACACAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM69 TTTGTGTAAATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
GM70 GTCGTATTAATTTAAAAAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM71 TTTTTTTAAATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
GM72 GTCGTATTAATTTATATAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM73 TTTATATAAATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
GM74 GTCGTATTAATTTTAATTAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
GM75 TTAATTAAAATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
GM76 ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT
GM77 GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACAGCTAGCTCAG
GM78 AAAGCACATCCGGTGACAGCTGCTAGCATTATACCTAGGACTGAGCTAGCTGTCAATGAG
GM79 GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA
GM80 GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC
GM81 ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA
GM82 GCATTTGTATGTTCACCTATATTTGCTGCTCGTAAAAAAAAAAACTTGACAGCTAGCTCAG
GM83 AAAGCACATCCGGTGACAGCTGCTAGCATTATACCTAGGACTGAGCTAGCTGTCAAGTTT
GM84 GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACAGCTAGCTCAG
GM85 AAAGCACATCCGGTGACAGCTGCTAGCATTATACCTAGGACTGAGCTAGCTGTCAAGTTT
GM86 GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACGGCTAGCTCAG
GM87 AAAGCACATCCGGTGACAGCTGCTAGCACTGTACCTAGGACTGAGCTAGCCGTCAATGAG
GM88 GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACGGCTAGCTCAG
GM89 AAAGCACATCCGGTGACAGCTGCTAGCACTGTACCTAGGACTGAGCTAGCCGTCAAGTTT
GM90 GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACAGCTAGCTCAG
GM91 AAAGCACATCCGGTGACAGCTGCTAGCACAGTACCTAGGACTGAGCTAGCTGTCAATGAG
GM92 GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACAGCTAGCTCAG
GM93 AAAGCACATCCGGTGACAGCTGCTAGCACAGTACCTAGGACTGAGCTAGCTGTCAAGTTT
GM94 GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACAGCTAGCTCAG
GM95 AAAGCACATCCGGTGACAGCTGCTAGCACAATACCTAGGACTGAGCTAGCTGTCAATGAG

```
GM96   GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACAGCTAGCTCAG
GM97   AAAGCACATCCGGTGACAGCTGCTAGCACAATACCTAGGACTGAGCTAGCTGTCAAGTTT
GM98   GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACAGCTAGCTCAG
GM99   AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAATGAG
GM100  GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACAGCTAGCTCAG
GM101  AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAGTTT
GM102  GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACGGCTAGCTCAG
GM103  AAAGCACATCCGGTGACAGCTGCTAGCACTATACCTAGGACTGAGCTAGCCGTCAATGAG
GM104  GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACGGCTAGCTCAG
GM105  AAAGCACATCCGGTGACAGCTGCTAGCACTATACCTAGGACTGAGCTAGCCGTCAAGTTT
GM106  GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACGGCTAGCTCAG
GM107  AAAGCACATCCGGTGACAGCTGCTAGCACAATACCTAGGACTGAGCTAGCCGTCAATGAG
GM108  GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACGGCTAGCTCAG
GM109  AAAGCACATCCGGTGACAGCTGCTAGCACAATACCTAGGACTGAGCTAGCCGTCAAGTTT
GM110  GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCACTGACAGCTAGCTCAG
GM111  AAAGCACATCCGGTGACAGCTGCTAGCATTATACCTAGGACTGAGCTAGCTGTCAGTGAG
GM112  GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACCTGACAGCTAGCTCAG
GM113  AAAGCACATCCGGTGACAGCTGCTAGCATTATACCTAGGACTGAGCTAGCTGTCAGGTTT
GM114  GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACGGCTAGCTCAG
GM115  AAAGCACATCCGGTGACAGCTGCTAGCACTATACCTAGGACTGAGCTAGCCGTAAATGAG
GM116  GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACGGCTAGCTCAG
GM117  AAAGCACATCCGGTGACAGCTGCTAGCACTATACCTAGGACTGAGCTAGCCGTAAAGTTT
GM118  GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACGGCTAGCTCAG
GM119  AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGGCTGAGCTAGCCGTAAATGAG
GM120  GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACGGCTAGCTCAG
GM121  AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGGCTGAGCTAGCCGTAAAGTTT
GM122  GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACGGCTAGCTCAG
GM123  AAAGCACATCCGGTGACAGCTGCTAGCATTGTACCTAGGACTGAGCTAGCCGTAAATGAG
GM124  GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACGGCTAGCTCAG
GM125  AAAGCACATCCGGTGACAGCTGCTAGCATTGTACCTAGGACTGAGCTAGCCGTAAAGTTT
GM126  GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACGGCTAGCTCAG
GM127  AAAGCACATCCGGTGACAGCTGCTAGCATAGTACCTAGGACTGAGCTAGCCGTAAATGAG
GM128  GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACGGCTAGCTCAG
GM129  AAAGCACATCCGGTGACAGCTGCTAGCATAGTACCTAGGACTGAGCTAGCCGTAAAGTTT
GM130  GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACAGCTAGCTCAG
GM131  AAAGCACATCCGGTGACAGCTGCTAGCATAGTCCCTAGGACTGAGCTAGCTGTCAATGAG
GM132  GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACAGCTAGCTCAG
GM133  AAAGCACATCCGGTGACAGCTGCTAGCATAGTCCCTAGGACTGAGCTAGCTGTCAAGTTT
GM134  GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTATAGCTAGCTCAG
GM135  AAAGCACATCCGGTGACAGCTGCTAGCATTGTACCAAGGGCTGAGCTAGCTATAAATGAG
GM136  GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTATAGCTAGCTCAG
GM137  AAAGCACATCCGGTGACAGCTGCTAGCATTGTACCAAGGGCTGAGCTAGCTATAAAGTTT
GM138  GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTATGGCTAGCTCAG
GM139  AAAGCACATCCGGTGACAGCTGCTAGCATTGTACCTAGGACTGAGCTAGCCATAAATGAG
GM140  GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTATGGCTAGCTCAG
GM141  AAAGCACATCCGGTGACAGCTGCTAGCATTGTACCTAGGACTGAGCTAGCCATAAAGTTT
GM142  GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACAGCTAGCTCAG
GM143  AAAGCACATCCGGTGACAGCTGCTAGCACAATCCCTAGGACTGAGCTAGCTGTCAATGAG
GM144  GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACAGCTAGCTCAG
GM145  AAAGCACATCCGGTGACAGCTGCTAGCACAATCCCTAGGACTGAGCTAGCTGTCAAGTTT
```

GM146 GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACAGCTAGCTCAG
GM147 AAAGCACATCCGGTGACAGCTGCTAGCACAGTCCCTAGGACTGAGCTAGCTGTAAATGAG
GM148 GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACAGCTAGCTCAG
GM149 AAAGCACATCCGGTGACAGCTGCTAGCACAGTCCCTAGGACTGAGCTAGCTGTAAAGTTT
GM150 GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCACTGATGGCTAGCTCAG
GM151 AAAGCACATCCGGTGACAGCTGCTAGCATAATCCCTAGGACTGAGCTAGCCATCAGTGAG
GM152 GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACCTGATGGCTAGCTCAG
GM153 AAAGCACATCCGGTGACAGCTGCTAGCATAATCCCTAGGACTGAGCTAGCCATCAGGTTT
GM154 GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCACTGATAGCTAGCTCAG
GM155 AAAGCACATCCGGTGACAGCTGCTAGCATAATCCCTAGGACTGAGCTAGCTATCAGTGAG
GM156 GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACCTGATAGCTAGCTCAG
GM157 AAAGCACATCCGGTGACAGCTGCTAGCATAATCCCTAGGACTGAGCTAGCTATCAGGTTT
GM158 GAGCAGCAAATATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAAT
GM159 TTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACAGCTAGCTCAGTCCTAGGTATT
GM160 GGTGACAGCTAAATTGTTATCCGCTCACAATTGCTAGCATAATACCTAGGACTGAGCTAG
GM161 ATAACAATTTAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAAC
GM162 TTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCCTCACGGACTCA
GM163 TTTAAAAATATACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAA
GM164 AAAAATTTTCCATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAAT
GM165 TTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACAGCTAGCTCAGTCCTAGGTATT
GM166 TTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACATTGTGAGCGGATAACAATATT
GM167 GGTGACAGCTAAATTGTTATCCGCTCACAATTGCTAGCATAATATTGTTATCCGCTCACA
GM168 TTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACATTGTGAGCGGATAACAATATT
GM169 AAAACCGTTGTGATTGGCGC
GM170 CTGACCCACGGTATTACGCC
GM171 TAAGGTAAAAGCGGCGGGAG
GM172 ACTAAAAGACAATTCAGCGGGT
GM173 GGGAAAATCGACGCTGGTCA
GM174 AGGTAAAGCCCTGGTCATGC
GM175 CTCAAGGGCGTGGATCAGG
GM176 GGGACGCTGCCAAAGACC
GM177 TCGCAAGGCTGACGATATCC
GM178 AGGCGATTTTCATGACCGGT
GM179 ATGAAAAAAACCGTTGTGATTGGCGCA
GM180 TCATTGCAGATCCTCAATCATCAGGCTG
GM181 ATGAGCCAACCGCCGCT
GM182 CTAAACGGGACGCTGCCAAAGAC
GM183 ATGGTGAGTGGCAGTAAAGCGG
GM184 TCAGGCGATTTTCATGACCGGTG
GM185 CTAAGCAGAAGGCCCCTGAC
GM186 CTTTCAGTTTGGCGGTCTGG
GM187 CGCCGCTGAGAAAAAGCGAAG
GM188 ACGAAAATTGCTTATCACGCGTTGC
GM189 taaattttgtgtcgcccTAATACGACTCACTATAGGGAG
GM190 cgcacttggaaaccatTTTTTTCGGAACACGAAAATTG
GM191 GGGCGACACAAAATTTATTCTAAATGCATAATAAATAC
GM192 taaattttgtgtcgcccTAATACGACTCACTATAGGGAG
GM193 cgcacttggaaaccatTTTTTTCGGAACTTATTACGC
GM194 ATGGTTTCCAAGTGCGAG
GM195 GCATTTGTATGTTCACCTATGGAAATTTTTTTTTAAAAAAAAACTTTACAGCTAGCTCAG

GM196 GCATTTGTATGTTCACCTATGTGGCAACAATTTCTACAAAACACTTTACAGCTAGCTCAG
GM197 AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAGTGT
GM198 GCATTTGTATGTTCACCTATTTATAGATTTTTTTTAAAAAACTATTTACAGCTAGCTCAG
GM199 AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAATAGT
GM200 GCATTTGTATGTTCACCTATCAGAAAAAAAGATCAAAAAAATACTTTACAGCTAGCTCAG
GM201 AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAGTAT
GM202 GCATTTGTATGTTCACCTATGGAAAATTCTTTTAAAAAAAACACTTTACAGCTAGCTCAG
GM203 GCATTTGTATGTTCACCTATGGAAAATTCTTTCAAAAAAAAAACTTTACAGCTAGCTCAG
GM204 GCATTTGTATGTTCACCTATGGAAAATTTTTTCAAAAAAAACACTTTACAGCTAGCTCAG
GM205 GCATTTGTATGTTCACCTATGGAAAATTTTTTCAACCAAAACACTTTACAGCTAGCTCAG
GM206 GCATTTGTATGTTCACCTATGGAAAATCTTTTCAACCAGCACACTTTACAGCTAGCTCAG
GM207 GCATTTGTATGTTCACCTATGGCACAGCTTTTCAACCAGCACACTTTACAGCTAGCTCAG
GM208 ATAGGTGAACATACAAATGCGGAACTCTAGAAGCGGCCGCGAATTCCAGAAATCATCCTT
GM209 ACAAATTAAGGAAGGTTCACGAGATGGCTTCCTCCGAAGACGTTATCAAAGAGTTCATGC
GM210 CTCTAGAAGCGGCCGCGAA
GM211 ATGGCTTCCTCCGAAGACGTTATCA
GM212 CCTGAAAATCTCGATAACTCAAAAAATACGCCCG
GM213 CGTTAACGGAACCTTCCATACGAACTTTGAA
GM214 AATTTTGTGTCGCCCTTTACACGTACTTAGTC
GM215 CCTCCTCGCACTTGGAAACCATATGTTTTA
GM216 GCGTCCCGTTTAGGGTCTCACCCCAAGGGCG
GM217 TGCCACTCACCATTTTTTTCCTCCTGCATTATACCTAGGACTGAGCTAGCTGTCAAGGGCGACACAAAATT
GM218 CAGGAGGAAAAAAATGGTGAGTGGCAGTAAAGC
GM219 CGGTTTTTTTCATTTTTTTCCTCCTTCAGGCGATTTTCATGACCG
GM220 TGAAGGAGGAAAAAAATGAAAAAAACCGTTGTGATTGGCG
GM221 GGTTGGCTCATTTTTTTCCTCCTTCAGGCTGGCGGTGGCTTTC
GM222 CCTGAAGGAGGAAAAAAATGAGCCAACCGCCGCTG
GM223 CTTGGGGTGAGACCCTAAACGGGACGCTGCCAAAG
GM224 GCGTCCCGTTTAGGGTGGCATGGATGAGCTG
GM225 GCCACTCACCATTTTTTTCCTCCTTCTCCCTATAGTGAGTCGTATTAAGTACGTGTAAAGGGCGACAC
GM226 GAAGGAGGAAAAAAATGGTGAGTGGCAGTAAAGC
GM227 CGGTTTTTTTCATTTTTTTCCTCCTTTGAACCCAAAAGGGCGG
GM228 AAGGAGGAAAAAAATGAAAAAAACCGTTGTGATTGGC
GM229 GGTTGGCTCATTTTTTTCCTCCTTGCAGATCCTCAATCATCAGGC
GM230 GCAAGGAGGAAAAAAATGAGCCAACCGCCGCTG
GM231 CATCCATGCCACCCTAAACGGGACGCTGCCAAAG
GM232 GCCACTCACCATTTTTTTCGGAACTCTCCCTATAGTGAGTCGTATTAAGTACGTGTAAAGGGCGACAC
GM233 GAGTTCCGAAAAAAAATGGTGAGTGGCAGTAAAGC
GM234 CGGTTTTTTTCATTTTTTTCGGAACTTGAACCCAAAAGGGCGG
GM235 AGTTCCGAAAAAAATGAAAAAAACCGTTGTGATTGGC
GM236 GGTTGGCTCATTTTTTTCGGAACTGCAGATCCTCAATCATCAGGC
GM237 GCAGTTCCGAAAAAAATGAGCCAACCGCCGCTG

Designed regulatory units inserted in pJ401

>P1 $INXC_{24}$-UP_{FALSE}-σ70P_{J23119}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATTATACCTAGGACTGAGCTAGCTGTCAATGAG
GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA
GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC
ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

>P2 $INXC_{24}$-$UP_{PROXIMAL}$-σ70P_{J23119}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTAAAAAAAAAACTTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTAAAAAAAAAACTTGACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATTATACCTAGGACTGAGCTAGCTGTCAAGTTT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P3 $INXC_{24}$-UP_{FULL}-σ70P_{J23119}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTTT
TTAAAAAAAAAAACTTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTATTATAGCTAGCCCTAGGACTGAGCTAGCTGTCAAGTTT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P4 $INXC_{24}$-UP_{FALSE}-P_{J23100}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTGACGGCTAGCTCAGTCCTAGGTACAGTGCTAGCAGCTGTCACCGGATGTGCTTTCC

GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACTGTACCTAGGACTGAGCTAGCCGTCAATGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P5 $INXC_{24}$-UP_{FULL}-P_{J23100}-$INXL_{RiboJ}$-$RBS_{80,000}$
ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAACTTGACGGCTAGCTCAGTCCTAGGTACAGTGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACTGTACCTAGGACTGAGCTAGCCGTCAAGTTT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P6 $INXC_{24}$-UP_{FALSE}-P_{J23102}-$INXL_{RiboJ}$-$RBS_{80,000}$
ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTGACAGCTAGCTCAGTCCTAGGTACTGTGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACAGTACCTAGGACTGAGCTAGCTGTCAATGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P7 $INXC_{24}$-UP_{FULL}-P_{J23102}-$INXL_{RiboJ}$-$RBS_{80,000}$
ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAACTTGACAGCTAGCTCAGTCCTAGGTACTGTGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACAGTACCTAGGACTGAGCTAGCTGTCAAGTTT

~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P8 $INXC_{24}$-UP_{FALSE}-P_{J23104}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTGACAGCTAGCTCAGTCCTAGGTATTGTGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACAATACCTAGGACTGAGCTAGCTGTCAATGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P9 $INXC_{24}$-UP_{FULL}-P_{J23104}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAAACTTGACAGCTAGCTCAGTCCTAGGTATTGTGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACAATACCTAGGACTGAGCTAGCTGTCAAGTTT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P10 $INXC_{24}$-UP_{FALSE}-P_{J23101}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAATGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P11 $INXC_{24}$-UP_{FULL}-P_{J23101}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAAACTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAGTTT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P12 $INXC_{24}$-UP_{FALSE}-P_{J23111}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACTATACCTAGGACTGAGCTAGCCGTCAATGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P13 $INXC_{24}$-UP_{FULL}-P_{J23111}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAAACTTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACTATACCTAGGACTGAGCTAGCCGTCAAGTTT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P14 $INXC_{24}$-UP_{FALSE}-P_{J23118}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTGACGGCTAGCTCAGTCCTAGGTATTGTGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~

GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACAATACCTAGGACTGAGCTAGCCGTCAATGAG
~~GCTGTCACCGCATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P15 $INXC_{24}$-UP_{FULL}-P_{J23118}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTA**T**GGAAAATTT**T**T
TTAAAAAA**A**AAACTTGAC**G**GCTAGCTCAGTCCTAGGTAT**TG**TGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACAATACCTAGGACTGAGCTAGCCGTCAAGTTT
~~GCTGTCACCGCATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P16 $INXC_{24}$-UP_{FALSE}-P_{J23108}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTA**T**ATTTGCTGC**T**C
GTGTTTAC**T**CTCA**C**TGACAGCTAGCTCAGTCCTAGGTATAATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCACTGACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATTATACCTAGGACTGAGCTAGCTGTCAGTGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P17 $INXC_{24}$-UP_{FULL}-P_{J23108}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTA**T**GGAAAATTT**T**T
TTAAAAAA**A**AAAC**C**TGACAGCTAGCTCAGTCCTAGGTATAATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACCTGACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATTATACCTAGGACTGAGCTAGCTGTCAGGTTT
~~GCTGTCACCGCATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P18 $INXC_{24}$-UP_{FALSE}-P_{J23106}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTTACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACTATACCTAGGACTGAGCTAGCCGTAAATGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P19 $INXC_{24}$-UP_{FULL}-P_{J23106}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAAACTTTACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACTATACCTAGGACTGAGCTAGCCGTAAAGTTT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P20 $INXC_{24}$-UP_{FALSE}-P_{J23107}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTTACGGCTAGCTCAGCCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGGCTGAGCTAGCCGTAAATGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P21 $INXC_{24}$-UP_{FULL}-P_{J23107}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAAACTTTACGGCTAGCTCAGCCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGGCTGAGCTAGCCGTAAAGTTT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P22 INXC$_{24}$-UP$_{FALSE}$-P$_{J23110}$-INXL$_{RiboJ}$-RBS$_{80,000}$
ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTTACGGCTAGCTCAGTCCTAGGTACAATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATTGTACCTAGGACTGAGCTAGCCGTAAATGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P23 INXC$_{24}$-UP$_{FULL}$-P$_{J23110}$-INXL$_{RiboJ}$-RBS$_{80,000}$
ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAAACTTTACGGCTAGCTCAGTCCTAGGTACAATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATTGTACCTAGGACTGAGCTAGCCGTAAAGTTT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P24 INXC$_{24}$-UP$_{FALSE}$-P$_{J23105}$-INXL$_{RiboJ}$-RBS$_{80,000}$
ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTTACGGCTAGCTCAGTCCTAGGTACTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAGTACCTAGGACTGAGCTAGCCGTAAATGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~

ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

>P25 INXC$_{24}$-UP$_{FULL}$-P$_{J23105}$-INXL$_{RiboJ}$-RBS$_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAACTTTACGGCTAGCTCAGTCCTAGGTACTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAGTACCTAGGACTGAGCTAGCCGTAAAGTTT
GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA
GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC
ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

>P26 INXC$_{24}$-UP$_{FALSE}$-P$_{J23116}$-INXL$_{RiboJ}$-RBS$_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTGACAGCTAGCTCAGTCCTAGGGACTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAGTCCCTAGGACTGAGCTAGCTGTCAATGAG
GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA
GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC
ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

>P27 INXC$_{24}$-UP$_{FULL}$-P$_{J23116}$-INXL$_{RiboJ}$-RBS$_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAACTTGACAGCTAGCTCAGTCCTAGGGACTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAGTCCCTAGGACTGAGCTAGCTGTCAAGTTT
GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA
GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC
ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

>P28 INXC$_{24}$-UP$_{FALSE}$-P$_{J23115}$-INXL$_{RiboJ}$-RBS$_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTTATAGCTAGCTCAGCCCTTGGTACAATGCTAGCAGCTGTCACCGGATGTGCTTTCC

GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTATAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATTGTACCAAGGGCTGAGCTAGCTATAAATGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P29 INXC$_{24}$-UP$_{FULL}$-P$_{J23115}$-INXL$_{RiboJ}$-RBS$_{80,000}$
ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAACTTTATAGCTAGCTCAGCCCTTGGTACAATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTATAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATTGTACCAAGGGCTGAGCTAGCTATAAAGTTT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P30 INXC$_{24}$-UP$_{FALSE}$-P$_{J23114}$-INXL$_{RiboJ}$-RBS$_{80,000}$
ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTTATGGCTAGCTCAGTCCTAGGTACAATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTATGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATTGTACCTAGGACTGAGCTAGCCATAAATGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P31 INXC$_{24}$-UP$_{FULL}$-P$_{J23114}$-INXL$_{RiboJ}$-RBS$_{80,000}$
ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAACTTTATGGCTAGCTCAGTCCTAGGTACAATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTATGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATTGTACCTAGGACTGAGCTAGCCATAAAGTTT

~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P32 $INXC_{24}$-UP_{FALSE}-P_{J23117}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTGACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACAATCCCTAGGACTGAGCTAGCTGTCAATGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P33 $INXC_{24}$-UP_{FULL}-P_{J23117}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAAACTTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTGACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACAATCCCTAGGACTGAGCTAGCTGTCAAGTTT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P34 $INXC_{24}$-UP_{FALSE}-P_{J23109}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCATTTACAGCTAGCTCAGTCCTAGGGACTGTGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACAGTCCCTAGGACTGAGCTAGCTGTAAATGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P35 $INXC_{24}$-UP_{FULL}-P_{J23109}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAAACTTTACAGCTAGCTCAGTCCTAGGGACTGTGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCACAGTCCCTAGGACTGAGCTAGCTGTAAAGTTT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P36 $INXC_{24}$-UP_{FALSE}-P_{J23113}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCACTGATGGCTAGCTCAGTCCTAGGGATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCACTGATGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAATCCCTAGGACTGAGCTAGCCATCAGTGAG
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P37 $INXC_{24}$-UP_{FULL}-P_{J23113}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TTAAAAAAAAAAACCTGATGGCTAGCTCAGTCCTAGGGATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACCTGATGGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAATCCCTAGGACTGAGCTAGCCATCAGGTTT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P38 $INXC_{24}$-UP_{FALSE}-P_{J23112}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTC
GTGTTTACTCTCACTGATAGCTAGCTCAGTCCTAGGGATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~

GCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCTCACTGATAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAATCCCTAGGACTGAGCTAGCTATCAGTGAG
~~GCTGTCACCGCATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P39 INXC$_{24}$-UP$_{FULL}$-P$_{J23112}$-INXL$_{RiboJ}$-RBS$_{80,000}$
ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTA**T**GGAAAATTT**T**T
TTAAAAAA**A**AAACCTGA**T**AGCTAGCTCAGTCCTAGG**G**AT**T**ATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTTAAAAAAAAAACCTGATAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAATCCCTAGGACTGAGCTAGCTATCAGGTTT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

Lac-inducible sig70 promoters

>P40 INXC$_{24}$-UP$_{FALSE}$-P$_{J23101}$-LacO-INXL$_{RiboJ}$-RBS$_{80,000}$
ATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTA**T**ATTTGCTGC**T**CGTGTTTAC**T**CT
CATT**T**ACAGCTAGCTCAGTCCTAGGTAT**T**ATGCTAGCAATTGTGAGCGGATAACAATTTAGCTGTCACCGG
ATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATAC
AAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAA

GAGCAGCAAATATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAAT
TTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACAGCTAGCTCAGTCCTAGGTATT
GGTGACAGCTAAATTGTTATCCGCTCACAATTGCTAGCATAATACCTAGGACTGAGCTAG
ATAACAATTTAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAAC
TTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCCTCACGGACTCA
TTTAAAAATATACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAA

>P41 INXC$_{24}$-UP$_{FULL}$-P$_{J23101}$-LacO-INXL$_{RiboJ}$-RBS$_{80,000}$
ATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTA**T**GGAAAATTT**T**TTTAAAAAA**A**AA
ACTT**T**ACAGCTAGCTCAGTCCTAGGTAT**T**ATGCTAGCAATTGTGAGCGGATAACAATTTAGCTGTCACCGG
ATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATAC
AAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAA

AAAAATTTTCCATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAAT
TTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACAGCTAGCTCAGTCCTAGGTATT
~~GGTGACAGCTAAATTGTTATCCGCTCACAATTGCTAGCATAATACCTAGGACTGAGCTAG~~
~~ATAACAATTTAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAAC~~

TTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCCTCACGGACTCA
TTTAAAAATATACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAA

>P42 INXC$_{24}$-INTC-UP$_{FALSE}$-P$_{J23101\,(LacO)}$-LacO-INXL$_{RiboJ}$-RBS$_{80,000}$

ATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATATTTGCTGCTCGTGTTTACTCT
CATTTACATTGTGAGCGGATAACAATATTATGCTAGCAATTGTGAGCGGATAACAATTTAGCTGTCACCGG
ATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATAC
AAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAA

GAGCAGCAAATATAGGTGAACATACAAATGCGGAAGGCGACACAAAATTTATTCTAAAT
TTCACCTATATTTGCTGCTCGTGTTTACTCTCATTTACATTGTGAGCGGATAACAATATT
GGTGACAGCTAAATTGTTATCCGCTCACAATTGCTAGCATAATATTGTTATCCGCTCACA
ATAACAATTTAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAAC
TTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCCTCACGGACTCA
TTTAAAAATATACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAA

>P43 INXC$_{24}$-INTC-UP$_{FULL}$-P$_{J23101\,(LacO)}$-LacO-INXL$_{RiboJ}$-RBS$_{80,000}$

ATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTTTTTAAAAAAAAA
ACTTTACATTGTGAGCGGATAACAATATTATGCTAGCAATTGTGAGCGGATAACAATTTAGCTGTCACCGG
ATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATAC
AAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAA

AAAAATTTTCCATAGGTGAACATACAAATGCGGAAGGCGACACAAAATTTATTCTAAAT
TTCACCTATGGAAAATTTTTTTAAAAAAAAAACTTTACATTGTGAGCGGATAACAATATT
GGTGACAGCTAAATTGTTATCCGCTCACAATTGCTAGCATAATATTGTTATCCGCTCACA
ATAACAATTTAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAAC
TTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCCTCACGGACTCA
TTTAAAAATATACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAA

>P44 INXC$_{24}$-UP$_{GM2}$-P$_{J23101}$-INXL$_{RiboJ}$-RBS$_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAATTTTTT
TTTAAAAAAAAACTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

ATAGGTGAACATACAAATGCGGAAGGCGACACAAAATTTATTCTAAATGCATAATAAAT
GCATTTGTATGTTCACCTATGGAAATTTTTTTTTAAAAAAAAACTTTACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAGTTT
GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA
GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC
ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

>P45 INXC$_{24}$-UP$_{recA}$-P$_{J23101}$-INXL$_{RiboJ}$-RBS$_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGTGGCAACAAT
TTCTACAAAACACTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGTGGCAACAATTTCTACAAAACACTTTACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAGTGT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P46 INXC$_{24}$-UP$_{rrnB-P1}$-P$_{J23101}$-INXL$_{RiboJ}$-RBS$_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTTATAGATTTT
TTTTAAAAAACTATTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATTTATAGATTTTTTTTAAAAAACTATTTACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAATAGT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P47 INXC$_{24}$-UP$_{rrnD-P1}$-P$_{J23101}$-INXL$_{RiboJ}$-RBS$_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATCAGAAAAAAAG
ATCAAAAAAATACTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATCAGAAAAAAAGATCAAAAAAATACTTTACAGCTAGCTCAG
AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAGTAT
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P48 INXC$_{24}$-UP$_{GM3}$-P$_{J23101}$-INXL$_{RiboJ}$-RBS$_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTCTT
TTAAAAAAAACACTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~

GCATTTGTATGTTCACCTATGGAAAATTCTTTTAAAAAAAACACTTTACAGCTAGCTCAG
~~AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAGTGT~~
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P49 $INXC_{24}$-UP_{GM4}-P_{J23101}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTCTT
TCAAAAAAAAAACTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTCTTTCAAAAAAAAAACTTTACAGCTAGCTCAG
~~AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAGTTT~~
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P50 $INXC_{24}$-UP_{GM5}-P_{J23101}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TCAAAAAAAACACTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTCAAAAAAAACACTTTACAGCTAGCTCAG
~~AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAGTGT~~
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P51 $INXC_{24}$-UP_{GM6}-P_{J23101}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATTTTT
TCAACCAAAACACTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATTTTTTCAACCAAAACACTTTACAGCTAGCTCAG
~~AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAGTGT~~
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P52 $INXC_{24}$-UP_{GM7}-P_{J23101}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGAAAATCTTT
TCAACCAGCACACTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGCGCACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGAAAATCTTTTCAACCAGCACACTTTACAGCTAGCTCAG
~~AAACCACATCCCGTGACACCTGCTAGCATAATACCTAGCACTGACCTAGCTGTAAAGTGT~~
~~GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>P53 $INXC_{24}$-UP_{GM8}-P_{J23101}-$INXL_{RiboJ}$-$RBS_{80,000}$

ATTTATTATGCATTTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGCACAGCTTT
TCAACCAGCACACTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAAATATACAAATTAAGGAA
GGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA

~~ATAGGTGAACATACAAATGCGGAAGGCGCACACAAAATTTATTCTAAATGCATAATAAAT~~
GCATTTGTATGTTCACCTATGGCACAGCTTTTCAACCAGCACACTTTACAGCTAGCTCAG
~~AAAGCACATCCGGTGACAGCTGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAGTGT~~
~~GCTGTCACCGCATGTGCTTTCCGGTCTGATGAGTCCGTCAGGACGAAACAGCCTCTACAA~~
~~GTGAACCTTCCTTAATTTGTATATTTTTAAACAAAATTATTTGTAGAGGCTGTTTCGTCC~~
~~ACAAATTAAGGAAGGTTCACGAGATGGTTTCCAAGTGCGAGGAGGATAACATGGCTATCA~~

>T7P1 T7P01-RiboJ-$RBS_{75,594}$

TTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTAATACGACTCACTATAGGGAGAA
GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTT
TAAAGGAGGAAAAAAATGGTTTCCAAGTGCGAGGAGGA

GTGAGTCGTATTAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAA
CACCTATTAATACGACTCACTATAGGGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGA
CAAAATTATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCC
AGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGAGGAGGA

>T7P01.NNN T7P01.XX

TTAGAATAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTAATACGACTCACTATWGGRNNNA
GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTT
TAAAGGAGGAAAAAAATGGTTTCCAAGTGCGAGGAGGA

~~GTGAGTCGTATTAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTATTCTAA~~
CACCTATTAATACGACTCACTATWGGRNNNAGCTGTCACCGGATGTGCTTTCCGGTCTGA
~~CAAAATTATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCC~~

~~AGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGAGGA~~

>T7P2 T7P$_{PhiOL}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTTGTCTTTATTAATACAACTCACTATA**A**GG
AGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTTGTATTAATAAAGACAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTTGTCTTTATTAATACAACTCACTATAAGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC
CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

>T7P03 T7P$_{phi1.1A}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATAACGCCAAATCAATACGACTCACTATA**G**AG
GGACAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTGATTTGGCGTTATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TAACGCCAAATCAATACGACTCACTATAGAGGGACAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P04 T7P$_{phi1.1B}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTTCTTCCGGTTAATACGACTCACTATA**G**GA
GGACCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAACCGGAAGAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTTCTTCCGGTTAATACGACTCACTATAGGAGGACCAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P05 T7P$_{phi1.3}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGACTGGAAGTAATACGACTCAGTATA**G**GG
ACAATAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTACTTCCAGTCCATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TGGACTGGAAGTAATACGACTCAGTATAGGGACAATAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~

~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P06 T7P$_{phi1.5}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATAGTTAACTGGTAATACGACTCACTAAA**G**GA
GGTACAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTACCAGTTAACTATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TAGTTAACTGGTAATACGACTCACTAAAGGAGGTACAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P07 T7P$_{phi1.6}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTGGTCACGCTTAATACGACTCACTAAA**G**GA
GACACAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAAGCGTGACCAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTGGTCACGCTTAATACGACTCACTAAAGGAGACACAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P08 T7P$_{phi2.5}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATAGCACCGAAGTAATACGACTCACTATT**A**GG
GAAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTACTTCGGTGCTATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TAGCACCGAAGTAATACGACTCACTATTAGGGAAGAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P09 T7P$_{phi3.8}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATCGTGGATAATTAATTGAACTCACTAAA**G**GG
AGACCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTTCAATTAATTATCCACGATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TCGTGGATAATTAATTGAACTCACTAAAGGGAGACCAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~

~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P10 T7P$_{phi4c}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATCCGACTGAGACAATCCGACTCACTAAA**G**AG
AGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGGATTGTCTCAGTCGGATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TCCGACTGAGACAATCCGACTCACTAAAGAGAGAGAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P11 T7P$_{phi4.3}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATAGTCCCATTCTAATACGACTCACTAAA**G**GA
GACACAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAGAATGGGACTATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TAGTCCCATTCTAATACGACTCACTAAAGGAGACACAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P12 T7P$_{phi4.7}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTTCATGAATACTATTCGACTCACTATA**G**GA
GATATAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGAATAGTATTCATGAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTTCATGAATACTATTCGACTCACTATAGGAGATATAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P13 T7P$_{phi6.5}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGTCCCTAAATTAATACGACTCACTATA**G**GG
AGATAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAATTTAGGGACATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TGTCCCTAAATTAATACGACTCACTATAGGGAGATAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~

~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P14 T7P$_{phi9}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGCCGGGAATTTAATACGACTCACTATA**G**GG
AGACCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAAATTCCCGGCATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TGCCGGGAATTTAATACGACTCACTATAGGGAGACCAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P15 T7P$_{phi10}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATACTTCGAAATTAATACGACTCACTATA**G**GG
AGACCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAATTTCGAAGTATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TACTTCGAAATTAATACGACTCACTATAGGGAGACCAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P16 T7P$_{phi13}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGGCTCGAAATTAATACGACTCACTATA**G**GG
AGAACAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAATTTCGAGCCATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TGGCTCGAAATTAATACGACTCACTATAGGGAGAACAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P17 T7P$_{phi17}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATGCGTAGGAAATAATACGACTCACTATA**G**GG
AGAGGAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTATTTCCTACGCATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TGCGTAGGAAATAATACGACTCACTATAGGGAGAGGAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~

~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P18 INXC$_{24}$-T7P$_{phiOR}$INXL$_{RiboJ}$-RBS$_{10,915}$

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATCACGATAAATTAATACGACTCACTATA**G**GG
AGAGGAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAATTTATCGTGATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TCACGATAAATTAATACGACTCACTATAGGGAGAGGAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P19 INXC$_{24}$-T7UP$_{GM1}$-T7P$_{1}$-INXL$_{RiboJ}$-RBS$_{10,915}$
TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTGCGCTCGGCTAATACGACTCACTATA**G**GG
AGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAGCCGAGCGCAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTGCGCTCGGCTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P20 INXC$_{24}$-T7UP$_{GM1}$-T7P$_{1}$-INXL$_{RiboJ}$-RBS$_{10,915}$
TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTGCGCTAGGCTAATACGACTCACTATA**G**GG
AGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAGCCTAGCGCAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTGCGCTAGGCTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P21
TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTGCGCTAGGTTAATACGACTCACTATA**G**GG
AGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAACCTAGCGCAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTGCGCTAGGTTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P22

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTGCGCTAGATTAATACGACTCACTATA**G**GG
AGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAATCTAGCGCAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTGCGCTAGATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P23

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTGCGCTAAATTAATACGACTCACTATA**G**GG
AGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAATTTAGCGCAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTGCGCTAAATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P24

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTTCGCTAAATTAATACGACTCACTATA**G**GG
AGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAATTTAGCGAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTTCGCTAAATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P25

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTTCTCTAAATTAATACGACTCACTATA**G**GG
AGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAATTTAGAGAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTTCTCTAAATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P26

TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTTGTGTAAATTAATACGACTCACTATA**G**GG
AGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAATTTACACAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTTGTGTAAATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P27
TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTTTTTTAAATTAATACGACTCACTATA**G**GG
AGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAATTTAAAAAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTTTTTTAAATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P28
TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTTATATAAATTAATACGACTCACTATA**G**GG
AGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAATTTATATAAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTTATATAAATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

>T7P29
TAAATTTTGTGTCGCCCTTCCGCATTTGTATGTTCACCTATTAATTAAAATTAATACGACTCACTATA**G**GG
AGAGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAAT
TTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA

GTCGTATTAATTTTAATTAATAGGTGAACATACAAATGCGGAAGGGCGACACAAAATTTA
TTAATTAAAATTAATACGACTCACTATAGGGAGAGAAGCTGTCACCGGATGTGCTTTCCG
~~TATTTGTAGAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGAAAGCACATCCGGTGAC~~
~~CGAAACAGCCTCTACAAATAATTTTGTTTAAAGGAGGAAAAAAATGGTTTCCAAGTGCGA~~

Scheme for isothermal DNA assembly of inserts from 60mer oligos

Promoter	OLIGO 1	OLIGO 2	OLIGO 3	OLIGO 4
T7P1	GM13	GM14	GM15	GM16
T7P1.NNN	GM13	GM17	GM15	GM16
T7P2	GM18	GM19	GM20	GM21
T7P3	GM22	GM23	GM20	GM21
T7P4	GM24	GM25	GM20	GM21
T7P5	GM26	GM27	GM20	GM21
T7P6	GM28	GM29	GM20	GM21
T7P7	GM30	GM31	GM20	GM21
T7P8	GM32	GM33	GM20	GM21
T7P9	GM34	GM35	GM20	GM21
T7P10	GM36	GM37	GM20	GM21
T7P11	GM38	GM39	GM20	GM21
T7P12	GM40	GM41	GM20	GM21
T7P13	GM42	GM43	GM20	GM21
T7P14	GM44	GM45	GM20	GM21
T7P15	GM46	GM47	GM20	GM21
T7P16	GM48	GM49	GM20	GM21
T7P17	GM50	GM51	GM20	GM21
T7P18	GM52	GM53	GM20	GM21
T7P19	GM54	GM55	GM20	GM21
T7P20	GM56	GM57	GM20	GM21
T7P21	GM58	GM59	GM20	GM21
T7P22	GM60	GM61	GM20	GM21
T7P23	GM62	GM63	GM20	GM21
T7P24	GM64	GM65	GM20	GM21
T7P25	GM66	GM67	GM20	GM21
T7P26	GM68	GM69	GM20	GM21
T7P27	GM70	GM71	GM20	GM21
T7P28	GM72	GM73	GM20	GM21
T7P29	GM74	GM75	GM20	GM21

Promoter	OLIGO 1	OLIGO 2	OLIGO 3	OLIGO 4	OLIGO 5	OLIGO 6
P1	GM76	GM77	GM78	GM79	GM80	GM81
P2	GM76	GM82	GM83	GM79	GM80	GM81
P3	GM76	GM84	GM85	GM79	GM80	GM81
P4	GM76	GM86	GM87	GM79	GM80	GM81
P5	GM76	GM88	GM89	GM79	GM80	GM81
P6	GM76	GM90	GM91	GM79	GM80	GM81
P7	GM76	GM92	GM93	GM79	GM80	GM81
P8	GM76	GM94	GM95	GM79	GM80	GM81
P9	GM76	GM96	GM97	GM79	GM80	GM81
P10	GM76	GM98	GM99	GM79	GM80	GM81
P11	GM76	GM100	GM101	GM79	GM80	GM81
P12	GM76	GM102	GM103	GM79	GM80	GM81
P13	GM76	GM104	GM105	GM79	GM80	GM81
P14	GM76	GM106	GM107	GM79	GM80	GM81
P15	GM76	GM108	GM109	GM79	GM80	GM81
P16	GM76	GM110	GM111	GM79	GM80	GM81
P17	GM76	GM112	GM113	GM79	GM80	GM81
P18	GM76	GM114	GM115	GM79	GM80	GM81
P19	GM76	GM116	GM117	GM79	GM80	GM81
P20	GM76	GM118	GM119	GM79	GM80	GM81
P21	GM76	GM120	GM121	GM79	GM80	GM81
P22	GM76	GM122	GM123	GM79	GM80	GM81
P23	GM76	GM124	GM125	GM79	GM80	GM81
P24	GM76	GM126	GM127	GM79	GM80	GM81
P25	GM76	GM128	GM129	GM79	GM80	GM81
P26	GM76	GM130	GM131	GM79	GM80	GM81
P27	GM76	GM132	GM133	GM79	GM80	GM81
P28	GM76	GM134	GM135	GM79	GM80	GM81
P29	GM76	GM136	GM137	GM79	GM80	GM81
P30	GM76	GM138	GM139	GM79	GM80	GM81
P31	GM76	GM140	GM141	GM79	GM80	GM81
P32	GM76	GM142	GM143	GM79	GM80	GM81
P33	GM76	GM144	GM145	GM79	GM80	GM81
P34	GM76	GM146	GM147	GM79	GM80	GM81
P35	GM76	GM148	GM149	GM79	GM80	GM81
P36	GM76	GM150	GM151	GM79	GM80	GM81
P37	GM76	GM152	GM153	GM79	GM80	GM81
P38	GM76	GM154	GM155	GM79	GM80	GM81
P39	GM76	GM156	GM157	GM79	GM80	GM81

P40	GM158	GM159	GM160	GM161	GM162	GM163
P41	GM164	GM165	GM160	GM161	GM162	GM163
P42	GM158	GM166	GM167	GM161	GM162	GM163
P43	GM158	GM168	GM160	GM161	GM162	GM163
P44	GM76	GM195	GM101	GM79	GM80	GM81
P45	GM76	GM196	GM197	GM79	GM80	GM81
P46	GM76	GM198	GM199	GM79	GM80	GM81
P47	GM76	GM200	GM201	GM79	GM80	GM81
P48	GM76	GM202	GM197	GM79	GM80	GM81
P49	GM76	GM203	GM101	GM79	GM80	GM81
P50	GM76	GM204	GM197	GM79	GM80	GM81
P51	GM76	GM205	GM197	GM79	GM80	GM81
P52	GM76	GM206	GM197	GM79	GM80	GM81
P53	GM76	GM207	GM197	GM79	GM80	GM81

Genbank files for base plasmids

```
LOCUS       pSB1C3                  3139 bp ds-DNA     circular     11-APR-2013
ACCESSION   pSB1C3
VERSION     pSB1C3
KEYWORDS    .
FEATURES             Location/Qualifiers
     misc_binding    1..21
                     /label="BioBrick suffix"
                     /ApEinfo_fwdcolor=cyan
                     /ApEinfo_revcolor=green
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     promoter        2071..2270
                     /label="R0010"
                     /ApEinfo_fwdcolor=#000eff
                     /ApEinfo_revcolor=#000eff
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     terminator      22..93
                     /label="Eco His Term."
                     /ApEinfo_fwdcolor=#00d5ff
                     /ApEinfo_revcolor=#00d5ff
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     CDS             2297..3002
                     /label="E1010"
                     /ApEinfo_fwdcolor=#008226
                     /ApEinfo_revcolor=#008226
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     primer_bind     156..176
                     /label="VR primer site"
                     /ApEinfo_fwdcolor=#8d8b8c
                     /ApEinfo_revcolor=#8d8b8c
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     rep_origin      261..875
                     /label="pMB1 rep origin"
                     /ApEinfo_fwdcolor=cyan
                     /ApEinfo_revcolor=green
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     terminator      1042..1147
                     /label="T0 terminator"
                     /ApEinfo_fwdcolor=#00d5ff
                     /ApEinfo_revcolor=#00d5ff
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     RBS             2279..2290
                     /label="B0034"
                     /ApEinfo_fwdcolor=#ff1e00
                     /ApEinfo_revcolor=#ff1e00
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     CDS             1160..1829
                     /label="CmR"
                     /ApEinfo_fwdcolor=#008226
                     /ApEinfo_revcolor=#008226
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     terminator      3011..3090
                     /label="B1010"
                     /ApEinfo_fwdcolor=#00d5ff
                     /ApEinfo_revcolor=#00d5ff
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     primer_bind     1943..1962
```

```
                     /label="VF2 primer site"
                     /ApEinfo_fwdcolor=#8d8b8c
                     /ApEinfo_revcolor=#8d8b8c
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     terminator      3099..3139
                     /label="B0012"
                     /ApEinfo_fwdcolor=#00d5ff
                     /ApEinfo_revcolor=#00d5ff
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     misc_binding    2049..2070
                     /label="BioBrick prefix"
                     /ApEinfo_fwdcolor=cyan
                     /ApEinfo_revcolor=green
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     primer_bind     12..33
                     /label="SB-prep-3P-1 primer site"
                     /ApEinfo_fwdcolor=#8d8b8c
                     /ApEinfo_revcolor=#8d8b8c
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     primer_bind     2031..2054
                     /label="SB-prep_2Ea primer site"
                     /ApEinfo_fwdcolor=#8d8b8c
                     /ApEinfo_revcolor=#8d8b8c
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
ORIGIN
        1 tactagtagc ggccgctgca gtccggcaaa aaagggcaag gtgtcaccac cctgcccttt
       61 ttctttaaaa ccgaaaagat tacttcgcgt tatgcaggct tcctcgctca ctgactcgct
      121 gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt
      181 atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc
      241 caggaaccgt aaaaaggccg cgttgctggc gtttttccac aggctccgcc cccctgacga
      301 gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata
      361 ccaggcgttt cccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac
      421 cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg
      481 taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc
      541 cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag
      601 acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt
      661 aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaagaacagt
      721 atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg
      781 atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac
      841 gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca
      901 gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac
      961 ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac
     1021 ttggtctgac agctcgaggc ttggattctc accaataaaa aacgcccggc ggcaaccgag
     1081 cgttctgaac aaatccagat ggagttctga ggtcattact ggatctatca acaggagtcc
     1141 aagcgagctc gatatcaaat tacgccccgc cctgccactc atcgcagtac tgttgtaatt
     1201 cattaagcat tctgccgaca tggaagccat cacaaacggc atgatgaacc tgaatcgcca
     1261 gcggcatcag caccttgtcg ccttgcgtat aatatttgcc catggtgaaa acgggggcga
     1321 agaagttgtc catattggcc acgtttaaat caaaactggt gaaactcacc cagggattgg
     1381 ctgagacgaa aaacatattc tcaataaacc ctttagggaa ataggccagg ttttcaccgt
     1441 aacacgccac atcttgcgaa tatatgtgta gaaactgccg gaaatcgtcg tggtattcac
     1501 tccagagcga tgaaaacgtt tcagtttgct catggaaaac ggtgtaacaa gggtgaacac
     1561 tatcccatat caccagctca ccgtctttca ttgccatacg aaattccgga tgagcattca
     1621 tcaggcgggc aagaatgtga ataaaggccg gataaaactt gtgcttattt ttctttacgg
     1681 tctttaaaaa ggccgtaata tccagctgaa cggtctggtt ataggtacat tgagcaactg
     1741 actgaaatgc ctcaaaatgt tctttacgat gccattggga tatatcaacg gtggtatatc
     1801 cagtgatttt tttctccatt ttagcttcct tagctcctga aaatctcgat aactcaaaaa
     1861 atacgcccgg tagtgatctt atttcattat ggtgaaagtt ggaacctctt acgtgcccga
     1921 tcaactcgag tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa
     1981 aataggcgta tcacgaggca gaatttcaga taaaaaaaat ccttagcttt cgctaaggat
     2041 gatttctgga attcgcggcc gcttctagag caatacgcaa accgcctctc cccgcgcgtt
     2101 ggccgattca ttaatgcagc tggcacgaca ggtttcccga ctggaaagcg ggcagtgagc
     2161 gcaacgcaat taatgtgagt tagctcactc attaggcacc ccaggcttta cactttatgc
     2221 ttccggctcg tatgttgtgt ggaattgtga gcggataaca atttcacaca tactagagaa
     2281 agaggagaaa tactagatgg cttcctccga agacgttatc aaagagttca tgcgtttcaa
     2341 agttcgtatg gaaggttccg ttaacggtca cgagttcgaa atcgaaggtg aaggtgaagg
     2401 tcgtccgtac gaaggtaccc agaccgctaa actgaaagtt accaaaggtg gtccgctgcc
```

```
     2461 gttcgcttgg gacatcctgt ccccgcagtt ccagtacggt tccaaagctt acgttaaaca
     2521 cccggctgac atcccggact acctgaaact gtccttcccg gaaggtttca aatgggaacg
     2581 tgttatgaac ttcgaagacg gtggtgttgt taccgttacc caggactcct ccctgcaaga
     2641 cggtgagttc atctacaaag ttaaactgcg tggtaccaac ttcccgtccg acggtccggt
     2701 tatgcagaaa aaaaccatgg gttgggaagc ttccaccgaa cgtatgtacc cggaagacgg
     2761 tgctctgaaa ggtgaaatca aaatgcgtct gaaactgaaa gacggtggtc actacgacgc
     2821 tgaagttaaa accacctaca tggctaaaaa accggttcag ctgccgggtg cttacaaaac
     2881 cgacatcaaa ctggacatca cctcccacaa cgaagactac accatcgttg aacagtacga
     2941 acgtgctgaa ggtcgtcact ccaccggtgc ttaataacgc tgatagtgct agtgtagatc
     3001 gctactagag ccaggcatca aataaaacga aaggctcagt cgaaagactg ggcctttcgt
     3061 tttatctgtt gtttgtcggt gaacgctctc tactagagtc acactggctc accttcgggt
     3121 gggcctttct gcgtttata
//
```

```
LOCUS       pJ401_sequenced         4678 bp ds-DNA     linear       11-APR-2013
ACCESSION   pJ401_sequenced
VERSION     pJ401_sequenced
KEYWORDS    .
FEATURES             Location/Qualifiers
     source          join(1..1414,1436..1474,1476..1482,1484..4678)
                     /label="source:synthetic"
                     /ApEinfo_fwdcolor=pink
                     /ApEinfo_revcolor=pink
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     misc_feature    2..805
                     /label="pUC_ori"
                     /product=pUC_ori
                     /ApEinfo_fwdcolor=#ff8400
                     /ApEinfo_revcolor=#ff8400
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     misc_feature    complement(977..1090)
                     /label="rpn txn terminator"
                     /product=rpn txn terminator
                     /ApEinfo_fwdcolor=#ff8400
                     /ApEinfo_revcolor=#ff8400
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     misc_feature    complement(1097..1397)
                     /label="bla txn terminator"
                     /product=bla txn terminator
                     /ApEinfo_fwdcolor=#ff8400
                     /ApEinfo_revcolor=#ff8400
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     misc_feature    2351..2525
                     /label="rrnB1 B2 T1 txn terminator"
                     /product=rrnB1 B2 T1 txn terminator
                     /ApEinfo_fwdcolor=#ff8400
                     /ApEinfo_revcolor=#ff8400
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     misc_feature    1281..1306
                     /label="pTF3"
                     /product=pTF3
                     /ApEinfo_fwdcolor=#ff8400
                     /ApEinfo_revcolor=#ff8400
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     misc_feature    complement(2434..2450)
                     /label="pTR"
                     /product=pTR
                     /ApEinfo_fwdcolor=#ff8400
                     /ApEinfo_revcolor=#ff8400
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     misc_feature    join(1476..1482,1484..1523)
                     /label="T5 promoter"
                     /product=T5 promoter
                     /ApEinfo_fwdcolor=#ff8400
                     /ApEinfo_revcolor=#ff8400
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     misc_feature    2710..3504
                     /label="Kanamycin"
                     /product=Kanamycin
                     /ApEinfo_fwdcolor=#ff8400
                     /ApEinfo_revcolor=#ff8400
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     misc_feature    3599..4669
                     /label="lacI"
                     /product=lacI
                     /ApEinfo_fwdcolor=#ff8400
```

```
                     /ApEinfo_revcolor=#ff8400
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     protein_bind    1436..1456
                     /label="lacO"
                     /ApEinfo_fwdcolor=cyan
                     /ApEinfo_revcolor=green
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     promoter        join(1476..1482,1484..1523)
                     /label="T5 promoter(1)"
                     /ApEinfo_label=T5 promoter
                     /ApEinfo_fwdcolor=#000eff
                     /ApEinfo_revcolor=#000eff
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     protein_bind    1524..1553
                     /label="lacOi"
                     /ApEinfo_fwdcolor=cyan
                     /ApEinfo_revcolor=green
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     CDS             1616..2323
                     /label="RFP(-ATG)"
                     /ApEinfo_fwdcolor=#2ccb22
                     /ApEinfo_revcolor=green
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     RBS             1590..1615
                     /label="5UTR-RFP"
                     /ApEinfo_fwdcolor=#fff917
                     /ApEinfo_revcolor=green
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     misc_feature    1415..1435
                     /label="Insertion"
                     /ApEinfo_fwdcolor=#ff0d14
                     /ApEinfo_revcolor=green
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     misc_feature    1475
                     /label="Mutation"
                     /ApEinfo_fwdcolor=#ff0808
                     /ApEinfo_revcolor=green
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
     misc_feature    1483
                     /label="Mutation(1)"
                     /ApEinfo_label=Mutation
                     /ApEinfo_fwdcolor=#ff0b0c
                     /ApEinfo_revcolor=green
                     /ApEinfo_graphicformat=arrow_data {{0 1 2 0 0 -1} {} 0}
                     width 5 offset 0
ORIGIN
        1 ctcatgacca aaatccctta acgtgagtta cgcgcgcgtc gttccactga gcgtcagacc
       61 ccgtagaaaa gatcaaagga tcttcttgag atccttttt tctgcgcgta atctgctgct
      121 tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa
      181 ctctttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact gttcttctag
      241 tgtagccgta gttagcccac cacttcaaga actctgtagc accgcctaca tacctcgctc
      301 tgctaatcct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg
      361 actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca
      421 cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat
      481 gagaaagcgc cacgcttccc gaagggagaa aggcggacag gtatccggta agcggcaggg
      541 tcggaacagg agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc
      601 ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc
      661 ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc
      721 cttttgctca catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg
      781 cctttgagtg agctgatacc gctcgccgca gccgaacgac cgagcgcagc gagtcagtga
      841 gcgaggaagc ggaaggcgag agtagggaac tgccaggcat caaactaagc agaaggcccc
      901 tgacggatgg ccttttgcg tttctacaaa ctctttctgt gttgtaaaac gacggccagt
      961 cttaagctcg ggcccccctgg gcggttctga taacgagtaa tcgttaatcc gcaaataacg
```

```
1021 taaaaacccg cttcggcggg ttttttttatg gggggagttt agggaaagag catttgtcag
1081 aatatttaag ggcgcctgtc actttgcttg atatatgaga attatttaac cttataaatg
1141 agaaaaaagc aacgcacttt aaataagata cgttgctttt tcgattgatg aacacctata
1201 attaaactat tcatctatta tttatgattt tttgtatata caatatttct agtttgttaa
1261 agagaattaa gaaaataaat ctcgaaaata ataaagggaa aatcagtttt tgatatcaaa
1321 attatacatg tcaacgataa tacaaaatat aatacaaact ataagatgtt atcagtattt
1381 attatgcatt tagaataaat tttgtgtcgc cctttacacg tacttagtcg ctgaaaattg
1441 tgagcggata acaattacga gcttcatgca cagttaaatc ataaaaaatt tatttgcttt
1501 gtgagcggat aacaattata atatgtggaa ttgtgagcgc tcacaattcc acaacggttt
1561 ccctctagaa ataatttttgt ttaacttttc gagaccttag gaggtaaaac atatggtttc
1621 caagtgcgag gaggataaca tggctatcat taaagagttc atgcgcttca aagttcacat
1681 ggagggttct gttaacggtc acgagttcga gatcgaaggc gaaggcgagg gccgtccgta
1741 tgaaggcacc cagaccgcca aactgaaagt gactaaaggc ggcccgctgc cttttgcgtg
1801 ggacatcctg agcccgcaat ttatgtacgg ttctaaagct tatgttaaac acccagcgga
1861 tatcccggac tatctgaagc tgtcttttcc ggaaggtttc aagtgggaac gcgtaatgaa
1921 ttttgaagat ggtggtgtcg tgaccgtcac tcaggactcc tccctgcagg atggcgagtt
1981 catctataaa gttaaactgc gtggtactaa ttttccatct gatggcccgg tgatgcagaa
2041 gaagacgatg ggttgggagg cgtctagcga acgcatgtat ccggaagatg gtgcgctgaa
2101 aggcgaaatt aaacagcgcc tgaaactgaa agatggcggc cattatgacg ctgaagtgaa
2161 aaccacgtac aaagccaaga aacctgtgca gctgcctggc gcgtacaatg tgaatattaa
2221 actggacatc acctctcata atgaagatta tacgatcgta gagcaatatg agcgcgcgga
2281 gggtcgtcat tctaccggtg gcatggatga gctgtacaaa taaggtctca ccccaagggc
2341 gacacccccct aattagcccg ggcgaaaggc ccagtctttc gactgagcct ttcgttttat
2401 ttgatgcctg gcagttccct actctcgcat ggggagtccc cacactacca tcggcgctac
2461 ggcgtttcac ttctgagttc ggcatggggt caggtgggac caccgcgcta ctgccgccag
2521 gcaaacaagg ggtgttatga gccatattca ggtataaatg ggctcgcgat aatgttcaga
2581 attggttaat tggttgtaac actgacccct atttgtttat ttttctaaat acattcaaat
2641 atgtatccgc tcatgagaca ataaccctga taaatgcttc aataatattg aaaaaggaag
2701 aatatgagcc atattcaacg ggaaacgtcg aggccgcgat taaattccaa catggatgct
2761 gatttatatg ggtataaatg ggctcgcgat aatgtcgggc aatcaggtgc gacaatctat
2821 cgcttgtatg ggaagcccga tgcgccagag ttgtttctga aacatggcaa aggtagcgtt
2881 gccaatgatg ttacagatga gatggtcaga ctaaactggc tgacggaatt tatgccactt
2941 ccgaccatca agcattttat ccgtactcct gatgatgcat ggttactcac cactgcgatc
3001 cccggaaaaa cagcgttcca ggtattagaa gaatatcctg attcaggtga aaatattgtt
3061 gatgcgctgg cagtgttcct gcgccggttg cactcgattc ctgtttgtaa ttgtcctttt
3121 aacagcgatc gcgtatttcg cctcgctcag gcgcaatcac gaatgaataa cggtttggtt
3181 gatgcgagtg attttgatga cgagcgtaat ggctggcctg ttgaacaagt ctggaaagaa
3241 atgcataaac ttttgccatt ctcaccggat tcagtcgtca ctcatggtga tttctcactt
3301 gataacctta tttttgacga ggggaaatta ataggttgta ttgatgttgg acgagtcgga
3361 atcgcagacc gataccagga tcttgccatc ctatggaact gcctcggtga gttttctcct
3421 tcattacaga aacggctttt tcaaaaatat ggtattgata atcctgatat gaataaattg
3481 cagtttcatt tgatgctcga tgagtttttc taagcggcgc gccatcgaat ggcgcaaaac
3541 ctttcgcggt atggcatgat agcgcccgga agagagtcaa ttcagggtgg tgaatatgaa
3601 accagtaacg ttatacgatg tcgcagagta tgccggtgtc tcttatcaga ccgtttcccg
3661 cgtggtgaac caggccagcc acgtttctgc gaaaacgcgg gaaaaagtgg aagcggcgat
3721 ggcggagctg aattacattc ccaaccgcgt ggcacaacaa ctggcgggca aacagtcgtt
3781 gctgattggc gttgccacct ccagtctggc cctgcacgcg ccgtcgcaaa ttgtcgcggc
3841 gattaaatct cgcgccgatc aactgggtgc cagcgtggtg gtgtcgatgg tagaacgaag
3901 cggcgtcgaa gcctgtaaag cggcggtgca caatcttctc gcgcaacgcg tcagtgggct
3961 gatcattaac tatccgctgg atgaccagga tgccattgct gtggaagctg cctgcactaa
4021 tgttccggcg ttatttcttg atgtctctga ccagacaccc atcaacagta ttattttctc
4081 ccatgaggac ggtacgcgac tgggcgtgga gcatctggtc gcattgggtc accagcaaat
4141 cgcgctgtta gcgggcccat taagttctgt ctcggcgcgt ctgcgtctgg ctggctggca
4201 taaatatctc actcgcaatc aaattcagcc gatagcggaa cgggaaggcg actggagtgc
4261 catgtccggt tttcaacaaa ccatgcaaat gctgaatgag ggcatcgttc ccactgcgat
4321 gctggttgcc aacgatcaga tggcgctggg cgcaatgcgc gccattaccg agtccgggct
4381 gcgcgttggt gcggatatct cggtagtggg atacgacgat accgaagata gctcatgtta
4441 tatcccgccg ttaaccacca tcaaacagga ttttcgcctg ctggggcaaa ccagcgtgga
4501 ccgcttgctg caactctctc agggccaggc ggtgaagggc aatcagctgt tgccagtctc
4561 actggtgaaa agaaaaacca ccctggcgcc caatacgcaa accgcctctc cccgcgcgtt
4621 ggccgattca ttaatgcagc tggcacgaca ggtttcccga ctggaaagcg ggcagtga
//
```

```
LOCUS       pAC_LYC                 9476 bp ds-DNA     circular     11-APR-2013
ACCESSION   pAC_LYC
VERSION     pAC_LYC
KEYWORDS    .
FEATURES             Location/Qualifiers
     source          join(1..1869,7107..9476)
                     /label="source:synthetic construct"
                     /mol_type="other DNA"
                     /db_xref="taxon:32630"
                     /plasmid="pACYC184"
                     /ApEinfo_fwdcolor=pink
                     /ApEinfo_revcolor=pink
     source          1870..7106
                     /label="source:Pantoea agglomerans"
                     /mol_type="genomic DNA"
                     /strain="Eho10"
                     /db_xref="taxon:549"
                     /tissue_lib="ATCC 39368"
                     /ApEinfo_fwdcolor=pink
                     /ApEinfo_revcolor=pink
     CDS             complement(6518..6732)
                     /label="beta-carotene hydroxylase"
                     /note="64 bases overlap crt B in opposite direction"
                     /citation=[1]
                     /citation=[3]
                     /codon_start=1
                     /transl_table=11
                     /product="beta-carotene hydroxylase"
                     /protein_id="AAA64983.1"
                     /db_xref="GI:148414"
                     /translation="MLVNSLIVILSVIAMEGIAAFTHRYIMHGWGWRWHESHHTPRKG
                     VFELNDLFAVVFAGVAIALIAVGTAGVWPLQWIGCGMTVYGLLYFLVHDGLVHQRWPF
                     HWIPRRGYLKRLYVAHRLHHAVRGREGCVSFGFIYARKPADLQAILRERHGRPPKRDA
                     AKDRPDAASPSSSSPE"
                     /ApEinfo_fwdcolor=#0000ff
                     /ApEinfo_revcolor=#77ff5d
     misc_feature    join(219..1869,7107..9036)
                     /label="chloramphenicol resistance gene"
                     /ApEinfo_fwdcolor=#cccccc
                     /ApEinfo_revcolor=#ff8468
     gene            4178..5636
                     /label="crtI"
                     /ApEinfo_fwdcolor=pink
                     /ApEinfo_revcolor=pink
     CDS             complement(6733..7106)
                     /label="unknown"
                     /citation=[1]
                     /codon_start=3
                     /transl_table=11
                     /product="unknown"
                     /protein_id="AAA64984.1"
                     /db_xref="GI:765097"
                     /translation="IYHLRALGVDVRLTRSDVTKGHPEYYQDYARRLAEETPGAYYID
                     QFSNPANPLAHATTTAPELFDQLAGRVDAVVVGVGSGGTLGGLQAWFAKHSPHTEFVL
                     ADPVGSILADQVETGQHGEAGSWRVEGIGEDFIPPLAQLENVRQAYRVTDAEAFATAR
                     QLLQTEGVLAGSSTGTLLAAALRYCQAQTTPKRVVTFACDSGNKYLSRMFNDEWLRQQ
                     GLQDDDRPAPAAPASALATLSVHSGAFNDQHGAVMPPIYATSTFAQPAPGQHTGYEYS
                     RSGNPTRHALESAIAELEGGTRGYAFASGLAAISTVLELLDKDSHIVAIDDVYGGTYR
                     LIENVRRRSTGLQVSWVKPDDLAGLEAAIRPDTRMIWVETPTNPLLKLADLAAIAEIA
                     RRHNVISVADNTFASPLIHRPLTLGFDIVVHSATKYLNGHSDVVAGLAVVGDRPALAE
                     KLGYLQNAVGGVLDPFSSFLTLRGIRTLALRIERHSSNALALARWLEQQPQVEKVWFP
                     WLTSHPQYALARQQMAQPGGMISIVIKGDDQRAAEVISKLKLFTLAESLGGVESLVSQ
                     PFSMTHASIPLAQRLANGITPQLIRLSVGIEDEKDLLADLQQALA"
                     /ApEinfo_fwdcolor=#0000ff
                     /ApEinfo_revcolor=#77ff5d
     misc_feature    443..583
                     /label="part of IS1 from Tn9"
                     /ApEinfo_fwdcolor=#cccccc
                     /ApEinfo_revcolor=#ff8468
     gene            1993..2916
                     /label="crtE"
```

```
                     /ApEinfo_fwdcolor=pink
                     /ApEinfo_revcolor=pink
     misc_feature    1993..2916
                     /label="crtE(1)"
                     /ApEinfo_label=crtE
                     /ApEinfo_fwdcolor=cyan
                     /ApEinfo_revcolor=green
     rep_origin      581..1493
                     /label="p15A origin of replication"
                     /ApEinfo_fwdcolor=cyan
                     /ApEinfo_revcolor=green
     gene            5653..6582
                     /label="crtB"
                     /ApEinfo_fwdcolor=pink
                     /ApEinfo_revcolor=pink
     misc_feature    4178..5656
                     /label="crtI(1)"
                     /ApEinfo_label=crtI
                     /ApEinfo_fwdcolor=cyan
                     /ApEinfo_revcolor=green
     misc_feature    complement(join(7107..8762,585..1869))
                     /label="Tn9 HaeII fragment"
                     /ApEinfo_fwdcolor=#cccccc
                     /ApEinfo_revcolor=#ff8468
     misc_feature    5653..6582
                     /label="crtB(1)"
                     /ApEinfo_label=crtB
                     /ApEinfo_fwdcolor=#cccccc
                     /ApEinfo_revcolor=green
     conflict        944..946
                     /label="conflict"
                     /citation=[1]
                     /citation=[4]
                     /replace="tc"
                     /ApEinfo_fwdcolor=pink
                     /ApEinfo_revcolor=pink
     misc_feature    join(1495..1869,7107..8507)
                     /label="pSC101 fragment"
                     /ApEinfo_fwdcolor=#cccccc
                     /ApEinfo_revcolor=#ff8468
     misc_feature    join(1581..1869,7107..8002)
                     /label="tetracycline resistance gene"
                     /ApEinfo_fwdcolor=#cccccc
                     /ApEinfo_revcolor=#ff8468
     gene            3797..4181
                     /label="crtY"
                     /ApEinfo_fwdcolor=pink
                     /ApEinfo_revcolor=pink
     gene            complement(6518..6732)
                     /label="crtZ"
                     /ApEinfo_fwdcolor=pink
                     /ApEinfo_revcolor=pink
     misc_feature    2979..4181
                     /label="unknown(1)"
                     /ApEinfo_label=unknown
                     /ApEinfo_fwdcolor=#ff0000
                     /ApEinfo_revcolor=green
ORIGIN
        1 gaattccgga tgagcattca tcaggcgggc aagaatgtga ataaaggccg gataaaactt
       61 gtgcttattt ttctttacgg tctttaaaaa ggccgtaata tccagctgaa cggtctggtt
      121 ataggtacat tgagcaactg actgaaatgc ctcaaaatgt tctttacgat gccattggga
      181 tatatcaacg gtggtatatc cagtgatttt tttctccatt ttagcttcct tagctcctga
      241 aaatctcgat aactcaaaaa atacgcccgg tagtgatctt atttcattat ggtgaaagtt
      301 ggaacctctt acgtgccgat caacgtctca ttttcgccaa aagttggccc agggcttccc
      361 ggtatcaaca gggacaccag gatttattta ttctgcgaag tgatcttccg tcacaggtat
      421 ttattcggcg caaagtgcgt cgggtgatgc tgccaactta ctgatttagt gtatgatggt
      481 gtttttgagg tgctccagtg gcttctgttt ctatcagctg tccctcctgt tcagctactg
      541 acggggtggt gcgtaacggc aaaagcaccg ccggacatca gcgctagcgg agtgtatact
      601 ggcttactat gttggcactg atgagggtgt cagtgaagtg cttcatgtgg caggagaaaa
      661 aaggctgcac cggtgcgtca gcagaatatg tgatacagga tatattccgc ttcctcgctc
      721 actgactcgc tacgctcggt cgttcgactg cggcgagcgg aaatggctta cgaacggggc
```

```
 781 ggagatttcc tggaagatgc caggaagata cttaacaggg aagtgagagg gccgcggcaa
 841 agccgttttt ccataggctc cgccccctg acaagcatca cgaaatctga cgctcaaatc
 901 agtggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct ggcggctccc
 961 tcgtgcgctc tcctgttcct gcctttcggt ttaccggtgt cattccgctg ttatggccgc
1021 gtttgtctca ttccacgcct gacactcagt tccgggtagg cagttcgctc caagctggac
1081 tgtatgcacg aacccccgt tcagtccgac cgctgcgcct tatccggtaa ctatcgtctt
1141 gagtccaacc cggaaagaca tgcaaaagca ccactggcag cagccactgg taattgattt
1201 agaggagtta gtcttgaagt catgcgccgg ttaaggctaa actgaaagga caagttttgg
1261 tgactgcgct cctccaagcc agttacctcg gttcaaagag ttggtagctc agagaacctt
1321 cgaaaaaccg ccctgcaagg cggttttttc gttttcagag caagagatta cgcgcagacc
1381 aaaacgatct caagaagatc atcttattaa tcagataaaa tatttctaga tttcagtgca
1441 atttatctct tcaaatgtag cacctgaagt cagccccata cgatataagt tgtaattctc
1501 atgtttgaca gcttatcatc gataagcttt aatgcggtag tttatcacag ttaaattgct
1561 aacgcagtca ggcaccgtgt atgaaatcta acaatgcgct catcgtcatc ctcggcaccg
1621 tcaccctgga tgctgtaggc ataggcttgg ttatgccggt actgccgggc ctcttgcggg
1681 atatcgtcca ttccgacagc atcgccagtc actatggcgt gctgctagcg ctatatgcgt
1741 tgatgcaatt tctatgcgca cccgttctcg gagcactgtc cgaccgcttt ggccgccgcc
1801 cagtcctgct cgcttcgcta cttggagcca ctatcgacta cgcgatcatg gcgaccacac
1861 ccgtcctgta gatctaaagg cacagcgtct catgcttcgc acaatgtaaa actgcttcag
1921 aacctggcga gagctatccg cgcggtctac ggttaactga tactaaaaga caattcagcg
1981 ggtaaccttg caatggtgag tggcagtaaa gcgggcgttt cgcctcatcg cgaaatagaa
2041 gtaatgagac aatccattga cgatcacctg gctggcctgt tacctgaaac cgacagccag
2101 gatatcgtca gccttgcgat gcgtgaaggc gtcatggcac ccggtaaacg gatccgtccg
2161 ctgctgatgc tgctggccgc ccgcgacctc cgctaccagg gcagtatgcc tacgctgctc
2221 gatctcgcct gcgccgttga actgacccat accgcgtcgc tgatgctcga cgacatgccc
2281 tgcatggaca acgccgagct gcgccgcggt cagcccacta cccacaaaaa atttggtgag
2341 agcgtggcga tccttgcctc cgttgggctg ctctctaaag cctttggtct gatcgccgcc
2401 accggcgatc tgccggggga gaggcgtgcc caggcggtca acgagctctc taccgccgtg
2461 ggcgtgcagg gcctggtact ggggcagttt cgcgatctta acgatgccgc cctcgaccgt
2521 acccctgacg ctatcctcag caccaaccac ctcaagaccg gcattctgtt cagcgcgatg
2581 ctgcagatcg tcgccattgc ttccgcctcg tcgccgagca cgcgagagac gctgcacgcc
2641 ttcgccctcg acttcggcca ggcgtttcaa ctgctggacg atctgcgtga cgatcacccg
2701 gaaaccggta aagatcgcaa taaggacgcg ggaaaatcga cgctggtcaa ccggctgggc
2761 gcagacgcgg cccggcaaaa gctgcgcgag catattgatt ccgccgacaa acacctcact
2821 tttgcctgtc cgcagggcgg cgccatccga cagtttatgc atctgtggtt tggccatcac
2881 cttgccgact ggtcaccggt catgaaaatc gcctgatacc gccctttgg gttcaagcag
2941 tacataacga tggaaccaca ttacaggagt agtgatgaat gaaggacgag cgccttgttc
3001 agcgtaagaa cgatcatctg gatatcgttc tcgacccccg tcgcgccgta actcaggcta
3061 gcgcaggttt tgagcgctgg cgctttaccc actgcgccct gccagagctg aattttagcg
3121 acatcacgct ggaaaccacc ttcctgaatc ggcagctaca ggctccgctg ctgatcagct
3181 ccatgaccgg cggcgttgag cgctcgcgcc atatcaaccg ccacctcgcc gaggcggcgc
3241 aggtgctaaa aattgcgatg ggggtgggct cccagcgcgt cgccattgag agcgacgcgg
3301 gcttagggct ggataaaacc ctgcggcagc tggctccgga cgtgccgctg ctggcgaacc
3361 tcggcgcggc gcagctgacc ggcagaaaag gtattgatta cgcccgacgg gccgtggaga
3421 tgatcgaggc ggatgcgctg attgtgcacc taaacccgct gcaggaggcg ctacagcccg
3481 gcggcgatcg cgactggcgc ggacggctgg cggctattga aactctggtc cgcgagctgc
3541 ccgttccgct ggtggtgaaa gaggtgggag ccggtatctc ccgaaccgtg gccgggcagc
3601 tgatcgatgc cggcgttacc gtgattgacg tcgcgggcgc gggcggcacc agctgggccg
3661 ccgttgaagg cgagcgggcg gccaccgagc agcagcgcag cgtggccaac gtctttgccg
3721 actgggggat ccccaccgct gaggcgctgg ttgacattgc cgaggcctgg ccgcagatgc
3781 cccttattgc ctcgggaatg cgggctgggc tatttcaccc taccactggc tattcgctgc
3841 cgctggcggt ggcccttgcc gacgcgattg ccgacagccc gcggctgggc agcgttccgc
3901 tctatcagct cacccggcag tttgccgaac gccactggcg caggcaggga ttcttccgcc
3961 tgctgaaccg gatgcttttc ctggccgggc gcgaggagaa ccgctggcgg gtgatgcagc
4021 gcttttatgg gctgccggag cccaccgtag agcgctttta cgccggtcgg ctctctctct
4081 ttgataaggc ccgcattttg acgggcaagc caccggttcc gctgggcgaa gcctggcggg
4141 cggcgctgaa ccattttcct gacagacgag ataaaggatg aaaaaaaccg ttgtgattgg
4201 cgcaggcttt ggtggcctgg cgctggcgat tcgcctgcag gcggcaggga tcccaaccgt
4261 actgctggag cagcgggaca agcccggcgg tcgggcctac gtctggcatg accagggctt
4321 tacctttgac gccgggccga cggtgatcac cgatcctacc gcgcttgagg cgctgttcac
4381 cctggccggc aggcgcatgg aggattacgt caggctgctg ccggtaaaac ccttctaccg
4441 actctgctgg gagtccggga agaccctcga ctatgctaac gacagcgccg agcttgaggc
4501 gcagattacc cagttcaacc cccgcgacgt cgagggctac cggcgctttc tggcttactc
4561 ccaggcggta ttccaggagg gatatttgcg cctcggcagc gtgccgttcc tctcttttcg
4621 cgacatgctg cgcgccgggc cgcagctgct taagctccag gcgtggcaga gcgtctacca
4681 gtcggtttcg cgctttattg aggatgagca tctgcggcag gccttctcgt tccactccct
4741 gctggtaggc ggcaacccct tcaccacctc gtccatctac accctgatcc acgcccttga
4801 gcgggagtgg ggggtctggt tccctgaggg cggcaccggg gcgctggtga acggcatggt
4861 gaagctgttt accgatctgg gcggggagat cgaactcaac gcccgggtcg aagagctggt
4921 ggtggccgat aaccgcgtaa gccaggtccg gctggcggat ggtcggatct ttgacaccga
4981 cgccgtagcc tcgaacgctg acgtggtgaa cacctataaa aagctgctcg gccaccatcc
```

5041 ggtggggcag aagcgggcgg cagcgctgga gcgcaagagc atgagcaact cgctgtttgt
5101 gctctacttc ggcctgaacc agcctcattc ccagctggcg caccatacca tctgttttgg
5161 tccccgctac cgggagctga tcgacgagat ctttaccggc agcgcgctgg cggatgactt
5221 ctcgctctac ctgcactcgc cctgcgtgac cgatccctcg ctcgcgcctc ccggctgcgc
5281 cagcttctac gtgctggccc cggtgccgca tcttggcaac gcgccgctgg actgggcgca
5341 ggaggggccg aagctgcgcg accgcatctt tgactacctt gaagagcgct atatgcccgg
5401 cctgcgtagc cagctggtga cccagcggat ctttaccccg gcagacttcc acgacacgct
5461 ggatgcgcat ctgggatcgg ccttctccat cgagccgctg ctgacccaaa gcgcctggtt
5521 ccgcccgcac aaccgcgaca gcgacattgc caacctctac ctggtgggcg caggtactca
5581 ccctggggcg ggcattcctg gcgtagtggc ctcggcgaaa gccaccgcca gcctgatgat
5641 tgaggatctg caatgagcca accgccgctg cttgaccacg ccacgcagac catggccaac
5701 ggctcgaaaa gttttgccac cgctgcgaag ctgttcgacc cggccacccg ccgtagcgtg
5761 ctgatgctct acacctggtg ccgccactgc gatgacgtca ttgacgacca gacccacggc
5821 ttcgccagcg aggccgcggc ggaggaggag gccacccagc gcctggcccg gctgcgcacg
5881 ctgaccctgg cggcgtttga aggggccgag atgcaggatc cggccttcgc tgcctttcag
5941 gaggtggcgc tgacccacgg tattacgccc cgcatggcgc tcgatcacct cgacggcttt
6001 gcgatggacg tggctcagac ccgctatgtc acctttgagg atacgctgcg ctactgctat
6061 cacgtggcgg gcgtggtggg tctgatgatg gccagggtga tgggcgtgcg ggatgagcgg
6121 gtgctggatc gcgcctgcga tctggggctg gccttccagc tgacgaatat cgcccgggat
6181 attattgacg atgcggctat tgaccgctgc tatctgcccg ccgagtggct gcaggatgcc
6241 gggctgaccc cggagaacta tgccgcgcgg gagaatcggg ccgcgctggc gcgggtggcg
6301 gagcggctta ttgatgccgc agagccgtac tacatctcct cccaggccgg gctacacgat
6361 ctgccgccgc gctgcgcctg ggcgatcgcc accgcccgca gcgtctaccg ggagatcggt
6421 attaaggtaa aagcggcggg aggcagcgcc tgggatcgcc gccagcacac cagcaaaggt
6481 gaaaaaattg ccatgctgat ggcggcaccg gggcaggtta ttcgggcgaa gacgacgagg
6541 gtgacgccgc gtccggccgg tctttggcag cgtcccgttt aggcgggcgg ccatgacgtt
6601 cacgcaggat cgcctgtagg tcggcaggct tgcgggcgta aataaaaccg aaggagacgc
6661 agccctcccg gccgcgcacc gcgtggtgca ggcggtgggc gacgtagagc cgcttcaggt
6721 agccccggcg cgggatcccg cctcgccgtg ctgtccggtc tcaacctgat ccgccagaat
6781 cgagccaacg ggatcggcca gcacgaattc ggtatgcgga gagtgtttgg caaaccatgc
6841 ctgcaagcca cccagcgtgc cgccggagcc aacgcctact accacggcat caaccctgcc
6901 tgccagctgg tcgaacagct ccggtgcggt ggtggtggcg tgcgccagcg ggttggccgg
6961 gttagagaac tgatcaatat agtaagcacc cggcgtctcc tctgccaggc ggcgggcgta
7021 gtcctgatag tactccgggt ggccctttgt cacgtcggag cgggtcaggc gcacatcaac
7081 acccagcgca cgcaggtggt agatcttcta cgccggacgc atcgtggccg gcatcaccgg
7141 cgccacaggt gcggttgctg gcgcctatat cgccgacatc accgatgggg aagatcgggc
7201 tcgccacttc gggctcatga gcgcttgttt cggcgtgggt atggtggcag gccccgtggc
7261 cgggggactg ttgggcgcca tctccttgca tgcaccattc cttgcggcgg cggtgctcaa
7321 cggcctcaac ctactactgg gctgcttcct aatgcaggag tcgcataagg gagagcgtcg
7381 accgatgccc ttgagagcct tcaacccagt cagctccttc cggtgggcgc ggggcatgac
7441 tatcgtcgcc gcacttatga ctgtcttctt tatcatgcaa ctcgtaggac aggtgccggc
7501 agcgctctgg gtcatttttcg gcgaggaccg ctttcgctgg agcgcgacga tgatcggcct
7561 gtcgcttgcg gtattcggaa tcttgcacgc cctcgctcaa gccttcgtca ctggtcccgc
7621 caccaaacgt ttcggcgaga agcaggccat tatcgccggc atggcggccg acgcgctggg
7681 ctacgtcttg ctggcgttcg cgacgcgagg ctggatggcc ttccccatta tgattcttct
7741 cgcttccggc ggcatcggga tgcccgcgtt gcaggccatg ctgtccaggc aggtagatga
7801 cgaccatcag ggacagcttc aaggatcgct cgcggctctt accagcctaa cttcgatcac
7861 tggaccgctg atcgtcacgg cgatttatgc cgcctcggcg agcacatgga acgggttggc
7921 atggattgta ggcgccgccc tataccttgt ctgcctcccc gcgttgcgtc gcggtgcatg
7981 gagccgggcc acctcgacct gaatggaagc cggcggcacc tcgctaacgg attcaccact
8041 ccaagaattg gagccaatca attcttgcgg agaactgtga atgcgcaaac caacccttgg
8101 cagaacatat ccatcgcgtc cgccatctcc agcagccgca cgcggcgcat ctcgggcagc
8161 gttgggtcct ggccacgggt gcgcatgatc gtgctcctgt cgttgaggac ccggctaggc
8221 tggcggggtt gccttactgg ttagcagaat gaatcaccga tacgcgagcg aacgtgaagc
8281 gactgctgct gcaaaacgtc tgcgacctga gcaacaacat gaatggtctt cggtttccgt
8341 gtttcgtaaa gtctggaaac gcggaagtcc cctacgtgct gctgaagttg cccgcaacag
8401 agagtggaac caaccggtga taccacgata ctatgactga gagtcaacgc catgagcggc
8461 ctcatttctt attctgagtt acaacagtcc gcaccgctgt ccggtagctc cttccggtgg
8521 gcgcggggca tgactatcgt cgccgcactt atgactgtct tctttatcat gcaactcgta
8581 ggacaggtgc cggcagcgcc caacagtccc ccggccacgg ggcctgccac catacccacg
8641 ccgaaacaag cgccctgcac cattatgttc cggatctgca tcgcaggatg ctgctggcta
8701 ccctgtggaa cacctacatc tgtattaacg aagcgctaac cgtttttatc aggctctggg
8761 aggcagaata aatgatcata tcgtcaatta ttacctccac ggggagagcc tgagcaaact
8821 ggcctcaggc atttgagaag cacacggtca cactgcttcc ggtagtcaat aaaccggtaa
8881 accagcaata gacataagcg gctatttaac gaccctgccc tgaaccgacg accgggtcga
8941 atttgctttc gaatttctgc cattcatccg cttattatca cttattcagg cgtagcacca
9001 ggcgtttaag ggcaccaata actgccttaa aaaaattacg ccccgccctg ccactcatcg
9061 cagtactgtt gtaattcatt aagcattctg ccgacatgga agccatcaca gacggcatga
9121 tgaacctgaa tcgccagcgg catcagcacc ttgtcgcctt gcgtataata tttgcccatg
9181 gtgaaaacgg gggcgaagaa gttgtccata ttggccacgt ttaaatcaaa actggtgaaa
9241 ctcacccagg gattggctga gacgaaaaac atattctcaa taaacccttt agggaaatag

```
     9301 gccaggtttt caccgtaaca cgccacatct tgcgaatata tgtgtagaaa ctgccggaaa
     9361 tcgtcgtggt attcactcca gagcgatgaa aacgtttcag tttgctcatg gaaaacggtg
     9421 taacaagggt gaacactatc ccatatcacc agctcaccgt ctttcattgc catacg
//
```

```
LOCUS       pGM-O-RFP               7819 bp ds-DNA     circular     11-APR-2013
ACCESSION   pGM-O-RFP
VERSION     pGM-O-RFP
KEYWORDS    .
FEATURES             Location/Qualifiers
     rep_origin      complement(6137..6804)
                     /label="Col E1 origin"
                     /vntifkey="33"
     promoter        complement(7750..7819)
                     /label="bla promoter"
                     /vntifkey="30"
     CDS             complement(6952..7749)
                     /label="kanR"
                     /vntifkey="4"
     terminator      1..44
                     /label="tR2 lambda phage terminator"
                     /vntifkey="43"
     terminator      complement(6093..6136)
                     /label="rrnB T1 transcription terminator"
                     /vntifkey="43"
     misc_feature    448..456
                     /label="vector ID"
                     /vntifkey="21"
     misc_feature    5938..5946
                     /label="vector ID"
                     /vntifkey="21"
     CDS             472..5922
                     /label="self-generator"
ORIGIN
        1 gctaataaca ggcctgctgg taatcgcagg ccttttttatt tggtgtattt agaaaaataa
       61 acaaataggg gttccgcgca catttccccg aaaagtgcca cctaaattgt aagcgttaat
      121 attttgttaa aattcgcgtt aaatttttgt taaatcagct catttttaa ccaataggcc
      181 gaaatcggca aaatccctta taaatcaaaa gaatagaccg agatagggtt gagtggccgc
      241 tacagggcgc tcccattcgc cattcaggct gcgcaactgt tgggaagggc gtttcggtgc
      301 gggcctcttc gctattacgc cagctggcga aaggggatg tgctgcaagg cgattaagtt
      361 gggtaacgcc agggttttcc cagtcacgac gttgtaaaac gacggccagt gagcgcgacg
      421 taatacgact cactataggg cgaattgtag aaggaaggcc gtcaaggcct aggcgcgcct
      481 aatacgactc actataggga gacgccgctg agaaaaagcg aagcggcact gctctttaac
      541 aatttatcag acaatctgtg tgggcactcg aagatacgga ttcttaacgt cgcaagacga
      601 aaaatgaata ccaagtctca agagtgaaca cgtaattcat tacgaagttt aattctttga
      661 gcgtcaaact tttaaattga agagtttgat catggctcag attgaacgct ggcggcaggc
      721 ctaacacatg caagtcgaac ggtaacagga agaagcttgc ttctttgctg acgagtggcg
      781 gacgggtgag taatgtctgg gaaactgcct gatggagggg gataactact ggaaacggta
      841 gctaataccg cataacgtcg caagaccaaa gaggggacc ttcgggcctc ttgccatcgg
      901 atgtgcccag atgggattag ctagtaggtg gggtaacggc tcacctaggc gacgatccct
      961 agctggtctg agaggatgac cagccacact ggaactgaga cacggtccag actcctacgg
     1021 gaggcagcag tggggaatat tgcacaatgg gcgcaagcct gatgcagcca tgccgcgtgt
     1081 atgaagaagg ccttcgggtt gtaaagtact ttcagcgggg aggaagggag taaagttaat
     1141 acctttgctc attgacgtta cccgcagaag aagcaccggc taactccgtg ccagcagccg
     1201 cggtaatacg gagggtgcaa gcgttaatcg gaattactgg gcgtaaagcg cacgcaggcg
     1261 gtttgttaag tcagatgtga aatccccggg ctcaacctgg gaactgcatc tgatactggc
     1321 aagcttgagt ctcgtagagg ggggtagaat tccaggtgta gcggtgaaat gcgtagagat
     1381 ctggaggaat accggtggcg aaggcggccc cctggacgaa gactgacgct caggtgcgaa
     1441 agcgtgggga gcaaacagga ttagataccc tggtagtcca cgccgtaaac gatgtcgact
     1501 tggaggttgt gcccttgagg cgtggcttcc ggagctaacg cgttaagtcg accgcctggg
     1561 gagtacggcc gcaaggttaa aactcaaatg aattgacggg ggcccgcaca agcggtggag
     1621 catgtggttt aattcgatgc aacgcgaaga accttacctg gtcttgacat ccacggaagt
     1681 tttcagagat gagaatgtgc cttcgggaac cgtgagacag gtgctgcatg gctgtcgtca
     1741 gctcgtgttg tgaaatgttg ggttaagtcc cgcaacgagc gcaacccttа tcctttgttg
     1801 ccagcggtcc ggccgggaac tcaaaggaga ctgccagtga gaaactggag gaaggtgggg
     1861 atgacgtcaa gtcatcatgg cccttacgac cagggctaca cacgtgctac aatggcgcat
     1921 acaaagagaa gcgacctcgc gagagcaagc ggacctcata aagtgcgtcg tagtccggat
     1981 tggagtctgc aactcgactc catgaagtcg gaatcgctag taatcgtgga tcagaatgcc
     2041 acggtgaata cgttcccggg ccttgtacac accgcccgtc acaccatggg agtgggttgc
     2101 aaaagaagta ggtagcttaa ccttcgggag ggcgcttacc actttgtgat tcatgactgg
     2161 ggtgaagtcg taacaaggta accgtagggg aacctgcggt tggatcacgg aactacctta
     2221 aagaagcgta ctttgtagtg ctcacacaga ttgtctgata gaaagtgaaa agcaaggcgt
     2281 ttacgcgttg ggagtgaggc tgaagagaat aaggccgttc gctttctatt aatgaaagct
     2341 caccctacac gaaaatatca cgcaacgcgt gataagcaat tttcgtgttc cgaaaaaaat
     2401 gaacacgatt aacatcgcta agaacgactt ctctgacatc gaactggctg ctatcccgtt
     2461 caacactctg gctgaccatt acggtgagcg tttagctcgc gaacagttgg cccttgagca
```

2521 tgagtcttac gagatgggtg aagcacgctt ccgcaagatg tttgagcgtc aacttaaagc
2581 tggtgaggtt gcggataacg ctgccgccaa gcctctcatc actaccctac tccctaagat
2641 gattgcacgc atcaacgact ggtttgagga agtgaaagct aagcgcggca agcgcccgac
2701 agccttccag ttcctgcaag aaatcaagcc ggaagccgta gcgtacatca ccattaagac
2761 cactctggct tgcctaacca gtgctgacaa tacaaccgtt caggctgtag caagcgcaat
2821 cggtcgggcc attgaggacg aggctcgctt cggtcgtatc cgtgaccttg aagctaagca
2881 cttcaagaaa aacgttgagg aacaactcaa caagcgcgta gggcacgtct acaagaaagc
2941 atttatgcaa gttgtcgagg ctgacatgct ctctaagggt ctactcggtg gcgaggcgtg
3001 gtcttcgtgg cataaggaag actctattca tgtaggagta cgctgcatcg agatgctcat
3061 tgagtcaacc ggaatggtta gcttacaccg ccaaaatgct ggcgtagtag gtcaagactc
3121 tgagactatc gaactcgcac ctgaatacgc tgaggctatc gcaacccgtg caggtgcgct
3181 ggctggcatc tctccgatgt tccaaccttg cgtagttcct cctaagccgt ggactggcat
3241 tactggtggt ggctattggg ctaacggtcg tcgtcctctg gcgctggtgc gtactcacag
3301 taagaaagca ctgatgcgct acgaagacgt ttacatgcct gaggtgtaca aagcgattaa
3361 cattgcgcaa aacaccgcat ggaaaatcaa caagaaagtc ctagcggtcg ccaacgtaat
3421 caccaagtgg aagcattgtc cggtcgagga catccctgcg attgagcgtg aagaactccc
3481 gatgaaaccg gaagacatcg acatgaatcc tgaggctctc accgcgtgga aacgtgctgc
3541 cgctgctgtg taccgcaagg acaaggctcg caagtctcgc cgtatcagcc ttgagttcat
3601 gcttgagcaa gccaataagt ttgctaacca taaggccatc tggttccctt acaacatgga
3661 ctggcgcggt cgtgtttacg ctgtgtcaat gttcaacccg caaggtaacg atatgaccaa
3721 aggactgctt acgctggcga aaggtaaacc aatcggtaag gaaggttact actggctgaa
3781 aatccacggt gcaaactgtg cgggtgtcga taaggttccg ttccctgagc gcatcaagtt
3841 cattgaggaa aaccacgaga acatcatggc ttgcgctaag tctccactgg agaacacttg
3901 gtgggctgag caagattctc cgttctgctt ccttgcgttc tgctttgagt acgctggggt
3961 acagcaccac ggcctgagct ataactgctc ccttccgctg gcgtttgacg ggtcttgctc
4021 tggcatccag cacttctccg cgatgctccg agatgaggta ggtggtcgcg cggttaactt
4081 gcttcctagt gaaaccgttc aggacatcta cgggattgtt gctaagaaag tcaacgagat
4141 tctacaagca gacgcaatca atgggaccga taacgaagta gttaccgtga ccgatgagaa
4201 cactggtgaa atctctgaga aagtcaagct gggcactaag gcactggctg gtcaatggct
4261 ggcttacggt gttactcgca gtgtgactaa gcgttcagtc atgacgctgg cttacgggtc
4321 caaagagttc ggcttccgtc aacaagtgct ggaagatacc attcagccag ctattgattc
4381 cggcaagggt ctgatgttca ctcagccgaa tcaggctgct ggatacatgg ctaagctgat
4441 ttgggaatct gtgagcgtga cggtggtagc tgcggttgaa gcaatgaact ggcttaagtc
4501 tgctgctaag ctgctggctg ctgaggtcaa agataagaag actggagaga ttcttcgcaa
4561 gcgttgcgct gtgcattggg taactcctga tggtttccct gtgtggcagg aatacaagaa
4621 gcctattcag acgcgcttga acctgatgtt cctcggtcag ttccgcttac agcctaccat
4681 taacaccaac aaagatagcg agattgatgc acacaaacag gagtctggta tcgctcctaa
4741 ctttgtacac agccaagacg gtagccacct tcgtaagact gtagtgtggg cacacgagaa
4801 gtacggaatc gaatcttttg cactgattca cgactccttc ggtaccattc cggctgacgc
4861 tgcgaacctg ttcaaagcag tgcgcgaaac tatggttgac acatatgagt cttgtgatgt
4921 actggctgat ttctacgacc agttcgctga ccagttgcac gagtctcaat tggacaaaat
4981 gccagcactt ccggctaaag gtaacttgaa cctccgtgac atcttagagt cggacttcgc
5041 gttcgcgtaa taagttccga aaaaaatggc ttcctccgaa gacgttatca aagagttcat
5101 gcgtttcaaa gttcgtatgg aaggttccgt taacggtcac gagttcgaaa tcgaaggtga
5161 aggtgaaggt cgtccgtacg aaggtaccca gaccgctaaa ctgaaagtta ccaaaggtgg
5221 tccgctgccg ttcgcttggg acatcctgtc cccgcagttc cagtacggtt ccaaagctta
5281 cgttaaacac ccggctgaca tcccggacta cctgaaactg tccttcccgg aaggtttcaa
5341 atgggaacgt gttatgaact tcgaagacgg tggtgttgtt accgttaccc aggactcctc
5401 cctgcaagac ggtgagttca tctacaaagt taaactgcgt ggtaccaact tcccgtccga
5461 cggtccggtt atgcagaaaa aaaccatggg ttgggaagct tccaccgaac gtatgtaccc
5521 ggaagacggt gctctgaaag gtgaaatcaa aatgcgtctg aaactgaaag acggtggtca
5581 ctacgacgct gaagttaaaa ccacctacat ggctaaaaaa ccggttcagc tgccgggtgc
5641 ttacaaaacc gacatcaaac tggacatcac ctcccacaac gaagactaca ccatcgttga
5701 acagtacgaa cgtgctgaag gtcgtcactc caccggtgct aggcctgctg caaacgacga
5761 aaactacgct ttagtagctt aataaccagg catcaaataa aacgaaaggc tcagtcgaaa
5821 gactgggcct ttcgttttat ctgttgtttg tcggtgaacg ctctctacta gagtcacact
5881 ggctcacctt cgggtgggcc tttctgcgtt tatattaatt aactggcctc atgggccttc
5941 cttctatcac tgcccgcttt ccagtcggga aacctgtcgt gccagctgca ttaacatggt
6001 catagctgtt tccttgcgta ttgggcgctc tccgcttcct cgctcactga ctcgctgcgc
6061 tcggtcgttc gggtaaagcc tggggtgcct aaataaaacg aaaggcccag tcttccgact
6121 gagcctttcg ttttattgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc
6181 gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa atcgacgctc
6241 aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag
6301 ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct
6361 cccttcggga agcgtggcgc tttctcatag ctcacgctgt aggtatctca gttcggtgta
6421 ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccg accgctgcgc
6481 cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactggc
6541 agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt
6601 gaagtggtgg cctaactacg gctacactag aagaacagta tttggtatct gcgctctgct
6661 gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc
6721 tggtagcggt ggttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca

```
     6781 agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta
     6841 agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa
     6901 atgaagtttt aaatcaatct aaagtatata tgagtaaact tggtctgaca gttattagaa
     6961 aaattcatcc agcagacgat aaaacgcaat acgctggcta tccggtgccg caatgccata
     7021 cagcaccaga aaacgatccg cccattcgcc gcccagttct tccgcaatat cacgggtggc
     7081 cagcgcaata tcctgataac gatccgccac gcccagacgg ccgcaatcaa taaagccgct
     7141 aaaacggcca ttttccacca taatgttcgg caggcacgca tcaccatggg tcaccaccag
     7201 atcttcgcca tccggcatgc tcgctttcag acgcgcaaac agctctgccg gtgccaggcc
     7261 ctgatgttct tcatccagat catcctgatc caccaggccc gcttccatac gggtacgcgc
     7321 acgttcaata cgatgtttcg cctgatgatc aaacggacag gtcgccgggt ccagggtatg
     7381 cagacgacgc atggcatccg ccataatgct cacttttct gccggcgcca gatggctaga
     7441 cagcagatcc tgacccggca cttcgcccag cagcagccaa tcacggcccg cttcggtcac
     7501 cacatccagc accgccgcac acggaacacc ggtggtggcc agccagctca gacgcgccgc
     7561 ttcatcctgc agctcgttca gcgcaccgct cagatcggtt ttcacaaaca gcaccggacg
     7621 accctgcgcg ctcagacgaa acaccgccgc atcagagcag ccaatggtct gctgcgccca
     7681 atcatagcca aacagacgtt ccacccacgc tgccgggcta cccgcatgca ggccatcctg
     7741 ttcaatcata ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat
     7801 gagcggatac atatttgaa
//
```

Vita

George Howard McArthur IV was born on 2 April 1984 in Roanoke, Virginia and is an American citizen. He grew up in the Shenandoah Valley and graduated from Turner Ashby High School in 2002. George received a Bachelor of Science from the University of Virginia (Charlottesville, VA) in 2008, with majors in Chemical Engineering and Music and a minor in Biology. In 2009, he earned a Master of Science in Chemical Engineering from Virginia Commonwealth University (Richmond, VA). George was a Whitaker International Fellow from 2011-12 in the laboratory of Travis S. Bayer at Imperial College London's Centre for Synthetic Biology and Innovation. He is completing his Doctor of Philosophy under the supervision of Stephen S. Fong at VCU.