Original citation:

Permanent WRAP URL:
http://wrap.warwick.ac.uk/91282

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher’s statement:
http://dx.doi.org/10.1038/nprot.2017.084

A note on versions:
The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher’s version. Please see the ‘permanent WRAP URL’ above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk
Intracellular directed evolution of proteins from combinatorial libraries based on conditional phage replication

Andreas K. Brödel\textsuperscript{1}, Alfonso Jaramillo\textsuperscript{2,3,4}, Mark Isalan\textsuperscript{1*}

\textsuperscript{1}Department of Life Sciences, Imperial College London, London SW7 2AZ, UK.

\textsuperscript{2}Warwick Integrative Synthetic Biology Centre and School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK.

\textsuperscript{3}CNRS-UMR8030, Laboratoire iSSB and Université Paris-Saclay and Université d’Évry and CEA, DRF, IG, Genoscope, Évry 91000, France.

\textsuperscript{4}Institute for Integrative Systems Biology (I2SysBio), University of Valencia-CSIC, 46980 Paterna, Spain.

* To whom correspondence should be addressed: Mark Isalan, Department of Life Sciences, South Kensington Campus, Imperial College London, London SW7 2AZ, UK, m.isalan@imperial.ac.uk, Tel: +44 (0)20 7594 6482
**ABSTRACT**

Directed evolution is a powerful tool to improve the characteristics of biomolecules. Here we present a protocol for the intracellular evolution of proteins with distinct differences and advantages to established techniques. These include the ability to select for a particular function from a library of protein variants inside cells, minimizing undesired co-evolution and propagation of non-functional library members, as well as allowing positive and negative selection logics using basally-active promoters. A typical evolution experiment comprises the following steps: (i) Preparation of a combinatorial M13 phagemid library expressing variants of the gene of interest and the *E. coli* host cells; (ii) Multiple rounds of an intracellular selection process towards a desired activity; (iii) The characterization of the evolved target proteins. The system has been developed for the selection of new orthogonal transcription factors\(^1\) (TFs) but is capable of evolving any gene – or gene circuit function – that can be linked to conditional M13 phage replication. Here we demonstrate our approach by the directed evolution of TFs based on \(\lambda\) cl against two synthetic bidirectional promoters. The evolved TF variants enable simultaneous activation and repression against their engineered promoters and do not cross-react with the wild-type promoter, thus ensuring orthogonality. This protocol requires no special equipment, allowing synthetic biologists and general users to evolve improved biomolecules within ~7 weeks.

**Keywords:** Directed evolution, Gene circuit engineering, Synthetic biology, Protein engineering, M13 bacteriophage, Gene networks, Logic gates, Artificial transcription factors

**EDITORIAL SUMMARY** This protocol describes a phagemid-based intracellular evolution approach to generate and select for proteins with improved biological characteristics.

**TWEET** Improving biomolecule function using an intracellular directed evolution approach.

**COVER TEASER** Improving biomolecules by intracellular directed evolution.

**Key references:**


INTRODUCTION

Directed evolution has emerged as a powerful tool to improve the characteristics of biomolecules\(^2\text{-}^4\). The approach mimics natural selection to evolve biomolecules towards a desired activity\(^5\). One efficient and commonly-used strategy to achieve this in a laboratory environment is to employ filamentous bacteriophages such as M13, to link a mutable genotype to a selectable phenotype. In this way, a number of M13 phage-assisted methods, such as the widely-used phage display technology\(^6\), have been developed and applied to improve a wide variety of proteins, including antibodies\(^7\text{-}^9\), DNA-binding proteins\(^10\text{-}^11\), and enzymes\(^12\text{-}^13\). These systems are characterized by an extracellular (in vitro) or intracellular (in vivo) mode of operation. In vitro systems are generally easier to engineer in terms of selection stringency adjustments\(^14\), but possess certain limitations that can only be overcome by applying intracellular processes. For example, selection from combinatorial libraries in vivo ensures compatibility with the host cell machinery. This facilitates the optimization of synthetic proteins and gene circuits\(^15\text{-}^17\), which ultimately have to function in a host cell context. In vivo methods promote selection for orthogonality\(^18\text{-}^19\), – a lack of cross-reactions – by intrinsically counter-selecting against adverse effects inside the cell. To further broaden the applications of in vivo directed evolution, we recently developed an M13 phage-based method\(^1\) for the intracellular selection of proteins from combinatorial libraries with distinct differences and advantages to established techniques.

Overview of the protocol

This protocol describes a general approach for the directed evolution of proteins from combinatorial libraries on phagemids (Fig. 1). The selection process takes place inside E. coli cells by linking the target protein’s activity to conditional phage production, thus allowing enrichment of functional library members. This is exemplified here by the directed evolution of orthogonal dual transcription factors (TFs) based on bacteriophage λ cl variants\(^1\), selecting against synthetic promoters. However, the method can be readily adapted for other target biomolecules (see Applications of the method). A typical evolution experiment consists of: 1) The preparation of a combinatorial M13 phage library (steps 1-38) and E. coli host cells (steps 39-47); 2) The selection process towards a desired activity (steps 48-61); and 3) The characterization of the selected target proteins (steps 62-68) (Fig. 2).

The system is based on E. coli cultures and three compatible plasmids (available from Addgene; see MATERIALS). Together, these conditionally produce phage (containing the evolving gene) in correlation to the activity of a library member. A selection experiment
always begins with an *E. coli* culture that contains the first two plasmids: a modified helper phage plasmid (HP) and an accessory plasmid (AP) (Fig. 3a). The HP provides almost all that is needed for phage propagation, except for two essential genes (gIII and gVI). Furthermore, the weak M13 packaging signal (PS) is removed from the original M13KO7 HP to obtain the final M13KO7-ΔPS-ΔgIII-ΔgVI HP. The second plasmid, AP, contains a conditional gene circuit that links an inducible input (e.g. a promoter with a novel operator) to gVI expression. The evolving gene or gene circuit is placed on the third plasmid, termed a phagemid (PM), which is packaged into an infectious phage particle only when all phage genes are expressed. The PM contains the second missing gene (gIII) and a combinatorially-randomized gene of interest (GOI) and is provided to the *E. coli* culture in the form of an infectious phage library (Fig. 3b). Crucially, our system moves Gene III onto the phagemid so that phage replication occurs only after initial infection, thus circumventing infection resistance^{20,21}, and decreasing the chances of propagating non-functional library members due to multiple infections. A GOI with the desired characteristics upregulates gene VI expression on the AP, completing the phage life cycle. For example, a randomized TF library member that activates an artificial promoter upstream of gVI will increase its own phage production (Fig. 4a). In this way, a protein with novel desired properties can be selected after several rounds of reinfection.

**Applications of the method**

The method has been used to evolve a set of dual activator-repressor switches for orthogonal logic gates, based on bacteriophage λ cl variants, and multi-input promoter architectures, and these switches have been successfully applied in downstream synthetic gene circuits^{1}. In general, the method is capable of evolving any gene – or gene circuit function – on the phagemid that can be linked to pVI production. This is analogous to previous uses of phage-assisted continuous evolution (PACE)^{22} (Fig. 4). With PACE, a wide range of medically and biotechnologically relevant biomolecules, including polymerases^{22}, proteases^{23}, genome-editing proteins^{11} as well as protein-protein interactions^{24} were linked to conditional M13 phage propagation. In principle, any application where directed evolution approaches have been proposed (e.g. biosensors^{25} or hybrids with chemical evolution^{26}) can be adapted to this method if the target protein’s activity can be linked to conditional M13 phage production. Although certain applications (e.g. membrane proteins) would be harder to adapt, which is why other methods such as liposome display^{27} have been developed.
Comparison with other methods

Several bacterial directed evolution methods have been developed based on phage replication\textsuperscript{22}, display technologies\textsuperscript{6,27-29}, genome engineering\textsuperscript{30}, as well as conditional cell growth\textsuperscript{31,32}. Linking a target protein’s activity to cell growth is a widely-used strategy and is particularly suitable when the evolving gene directly improves cellular fitness\textsuperscript{33,34}. The use of bacteriophage offers a convenient way to uncouple the fitness function of a cell with target protein activity. This is achieved by linking a target gene’s activity to phage replication using a conditional gene circuit. The main advantage of conditional phage production over display technologies is the compatibility of target genes or gene circuits with the host cell machinery as these have to function in a host cell context. In contrast to PACE (which uses gIII as the sole conditional gene), our phagemid-based approach facilitates the selection of large combinatorial libraries and enables positive and negative selection logics using promoters with basal gene expression. Our system also minimizes the undesired co-evolution of phage genes as only the packaged phagemid is evolving and not the helper phage itself. In comparison to PACE, the protocol is performed in batch mode and therefore requires no special equipment for reactor assembly but instead relies on a daily researcher intervention during selections. Moreover, the batch process facilitates the performance of multiple selections in parallel, enabling the scalability of each individual selection and easy handling. Continuous culture evolution systems can suffer from 'phage washout' (loss of phage) when conditional phage production rates are not compatible with the flow rates. By contrast, batch modes are not as sensitive to loss of phage. On the other hand, dozens of rounds of reinfections occur in a single day of PACE whereas our system is currently limited to one round per overnight cycle. In addition, combinatorial libraries have to be designed and cloned because, unlike PACE, our system does not include a random mutagenesis plasmid\textsuperscript{35}. This means that structural information or a partial understanding of how a set of amino acid changes will affect the target protein’s activity is required to run our system.

Limitations of the phagemid-based system

The main limitation of the system is the combinatorial size of the library which is linked to transformation efficiency (\(10^5\)-\(10^{10}\) variants)\textsuperscript{36}. The selection process itself is not limited to a certain number of gene variants but it has to be noted that the use of larger libraries comes with the cost of prolonged experiment times. Another limitation can be the linkage of the target protein’s activity to conditional M13 phage replication as this depends on the individual protein’s characteristics. This is certainly more complicated for complex proteins such as membrane proteins than it is for cytosolic proteins. Furthermore, general limitations of...
bacterial expression over mammalian expression (e.g. protein solubility, disulfide bonds, posttranslational modifications) need to be considered for individual target proteins. For instance, our system would need to be adapted to enable the selection of proteins that require disulfide bonds for proper folding in bacterial cells.

**Experimental Design**

**Combinatorial library cloning on phagemid (PM).** Choosing which positions to randomize in the protein of interest is a critical step as this affects the library size, the cloning strategy, and ultimately the overall selection results. Small libraries with only one or two randomized positions can easily be obtained by round-the-world PCR whereas bigger libraries require overlap extension PCR or end-to-end ligation. Round-the-world PCR means in this context, that single base pair mutations are inserted into the target region by amplification of the whole plasmid DNA with randomized primers so that no additional step for plasmid ligation is required. For round-the-world PCR, both randomized primers must contain the mutations and bind to the same DNA sequence on opposite strands of the plasmid. Primers are generally 30-60 nucleotides long (N) and contain mutations in the middle of the randomized primers, flanked with 15-20 bases of correct sequence on both sides. These primers should ideally have a minimum GC content of 40% (%GC), end with one or more C or G bases and are purified by polyacrylamide gel electrophoresis (PAGE). The annealing region should have a melting temperature ($T_m$) of $ \geq 78°C$ using the following formula: $T_m = 81.5 + 0.41(\%GC) - 675/N - \%\text{mismatch}$. In this protocol, we focus on an overlap PCR approach prior to Gibson Assembly as this has been our method of choice for building λ $cI_{\text{opt}}$ libraries with a combinatorial space of $>10^6$ variants (Fig. 5a). These libraries are based on a λ $cI_{\text{opt}}$ mutant ($cI_{\text{opt}}$) with a strong activation region. The protocol presented here is optimized for the construction of combinatorial libraries using Gibson Assembly. It is our method of choice because it bypasses the need for restriction sites inside target genes which makes it much easier to construct sequence-targeted libraries. However, the selection system itself is compatible with any other library generation method as long as our phagemid vector backbone is used. The design of randomized oligonucleotides for overlap PCR is similar to conventional Gibson primer design. Briefly, PCR primers for insert amplification require a 15-25 bp overlap with each other, as well as a 15-25 bp overlap with the amplified PM vector backbone. Randomized positions should be avoided in the annealing regions and primers should ideally have a $T_m$ of 50-60°C using the following formula: $T_m = 4(G + C) + 2(A + T)$ (where A, C, G and T are the numbers of each base in the primer). The temperature difference of the primer pairs should be matched and lie within a 5°C range. The maximum insert size is limited by oligonucleotide synthesis (currently about 120 bp; desalted...
200 oligonucleotides are sufficiently pure). For evolving a novel protein, the user should ideally
201 start with a crystal structure of the target molecule (if available) and randomize positions
202 known to affect the desired activity (e.g. change positions of the binding interface in order to
203 alter the protein binding interaction). In other cases, biochemical information might also be
204 sufficient to guide library construction.
205
206 **Accessory plasmid (AP) design.** The conditional gene circuit that links an inducible input to
207 gVI expression has to be adapted to individual needs. This is achieved by replacing the λ
208 PRM promoter (pJPC12-ΔPS-PRM-B0034-gVI) with a different promoter or inducible input
209 depending on the desired application (Fig. 5b). The bidirectional promoter PR/PRM consists of
210 three operator sites (O1-O2-O3) where λ cl binding to O1-O2 leads to PRM activation41.
211 Counterselection via repression is achieved by putting a specific DNA sequence at operator
212 position O3 which is located between the -35 and -10 regions. For example, the O3 site of
213 the PRM promoter can be replaced with the consensus wild-type (WT) sequence called OCS.
214 Thus, binding of a clOPT library member to O1-O2 of an engineered PRM promoter activates
215 gene VI expression (and so promotes selection) while simultaneous binding to WT O3
216 represses gene VI, enabling counterselection against unwanted WT activity. Positive and
217 negative selections against the synthetic promoters PMLG6G and PMSST6 are depicted as
218 examples for this protocol (Supplementary Fig. 1). The engineered promoters are
219 designated according to the positions of the base substitutions in the consensus half-site of
220 O1 and O2.

221 **Reporter plasmid (RP) design.** This protocol describes the downstream functional
222 characterization of evolved TFs by fluorescence analysis. It has to be noted that a suitable
223 reporter assay needs to be adapted to the target protein’s properties according to the user’s
224 needs. To achieve this, the bidirectional PR/PRM promoter on the RP plasmid (pJPC12-ΔPS-
225 mCherry-PR/PRM-GFP) has to be replaced by the same inducible input used on the AP for
226 selection (Fig. 5c). The insertion of the bidirectional promoters P/PMLG6G and P/PMSST6 into
227 the RP is depicted as examples for this protocol (Supplementary Fig. 2). For other target
228 proteins, it might be sufficient to use one of the two reporters to analyse activity of the
229 selected proteins.
230
231 **Control selections.** Enrichment assays can be performed to test the efficiency of the
232 selection process. Mix plasmids containing λ clOPT (Addgene plasmid ID: 80852) and one of
233 the orthogonal cl variants (e.g. cl5G6G,P; Addgene plasmid ID: 80861) in different ratios (e.g.
234 10⁻³ and 10⁻⁶). Then transform these into TOP10 cells with the modified helper phage
235 M13KO7-ΔPS-Δgenelll-ΔgeneVI and the accessory plasmid pJPC12-ΔPS-PRM-B0034-
236 geneVI. This will allow the production of a phage stock packaged with clOPT and cl5G6G,P (steps
Use the obtained phage population and run a batch selection using the accessory plasmid pJPC12-ΔPS-PrM-B0034-geneVI (steps 41-61). Enrichment of λ cl<sub>opt</sub> can be monitored by infecting TG1 cells (containing the plasmid pJPC12-ΔPS-mCherry-PrM-GFP; Addgene plasmid ID: 80859) with the phage titer obtained after each round of selection. Streak out infected cells on agar plates supplemented with chloramphenicol and ampicillin and grow overnight at 37°C. The next day, analyse plates under the UV light of a gel documentation system. The ratio of green to red colonies should increase over time because the non-active TF cI<sub>G6G,P</sub> results in red colonies while the enriched active cl<sub>opt</sub> leads to green colonies due to GFP activation and mCherry repression. As an alternative control selection, the transcription factor cl<sub>G6G,P</sub> can be replaced by a reporter (e.g. a red fluorescent protein, RFP) on the PM and the selection process can be monitored by infecting TG1 cells and counting the ratio of red to white colonies after each round of selection<sup>1</sup>.

**MATERIALS**

**REAGENTS**

**Cloning and plasmid construction**

- Plasmids: M13KO7-ΔPS-ΔgeneIII-ΔgeneVI (Addgene plasmid ID: 80840), pLITMUS-rpoN-cl<sub>opt</sub>-J23106-geneIII (Addgene plasmid ID: 80852), pJPC12-ΔPS-PrM-B0034-geneVI (Addgene plasmid ID: 80858), optional: pJPC12-ΔPS-mCherry-PrM-GFP (Addgene plasmid ID: 80859) (Supplementary Fig. 3, Supplementary Table 1). Sequences of all plasmids are listed in Supplementary Data 1-5.

- Oligonucleotides (Sigma). Primers used for cloning are listed in Supplementary Table 2.

- KOD Hot Start DNA Polymerase (Merck Millipore, cat. no. 71086). PCR reaction components are listed in the Equipment Setup.

- Diethyl pyrocarbonate (DEPC)-treated and sterile filtered water (Sigma, cat. no. 95284)

- Gibson Assembly Master Mix (New England BioLabs, cat. no. E2611)

- DpnI endonuclease (New England BioLabs, cat. no. R0176)

- Super Optimal broth with Catabolite repression (S.O.C.) medium (Sigma, cat. no. 15544034)

- DNA Gel Loading Dye, 6× (Thermo Scientific, cat. no. R0611)

- 1 kb Plus DNA Ladder (Thermo Scientific, cat. no. 10787026)

- SYBR Safe DNA Gel Stain (Life Technologies, cat. no. S33102)

- Tris-borate-ethylenediaminetetraacetic acid (TBE) Buffer, 10× (Sigma, cat. no. T4415)

- Agarose for gel electrophoresis (Sigma, cat. no. A9539)
• QIAquick Gel Extraction Kit (Qiagen, cat. no. 28704)
• QIAquick PCR Purification Kit (Qiagen, cat. no. 28104)
• MinElute PCR Purification Kit (Qiagen, cat. no. 28004)
• QIAprep Spin Miniprep Kit (Qiagen, cat. no. 27104)
• HiSpeed Plasmid Maxi Kit (Qiagen, cat. no. 12663)

Strains, buffers and media
• One Shot Chemically Competent TOP10 E. coli (Fisher Scientific, cat. no. C404010)
• Mix & Go Competent Cells - Strain TG1 (Zymo Research, cat. no. T3017)
• 5-alpha Electrocompetent E. coli, optional (New England BioLabs, cat. no. C2989K)
• Lysogeny broth (LB) with agar (Sigma, cat. no. L2897)
• Ampicillin (Sigma, cat. no. A0166), chloramphenicol (Sigma, cat. no. C0378), kanamycin (Sigma, cat. no. K4000), carbenicillin disodium salt (Sigma, cat. no. C1389)
• 2× tryptone yeast extract (2×TY): NaCl (Sigma, cat. no. S9888), yeast extract (Sigma, cat. no. Y1625), tryptone (Sigma, cat. no. T7293)
• Glycerol (Sigma, cat. no. G5516)
• Ethanol (≥ 99.8%) for molecular biology (Merck Millipore, cat. no. 1085430250)
• M9 Minimal Salts, 5× (Sigma, cat. no. M6030)
• M9 plates: bacteriological agar (Sigma, cat. no. A5306), MgSO₄ (Sigma, cat. no. M7506), D-(+)-glucose (Sigma, cat. no. G8270), CaCl₂ (Sigma, cat. no. C1016), thiamine-HCl (Sigma, cat. no. T1270)

EQUIPMENT
• Polymerase chain reaction (PCR) tubes (VWR, cat. no. 732-0545)
• Microcentrifuge tubes (1.5 ml; Thermo Scientific, cat. no. 05-408-129)
• Conical centrifuge tubes, polypropylene, 15 ml (BD Falcon, cat. no. 352097)
• Conical centrifuge tubes, polypropylene, 50 ml (Corning, cat. no. 430829)
• Schott culture flasks, 250 ml (Sigma, cat. no. Z620033)
• Nunc Cryo Tubes (Thermo Scientific, cat. no. 366656)
• Serological pipettes (5 ml, 10 ml, and 25 ml; Fisher Scientific, cat. nos. 13-678-11D, 13-678-11E and 13-678-11)
• Sterile filters (0.22 µm pore size, Millex-GV, cat. no. SLGV033RS)
• L-shaped cell spreaders (Fisher Scientific, cat. no. 14-665-231)
• Cell culture centrifuge Avanti J-26XP (Beckman Coulter, cat. no. 393124)
• Microcentrifuge (Eppendorf, 5415D)
• Dri-block heater (Techne, DB100/2)
• Eppendorf Thermomixer Compact (Sigma, cat. no. T1317)
• Balance Sartorius Excellence (Sartorius)
• NanoDrop Lite Spectrophotometer (Thermo Scientific)
• Biophotometer (Eppendorf)
• Biophotometer cuvettes (Sigma, cat. no. Z605050)
• Horizontal gel electrophoresis systems (Bio-Rad)
• Gel documentation system (InGenius 3, Syngene)
• Gene Pulser Cuvette, 0.1 cm electrode (Bio-Rad, cat. no. 165-2089)
• Gene Pulser Xcell Microbial System (Bio-Rad, cat. no. 1652662)
• PCR thermocycler (Bio-Rad S1000, cat. no. 1852196)
• Petri dishes, 57 cm² (Sigma, cat. no. P7741)
• Mini Incubator (Labnet International, I5110A)
• Nunc Square BioAssay Dishes, 24.1 cm × 24.1 cm (Thermo Scientific, cat. no. 10570502)
• Shaking Incubator SI500 (Stuart)
• Cell culture microplate, 96 well, optional (Greiner Bio-One, cat. no. 655090)
• Infinite M200 plate reader, optional (Tecan)
• Research pipettes: 10 µl, 100 µl, 1000 µl (Sigma, cat. no. Z683884)
• Tips: 10 µl, 200 µl, 1000 µl (Starlab, cat. nos. S1111-3700-C, S1113-1700-C, S1111-6701-C)

REAGENT SETUP

Antibiotic stocks Prepare 100 mg ml⁻¹ ampicillin in H₂O (sterile filtered), 100 mg ml⁻¹ kanamycin in H₂O (sterile filtered) and 100 mg ml⁻¹ chloramphenicol in ethanol. Aliquot stocks into sterile 1.5 ml tubes and store at -20°C for up to 6 months. The final concentrations, if not stated otherwise, are 100 µg ml⁻¹ ampicillin (1:1000), 50 µg ml⁻¹ kanamycin (1:2000), and 25 µg ml⁻¹ chloramphenicol (1:4000).

Glycerol Prepare a sterile 10% (v/v) glycerol solution in H₂O for making electrocompetent cells. Glycerol stocks are obtained by preparing a sterile 50% (v/v) glycerol solution in H₂O and adding glycerol to the cell culture (f.c. 20% (v/v)) prior to freezing at -80°C. Store glycerol stock at 4°C for up to 3 months.

TBE electrophoresis buffer Dilute TBE buffer in distilled water to a 1× working solution and store at room temperature (15-25°C) for up to 6 months.
**Culture medium** Autoclave 2×TY medium (5 g l\(^{-1}\) NaCl, 10 g l\(^{-1}\) yeast extract, 16 g l\(^{-1}\) tryptone) and add antibiotics where appropriate before use. Store medium at 4°C for up to several months.

**LB plates** Add 35 g LB powder in 1 l water and autoclave. Add antibiotics where appropriate, pour into petri dishes and allow to solidify. Store plates at 4°C for up to several weeks. Note that antibiotics degrade over time which might affect the concentration when stored for prolonged times.

**M9 minimal medium plates** Autoclave 7 g bacteriological agar in 500 ml 1x M9 medium. Add 1 ml 1M MgSO\(_4\) (autoclaved), 5 ml 20% (w/v) D-(+)-glucose (sterile filtered), 50 µl 1M CaCl\(_2\) (autoclaved) and 500 µl 1M thiamine-HCl (sterile filtered) to M9 agar just before use. Add antibiotics where appropriate. Store plates at 4°C for up to several months.

**EQUIPMENT SETUP**

**PCR thermocycler** The PCR reaction components are listed below.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume [µl]</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>10× Buffer</td>
<td>5</td>
<td>1×</td>
</tr>
<tr>
<td>25 mM MgSO(_4)</td>
<td>3</td>
<td>1.5 mM</td>
</tr>
<tr>
<td>dNTPs (2 mM each)</td>
<td>5</td>
<td>0.2 mM (each)</td>
</tr>
<tr>
<td>H(_2)O</td>
<td>varies</td>
<td></td>
</tr>
<tr>
<td>Forward primer (5 µM)</td>
<td>3</td>
<td>0.3 µM</td>
</tr>
<tr>
<td>Reverse primer (5 µM)</td>
<td>3</td>
<td>0.3 µM</td>
</tr>
<tr>
<td>Template DNA</td>
<td>varies</td>
<td>0.02-0.2 ng/µl</td>
</tr>
<tr>
<td>KOD DNA Polymerase (1 Unit/µl)</td>
<td>1</td>
<td>0.02 U/µl</td>
</tr>
<tr>
<td><strong>Total reaction volume</strong></td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

**CRITICAL:** For targets greater than 2 kb, final Mg\(^{2+}\) concentrations are adjusted to 2mM.

The following conditions are used for all PCR reactions:

<table>
<thead>
<tr>
<th>Step</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Polymerase activation</td>
<td>95°C, 2 min</td>
</tr>
<tr>
<td>2. Denature</td>
<td>95°C, 30 s</td>
</tr>
<tr>
<td>3. Annealing</td>
<td>temperature varies, 30 s</td>
</tr>
<tr>
<td>4. Extension</td>
<td>70°C, time varies</td>
</tr>
<tr>
<td>Repeat steps 2.-4.</td>
<td>Number of cycles vary</td>
</tr>
<tr>
<td>Final extension</td>
<td>70°C, 10 min</td>
</tr>
<tr>
<td>Infinite hold</td>
<td>4°C</td>
</tr>
</tbody>
</table>

See Supplementary Table 3 for PCR conditions specific for individual reactions.
Infinite M200 plate reader Temperature: 37°C, duration: 10 h, shaking: 281 r.p.m., absorbance: 600 nm +/- 9 nm, fluorescence mCherry: excitation 585 nm +/- 9 nm; emission 625 nm +/- 20 nm; gain value 70, fluorescence GFP: excitation 485 nm +/- 9 nm; emission 520 nm +/- 20 nm; gain value 40.

PROCEDURE

Phagemid construction by Gibson Assembly • TIMING 2 weeks

1. Design and order generic forward and reverse primers (e.g. pLITMUS-F and pLITMUS-R; Supplementary Table 2) for the amplification of the PM vector backbone (pLITMUS-rpoN-clopt-J23106-genIII) upstream of clopt and downstream the terminator BBa_B0015. Note that the terminator BBa_B0015 occurs twice in the parental plasmid. The "medium strength" rpoN promoter is used to express the evolving gene to achieve a balance between functional expression and any potential metabolic load. The levels of the expressed target gene may need to be adjusted to the function in other cases.

2. Design and order user-specific primers (e.g. cl-F and cl-R; Supplementary Table 2) for the gene of interest plus terminator of choice (e.g. BBa_B0015) with a 15-25 bp overlap to the PM vector backbone.

3. Amplify the gene of interest and vector backbone by PCR (See Equipment Setup) and purify the samples using the QIAquick PCR Purification Kit. ▲ CRITICAL STEP If the PCR reactions contain unwanted by-products, gel extraction should be performed throughout the protocol using the QIAquick Gel Extraction Kit. Use a DNA polymerase with proof-reading activity (e.g. KOD DNA Polymerase) for all PCR reactions throughout the protocol.

4. Remove the parental plasmid by adding 1 µl DpnI per 50 µl PCR reaction product and incubating for 1-2 h at 37°C and 400 r.p.m. (Thermomixer Compact).

5. Fuse the two fragments by Gibson Assembly according to the manufacturer’s instructions. Note that Gibson reactions can be downscaled to 5 µl per reaction.

6. Dilute the assembled products 4-fold with H2O, add 2 µl of the diluted product to 50 µl chemically competent Top10 cells and transform the cells according to the manufacturer’s instructions.

7. Incubate the cells for 1 h at 37°C, 220 r.p.m. (incubator SI500) and spread them onto LB plates supplemented with 100 µg ml⁻¹ ampicillin.

8. Grow the cells overnight at 37°C.

9. The next day, pick single colonies and grow them in 5 ml 2×TY supplemented with 100 µg ml⁻¹ ampicillin overnight at 37°C and 220 r.p.m. (incubator SI500).
10. Extract the phagemid DNA (QIAprep Spin Miniprep Kit) according to the manufacturer’s instructions and confirm the nucleotide sequences by DNA sequencing using the primers pLITMUS-F and pLITMUS-R (Table 1).

Combinatorial library cloning on phagemids • TIMING 2 weeks

11. Design and order user-specific forward and reverse primers for the amplification of the PM vector backbone (e.g. pLITMUS-Lib-F and pLITMUS-Lib-R; Supplementary Table 2) and the insertion of the randomized target sequence (e.g. Library 1-F, Library 1-R; Supplementary Table 2). PCR primers for insert amplification require a 15-25 bp overlap with each other as well as a 15-25 bp overlap with the amplified vector backbone. ▲CRITICAL STEP Avoid randomized library positions within the primer overlap regions.

12. Amplify the PCR fragments (see Equipment Setup) and purify the samples using the QIAquick PCR Purification Kit or the MinElute PCR Purification Kit (for samples <100 bp). ▲CRITICAL STEP The PCR product concentration affects the efficiency of the assembly reaction. Optimized cloning efficiency requires at least 20 ng µl⁻¹ of the PM vector backbone.

13. Add 1 µl DpnI per 50 µl PCR reaction product and incubate for 1-2 h at 37°C and 400 r.p.m. (Thermomixer Compact).

14. Fuse the DNA fragments by Gibson Assembly³⁸. Upscale Gibson reactions (e.g. 4 × 20 µl) to increase the total plasmid concentration.

15. Pool the Gibson reactions, purify the assembled plasmid using the QIAquick PCR Purification Kit and elute in 30 µl H₂O. ▲CRITICAL STEP Note that purification is important to decrease the salt concentration and to decrease the Gibson reaction components as these are toxic to the cells at high concentrations.

16. Measure the plasmid concentration with a spectrophotometer (NanoDrop Lite). ▲CRITICAL STEP DNA concentrations should be >10 ng µl⁻¹ for high transformation efficiency.

17. Transform 1-2 µl of DNA into 50 µl electrocompetent cells (DH5-alpha or TG1) and add 950 µl S.O.C medium. ▲CRITICAL STEP Use electroporation as the method of choice for transformation as it allows much larger library sizes.

18. Incubate for 1 h at 37°C, 220 r.p.m. (incubator SI500).

19. Plate the transformation reaction on Nunc Square BioAssay Dishes (24.1 cm × 24.1 cm) supplemented with 100 µg ml⁻¹ ampicillin and incubate overnight at 37°C.

20. The next day, harvest the cells with a cell spreader. ▲CRITICAL STEP Only use plates with more than 10⁵ clones. The transformation efficiency should be estimated by plating
serial dilutions ($10^{-2}$ and $10^{-4}$ in 2×TY) on additional petri dishes (57 cm$^2$) supplemented with 100 µg ml$^{-1}$ ampicillin and colony counting the following day. Ideally, to cover the whole library space, at least a 3-fold excess of colonies relative to the theoretical library size is desired. **TROUBLESHOOTING**

21. Purify the combinatorial DNA library using the HiSpeed Plasmid Maxi Kit and elute in 0.5 ml TE buffer. Measure the plasmid concentration with a spectrophotometer (NanoDrop Lite). The obtained plasmid concentration should ideally be >50 ng ml$^{-1}$.

22. Pick individual colonies (10-100 clones of a library depending on the library size and quality control desired) from the petri dishes which were used to estimate the transformation efficiency (see Step 20) and culture each in 5 ml 2×TY supplemented with 100 µg ml$^{-1}$ ampicillin overnight at 37°C, 220 r.p.m. (incubator SI500). The next day, extract phagemid DNA (QIAprep Spin Miniprep Kit) and sequence the gene of interest using the primers pLITMUS-F and/or pLITMUS-R to confirm library diversity (Table 1). **TROUBLESHOOTING**

**Production of M13 phage from a combinatorial phagemid library • TIMING 1 week**

23. Transform 50 µl chemically competent Top10 cells with equal moles of HP (M13KO7-ΔPS-ΔgeneIII-ΔgeneVI) and AP (pJPC12-ΔPS-P$_{RM}$-B0034-geneVI) (10-20 fmol per plasmid, typically 1-2 µl in total). Note that the P$_{RM}$ promoter can be replaced by an alternative promoter (e.g. T7) to obtain higher phage titers in the absence of the activator λ cl.

24. Add 250 µl S.O.C. medium to the samples and incubate for 1 h at 37°C, 220 r.p.m. (incubator SI500).

25. Spread the cells on LB plates supplemented with 25 µg ml$^{-1}$ chloramphenicol and 50 µg ml$^{-1}$ kanamycin. Grow overnight at 37°C.

26. The next day, pick a single colony, grow in 2×TY supplemented with 12.5 µg ml$^{-1}$ chloramphenicol and 25 µg ml$^{-1}$ kanamycin at 37°C, 250 r.p.m. (incubator SI500) until the OD$_{600}$ reaches 0.4-0.6 (mid-exponential phase) and make cells electrocompetent as described in Gonzales *et al.*

■ PAUSE POINT Stored electrocompetent cells can be used for the construction of any phage library.

27. Transfer 50 µl of the electrocompetent cells to a prechilled 1.5 ml tube on ice and add 1-2 µl of the cloned combinatorial phagemid library.

28. Electroporate cells, add immediately 950 µl S.O.C. medium and incubate for 1 h at 37°C, 220 r.p.m. (incubator SI500).
29. Estimate the actual phage library by colony counting of serial dilutions (10^{-2} and 10^{-4} in 2×TY) on LB plates supplemented with ampicillin (see step 20). ▲ CRITICAL STEP Make sure not to lose any library members through low transformation efficiencies.

? TROUBLESHOOTING

30. Add 3 ml 2×TY supplemented with 12.5 µg ml^{-1} chloramphenicol, 25 µg ml^{-1} kanamycin, and 50 µg ml^{-1} ampicillin to the transformation reaction and grow for 18-20 h at 30°C, 250 r.p.m. (incubator SI500). Note that the volume can be adjusted depending on the desired volume of the phage titer.

31. The next day, centrifuge sample for 5 min at 5,000 g.

32. Sterile-filter the phage supernatant (0.22 µm pore size). ▲ PAUSE POINT The phage library can be kept at 4°C for short-term storage (weeks) or at -20°C for the long-term (years).

Phage titer analysis • TIMING 3 d

33. Streak out TG1 cells from a glycerol stock (~1-5 µl) on an M9 minimal medium plate and incubate overnight at 37°C. Note that TG1 plates can be used for a minimum of two weeks when stored at 4°C. ▲ CRITICAL STEP Use M9 minimal medium plates to select F-pilus positive TG1 cells.

34. The next day, pick 1-4 single isolated colonies from the M9 minimal medium plate and inoculate in 10 ml 2×TY medium in a 50 ml conical centrifuge tube.

35. Incubate at 37°C and 250 r.p.m. (incubator SI500) until the OD_{600} reaches 0.4-0.6 (mid-exponential phase). It typically takes 4-6 hours for the culture to reach the desired OD_{600}.

▲ CRITICAL STEP Do not let the cells grow into stationary phase as TG1 cells tend to lose the F' episome and this lowers the overall infection rate.

36. In the meantime, prepare serial dilutions (10^{-2}, 10^{-4}, 10^{-6}, 10^{-8} in 2×TY) of the phage library in sterile 1.5 ml tubes. Phage stocks are diluted before infection to ensure that each cell is only infected by one phage particle (number of colonies on plates equal number of phage particles).

37. Add 100 µl of the phage dilutions to 900 µl TG1 cells in a sterile microcentrifuge tube. Mix gently and incubate the samples for 1 h at 37°C with no shaking. Plate 100 µl cell suspension on prewarmed LB plates supplemented with 100 µg ml^{-1} ampicillin and incubate overnight at 37°C.

38. The next day, count the number of colonies and calculate the phage titer (Equation 1). Ideally, use the plates containing 20 to 400 colonies. Note that the 100-fold dilution (step 37) has to be taken into consideration.

\[
\text{Phage titer per ml} = \text{dilution factor} \times 100 \times \text{number of colonies on plate} \quad (\text{Equation 1})
\]
The phage titer should lie between $10^8$-$10^{13}$ colony-forming units (cfu) per ml.

**Construction of accessory and reporter plasmids • TIMING 2 weeks**

39. Order user-specific forward and reverse primers to replace the $P_{RM}$ promoter on the AP plasmid (pJPC12-ΔPS-PRM-B0034-geneVI) with a different promoter or inducible input. For vector amplification, use primers B0034-gVI-F and gVI-R that bind upstream and downstream of $P_{RM}$ (Supplementary Table 2). For insert amplification make sure to add a 15-25 bp overlap for assembly. Optional: Clone the same inducible input into pJPC12-ΔPS-mCherry-P/RPRM-GFP to obtain a reporter for the functional characterization of selected proteins. Use primers GFP-F and mCherry-R for vector amplification (Supplementary Table 2).

40. Clone the accessory and reporter plasmids as described in steps 3-10. Use sequencing primers pJPC12-F and/or pJPC12-R for gene VI constructs and pJPC12-F2 for reporters (Table 1).

**Preparation of host cells for directed evolution • TIMING 3 d**

41. Transform 50 µl competent TG1 cells with equal moles of HP M13KO7-ΔPS-ΔgeneIII-ΔgeneVI and the cloned accessory plasmid (AP) (10-20 fmol per plasmid, typically 1-2 µl in total). ▲CRITICAL STEP Always use an E. coli strain that contains the F-factor needed for M13 phage infection.

42. Add 250 µl S.O.C. medium to the sample and incubate for 1 h at 37°C, 220 r.p.m. (incubator SI500).

43. Spread the cells on LB plates supplemented with 25 µg ml$^{-1}$ chloramphenicol and 50 µg ml$^{-1}$ kanamycin and grow cells overnight at 37°C.

44. The next day, pick a single colony and grow cells in 2 ml 2×TY supplemented with kanamycin and chloramphenicol for 4-6 h at 37°C, 250 r.p.m. (incubator SI500) until the cells reach the late-exponential phase.

45. (Optional) Make glycerol stock and store at -80°C (see Reagent Setup).

46. Make serial dilutions of the cell suspension (e.g. $10^{-6}$ or $10^{-8}$ in M9 medium) and spread the diluted cells on an M9 minimal medium plate supplemented with 25 µg ml$^{-1}$ chloramphenicol and 50 µg ml$^{-1}$ kanamycin. These conditions promote phage infectability by maintaining F’ pili.

47. Incubate plates for 30-48 h at 37°C. Note that bacteria grow much slower on minimal media than on rich media. This plate is used as a source of fresh colonies for selection experiments and can be used for up to two weeks when stored at 4°C.
Phage-assisted batch selection • TIMING 2 weeks

48. Inoculate 10-20 ml 2×TY containing 12.5 µg ml\(^{-1}\) chloramphenicol and 12.5 µg ml\(^{-1}\) kanamycin with 1-4 colonies from prepared M9 plate (step 47) in a 50 ml tube (Fig. 6).

49. Grow the starter culture for 6-8 h at 37°C and 250 r.p.m. (incubator SI500) until the OD\(_{600}\) reaches 0.4-0.6.

50. Infect 10 ml of the starter culture with the combinatorial phage library at a multiplicity of infection (MOI) of 0.5-5. An excess of cell culture can be chilled on ice and then be stored at 4°C for up to one week. This culture may be used for the next rounds of selection. Note that the selection volume can be easily up- or downscaled according to the user’s need.

51. Incubate the infected cells at 37°C without stirring for 5 min.

52. Incubate the sample for 18-20 h at 30°C and 250 r.p.m. (incubator SI500).

53. The next day, centrifuge the culture for 5 min at 5,000 g and transfer 1 ml of the supernatant into a sterile microfuge tube. This sample is used to start a new round of selection.

■ PAUSE POINT Phage supernatants for each round of selection can be stored at 4°C for short-term storage, or at -20°C for the long-term, to continue selection at a later time.

54. Infect the starter culture (step 49) at a ratio of 10\(^3\)-10\(^{-1}\) (e.g. 10-1000 µl phage supernatant in 10 ml culture) for the next round of selection.

55. Run selection cycle (steps 51-54) for several rounds until the target protein(s) are enriched. This usually takes four to eight rounds depending on the target protein’s activity and thus the conditional gene VI expression. ▲CRITICAL STEP The phage titer should ideally stay between 10\(^6\)-10\(^{12}\) cfu ml\(^{-1}\) after each round of the selection (see step 56). Very high infection rates (MOI >10) lead to multiple infections and thus propagation of non-functional library members (‘cheaters’) whereas very low rates (MOI <0.1) decrease the performance of the system. ? TROUBLESHOOTING

56. (Optional) During the selection process, monitor the phage titer for each round by phage titer analysis (steps 33-38). ? TROUBLESHOOTING

57. (Optional) Monitor the selection process by infecting reporter cells (TG1 with a suitable reporter plasmid, e.g. pJPC12-ΔPS-mCherry-P/P\(_{M.5G6G-GFP}\)) with the obtained phage titer for each round (analogous to steps 33-38). Streak out infected cells on LB plates supplemented with 25 µg ml\(^{-1}\) chloramphenicol and 100 µg ml\(^{-1}\) ampicillin and grow overnight at 37°C. The next day, analyse the plates under the UV light of a gel documentation system. Non-active library member result in red colonies while active library members lead to green colonies due to GFP activation and mCherry repression. Store plates at 4°C overnight for improved mCherry signals.
58. After selection, sterile-filter the phage supernatant (0.22 µm pore size) and serial dilute the sample with 2×TY medium before infecting TG1 cells with an OD600 of 0.4-0.6. Incubate the infected cells for 1 h at 37°C before plating (see step 37).

59. Select infected cells on 100 µg ml⁻¹ ampicillin plates overnight at 37°C.

60. The next day, pick at least three colonies per selection and grow each colony in 5 ml 2×TY supplemented with ampicillin overnight at 37°C and 250 r.p.m. (incubator SI500).

61. The next day, extract phagemid DNA (QIAprep Spin Miniprep Kit) and sequence the gene of interest using the primers pLITMUS-F and/or pLITMUS-R (Table 1).

**Characterization of evolved proteins (optional) • TIMING 3 d**

62. Transform 50 µl competent TG1 cells with equal moles of a selected phagemid and a suitable reporter plasmid (e.g. pJPC12-ΔPS-mCherry-P/P_{M,5G6G}-GFP) (10-20 fmol per plasmid, typically 1-2 µl in total). Transform the reporter plasmid into TG1 cells and use as a control. (Optional) Delete the expression cassette rpoN-clOpt-B0015 from the phagemid (e.g. pLITMUS-ΔclOpt-F, pLITMUS-ΔclOpt-R) and transform the obtained plasmid (pLITMUS-J23106-genIII) together with the reporter to compensate for growth effects between control and selected phagemids.

63. Spread the cells on LB plates supplemented with 25 µg ml⁻¹ chloramphenicol and 100 µg ml⁻¹ ampicillin and grow the cells overnight at 37°C.

64. The next day, pick single colonies and grow in 1 ml 2×TY supplemented with 5 µg ml⁻¹ chloramphenicol and 5 µg ml⁻¹ carbenicillin for 4-6 h at 37°C, 250 r.p.m. (incubator SI500). Analyse at least three replicates per transformation.

65. Measure OD600 of each replicate (150 µl) using the Tecan Infinite M200 plate reader.

66. Dilute cultures in 2×TY supplemented with 5 µg ml⁻¹ chloramphenicol and 5 µg ml⁻¹ carbenicillin to a final OD600 of 0.01 (150 µl) in a 96-well microplate.

67. Measure the absorbance at 600 nm, green fluorescence (excitation: 485 nm, emission: 520 nm), and red fluorescence (excitation: 585 nm, emission: 625 nm) every 10 min with the Infinite M200 plate reader (37°C, shaking between readings) until the cells reach stationary phase.

68. For data analysis use fluorescence readings in the mid-exponential phase (OD600 of 0.2) and correct absorbance and fluorescence against readings of a TG1 culture. Normalize the fluorescence for the number of cells by dividing by the absorbance.

**TROUBLESHOOTING**

Troubleshooting advice can be found in Table 2.
• TIMING

Phagemid construction by Gibson Assembly
Steps 1-2, design of primers and oligo synthesis by supplier: 1 week
Steps 3-10, cloning of phagemid: 1 week

Combinatorial library cloning on phagemids
Step 11, design of primers and oligo synthesis by supplier: 1 week
Steps 12-22, cloning of combinatorial library: 1 week

Production of M13 phage from a combinatorial phagemid library
Steps 23-32, transfer from plasmid library to phage library: 1 week

Phage titer analysis
Steps 33-38, analysis of phage concentration: 3 d

Construction of accessory and reporter plasmids (can be done in parallel with phage library cloning)
Step 39, design of primers and oligo synthesis by supplier: 1 week
Step 40, cloning of accessory plasmid and reporter plasmid: 1 week

Preparation of host cells for directed evolution (can be done in parallel after successful AP cloning)
Steps 41-47, transformation and plating of cells: 3 d

Phage-assisted batch selection
Steps 48-57, batch selections: 1 week
Steps 58-61, extraction and sequencing of selected genes: 1 week

Characterization of evolved proteins (optional)
Steps 62-68, functional characterization by reporter assay: 3 d

ANTICIPATED RESULTS

The first section of this protocol describes the construction of combinatorial libraries used for subsequent directed evolution experiments. As examples, we describe the construction of two cLopt libraries which contain five randomized positions: Library 1 (45S, 46G, 47V, 48G, 55N); Library 2 (45S, 46G, 48G, 49A, 55N). Quality control sequencing of 10-100 clones of a library may be performed to confirm diversity, depending on the library size and quality control desired. For example, ten individual clones of a constructed library should ideally result in ten different variants (Table 3).

The second section illustrates the directed evolution of proteins based on conditional M13 phage propagation. Libraries 1 and 2 are selected against engineered promoters for six to eight rounds leading to enrichment of TFs with binding activations against their novel
promoters (Table 4). We frequently obtain amino acid substitutions that occur spontaneously at certain positions not covered by the combinatorial space of the library. These mutations can originate either from mutations during library cloning or from the spontaneous error rate of M13 phage replication which is ~0.0046 mutation rate per genome per replication\(^4\). Such mutations can provide function\(^1\) and contribute towards directed evolution.

The last section of the protocol describes the characterization of selected TFs. The reporter assay is designed in a way that TF binding to the bidirectional promoter results in GFP activation and mCherry repression. For baseline comparison, GFP and mCherry expression is measured for each promoter in the absence of a TF. The evolved TF variants enable simultaneous activation and repression against their engineered bidirectional promoters. For the selected cl variant (cl\(_{5G6G,P}\)) against the bidirectional promoter P/P\(_{M,5G6G}\), GFP production is upregulated 10-fold and 94\% of mCherry is repressed (Fig. 7a,b). The evolved cl variant (cl\(_{5T6T,P}\)) against the bidirectional promoter P/P\(_{M,5T6T}\) results in a 9-fold activation and 98\% mCherry repression (Fig. 7c,d). This protocol further shows a method to analyse cross-reactivities for DNA-binding proteins. WT cl and cl\(_{opt}\) activate GFP 6-fold and 9-fold and simultaneously repress 90\% and 82\% of mCherry production on the WT P\(_R/P_{RM}\) promoter whereas this effect is not observed for any of the engineered promoter variants (Fig. 7e,f). The selected TFs also do not cross-react with each other, thus ensuring orthogonality.

**AUTHOR CONTRIBUTIONS**

AKB, AJ and MI developed the protocol. AKB performed the experiments. AKB and MI wrote the manuscript. MI and AJ supervised the project and contributed reagents, materials and analysis tools.

**ACKNOWLEDGMENTS**

This research was supported by the European Commission grant FP7-ICT-2013-10 (no 610730, EVOPROG). AJ was funded by FP7-KBBE (no 613745, PROMYS), H2020 Marie Sklodowska-Curie (no 642738, MetaRNA) and EPSRC-BBSRC (no BB/M017982/1, WISB center). MI is funded by New Investigator award no WT102944 from the Wellcome Trust U.K. The authors like to thank Tina Bartels and Marko Storch for their critical reading of the manuscript.

**COMPETING FINANCIAL INTERESTS**

The authors declare that they have no competing financial interests.
REFERENCES


Intracellular directed evolution of proteins from combinatorial libraries based on conditional phage replication.

1) Phagemid library generation: A combinatorial DNA library is generated from the target gene on a phagemid (PM), which also contains conditionally expressed M13 gene III and the M13 packaging signal. The DNA library members are then packaged into phage particles which are the starting point for selection.

2) Selection cycles: TG1 cells containing a modified helper phage (HP) and an accessory plasmid (AP) are infected with the constructed phage library. The HP provides all that is required for phage propagation, except for two essential genes (gIII and gVI). The AP contains a conditional gene circuit that links the target protein’s activity to conditional phage production (gene VI expression). Enrichment for a particular protein function occurs after several rounds of selection.

3) Gene isolation: Cells are infected with selected phages and the phagemid DNA is amplified and purified.

4) Protein characterization: The target protein’s activity needs to be analysed using a suitable reporter assay.
Figure 2. Flowchart and timeline of the directed evolution protocol. All major steps for design, cloning, selection and functional characterization are depicted. It takes ~7 weeks to select a candidate protein from a constructed combinatorial library.
Figure 3. Directed evolution of proteins from combinatorial libraries. (a) Plasmids needed to set up the phagemid-based selection system. The modified helper phage HP (M13KO7-ΔPS-ΔIII-ΔVI) contains the kanamycin resistance gene (KanR) and all phage genes required for phage replication except the genes III and VI. The weak packaging signal (PS) is removed to prevent helper phage propagation as a phage. The accessory plasmid AP contains the chloramphenicol resistance gene (CamR) and a conditional gene VI expression circuit, induced by an active library member on the phagemid (PM). The PM also provides the ampicillin resistance gene (AmpR), the M13 packaging signal (PS; to allow DNA packaging in phage) as well as constitutively expressed gene III. (b) Scheme of the intracellular selection process. An active library member on the packaged PM induces gene VI expression to complete the phage life cycle, thus enriching this variant over time.
Figure 4. Linkage of an evolving protein’s activity to conditional M13 phage propagation. (a) An evolving transcription factor (e.g. λ cI) activates gene VI expression downstream of a specific promoter (e.g. λ P_RM). This example is depicted in this protocol by selecting new activators against engineered synthetic promoters. (b) An evolving RNA polymerase (grey) enables transcription and hence gene VI expression. (c) An evolving DNA-binding protein (red) derived from genome-editing systems (transcription activator-like effector nucleases; TALENs) is linked to the ω subunit of bacterial RNA polymerase III (grey). Binding to a target DNA sequence (dark blue) upstream of a minimal lac promoter (black) induces transcription of gene VI. (d) The target protein (dark blue) is bound to the DNA upstream the promoter P_λacZ Opt (black) via a fused DNA-binding domain (orange) and the RNA polymerase omega subunit (RpoZ, yellow) is fused to the evolving protein (red). Target protein binding of the evolving protein enables the localization of RNA polymerase upstream gene VI, initiating gene expression from the P_λacZ Opt promoter. (e) The T7 polymerase (grey) is inhibited when bound to T7 lysozyme (dark blue) as it inhibits transcription initiation and the transition from initiation to elongation. Proteolysis of the target cleavage site (red) by an evolving protease activates the T7 RNA polymerase enabling gene VI expression downstream of the T7 promoter. Gene VI is annotated with an asterisk where originally conditional gene III was used instead of gene VI using PACE. Gene III and gene VI are both minor coat proteins, each present in 3-5 copies per phage particle.
Figure 5. Experimental design and cloning strategy. (a) Combinatorial library construction on phagemids (PM). Randomized oligonucleotides require a 15-25 bp overlap with each other as well as a 15-25 bp overlap with the amplified vector backbone and are fused by PCR. Randomized positions are marked with an ‘x’ and must be avoided within overlap regions. Primers 3 and 4 bind upstream and downstream of the randomized target region and are used for vector linearization. (b) Accessory plasmid (AP) design. The conditional gene circuit that links an inducible input to gVI expression has to be adapted for individual needs. This is achieved by replacing the $\lambda P_{RM}$ promoter with a different promoter or inducible input depending on the desired application. For example, an engineered promoter (e.g. $P_{M,5G6G}$) is constructed by overlap extension PCR and inserted into the linearized fragment by Gibson Assembly. Primers 3 and 4 bind upstream and downstream of $\lambda P_{RM}$ and are used to remove the $P_{RM}$ promoter. (c) Reporter plasmid (RP) design. The bidirectional $\lambda P_R/P_{RM}$ is replaced by the same inducible input used on the AP (e.g. $P/P_{M,5G6G}$). The fluorescent proteins mCherry and GFP on the RP are used to characterize the activity of the selected proteins on the PM. Note that the maximum insert size using overlap extension PCR is limited by oligonucleotide synthesis (currently about 120 bp).
**Figure 6. Phage-assisted batch selection.** A starter culture from TG1 cells containing the modified M13 helper phage HP (M13KO7-ΔPS-ΔgIII-ΔgVI) and an accessory plasmid AP is prepared and cells are grown for 6-8 h at 37°C until the OD$_{600}$ reaches 0.4-0.6. Starter cells are infected with the constructed phagemid library to start the first round of selection. Conditional phage production is performed in a shaking incubator for 16-18 h at 30°C and the resulting phage particles are separated from the cells by centrifugation. The obtained phage stock is used to start a new round of selection by infecting a fresh starter culture (= Round 2). After several rounds of reinfection and selection, a TG1 preculture is infected with the obtained phage stock and infected TG1 cells are selected on LB plates supplemented with ampicillin. Single colonies are picked and cells are grown overnight at 37°C in a shaking incubator. The next day, phagemid DNA is purified and the gene of interest is sequenced.
Figure 7. Dual activation and repression of engineered bidirectional $\lambda$ P/P\textsubscript{M} promoters by selected cI variants. The activity of the selected TFs needs to be verified by a reporter assay. (a,b) Basal promoter strength of the bidirectional promoter P/P\textsubscript{M,5G6G} and its dual activation and repression by the selected TF cI\textsubscript{5G6G,P}. (c,d) Basal promoter strength of the bidirectional promoter P/P\textsubscript{M,5T6T} and its dual activation and repression by the selected TF cI\textsubscript{5T6T,P}. (e,f) Basal promoter strength of P\textsubscript{R/P\textsubscript{RM}} and its activation-repression by WT $\lambda$ cI and cI\textsubscript{opt}. Cross-reactivity of TF variants is ruled out by reporter analysis. Basal mCherry expression varies between promoters due to base pair substitutions next to the -35 and -10 regions and y-axes are adjusted accordingly. GFP and mCherry expression was normalized to OD\textsubscript{600}. Four biological replicates were measured for each sample and each replicate is represented by one dot. Error bars show 1 standard deviation.
### TABLE 1. Oligonucleotides used for sequencing.

<table>
<thead>
<tr>
<th>Name</th>
<th>Oligonucleotide sequence</th>
<th>Step</th>
</tr>
</thead>
<tbody>
<tr>
<td>pLITMUS-F</td>
<td>5' GTC GAT TTT TGT GAT GCT CG 3'</td>
<td>10, 22, 61</td>
</tr>
<tr>
<td>pLITMUS-R</td>
<td>5' GGG TTA TTG TCT CAT GAG CGG ATA C 3'</td>
<td>10, 22, 61</td>
</tr>
<tr>
<td>pJPC12-F</td>
<td>5' AAA CGA CGG CCA GTG AGC 3'</td>
<td>40</td>
</tr>
<tr>
<td>pJPC12-F2</td>
<td>5' AGC CGT ACA TGA ACT GAG 3'</td>
<td>40</td>
</tr>
<tr>
<td>pJPC12-R</td>
<td>5' GAT AAC AAT TTC ACA CAG G 3'</td>
<td>40</td>
</tr>
</tbody>
</table>

### TABLE 2. Troubleshooting table.

<table>
<thead>
<tr>
<th>Step</th>
<th>Problem</th>
<th>Possible reason</th>
<th>Solution</th>
</tr>
</thead>
</table>
| 20   | The transformation efficiency is too low          | The DNA concentration of the cloned library too low    | - Harvest and pool several transformation reactions to increase the practical library size;  
|      |                                                  |                                                        | - Optimize the cloning procedure (e.g. PCR reactions) to increase library concentration;  
|      |                                                  |                                                        | - Check cell competency                                                  |
| 22   | The library contains the parental plasmid         | DpnI digestion was incomplete                          | - Gel extract the PCR-amplified pLITMUS vector backbone;  
|      |                                                  |                                                        | - Increase DpnI incubation time to ensure complete digestion             |
| 29   | The transformation efficiency is too low          | The competency of prepared cells is insufficient        | - Harvest and pool several transformation reactions to increase library size;  
|      |                                                  |                                                        | - Optimize the procedure for making competent cells (e.g. do not freeze cells before transformation) |
| 55   | A large number of phages is lost during selection | The infection rate is too low                          | - Increase the volume of supernatant for infection;  
|      |                                                  |                                                        | - Phage enrichment via polyethylene glycol (PEG) precipitation is generally not needed but can be performed to increase the phage titer for the next round of selection |
| 55   | The phage concentration is too high               | The infection rate is too high                         | Decrease the volume of supernatant to lower the MOI                     |
| 56   | No enrichment of target proteins                  | The phage library does not contain functional library members | - Redesign and reconstruct the combinatorial library;  
|      |                                                  |                                                        | - Check with a positive WT control diluted in non-functional phage (see control selections in Experimental design) |
TABLE 3. Sequencing results of a combinatorial cl_{opt} library. 10 clones are shown for illustration, below. Typically 10-100 clones of a library (e.g. Library 1: 45S, 46G, 47V, 48G, 55N) may be sequenced to confirm diversity, depending on the library size and quality control desired. The obtained base pairs at the randomized NNS motifs were translated into their corresponding amino acids. Wild-type amino acids are highlighted in blue. The library contains five randomized amino acid positions known to contact promoter DNA\textsuperscript{45-47}. This results in a combinatorial space of 3.2×10\textsuperscript{6} variants.

<table>
<thead>
<tr>
<th>Position</th>
<th>45</th>
<th>46</th>
<th>47</th>
<th>48</th>
<th>55</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clone 1</td>
<td>F</td>
<td>T</td>
<td>E</td>
<td>F</td>
<td>N</td>
</tr>
<tr>
<td>Clone 2</td>
<td>P</td>
<td>A</td>
<td>C</td>
<td>F</td>
<td>R</td>
</tr>
<tr>
<td>Clone 3</td>
<td>F</td>
<td>N</td>
<td>P</td>
<td>V</td>
<td>L</td>
</tr>
<tr>
<td>Clone 4</td>
<td>F</td>
<td>Y</td>
<td>L</td>
<td>S</td>
<td>M</td>
</tr>
<tr>
<td>Clone 5</td>
<td>G</td>
<td>C</td>
<td>L</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>Clone 6</td>
<td>I</td>
<td>P</td>
<td>M</td>
<td>P</td>
<td>T</td>
</tr>
<tr>
<td>Clone 7</td>
<td>F</td>
<td>N</td>
<td>P</td>
<td>F</td>
<td>N</td>
</tr>
<tr>
<td>Clone 8</td>
<td>S</td>
<td>I</td>
<td>G</td>
<td>L</td>
<td>Y</td>
</tr>
<tr>
<td>Clone 9</td>
<td>K</td>
<td>I</td>
<td>I</td>
<td>Y</td>
<td>L</td>
</tr>
<tr>
<td>Clone 10</td>
<td>I</td>
<td>T</td>
<td>S</td>
<td>I</td>
<td>T</td>
</tr>
<tr>
<td>Wild-type</td>
<td>S</td>
<td>G</td>
<td>V</td>
<td>G</td>
<td>N</td>
</tr>
</tbody>
</table>

TABLE 4. Sequencing results of selected TFs. Library 1 (45S, 46G, 47V, 48G, 55N) is selected against P_{\text{M,ST6T}} and Library 2 (45S, 46G, 48G, 49A, 55N) against P_{\text{M,SG6G}} while counterselecting against wild-type binding. Wild-type amino acids are highlighted in blue, randomized positions are underlined and the amino acid that is not part of the combinatorial library is annotated with an asterisk. Positions 35, 38, 39 illustrate the amino acid mutations in \lambda cl to obtain cl_{opt} and are denoted by a “P” for selected variants (e.g. cl_{SG6G,P}).

<table>
<thead>
<tr>
<th>Position</th>
<th>35</th>
<th>38</th>
<th>39</th>
<th>43</th>
<th>44</th>
<th>45</th>
<th>46</th>
<th>47</th>
<th>48</th>
<th>49</th>
<th>55</th>
</tr>
</thead>
<tbody>
<tr>
<td>cl</td>
<td>S</td>
<td>D</td>
<td>K</td>
<td>G</td>
<td>Q</td>
<td>S</td>
<td>G</td>
<td>V</td>
<td>G</td>
<td>A</td>
<td>N</td>
</tr>
<tr>
<td>cl_{opt}</td>
<td>L</td>
<td>Y</td>
<td>E</td>
<td>G</td>
<td>Q</td>
<td>S</td>
<td>G</td>
<td>V</td>
<td>G</td>
<td>A</td>
<td>N</td>
</tr>
<tr>
<td>cl_{SG6G,P}</td>
<td>L</td>
<td>Y</td>
<td>E</td>
<td>G</td>
<td>Q</td>
<td>S</td>
<td>A</td>
<td>V</td>
<td>S</td>
<td>E</td>
<td>W</td>
</tr>
<tr>
<td>cl_{ST6T,P}</td>
<td>L</td>
<td>Y</td>
<td>E</td>
<td>W*</td>
<td>Q</td>
<td>N</td>
<td>R</td>
<td>I</td>
<td>C</td>
<td>A</td>
<td>A</td>
</tr>
</tbody>
</table>
**Supplementary Information**

<table>
<thead>
<tr>
<th>Operator</th>
<th>Sequence</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1</td>
<td>GCAACCAT TATCACCGCCGAGGTAAAATAGTCAACACCCACGGTGTAAATAGTCAACACGCACGGTGT</td>
<td>Replaced by natural operator OR1 from cI434 in pJPC12-ΔPS-P_{RM}-B0034-geneVI (Addgene plasmid ID: 80858) to bypass autorepression at high cI concentration (mismatches shown in red). Synthetic promoters are derived from the consensus sequence CS (TATCACCGCCGGTGATA) of bacteriophage λ operators. Mutated base pairs in O1 and O2 of synthetic promoters are bold and the engineered promoters are named after the position of the base substitution in the consensus half-site. WT cI binding to O3 was restored for counterselections by inserting the CS at O3. For selections, the promoter variants are used as unidirectional promoters with M13 gene VI downstream P_{M} variants on the accessory plasmid. Operators are highlighted as follows: O1 blue, O2 green, O3 grey.</td>
</tr>
</tbody>
</table>
**Supplementary Figure 2. Sequences of synthetic promoters used for functional characterization of selected TF variants.** The natural operator O3 of $P_R$/P$_{RM}$ is replaced by an obliterated O3 site\(^1\) in order to bypass autorepression at high cI concentration. Synthetic promoters are derived from the consensus sequence CS (TATCACGCGCCGGTGATA) of bacteriophage $\lambda$ operators. Mutated base pairs in O1 and O2 of synthetic promoters are bold and the engineered promoters are named after the position of the base substitution in the consensus half-site. For functional characterization of selected TFs, the bidirectional promoter variants are used upstream of GFP and mCherry on the reporter plasmid. Operators are highlighted as follows: O1 blue, O2 green, O3 grey.
Supplementary Figure 3. Plasmid maps needed to set up the directed evolution system. (a) The modified helper phage HP (M13KO7-ΔPS-ΔgeneIII-ΔgeneVI) contains all phage genes required for phage replication except the genes III and VI. The weak packaging signal (PS) is removed to bypass helper phage propagation. (b) The phagemid pLITMUS-rpoN-clopt-J23106-gIII provides the gene of interest clopt, the M13 packaging signal (PS) as well as constitutively expressed gene III. (c) The accessory plasmid pJPC12-ΔPS-Prm-gVI contains a conditional gene VI expression circuit, induced by an active library member on the phagemid. (d) The reporter plasmid pJPC12-ΔPS-mCherry-PRm-GFP is used to characterize the selected library members. For additional information on the pJPC12 vector see reference 2. DNA sequences are shown in Supplementary Data 1-5.
**Supplementary Table 1. Plasmids used in this protocol.** A minimal set of four plasmids (annotated with an asterisk) is required to set up the entire selection system plus subsequent reporter analysis (Supplementary Fig. 3). Sequences and plasmid maps, as well as DNA plasmids have been deposited and are available from Addgene (https://www.addgene.org/). For a more comprehensive list of deposited plasmids see reference 3.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Class</th>
<th>Antibiotic resistance</th>
<th>Addgene ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>M13KO7-ΔPS-ΔgeneIII-ΔgeneVi*</td>
<td>HP</td>
<td>Kanamycin</td>
<td>80840</td>
</tr>
<tr>
<td>pLITMUS-J23106-geneIII</td>
<td>PM</td>
<td>Ampicillin</td>
<td>80843</td>
</tr>
<tr>
<td>pLITMUS-rpoN-cl-J23106-geneIII</td>
<td>PM</td>
<td>Ampicillin</td>
<td>80852</td>
</tr>
<tr>
<td>pLITMUS-rpoN-cl5G6G,P-J23106-geneIII*</td>
<td>PM</td>
<td>Ampicillin</td>
<td>80861</td>
</tr>
<tr>
<td>pLITMUS-rpoN-cl5T6T,P-J23106-geneIII</td>
<td>PM</td>
<td>Ampicillin</td>
<td></td>
</tr>
<tr>
<td>pJPC12-ΔPS-PRM-B0034-geneVi*</td>
<td>AP</td>
<td>Chloramphenicol</td>
<td>80858</td>
</tr>
<tr>
<td>pJPC12-ΔPS-PM5G6G-OCS3-B0034-geneVI</td>
<td>AP</td>
<td>Chloramphenicol</td>
<td></td>
</tr>
<tr>
<td>pJPC12-ΔPS-PM5T6T-OCS3-B0034-geneVI</td>
<td>AP</td>
<td>Chloramphenicol</td>
<td></td>
</tr>
<tr>
<td>pJPC12-ΔPS-mCherry-PRM-GFP*</td>
<td>RP</td>
<td>Chloramphenicol</td>
<td>80859</td>
</tr>
<tr>
<td>pJPC12-ΔPS-mCherry-P/PM5G6G-GFP</td>
<td>RP</td>
<td>Chloramphenicol</td>
<td>80911</td>
</tr>
<tr>
<td>pJPC12-ΔPS-mCherry-P/PM5T6T-GFP</td>
<td>RP</td>
<td>Chloramphenicol</td>
<td></td>
</tr>
</tbody>
</table>
Supplementary Table 2. Primers used for cloning. The complementary sequences used for Gibson Assembly are underlined. Oligonucleotides used for combinatorial library cloning obtain NNS codons (where S = G/C) at the randomized positions.

<table>
<thead>
<tr>
<th>Name</th>
<th>Oligonucleotide sequence</th>
<th>Use</th>
<th>Step</th>
</tr>
</thead>
<tbody>
<tr>
<td>pLITMUS-F</td>
<td>5’ GGT CTC CTG GCC TGC AGT AAA G 3’</td>
<td>PM linearization</td>
<td>1</td>
</tr>
<tr>
<td>pLITMUS-R</td>
<td>5’ GTT CAG AAT ACT CTC CTG C 3’</td>
<td>PM linearization</td>
<td>1</td>
</tr>
<tr>
<td>ci-F</td>
<td>5’ GGA GAG TAC GAT TCT GAA CAT GAG CAC AAA AAA GAA AC 3’</td>
<td>Amplification of target gene</td>
<td>2</td>
</tr>
<tr>
<td>ci-R</td>
<td>5’ CTT TAC TGC AGG CCA GGA GAC CTA TAA ACG CAG AAA GGC CC 3’</td>
<td>Amplification of target gene</td>
<td>2</td>
</tr>
<tr>
<td>pLITMUS-Lib-F</td>
<td>5’ GCA TTA AAT GCT TAT AAC GC 3’</td>
<td>PM linearization for library construction</td>
<td>11</td>
</tr>
<tr>
<td>pLITMUS-Lib-R</td>
<td>5’ CAT CCC CAT CTC GTA TGC 3’</td>
<td>PM linearization for library construction</td>
<td>11</td>
</tr>
<tr>
<td>Library 1-F</td>
<td>5’ GTC GCA TAC GAG ATG GGG ATG GGG CAG NNS NNS NNS NNS GCT TTA TTT AAT GGC ATC 3’</td>
<td>Randomized oligos for library construction</td>
<td>11</td>
</tr>
<tr>
<td>Library 1-R</td>
<td>5’ GCG TTA TAA GCA TTT AAT GCN NGN ATG CCA TTA AAT AAA GC 3’</td>
<td>Randomized oligos for library construction</td>
<td>11</td>
</tr>
<tr>
<td>Library 2-F</td>
<td>5’ GTC GCA TAC GAG ATG GGG ATG GGG CAG NNS NNS NNS CTT NNS NNS TTA TTT AAT GGC ATC NNS GCA TTA AAT GCT TAT AAC GC 3’</td>
<td>Randomized oligos for library construction</td>
<td>11</td>
</tr>
<tr>
<td>Library 2-R</td>
<td>5’ GCG TTA TAA GCA TTT AAT GC 3’</td>
<td>Randomized oligos for library construction</td>
<td>11</td>
</tr>
<tr>
<td>pLITMUS-Δclopt-F</td>
<td>5’ TTA GAA TTT GCT GAT GAG ACG GTC TCC TGG CCT GCA G 3’</td>
<td>Deletion of rpoN-clopt-B0015 cassette</td>
<td>62</td>
</tr>
<tr>
<td>pLITMUS-Δclopt-R</td>
<td>5’ CTG CAG GCC AGG AGA ACG GTC TCT CAG CGA ATT CTA A 3’</td>
<td>Deletion of rpoN-clopt-B0015 cassette</td>
<td>62</td>
</tr>
<tr>
<td>B0034-gVI-F</td>
<td>5’ AAA GAG GAG AAA TAC TAG ATG CC 3’</td>
<td>AP linearization</td>
<td>39</td>
</tr>
<tr>
<td>gVI-R</td>
<td>5’ GTC GCA TAC GAG ATG GGG ATG GGG CAG NNS NNS NNS NNS GCT TTA TTT AAT GGC ATC 3’</td>
<td>AP linearization</td>
<td>39</td>
</tr>
<tr>
<td>P₅ₛ₆₆₅-OCS3-gVI-F</td>
<td>5’ CTT GAT ATC GAA TTT GCA ACC ATT ATC GGC GCC GCC GAT AAA ATA GTC AAG GCC GGC GCC GAT AGA TAA ATA G 3’</td>
<td>Promoter insertion on AP</td>
<td>39</td>
</tr>
<tr>
<td>P₅₆₆₅₋OCS3-gVI-F</td>
<td>5’ CTT GAT ATC GAA TTT GCA ACC ATT ATC GGC GCC GCC GAT AAA ATA GTC AAG GCC GGC GCC GAT AGA TAA ATA G 3’</td>
<td>Promoter insertion on AP</td>
<td>39</td>
</tr>
<tr>
<td>OCS3-gVI-R</td>
<td>5’ CAT GTA TTT CTC CTC TTT ACG TTA AAT CTA TCA CCG CCG GTC ATA AAT ATC 3’</td>
<td>Promoter insertion on AP</td>
<td>39</td>
</tr>
<tr>
<td>GFP-F</td>
<td>5’ AAA GAG GAG AAA TAC TAG ATG 3’</td>
<td>RP linearization</td>
<td>39</td>
</tr>
<tr>
<td>mCherry-R</td>
<td>5’ GTC GCA AAG AGG AGA AAG 3’</td>
<td>RP linearization</td>
<td>39</td>
</tr>
<tr>
<td>P₅ₛ₆₆₅-GFP-F</td>
<td>5’ ATT CTT TCT CTT TCT GTC ACC ATT CAT TGG CCG CCG CCG CCG CAA AAT AGT CAA CGG CCG CCG CCG CAA TAA AAT AGT CAA CGG CCG CCG CCG CCG CAA GAT AGA TAA ATA TTT AAT AAT AGT GGT GAT AG 3’</td>
<td>Promoter insertion on RP</td>
<td>39</td>
</tr>
<tr>
<td>P₅₆₆₅₋GFP-F</td>
<td>5’ ATT CTT TCT CTT TCT GTC ACC ATT CAT TGG CCG CCG CCG CCG CAA AAT AGT CAA CGG CCG CCG CCG CCG CAA GAT AGA TAA ATA TTT AAT AAT AGT GGT GAT AG 3’</td>
<td>Promoter insertion on RP</td>
<td>39</td>
</tr>
<tr>
<td>P₅₋GFP-R</td>
<td>5’ CAT GTA TTT CTC CTC TTT ACG TTA AAT CTA TCA CCG CCA GTC TTA AAT AAT ATC 3’</td>
<td>Promoter insertion on RP</td>
<td>39</td>
</tr>
</tbody>
</table>
Supplementary Table 3. PCR conditions specific for individual reactions.

<table>
<thead>
<tr>
<th>Protocol step</th>
<th>Primers</th>
<th>Annealing temperature [°C]</th>
<th>Extension time [min]</th>
<th>Number of cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>pLITMUS-F, pLITMUS-R</td>
<td>54</td>
<td>2.5</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>cI-F, cI-R</td>
<td>46</td>
<td>1.0</td>
<td>25</td>
</tr>
<tr>
<td>12</td>
<td>pLITMUS-Lib-F, pLITMUS-Lib-R</td>
<td>50</td>
<td>2.5</td>
<td>25</td>
</tr>
<tr>
<td>12</td>
<td>Library 1-F, Library 1-R</td>
<td>45</td>
<td>0.5</td>
<td>20</td>
</tr>
<tr>
<td>39</td>
<td>B0034-gVI-F, gVI-R</td>
<td>48</td>
<td>3.0</td>
<td>25</td>
</tr>
<tr>
<td>39</td>
<td>P_{M,5G6G}-O_Cs3-gVI-F, O_Cs3-gVI-R</td>
<td>54</td>
<td>0.5</td>
<td>20</td>
</tr>
<tr>
<td>39</td>
<td>GFP-F, mCherry-R</td>
<td>45</td>
<td>3.0</td>
<td>25</td>
</tr>
<tr>
<td>39</td>
<td>P/P_{M,5G6G}-GFP-F, P/P_{M}GFP-R</td>
<td>46</td>
<td>0.5</td>
<td>20</td>
</tr>
</tbody>
</table>
Supplementary Data 1. Sequence of the modified helper phage (HP) M13KO7-ΔPS-Agenelli-AgeneVI. The total plasmid size is 6,910 bp. The plasmid is available from Addgene (ID 80840).

<table>
<thead>
<tr>
<th>FEATURES</th>
<th>Location/Qualifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>gene</td>
<td>1..831 /label=&quot;Gene II&quot;</td>
</tr>
<tr>
<td>gene</td>
<td>496..831 /label=&quot;Gene X&quot;</td>
</tr>
<tr>
<td>gene</td>
<td>B43..1106 /label=&quot;Gene V&quot;</td>
</tr>
<tr>
<td>gene</td>
<td>1108..1209 /label=&quot;Gene VII&quot;</td>
</tr>
<tr>
<td>gene</td>
<td>1206..1304 /label=&quot;Gene IX&quot;</td>
</tr>
<tr>
<td>gene</td>
<td>1301..1522 /label=&quot;Gene II&quot;</td>
</tr>
<tr>
<td>termina</td>
<td>4039..4071</td>
</tr>
<tr>
<td>origin</td>
<td>complement(4287..4832) /label=&quot;pB5A ori&quot;</td>
</tr>
</tbody>
</table>

| gene | 1 /label="Gene II" |

```
1 AAGCGCTACA CTATTTAAGT AATGATGCC ACCCTTTCAG CTGGCCCAAAC AATGAAAT
61 ATACGCTACTA AAGTTTTAGA CCGTTCGCA AATGTCGTA AATGACCAAAT TAAACTACT
121 CGTCGGCAGA AGTTGGAACT AACCTTTCAG TGGACTGAAA CTCGACAGCA CTGGATTTA
181 GTTGCAGATTAA TAAACATAGT TGAGACTGAC CACCGAGATG AGCCATTAGAA CTCGGCA
241 TCGCCAAAAA TGCTCCCTTA CCAAAAGGAG CAAATTAAGG TACTCTCTAA CTCGGCA
301 TGTGAGGTTG CTTGCGGGCT GTGCGGCGAA GAACTGCGA AGTAATAGAAG GTTAATGA
361 CTTTCCGCGG TCTTCCTTTAA TCTTTTTGAT GTCACTCTAT TTTCCTCCTG CTAGCAAT
421 CGATGGTTAAG AGCTGATTTT TATTTTTGAT TTCTTCTTGA CTAGCAAT
481 TTTGAGGGGG AGTTAAAGGA AATCTTCAAG CACCTTGTGA CAGAAAGTA GATCTTGGG
541 AAAGCTTTTAG CTATTTATGC CCGGCGACAA CTTTTCACAG CAGACCACCA CATCGACAT
601 GCTTTTTTATC GTGCGCGGTG AAACGGGAGT TATTTTTTTT TTTCCTTCTG ATCAAGAT
661 AATCTCCTTTT GCGCAATGTT ATCGCAGTTA GTGTCGATGG GTACAATCTG CTAGCAAT
721 AGCTACCTTTT TTCATGGTTA TTACGTTGGA GCTATGCAGC GTTCTGTTGC GCCTGAA
781 TCTCTCCCAAC GTCTCGAGCT GTATGATTTG TTTAAGGGGC CCAATCAACAT AATGAAAT
841 CTAGAGCTGCA AGTGGAGTTA AAACCAATGC TACCTGGGCA AAGGACCAAA ATATCCAG
901 CTCGTCGAGG CAAACCCATC TATCTGAACT AGCATCTTCT TATTTTTTCT TATCCAG
961 AAATCTGTTAAG AGCTGATTTT TATTTTTTCT TATTTTTTCT TATCCAG
1021 TTTGACGGCT TGCTACGNCTT CTTTACAAAC TTTGAGCTAT ATGCCTGAAA ATATCCAG
1081 TCTGCGCCCC GCAGGCCGATG GCAGGCCGATG GCAGGCCGATG GCAGGCCGATG ATATCCAG
1141 CAGGGCTATG AGCAGGCTG TACGGAGGG TATTGTCAGG ATGTTGTAAG GATTTTTTCT
1201 TAAAGGATGG GTTTTGTATG TATTTTTTTTT TTTTGTATG TATTTTTTTTT TTTTGTATG
1261 GTGCGCATTT AATTTCTACC GTATGTTAGT GAATTTGTTA AATTTCTACC GTATGTTAGT
1321 CAAGGCTTCT GTAGGGTTGTT GCTCGCCCTTT CTAATGGAGT ATGCTGAGCT ATGCTGAGCT
1381 CGATGCGCGA AAAGGGGCGT CAACTTGTAT TTTTGGGTTG GCTGAGCT ATGCTGAGCT
1441 TGGCCGAGG ATGGTTCGTT TATTCGGTAT CAGGACTCTC TATCTGAACT TAGCTGAGCT
1501 ATACACTCCT AAAGAACAGT GAAACACAGC TAGATTTGAT CAGGACTTCTT TATCTGAACT
1561 TTATTGGGAAA TTTTTCAAGT GAAACACTAT TATTGTCGAT TATCTGAACT TATCTGAACT
1621 TATTTCTACCT GCCTGAGAAC GCTTGAAGAG CTCTATATAA GATCTGAGCT TATCTGAACT
1681 GCAAAAGGCT TTTCAGGGTT GCTTGTTGAT ATGGGCAAGA TATCTGAACT TATCTGAACT
1741 CGGGTGGCAA ATATGGAACCT ATATTTGTAG CTATGTTGTA AATCTTCTGCA GTAGCTGAGCT
1801 GGTGGTCATTA ACTTTGTACC CTTTATAGGT AATGTATCTTTT TCATGTTGTA AATCTTCTGCA
1861 TGTCAGTTG GTGCGGGGTA TGGATTCGGT ATGTTGTCGG TAGTTGTCGG TAGTTGTCGG
1921 ATGAAATCTG TGGATTCGGT AATCGGCGGT GGCAGTGATG GGGCTGGGTA TGGATTCGGT
1891 TGATGAGGTTG TCTACAGCT CATGTAAGAT GCTATCGGAT CTTTCGCGG CTATGTAAGAT
2041 TATCTAGTTG TGGATTGAGC CAGCTTATTG ATGTTGTCGG TAGTTGTCGG TAGTTGTCGG
2101 GTGCGGAGAG ATATTTTTCT ATCCTTGGCG GTCTATTTTT ATGTTGTCGG TAGTTGTCGG
2151 AAATGCTTCT GTATGTTGTA GTTGTGTAAG CTTGTTGAGAT TAGTTGTCGG TAGTTGTCGG
2211 CACTATGGTT GAAGGGGTTTT TCTATGTTA AGATGAGATT ATGTTGTCGG TAGTTGTCGG
2261 CTATACGTTG GTGACTGCTT ATGGGTTTAT GCTATCAGT ATGTTGTCGG TAGTTGTCGG
2311 CACGTCGTTGA GTGACTGCTT ATGGGTTTAT GCTATCAGT ATGTTGTCGG TAGTTGTCGG
2281 AGCTTGGTTC TATGAAATCT ATGTTGTCGG GTTTTTTTTT TTTTGTATG TATTTTTTTTT
2341 AGCTGCGGGG GCAGGGGTTTT TTTTGTATG TATTTTTTTTT TTTTGTATG TATTTTTTTTT
2401 TGACAGGCTT ATGTTGTCGG TAGTTGTCGG TAGTTGTCGG TAGTTGTCGG TAGTTGTCGG
2461 ATGACTGTTA GGCAGGCTT ATGTTGTCGG TAGTTGTCGG TAGTTGTCGG TAGTTGTCGG
2521 ATAAATTTGTG TCTGAGCTG TCTCAGCCGCT TAAATTCAG TCTCAGCCGCT TAAATTCAG
```

Supplementary Data 2. Sequence of the phagemid (PM) pLITMUS-rpoN-cl*optJ23106-
genelll. Sequencing primers listed in Table 1 are underlined (pLITMUS-F, pLITMUS-R). The total plasmid size is 4,924 bp. The plasmid is available from Addgene (ID 80852).

**FEATURES**

<table>
<thead>
<tr>
<th>Location/Qualifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>gene</td>
</tr>
<tr>
<td>/label=&quot;Beta-lactamase (Ampicillin resistance)&quot;</td>
</tr>
<tr>
<td>promoter</td>
</tr>
<tr>
<td>/label=BBa_J23106</td>
</tr>
<tr>
<td>RBS</td>
</tr>
<tr>
<td>/label=BBa_B00034</td>
</tr>
<tr>
<td>gene</td>
</tr>
<tr>
<td>/label=&quot;Gene III&quot;</td>
</tr>
<tr>
<td>terminator</td>
</tr>
<tr>
<td>/label=pMB1 ori</td>
</tr>
<tr>
<td>origin</td>
</tr>
<tr>
<td>complement(2514..)</td>
</tr>
<tr>
<td>/label=&quot;M13 ori&quot;</td>
</tr>
<tr>
<td>origin</td>
</tr>
<tr>
<td>/label=&quot;pLITMUS&quot;</td>
</tr>
<tr>
<td>terminator</td>
</tr>
<tr>
<td>/label=BBa_B0015</td>
</tr>
</tbody>
</table>

```plaintext
1 GTTAACTAGC TCGAGTGGCA CTTTTCCGGG AAATGTCGCC GGAACCCCTA TTTGTTTATT
61 GTTAACTACG CATTACAA TGTATCCGCTCG TATGACAGCAA TAAACCTTTC TATCTTCTCA
121 TATAATATTGA AAAAGGAAAG ATAGGATGTAT TCAACATTTT CCGGTCGCCC TTATCTTCTT
181 TTTGGGCGCA TTTTGGCTCC TCTTTGTTGC TACCCGAGAA AGGCGTGGTA AAGTAAAAGA
241 TGCTGAAAGAT CAGTGGTGCT CAGAGTGGTG TATCATCGCA CTAAGTCTCTA ACAGCGTGGAA
301 GATCCTTGGGT AGCTCTGGCC GGACAGAAAG TTGCTCAGAG AGAAGAAGAG CTATCACC
361 GCATCTAGGC GCGGTAATTG CCGGCTGACT CGCCGACGAA AGCAGATCCG CGCCGGCAT
421 ACATCTTCTCT CAGAATGACT TGGTGTGACTG ATCACCGCTG AGAACAAAGG AACATCCGGA
481 TGGCTAGCA GAAAGAAGAT TATAGACGAC TGGTAAACAG ACCATGTAGT ACAGCGGCGC
541 CAACTCACCCT CTGCAAGGAG GCGGGAGCAG GAGAGAAGCA ACCGGTTTCT TGGAGACACAT
601 GGCGGATCC GAATACGGCC TGGTATCGTG CAGCGCGGAG CTTGAGTAAAC ATACCCAGAA
661 CGAGCGGGCC GACACACCGG TCGCTTGAGC AATGCGCACA AGCGTGGCAG AACTTAAAC
721 TGCCGAACTA CTTCTATCTGT CCTCAGGGCA ACAATATATA GACTGTAGTG AGGCGGATA
781 AGTGGCAGAG CACACTTACGT GCTGCGCCCCT TCGGTCCTTC TGTTTTATAG TGCTTTACAT
841 TGGACCGGCGG GACGGCTTGG GTCCGGCGAT CATTGAGCAG CCGGGGCGAG ATGTGAAAGC
901 CTCGCGGCATAT GTAGTTATTGT ACAGCGGCGG GACGGCGCAG ATACGGTATG AAGCGAAAT
961 CAACAGTCCG CCTGATAGGT CCTGACTGAT TACGACTTGT TACTTACGGA CAAAAGATTT
1021 CTCAATTTTA CGGCTACGCTTC GTGCACTAGGT ATAGCTGTAG TGGAGAAAAG CAGGAGAAT
1081 ACTAGTAGAA AAATATTAAT TTTCGCAATCT TTCTATGTTA TTTCTCAATCT TCTCATTCC
1141 TTTGGTATCAT TGAAGTATG TGCAGCAAACT CAGCTCACAA AAGATCTCATT TCTAAGTCA
1201 GGAGGACGCA AAAACCTTGA GATGTTTACAT GCTATATCTG GCGGCTGCTA GGAATGCTA
1261 CAGGGTTGTAT AGTTGTTACT GGGGAGGCAAA CTGAAGTGTAG TGGATACGT TTCACTTAC
1321 GCTTCTGCTAT CCTGAAAAAT GAGGGGGTGT GCTAGACTTG GCGGTGCTCT GAGGCTGCGG
1381 GGGTGAGTTG AGGCTCGATT AACGCTTCTTG AGTACGTAGC TGAACAGTCA ACGGCTGC
1441 CTATATCATT ACCCTCGGTAC CCGCTTGCAG CTGGGAAAAC CCGCCATCAT
1501 CTATATCATT TCTTGGAGAG TCTGACATCT TTATATACCT TGTAGTCAG AATATAGATG
1561 CCGTAAATAG GCAGCGGCCA TTAAGCGCTT ATAGCGGCAAG TCAATGGTACT GCACTGACC
1621 CGGTAAATAC TTTAATACCG TCTACTCCTC TATCATCAGA AGACGCTGTA GCACTGACC
1681 GGAACGGTAA ACCGCGGAG CTTGATCTGC ATACGGGACTA TAAATGATCT TCTGTTGTC
1741 TGTTGAATCA AGGCGCCATT CGCTGACGCT CTGGGAGATT CTGCTGCTG TGGCCGCGCT
1801 CTGTTGAGTG TCTTGGTGCT GGGCTGAGAG TGGTGCGCTG AGGCGGTCAG GAGTTGAGC
1861 GTGCGGCGC CGGAGGCGG GCTGCGGCGG GCTGCGGCGG TGGCCTGCGG TGGCCTGCGG
1921 AAGATGGGAG GCGGCGGATT AACGCGGGGA CGGGACTGAA TCTGCTGAA AAGCGGCGG
1981 AGTCTGCACT TAAACCGGGA CTTGATCTGC ATCTGCTGAC GCTGACGCTG TGGCCGCTG
2041 GTTTAGATTG CGGCTGCTCT GGGCTGCTAG ATGGTTAAGG TCGGACCTGT GGTGCCCCG
2101 GGCACTTTAT CCAAAAGCGCT AGATCGGCTT CTGAACTTG AAATGACTAT
2161 GCGCCATATA CTTCAGCCTT CTTGCTCAAT TGGTGATGAG TCCGTCCTTT GLGTGGCGC
2221 GTGAGACACG ATACGCTGTC ATATGAATTG CAATAGTATG CTTGCTGCTG
2281 GTGGTCTCTCT TTTATATTGT GCCACCCTTTA TGATAGTATT TTATGCTCTT GCTGATCAT
2341 TGCGTATATA GACGGCTTTA GAAGACCGAC GCATCAAAT AAGCGAAGG CTCAGCTGCA
2401 AGAATCGGCCG GCCGGTCTTA TCTGGTGTTT GTGCGGAC AGCTCTCCTC AGAGCTGAGC
2461 TGGCGCGGGC CGCGGTCTTT TCTTTACGTA TATATAGATG TGGTTAAATG TGGGCTGCGC
2521 TTTGATGTTA TATCTGCTCTT CTTGGTCTAA AGAATTTTATG TCTGCTGCTG
2581 AGCGTTATTG TTTCGTTTTA AATGTTGTTA ATACGCTGCA TCTGCTTCTC
2641 AATAGCGCAG GTCCGAGCAA ATCCCTTATG ACCAAGAAA ATACGCGAGG ATAGGGTGTA
2701 GTGGTGTCTT CTGGTGCAAG AAGATGCTCA TATTAACAGA GACGCTGCTG GACGCTGCA
2761 GCCGAAAACC CACGCGGAGC CACGGCGGAGC AACCACCAATG CACATGCGCT
2821 TTTGGGTCGG GACGGGCTGT AAGACGCTAA TACGCGCCAA TAAAGGCGG CCCCGCTTGA
2881 GAGCGGCTAC CGGAGGCGG ACGTGGCGGA AAGGATTAGG AAGAAAGGGA AAGAGGCGG
2941 CGGAAAGGGG CCTGCAAGGTT TACGGGCTGG TGGCGGGTAA ACCACCCACG CGCGGGCGG
```
TAATCGGGCC CGTACAGGGCC CTTAAAAGGA TCTAGGTCGA GACCTCTTTTT GTAATCTCAA

TAATGCGCCG CTACAGGGCC CTTAAAAGGA TCTAGGTCGA GACCTCTTTTT GTAATCTCAA

10
Supplementary Data 3. Sequence of the phagemid (PM) pLITMUS-rpoN-clG66P-J23106-genelll. Sequencing primers listed in Table 1 are underlined (pLITMUS-F, pLITMUS-R). The total plasmid size is 4,924 bp. The plasmid is available from Addgene (ID 80861).

<table>
<thead>
<tr>
<th>FEATURES</th>
<th>Location/Qualifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>gene</td>
<td>143..1003</td>
</tr>
<tr>
<td>promoter</td>
<td>1027..1061</td>
</tr>
<tr>
<td>RBS</td>
<td>1068..1079</td>
</tr>
<tr>
<td>gene</td>
<td>1086..2360</td>
</tr>
<tr>
<td>terminator</td>
<td>2367..2495</td>
</tr>
<tr>
<td>origin</td>
<td>complement(2514..3023)</td>
</tr>
<tr>
<td>gene</td>
<td>3031..4744</td>
</tr>
<tr>
<td>terminator</td>
<td>4751..4879</td>
</tr>
</tbody>
</table>

1 GTTAACTAGC TCGGCTGGCA CTTTTCGCGG AAATGTGCAG GGAAACCTTA TTTGTTTATT
61 DTATATATTA TATCATTCAT AGAGGTGAGAT GATCTATGCA TCTTCTATCA TCTCATTCC
121 ATAATATCAG CACAAAATTT CATGATGCAAT CATGATGCAAT CATGATGCAAT CATGATGCAAT
181 TTTTGGGCGA TTTTGGGCGA TTACCTTTGG ACACGTGGGA AAGTAAGAAG
241 TGTGCAAGGT CAGGGTTGTT CAGGAGTGGG AGTATCCTCA TCTATGCACT ACACGGTTAA
301 GACGGCTTCAG AGCCTCGGAG GAACAGGAGG TGCATGAGG TATGATGGA TATGATGGA
361 GCTTATGCT GCAGCGGCGA CGCTAGGGCG CTGGCAAGT TTTTGGGGTC GAGGTGCCG
421 ACACGTATCT CGTTATGAGT CGTATGATGG CAGAAAAGGC CAGAAAAGGC
481 TGCGAAGTCG TGGAGAAGAT TGGCAGTCGC TGGCAGTCGC TGGCAGTCGC TGGCAGTCGC
541 CAACGTATCT CGTACAGGCA TCGGGCAAGG GAGGAGGCTA AGGGCGTCTT TCGCAACAT
601 GGCGGATCTAG TGGAGGCTTC TGGGAGGCTTC TGGGAGGCTTC TGGGAGGCTTC TGGGAGGCTTC
661 CGACGGGCGG GCACAGCAGG CAGCGGCGG GCACAGCAGG CAGCGGCGG GCACAGCAGG
721 TGGGCGAACG CCATCAGGCA GCCGCTTGGT TCCGCGCCTT TCGTGGTGTG TTGTTTATGT TCGTGGTGTG
781 AGTGGCAGGA CACCATCAGG CTCGCGCCTT CTCGCGCCTT CTCGCGCCTT CTCGCGCCTT
841 TGGGCGGCGG GAGGCGGCGG CAGCGGCGG CAGCGGCGG CAGCGGCGG CAGCGGCGG
901 CGCGAGCTAT GCAGCTGACT ACAGTGAGGC ACAGTGAGGC ACAGTGAGGC ACAGTGAGGC
961 AGACAGGCGG GAGAGGCGG GCACAGCAGG GCACAGCAGG GCACAGCAGG GCACAGCAGG
1021 CTCAGATTTA CTCAGATTTA CTCAGATTTA CTCAGATTTA CTCAGATTTA CTCAGATTTA
1081 ACTATGATAT AAAATATATG TTACTTCACT TCACACTTAT TCACACTTAT TCACACTTAT
1141 GCATATGATT ATAGATTATG TATATATGAT TATATATGAT TATATATGAT TATATATGAT
1201 GAGGACTAGA AAAAAGTATT CCGCACTAGA TACCGAGGAA AGCGATGATT TATATATGAT
1261 GAGGACTAGA AAAAAGTATT CCGCACTAGA TACCGAGGAA AGCGATGATT TATATATGAT
1321 CGCGTGGATT GCTGGGGATT CAGCAGGGATT CAGCAGGGATT CAGCAGGGATT CAGCAGGGATT
1381 CGCGTGGATT GCTGGGGATT CAGCAGGGATT CAGCAGGGATT CAGCAGGGATT CAGCAGGGATT
1441 CGCTCACTTT CCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC TGGACGCGAC
1501 CTATACATCA GGGGTTTCTT GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
1561 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
1621 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
1681 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
1741 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
1801 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
1861 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
1921 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
1981 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2041 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2101 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2161 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2221 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2281 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2341 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2401 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2461 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2521 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2581 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2641 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2701 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2761 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2821 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2881 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC
2941 TTAATACGCA ACGTATTATG GCTGCTGCTT CAGCCTATCT TGGACGCGAC TGGACGCGAC

3001 TAATGCGCCG CTACAGGGCG GCTAAAGAGA TCTAGTGGAA GATCCCTTTT TGAATACCTCA
3061 TGGACCAAAAT CCGCTTTACGT GATGTTTCTGT TCCACGTGACC GTCAGACCCCT GTGAAACAGGA
3121 TCAAGAGGATCT GCCTTGGGACT GCCCTTTTCTT GCAGCTGTTATCT GTGACCTGTG TCCTGCTGCA
3181 AGCCGCCTCCT AGGCGCGGAG GCTTTGTTGCC CAATAGAGGA GCTACAACTT CTTTTGCGCA
3241 AGGTAAGTGT CCTACAGGCAA GCGAGATAC GAAATGCTTT TCTGTCTGATG TATGCCGTATT
3301 TACCCAGCCA TTTTACAGAAC TCTTGAACAC CGGCATACA CTCTGCTCTG TATCACTGTGT
3361 TACCTGAGGCC TCTGGGAGCT GCGGATAAGT CTTGCTTACG TCCTGTTAGAC TCAAGAGGATCT
3421 AGTACCCAGGA TAGCCACCAT CACGAGGAGG ATCTTACGGAG CTAGATGCTGG TAAAGGGGAG
3481 TGAGGAAACA GACCTACACC GACTAGGAGT ACTCTACGAG CGGCTTATGGA GAAAGAGGCA
3541 GCCCTCCAGGA AGGAGATACG GGCGGAGCTT ACGCGGTAAG TCGAGGGATAG GGAACAGGG
3601 AGGCGACAGG GAGGCTTACA GGGGAACAGC CCTGTATCTG TTTATGCTCT GTGGGCTTTTC
3661 GCCACGCTCTG ATCTGACCATT CTTGTTTGGCT GAGGGGGCGC AGGGTATTGCAA AGGCAGTCCA
3721 AAAGACCCAG GAAGCAGGCC TTCTTACGCT GCTGGCCCTT TTGGCTGGCAT TTTGCTGGGCA
3781 TGTAATGCTG GTAAGGCTCAAC TCTTGTAGAT GCTGTGAGAG ACCCTACGGTG ATCTAGGGAC
3841 ACAAGCTGGCG AGAAGACCTAG CGGTTTCTGTG ACGCCGCTTA TATCGCAAGT CAGGGCCATT
3901 TGATCCGCCCA CGCCGAGGCT ACAAGGAATCT TACAGAGAAGA ACACGTATAAG CTGTTTACCC
3961 TTGGGAGAGG TCTACAGACTG TAGCTAGGAT GAAGTTTGCC AGCTTTTACG AGAGAAGGACT
4021 GATTCGGACG AAGACGCAAA AAAGCAACCT ACAAGGCTGAT GGGCACTGAGG AGGAGGAGGAC
4081 TCCCTTTAAA GAAATGCTAA AAAAAAGAAGA AATGAACCTT GCCCTATCCCT AGGAAATGCTG
4141 GCAGATGAGG TAGGGGATGG GGCAGTGGCG GCTTTCCGAG TATATTAAGT GACATCGGCC
4201 ATTAAGCTCT TATAAGCCGG CATTGCTGAC AAAAATTCCT AAGTTAGGC TTGAAGATTT
4261 TAGCCCTTCA ATGCGAGGAG AATCTAACA GAGTGTAGAG AGCTTAAATG TCAAGGGTCT
4321 ACTTCAAGAGT GATGTGAGAA ACCCTTTTAC TCTGCTATGG CAGCCAGGTA GGTGCTTACC
4381 TGAGCTTTAGA ACCTTAAGGA AAGGGAGATT GGAAGAGTGG TGAACACACA CCCAAAAAGC
4441 CAGTATTTCT CCTTCTTGGA TTTGATTGGA ATGAAATCCC ATGACGCCAG AAACAGGCTT
4501 CAGGCCAAAAT CTTGCTTACC GAATTTGAAAT TCTGTTAGCC CTGAGACAGG CTGTGAGGCC
4561 AGGTGATTCG TCTAGAGAGA GACGTGGGGG TGAGATGTTT ACGTACAAGA AATCAAGCAG
4621 GGAATAGGCT GAGGTTTTCTT TACACACTT ACCACCACTA TACCAAACTG TCCACTGAGA
4681 TGAGATTTGC TCCCTTGTGG GCCAAATCTAT CGCTACGTGA TCGCGTGAAG AGAGGTTTGG
4741 CAATGAAAGAGC CAGGCGATCA AAAAAGACCA AAAGCTCAAGT CGAGAGACTG GGGCTTTCTGT
4801 TTAAATCTGGT GTTGTGCGGGT GAAGCGCTTCC GAGCTAGAAC CGACGCTGCT ACCAGGGGCT
4861 GGGCTTCTTCT GCGTATTAG GCTCTCXXGCT CTGCAATGAA GCCGGCTCTCG GCGGHCCTTTT
4921 TTGT
Supplementary Data 4. Sequence of the accessory plasmid (AP) pJPC12-ΔP5-PRM

B0034-geneVI. Sequencing primers listed in the Table 1 are underlined (pJPC12-F, pJPC12-R). The total plasmid size is 5,300 bp. The plasmid is available from Addgene (ID 80858).

<table>
<thead>
<tr>
<th>FEATURES</th>
<th>Location/Qualifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>gene</td>
<td>complement (1..146 and 4783..5300) /label=&quot;Chloramphenicol resistance&quot;</td>
</tr>
<tr>
<td>promoter</td>
<td>593..683 /label=&quot;PRM&quot;</td>
</tr>
<tr>
<td>RBS</td>
<td>684..695 /label=&quot;Bba_B0034&quot;</td>
</tr>
<tr>
<td>gene</td>
<td>702..1040 /label=&quot;gene 1&quot;</td>
</tr>
<tr>
<td>terminator</td>
<td>1041..1074 /label=&quot;Bba_B1002&quot;</td>
</tr>
<tr>
<td>origin</td>
<td>2794..3744 /label=&quot;repA&quot;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>bp.</th>
<th>TTTACGGCTTC TTTAAAAGGC CGGTTATTC GGAGTCGAC GTGCTCCTAA CTGCTGCTTC</th>
</tr>
</thead>
<tbody>
<tr>
<td>61</td>
<td>GCAATGACGTC GAAATGCTTC AAAAAAGTTT TATGAGATTTTT TTCCAGCTGTT</td>
</tr>
<tr>
<td>121</td>
<td>GATATCTCCAG TGGTTTCTT TTCTCATCTG TGCTCTGTTAT TTGCTGCTTT TTGCTGCTTT</td>
</tr>
<tr>
<td>181</td>
<td>CACCCGCGGTAG ATGATTTTTT TATGAGATTTTT TTGCTGCTTT TTGCTGCTTT</td>
</tr>
<tr>
<td>241</td>
<td>AGCTTTCAGTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>301</td>
<td>TTTAGAGTTT TTTAGAGTTT TTTAGAGTTT TTTAGAGTTT TTTAGAGTTT</td>
</tr>
<tr>
<td>361</td>
<td>CGTCTGGGTTA TAGTGGTGGG GCTCCGGGTAG CTGCTCTGTTAT TTGCTGCTTT TTGCTGCTTT</td>
</tr>
<tr>
<td>421</td>
<td>GAAATGCTTC ATGATCTTAA ATGATCTTAA ATGATCTTAA ATGATCTTAA</td>
</tr>
<tr>
<td>481</td>
<td>AGCTTTCAGTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>541</td>
<td>CGCGCGCTGTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>601</td>
<td>TATCCAGCTT TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>661</td>
<td>TATCCAGCTT TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>721</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>781</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>841</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>901</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>961</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1021</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1081</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1141</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1201</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1261</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1321</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1381</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1441</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1501</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1561</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1621</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1681</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1741</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1801</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1861</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1921</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>1981</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2041</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2101</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2161</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2221</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2281</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2341</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2401</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2461</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2521</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2581</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2641</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2701</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2761</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2821</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2881</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>2941</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>3001</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>3061</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>3121</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>3181</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
<tr>
<td>3241</td>
<td>CTTGCTGCTTC TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA TTTAGCTTAA</td>
</tr>
</tbody>
</table>

CTGTAACCG AACTTGAGA CAACCAGATA AAAATGAAT GTGACAAAT ACCAACAAC
ATTACATCAG ATTCCTACT ACAAATCCGA TAAAGAAAACA CACTACAGA TGCTTTAACA
GCAGAAATTG AGCTCAACAG TTGGAGGCCA AAAAAATGGA GTGACATGCA AAGTAAATAT
GAATGCAACG GTTCGTTTCC AGGCTCAACG CAAAAGAAGA GAACCACACT AGAGAACA
CTGATTAAGT AGGAAATGTT AGAAGGCTCT TATGCTCTCT GTAATCTACA GAAGACGATC
AACGCTAACAA AAAAAATGA AACAACGTT CTACGTTTAC AATACAAAGG GAAACTGTC
CATATGGA CAGAGCTTGC GAAAATAGT TATCCCAAGG GAAACTGTC
3841 CATATGGCA GATGAAACCG GTGTAAYAAA GATAGCTACA TACAGCTTT TACAGCTTTT
3901 TGTTGCAATTC AAGGCTTGTC ACCATGAAAT GATGCAAAAT GTAAACAGATG AACAAGCTAT
3961 AACACTTAAAT AGAAGCAGGT GAAACCAAGA AAAACAAAAA ACTAGACATG GAAGTGAAC
4021 ACTGAGACA ACTGGTTTACA GTCACACAGT CACACATAGA GACGCTAAAA GACGGATGC
4081 TGTTATATCTA AGTAAAGATCG CCAACAAACG AAGCCCGCAT GAGCTCTCCC GTGGAACAAA
4141 AATATGCGGA ATTTCGGAAG AATAGCGGCT TACGCAAGGC AAAAAAGGGG AACAGATGC
4201 TCGATATTTCTG ATAACAAAGCT AGCAACACCA GAAAGCAGCC TTTGCGGCC GCAAACCGC
4261 TACATTGGGA CTTTCCGGGT GTTTTTTGTT TGGTAAAGGG TGGCGGCC GGGCCAGGA
4321 TCCATATGTT CAAACCGGCG AGTTTACACT TCAAAACCGG CCGTTAACA CAATGCAAAA
4381 TCCATGACCG GATTTTAAAC ACCAACCCCA CCCGTAACACA CAAATACCC AACCAAGT
4441 GGGTTGTTGAT AAATAAACAGT ATTTGCTGTCGA GTATTTCCTCCT GAGCGGGGT GTGTGTCAG
4501 GCAATACCGAA GGTTCGGCAAT CTTGCTTATT GCCTGCCGTC TGCTACGATC
4561 GCCATACGGCT CGCAACCTGT ACGACTCTAG TCCGCAACGG CAGTGAAACC CCGCAACAG
4621 GATGTCTGACG TGCAAGGATG AACAAAATGAAC GTAAAGGATG TTTGTATAG AGACATATAC
4681 CCTGATATAAT GCTTCATCAA TCGCAGACCA AAAAGATCTG AGAAGATCAA TCATTAAAT
4741 CAGATAAAAAT ATTTCTTACC ACCAAATACCT GCCCTAAAAA AATTAGCGCG CGCCGCTCACA
4801 CACAGTTCGCA ATCTTTGGTG TATCTTAAGG AATATGCGGA CACGCTATTT GCATACAAAC
4861 GCCATGATGC CAAGACTCGT CACAGCCGAT CGACCTTCAG TCCGTTCGGG TAAATATTTT
4921 GCCCATGCTG AAAACCGGGG CAGAAGATT TGCATATTTT GCCAGGTATTT AATCAAACCT
4981 GGTGAAACTC ACGAGGCAAC TGGTGAAATAA AAAAAATAA TTTCTCATAA ACCCTTTTAG
5041 GAAAAATGGG CAGGATTTTG CAGTAAACAGC GAAACATCTT GGAATATTG TGTGAGAAA
5101 GCGGGCAAA GTCGTCTGCCT ATTCATCACA GGGCTAGAGA AAGCTCTTGAG TTTGCTCATG
5161 GAAACAGCTG TAAACAGGAT GAAACACTATG CCATACATCC AGCTACAGGC CTTCATGCC
5221 CATACGAAAT TCCGGATAGG CATTCCATCGG GGGGCAAGA ATGTGAATAT AGGCGGATA
5281 AAACTTGTCG TTTATTTTCC

Supplementary Data 5. Sequence of the reporter plasmid (RP) pJPC12-ΔPS-mCherry-P_{N/R_{PS,cm}}-GFP. The sequencing primer listed in Table 1 is underlined (pJPC12-F2). The total plasmid size is 6,238 bp. The plasmid is available from Addgene (ID 80859).

FEATURES | Location/Qualifiers
--- | ---
gene | complement (1..146 and 4783..5300)
/label="Chloramphenicol resistance"
complement terminator (593..697)
/label=T1
gene terminator (709..1419)
/label="mCherry"
RBS | complement (1426..1437)
gene complement (709..1419)
/label="PRM"
RBS | complement (1..146)
/label="RP"
origin | 3732..4682
/label="repA"

1 | TTTACGCGTCT TTTAAAAACG CGTAAATTACG AGCTGAAACG TCTGTTATTA GTGACATTGA
61 | GCAACTGCACT GAAGATGGCT AAAATGTGCT TTCAACTGCG ACGATGGGT ACTGAGCGCA
121 | GTAATTCCAC GGAATTTTCT TCTCATTATA GTCATCTGAA GTCCTCAGAA TCGCTGATAC
181 | GCAAGAGATTA AGAACATTTA TGATCTGATA TTTACGATGT ACGATGATGC GCGCTAAGTG
241 | AAAGGATCCA GGTAGAGATC TTTTTTGTAC ATCTAGATGC CGAATGACAT TGGTGATGTA
301 | TTATGAAAAA AAAACCAAATG GGTTCGCGCC GACTTTTCTC GCGAAGATGT CCAACCTGAC
361 | GCTGCACATG TCTTTTGAAC TCTATTACG GCCCATGATC AGAAGACGCA
421 | GAAAGGGGGA ACGAGACTGT TCTGCATAAG TCTGACAGCA TGGGATGTTA CGGAGCTGAC
481 | GGCCA GTGAGCGCGC GTAAGGACAG TGGCAGATT ACGATCCTCT CACTATTAGT
541 | TACCGGGCCC CCCCTTACG ACGAAAGAAG GCTGCGCTTC GAGGAAAGG GTTCTGAAGT
601 | CGGCGGATTT GTCCTGAGAC ACAGGGGACA ACGAGGACA ACGGGGAGA CGGAGGAGT
661 | CCAGTCTTTC GACTGCGCTG TACGCGACG CGAGGCTGGA GAAAGCAG GCCATTATCG
721 | GAATCTTG CACGCCCTCG GCTGGAGCGA GACCGCTTTC GACTGTCT TCTTTATCAT
781 | CTTCTCCTCC GGGTCCCTTG CTTCTTGGCC CTCTGACCG ACGGAAGGCA CGGAGGAGT
841 | TTTCGCC TCGCTCACTG ACCATAGTCT TGGTGTGTTA CTTGACCAG TGTATGTCCAA
901 | TGGCCGACG CGAGGCTGGA GAAAGCAG GCCATTATCG GCTGGAGCGA GACCGCTTTC
961 | ACGGA GAGGTTTCT CTTCACTCG CCCATGCCCT CGCTTGAGAC
1021 | CGGTGTTCCG CGGGAGCCCA GGGATGGGGT ACGATGTGTC
1081 | GCAACTGCTG ACCACCCGGG CTTCTCCAGA GTCATCAAG CGCCTCACTC GAAAACCTGC
1141 | GGGGAAGACG ACTTCAAGTT AGCGGGAAG TCTGCCGAGA GGCGGAGGA GCGCTTCAGA
1201 | GGCCTTACAG AATCGAGAGG AGACGAGAGA GAGGGGAGGA CGACCTCCGT GCCCTGCTG
1261 | CCTCTCCCGG TGGCCCTTAC GAGGAGGAG TCTGAGGAGA GGCGGAGGA GAGGGGAGA
1321 | TCTGCAACTG CGCGGCTTTC TGGAGGCGCG CTTCCCTCAG CAGATGGCAG
1381 | GATAATAACC ATCTAGATCC CCTCTTGCGT CCTCTGACCA GAAAATCCCA TGGATGACAG
1441 | ACCAATATCA CGCGCAGGAG TAAAAATGTC AACACGCACT GTGTGAATA TTTAATAATA
1501 | CCTGTTGACTA AGGCGTACAT TCTGCTACAG AGCAAGATCA AGAAGGAGA AATACGATGCT
1561 | GAAGAAGTCA GTACCGGTGT GTTTCGCTCA CGTTGAGGAT TTGGATGAGT
1621 | CATATAATCA GTGGTCCGCG CAGAAGTGAA GGGATGCGA CGGGAGCAG ACTGACCTGC
1681 | AAATATATCT GCACAAAGGC TAAAAAAGCC GTCGTTGATG GACGCTGAG GACGGCGGT
1741 | ACCATGATGG CGGTTAAGGT CTTCGAGCTC GCAGGAGAAA ACTGACCTGC
1801 | AAATGCGCA CGCAGGAGGA TTCTACTTTA AGCAGGGGAT
1861 | AGCTAATAAAA CCAGGAGGAT ATGGTATTCA AGAAGGTGGA GCTGTGGAGA
1921 | CTGGAAGATG CGTACCTGGA AGAAGGCGCC AATTTGCTGC GCATTAAACCT GGAATTAAC
1981 | TTTAACTTCC ACAAGCTTG TA CATACGACT GGCGGAGGAT TGACCAGAG GAAACCAAA
2041 | TTCAAAATCC GCCAAGTTG GAGAAGGAG TGGTGTGAGA TGGGCGGCG GCCCTCGGCG
2101 | AAGCCAGCGCA TGGTGTGAGA CCGCGCTTCT GCAGATCGAT GGATGAGAG
2161 | TCCGGCTGTG CAAAAGATCC GAAAGGAAAA ATGACGCAAC TGGAATGAGA
2221 | ACCGCGTCCG GTATACCAAC CGCAGATGCG GAACTGTATA AATAAACCAA AAAACCACCC
2281 | TCGCGCGGTA TTTTTTGCGG TCACGATCGA CTCCGATCGA CTCCGATCGA CGCGCGCGGA
2341 | GACGCGTACAG CAGCGGCGCTT CACGAGTTAC TACAGGATGA CGAAGATGAGA
2401 | CAGAAAGGAA CACGAGTATAG GTTACGATG CTCTACGAGT CGAAGATGAGA
2461 | CTCGAGTGCG GAAGCGGATG GCCCTTGGAGA CTTCCCATAC CGATGAGCT TCTCCGGGAG
2521 | GCGCGGCGCA TGACTATCG CCGCGACTT AGACTGATC TCTTATACAT GCACTGATGC
2581 | CGCAGGCGCGA TGGGCGGTTG TGCGGTGATG CGGAACTGCT CGGAACTGCT CGGAACTGCT
2641 | GACGATGATCC GCTGCTGTGT CGGAACTGCT CGGAACTGCT CGGAACTGCT
2701 | CGTCATCTTG CCGCGAACCA AAGGTTCGAC CGGAAGAAGC GCGAATGAGA
2761 | CGCGCCAGCC TGGCGTCAGC TTTCCGAGGC AAGGCGGCGA AGCCGTTAAG
2821 | CAGTGTATGC TCTTACGCGT CCGCGGCAAC CTTGCGACCG CTTGCGACCG
2881 | CAGGCGGCGA TCTGTGACAT GTGCTGCGGC GCCGTTACG GCCGTTACG
2941 | CCTACCTGAC CGTACGCTGG CGTGACTGAT CGGCAAAGCT TACGAGCTT CGGAACTGCT
3001 | ATGGACAGCC TGGCGATGTT TGTAGGCGCC CCGGTCATAC TTGCGCTC CGCGGGTGTT
3061 | GCCTGCGGCG GCATGGAACGG GCAGCACTCG GACTGATGAA GAGCGGCGCG GCGACCGCT
