(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
06 December 2018 (06.12.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/222891 A1

(51) International Patent Classification:
C12N 15/86 (2006.01) *C12N 15/90* (2006.01)
C12N 7/00 (2006.01)

(21) International Application Number:
PCT/US2018/035420

(22) International Filing Date:
31 May 2018 (31.05.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/513,707 01 June 2017 (01.06.2017) US

(71) Applicant: **THE CHARLES STARK DRAPER LABORATORY INC.** [US/US]; 555 Technology Square, Cambridge, Massachusetts 02139 (US).

(72) Inventors: **MCBRINE, Connor**; 344 Summer Street, Unit 1, Somerville, Massachusetts 02144 (US). **HOLDER, Jason**; 11 Bay View Drive, Swampscott, Massachusetts 01907 (US).

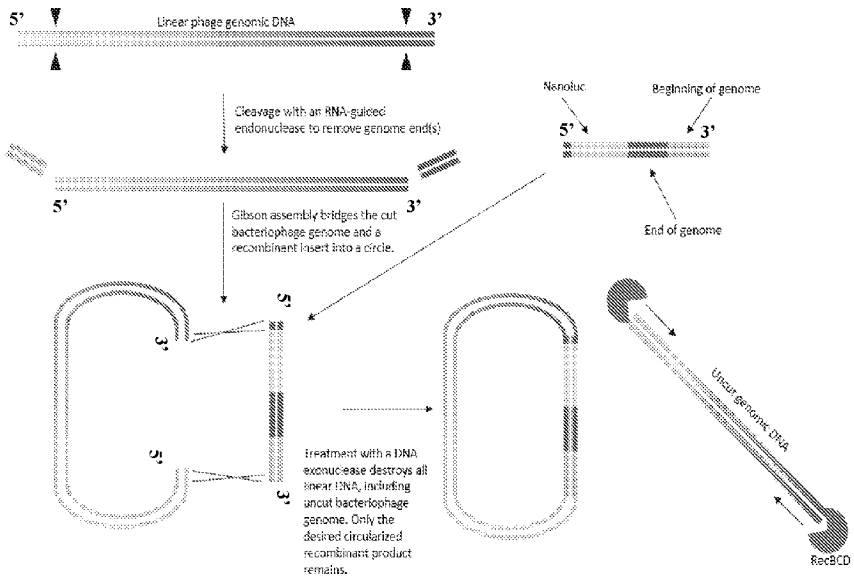
(74) Agent: **EWING, James F.** et al.; Foley & Lardner LLP, 3000 K. Street, N.W., Suite 600, Washington, District of Columbia 20007 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

(54) Title: PHAGE ENGINEERING: PROTECTION BY CIRCULARIZED INTERMEDIATE

Figure 1



(57) Abstract: The present disclosure provides methods of generating recombinant bacteriophage genomes. Specifically, the present technology provides methods of integrating a heterologous nucleic acid sequence into a linear bacteriophage DNA genome, and isolating recombinant bacteriophages that express the heterologous nucleic acid sequence.

WO 2018/222891 A1

[Continued on next page]

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

— *with international search report (Art. 21(3))*

PHAGE ENGINEERING: PROTECTION BY CIRCULARIZED INTERMEDIATE

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of and priority to US Provisional Appl. No. 62/513,707, filed June 1, 2017, the disclosures of which are incorporated by reference herein in their entireties.

TECHNICAL FIELD

[0002] The present technology relates generally to methods and kits for generating recombinant bacteriophage genomes. In particular, the present technology relates to methods of integrating a heterologous nucleic acid sequence into a linear bacteriophage DNA genome, and isolating recombinant bacteriophages that express the heterologous nucleic acid sequence.

BACKGROUND

[0003] The following description of the background of the present technology is provided simply as an aid in understanding the present technology and is not admitted to describe or constitute prior art to the present technology.

[0004] Model phages have been engineered using molecular biology techniques to deliver heterologous protein products to bacterial cells. *E.g.*, US 2009/0155215; M. J. Loessner *et. al.*, *Applied and Environmental Microbiology*, Vol. 62, No. 4, pp. 1133-40 (1996)). The natural host range of model phage engineered to date is limited. Methods for creating variations in phage genomes and engineering new phage genomes may lead to the identification of phages with varied properties (*e.g.*, varied host ranges) that are useful for diagnostic and therapeutic purposes.

[0005] Engineering diverse phage is generally made more difficult by the properties of phage genomes. For example, phage genomes have relatively few restriction sites and are heavily modified, making use of traditional cloning techniques with phage challenging. Phages also have compact genomes with very little non-coding DNA, which can make it challenging to find sites within the genome that are compatible with traditional engineering. Many existing phage engineering technologies that rely on *in vitro* strategies are generally inefficient and challenging to scale up. Further, engineering phages within bacteria can be problematic due to toxicity of phages to bacteria as well as the difficulty in maintaining the stability of large engineered genomes.

SUMMARY OF THE PRESENT TECHNOLOGY

[0006] In one aspect, the present disclosure provides a method for integrating a heterologous nucleic acid into a linear bacteriophage DNA genome comprising: (a) cleaving a first site at the 5' end of a plurality of linear bacteriophage DNA genomes and a second site at the 3' end of the plurality of linear bacteriophage DNA genomes with a CRISPR enzyme *in vitro*, wherein the plurality of linear bacteriophage DNA genomes is present in an isolated sample; (b) recombining *in vitro* the cleaved plurality of linear bacteriophage DNA genomes with the heterologous nucleic acid in the presence of a recombination system, wherein the heterologous nucleic acid comprises a 3' flanking region and a 5' flanking region that are homologous to the 5' and 3' ends of the cleaved plurality of linear bacteriophage DNA genomes respectively, thereby generating a plurality of circularized bacteriophage DNA genomes; and (c) enriching the plurality of circularized bacteriophage DNA genomes by incubating the sample with at least one exonuclease. The at least one exonuclease may be bidirectional or unidirectional. Additionally or alternatively, in some embodiments, the at least one exonuclease is selected from the group consisting of Lambda exonuclease, Exonuclease III, RecBCD, Exonuclease VIII truncated, T5 exonuclease, and T7 exonuclease. In some embodiments, the at least one exonuclease is RecBCD. In certain embodiments, the method further comprises propagating the plurality of circularized bacteriophage DNA genomes in a non-natural bacterial host. The plurality of linear bacteriophage DNA genomes may be recombinant or non-recombinant. In some embodiments, the linear bacteriophage DNA genome is a terminally redundant linear bacteriophage DNA genome.

[0007] In certain embodiments, the plurality of linear bacteriophage DNA genomes correspond to a bacteriophage family or order selected from the group consisting of *Myoviridae, Styloviridae, Siphoviridae, Pedoviridae, Tectiviridae, Leviviridae, Podoviridae,* and *Plasmaviridae*. In some embodiments, the plurality of linear bacteriophage DNA genomes are derived from one or more bacteriophage genuses (or genera) selected from the group consisting of T7-like phage, phiKMV-like phage, LUZ24-like phage, phiKZ-like phage, PB1-like phage, Felix-O1-like phage, T4-like phage, phi92-like phage, rV5-like phage, SP6-like phage, N4-like phage, phiEco32-like phage, T5-like phage, KP34-like phage, KP15-like phage, GAP227-like phage, AP22-like phage, phiFel-like phage, Sap6-like phage, Silvia-like phage, Kay-like phage, Twort-like phage, P68-like phage, and phiETA-like phage.

[0008] Additionally or alternatively, in some embodiments, the plurality of linear bacteriophage DNA genomes correspond to *Klebsiella* phage K11, lambda phage,

Enterobacteria phage T2, *Enterobacteria* phage T1, *Enterobacteria* phage T7, *Enterobacteria* phage T5, *Enterobacteria* phage P1, *Enterobacteria* phage PRD1, K1E phage, K1-5 phage, RB49 phage, RB16 phage, KP15 phage, KP27 phage, Miro phage, Matisse phage, phiEap-3 phage, ECP3 phage, EFDG1 phage, EFLK1 phage, vB_Efae230P-4 phage, vB_EfaP_IME195 phage, SA11 phage, Stau2 phage, K phage, G1 phage, SA12 phage, 812 phage, P68 phage, SAP-2 phage, 44AHJD phage, or SA97 phage.

[0009] Additionally or alternatively, in some embodiments of the method, the CRISPR enzyme is a Cas protein selected from the group consisting of Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, and Csf4. The CRISPR enzyme may be coupled to a sgRNA.

[0010] Additionally or alternatively, in some embodiments of the method, the recombination system comprises a 5'-3' exonuclease, a DNA polymerase, and a DNA ligase. In one embodiment, the 5'-3' exonuclease is T5 exonuclease, the DNA polymerase is Phusion® DNA polymerase (Thermo Fisher Scientific, Waltham, MA), and the DNA ligase is *Taq* ligase. In other embodiments, the recombination system comprises a 3'-5' exonuclease, a DNA polymerase, and a DNA ligase.

[0011] In any of the above embodiments, the heterologous nucleic acid comprises an open reading frame that encodes a bioluminescent protein, a fluorescent protein, a chemiluminescent protein, or any combination thereof. In certain embodiments, the open reading frame of the heterologous nucleic acid is operably linked to an expression control sequence that is capable of directing expression of the bioluminescent protein, the fluorescent protein, the chemiluminescent protein, or any combination thereof. In some embodiments, the expression control sequence is an inducible promoter or a constitutive promoter. The heterologous nucleic acid can be about 100-500 base pairs in length, about 500-1000 base pairs in length, 1000-1500 base pairs in length, about 1500-2000 base pairs in length, 2000-2500 base pairs in length, about 2500-3000 base pairs in length, 3000-3500 base pairs in length, or about 3500-4000 base pairs in length.

[0012] Examples of bioluminescent protein include, but are not limited to, Aequorin, firefly luciferase, *Renilla* luciferase, red luciferase, luxAB, and nanoluciferase. Examples of chemiluminescent protein include β-galactosidase, horseradish peroxidase (HRP), and

alkaline phosphatase. Examples of fluorescent protein include, but are not limited to, TagBFP, Azurite, EBFP2, mKalama1, Sirius, Sapphire, T-Sapphire, ECFP, Cerulean, SCFP3A, mTurquoise, monomeric Midoriishi-Cyan, TagCFP, mTFP1, EGFP, Emerald, Superfolder GFP, Monomeric Azami Green, TagGFP2, mUKG, mWasabi, EYFP, Citrine, Venus, SYFP2, TagYFP, Monomeric Kusabira-Orange, mKOκ, mKO2, mOrange, mOrange2, mRaspberry, mCherry, dsRed, mStrawberry, mTangerine, tdTomato, TagRFP, TagRFP-T, mApple, mRuby, mPlum, HcRed-Tandem, mKate2, mNeptune, NirFP, TagRFP657, IFP1.4, iRFP, mKeima Red, LSS-mKate1, LSS-mKate2, PA-GFP, PAmCherry1, PATagRFP, Kaede (green), Kaede (red), KikGR1 (green), KikGR1 (red), PS-CFP2, PS-CFP2, mEos2 (green), mEos2 (red), PSmOrange, or Dronpa.

[0013] In some embodiments of the method, the plurality of linear bacteriophage DNA genomes correspond to *Klebsiella* phage K11, wherein the first site at the 5' end of the plurality of linear bacteriophage DNA genomes is after the nucleotide located at position 232 of SEQ ID NO:1 and the second site at the 3' end of the plurality of linear bacteriophage DNA genomes is after the nucleotide located at position 40,715 of SEQ ID NO:1. In certain embodiments, the CRISPR enzyme is Cas9 and is coupled to a single-guide RNA (sgRNA) having the sequence 5' GCCACCUGAGGUUAGACCAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:2) or 5' GGCUACUUGGCGAUUUCUGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:3).

[0014] In another aspect, the present disclosure provides a method for making an enriched plurality of recombinant circularized bacteriophage DNA genomes from a plurality of linear bacteriophage DNA genomes comprising: (a) contacting a sample comprising the plurality of linear bacteriophage DNA genomes with a first sgRNA-CRISPR enzyme complex and a second sgRNA-CRISPR enzyme complex *in vitro* under conditions where (i) the first sgRNA-CRISPR enzyme complex cleaves a first site at the 5' end of the plurality of linear bacteriophage DNA genomes; and (ii) the second sgRNA-CRISPR enzyme complex cleaves a second site at the 3' end of the plurality of linear bacteriophage DNA genomes; (b) recombining *in vitro* the cleaved plurality of linear bacteriophage DNA genomes with a heterologous nucleic acid in the presence of a recombination system under conditions to

produce recombinant circularized bacteriophage DNA genomes; and (c) contacting the sample with at least one exonuclease to obtain an enriched plurality of recombinant circularized bacteriophage DNA genomes. The at least one exonuclease may be bidirectional or unidirectional. Additionally or alternatively, in some embodiments, the at least one exonuclease is selected from the group consisting of Lambda exonuclease, Exonuclease III, RecBCD, Exonuclease VIII truncated, T5 exonuclease, and T7 exonuclease. In some embodiments, the at least one exonuclease is RecBCD. The plurality of linear bacteriophage DNA genomes may be recombinant or non-recombinant.

[0015] The heterologous nucleic acid comprises a 3' flanking region and a 5' flanking region that are homologous to the 5' and 3' ends of the cleaved plurality of linear bacteriophage DNA genomes respectively. In certain embodiments, the method further comprises propagating the plurality of bacteriophage circularized genomes in a non-natural bacterial host.

[0016] In some embodiments, the plurality of linear bacteriophage DNA genomes correspond to *Klebsiella* phage K11. In certain embodiments, the first sgRNA-CRISPR enzyme complex comprises the sequence 5' GCCACCUGAGGUUAGACCAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:2), and the second gRNA-CRISPR enzyme complex comprises the sequence 5' GGCUACUUGGCGAUUUCUGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:3).

[0017] In another aspect, the present disclosure provides a method for integrating a heterologous nucleic acid into a non-terminally redundant linear bacteriophage DNA genome. In some embodiments, the method comprises: (a) contacting a sample comprising a linear bacteriophage DNA genome with a sgRNA-CRISPR enzyme complex *in vitro* under conditions where the sgRNA-CRISPR enzyme complex cleaves a first site at the 3' end of the linear bacteriophage DNA genome to produce a cleaved bacteriophage genome, wherein the cleaved bacteriophage genome comprises genomic sequence that is upstream of the first site; (b) recombining *in vitro* the cleaved bacteriophage genome with a heterologous nucleic acid in the presence of a recombination system under conditions to produce a circularized recombinant bacteriophage genome; and (c) contacting the sample with at least one exonuclease to enrich for the circularized recombinant bacteriophage genome. In other

embodiments, the method comprises: (a) contacting a sample comprising a linear bacteriophage DNA genome with a sgRNA-CRISPR enzyme complex *in vitro* under conditions where the sgRNA-CRISPR enzyme complex cleaves a first site at the 5' end of the linear bacteriophage DNA genome to produce a cleaved bacteriophage genome, wherein the cleaved bacteriophage genome comprises genomic sequence that is downstream of the first site; (b) recombining *in vitro* the cleaved bacteriophage genome with a heterologous nucleic acid in the presence of a recombination system under conditions to produce a circularized recombinant bacteriophage genome; and (c) contacting the sample with at least one exonuclease to enrich for the circularized recombinant bacteriophage genome. In any of the above embodiments of the method, the heterologous nucleic acid comprises a 5' flanking region that is homologous to the 3' end of the cleaved bacteriophage genome, and a 3' flanking region that is homologous to the 5' end of the cleaved bacteriophage genome. In any of the above embodiments, the linear bacteriophage DNA genome may be recombinant or non-recombinant.

[0018] Additionally or alternatively, in some embodiments of the methods of the present technology, the at least one exonuclease is selected from the group consisting of Lambda exonuclease, Exonuclease III, RecBCD, Exonuclease VIII truncated, T5 exonuclease, and T7 exonuclease. The at least one exonuclease may be bidirectional or unidirectional. In some embodiments, the at least one exonuclease is RecBCD. In certain embodiments, the methods disclosed herein further comprise propagating the circularized recombinant bacteriophage genome in a bacterial host. The non-terminally redundant linear bacteriophage DNA genome may correspond to a bacteriophage family or order selected from the group consisting of *Myoviridae, Styloviridae, Siphoviridae, Pedoviridae, Tectiviridae, Leviviridae, Podoviridae,* and *Plasmaviridae*.

[0019] Additionally or alternatively, in some embodiments of the methods of the present technology, the recombination system comprises a 5'-3' exonuclease, a DNA polymerase, and a DNA ligase. In one embodiment, the 5'-3' exonuclease is T5 exonuclease, the DNA polymerase is Phusion® DNA polymerase (Thermo Fisher Scientific, Waltham, MA), and the DNA ligase is *Taq* ligase. In other embodiments, the recombination system comprises a 3'-5' exonuclease, a DNA polymerase, and a DNA ligase.

[0020] Additionally or alternatively, in some embodiments of the methods of the present technology, the CRISPR enzyme is a Cas protein selected from the group consisting of Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Csy1, Csy2, Csy3, Cse1,

Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, and Csf4.

[0021] In any of the above embodiments of the methods disclosed herein, the heterologous nucleic acid comprises an open reading frame that encodes a bioluminescent protein, a fluorescent protein, a chemiluminescent protein, or any combination thereof. In certain embodiments, the open reading frame of the heterologous nucleic acid is operably linked to an expression control sequence that is capable of directing expression of the bioluminescent protein, the fluorescent protein, the chemiluminescent protein, or any combination thereof. In some embodiments, the expression control sequence is an inducible promoter or a constitutive promoter. The heterologous nucleic acid can be about 100-500 base pairs in length, about 500-1000 base pairs in length, 1000-1500 base pairs in length, about 1500-2000 base pairs in length, 2000-2500 base pairs in length, about 2500-3000 base pairs in length, 3000-3500 base pairs in length, or about 3500-4000 base pairs in length.

[0022] Examples of bioluminescent protein include, but are not limited to, Aequorin, firefly luciferase, *Renilla* luciferase, red luciferase, luxAB, and nanoluciferase. Examples of chemiluminescent protein include β-galactosidase, horseradish peroxidase (HRP), and alkaline phosphatase. Examples of fluorescent protein include, but are not limited to, TagBFP, Azurite, EBFP2, mKalama1, Sirius, Sapphire, T-Sapphire, ECFP, Cerulean, SCFP3A, mTurquoise, monomeric Midoriishi-Cyan, TagCFP, mTFP1, EGFP, Emerald, Superfolder GFP, Monomeric Azami Green, TagGFP2, mUKG, mWasabi, EYFP, Citrine, Venus, SYFP2, TagYFP, Monomeric Kusabira-Orange, mKOκ, mKO2, mOrange, mOrange2, mRaspberry, mCherry, dsRed, mStrawberry, mTangerine, tdTomato, TagRFP, TagRFP-T, mApple, mRuby, mPlum, HcRed-Tandem, mKate2, mNeptune, NirFP, TagRFP657, IFP1.4, iRFP, mKeima Red, LSS-mKate1, LSS-mKate2, PA-GFP, PAmCherry1, PATagRFP, Kaede (green), Kaede (red), KikGR1 (green), KikGR1 (red), PS-CFP2, PS-CFP2, mEos2 (green), mEos2 (red), PSmOrange, or Dronpa.

[0023] Also disclosed herein are kits for integrating a heterologous nucleic acid sequence into a linear bacteriophage DNA genome.

BRIEF DESCRIPTION OF THE DRAWINGS

[0024] Figure 1 shows the schematic representation of the circular payload integration method disclosed herein.

[0025] Figure 2 shows the efficient recovery of recombinant bacteriophages containing the 907 base pairs (bp) nanoluciferase sequence using the methods of the present technology. Junctional PCR amplified a 907 bp product expected only in recombinant phages (13 out of 16 isolates). Primers span from inside the nanoluciferase insertion site to downstream of the nanoluciferase insert. The three isolates that yielded amplicons of different sizes correspond to viable phage that were incorrectly assembled during Gibson assembly or had otherwise incorporated an expanded or truncated insert.

[0026] Figure 3 shows the complete genome sequence of *Klebsiella* phage K11 (SEQ ID NO: 1).

[0027] Figure 4 shows the luminescence activity profile of a recombinant K11 phage. The recombinant K11 phage was picked and mixed into 25 µL of Tris-HCl with 10 mM $MgSO_4$ and was subsequently diluted by 1:100. 1 µL of diluted phage was added to 100 µL LB media, alone or in combination with 100 µL log-phase *K. pneumoniae* strain Kp390. Both the LB and *K. pneumoniae* cultures contained 10 mM $MgSO_4$. Paired negative controls that lacked phage were also included. The infections/mock infections were incubated at 37°C for 1 hour.

[0028] Figure 5 shows the design of the K11 chimeric guide RNA expression construct used in the Break and Recombine 3.0 (BAR 3.0) experiments.

[0029] Figure 6 shows a representative gel image of K11 genomic DNA after cleavage with the Cas9/ sgRNA 4.5 complex.

[0030] Figure 7 shows the synthetic DNA construct used to introduce the nanoluciferase gene into the cleaved K11 genome *via* BAR 3.0.

DETAILED DESCRIPTION

[0031] It is to be appreciated that certain aspects, modes, embodiments, variations and features of the present methods are described below in various levels of detail in order to provide a substantial understanding of the present technology.

[0032] Manipulating phage genomes is more difficult compared to manipulating bacterial hosts. *In vitro* synthesis and assembly of phage genomes is inefficient and relies on the delivery of large DNA molecules across the cell membranes of a bacterial host. Some bacterial strains are recalcitrant to large DNA transformation across the membrane. Classic

in vivo recombination strategies are also inefficient and are complicated by the fact that lytic phage genomes have a comparatively short residence time in a host before lysis.

[0033] One of the most commonly used and well-established methods for engineering phage genomes is homologous recombination in their bacterial hosts, which can occur between two homologous DNA sequences as short as 23 bp (Alberts B *et al.*, MOLECULAR BIOLOGY OF THE CELL, 5th ed. Garland Science, New York, NY (2007); Snyder L *et al.*, MOLECULAR GENETICS OF BACTERIA, 4th ed. ASM Press, Washington, DC (2013)). Homologous recombination occurs between the plasmid and the phage genome, allowing the heterologous gene to be integrated into the phage genome and eventually packaged within the phage particle. However, homologous recombination only yields a small fraction of recombinant progeny phage. Reported recombination rates range from 10^{-10} to 10^{-4} (Loessner M. *et al.*, *Appl Environ Microbiol* 62:1133–1140 (1996); Le S. *et al.*, *PLoS One* 8:e68562 (2013); Mahichi F. *et al.*, *FEMS Microbiol Lett* 295:211–217 (2009)). One of the major challenges of generating recombinant bacteriophages is that the recombinant processes used to create such bacteriophages are inefficient, and often result in a low yield of recombinant bacteriophage genomes. Transformation of large bacteriophage genomes (*e.g.*, about or greater than 40-48 kb) is prohibitive in many bacterial strains and species, making it difficult to isolate viable bacteriophage particles post-transformation. *See e.g.*, Chauthaiwale *et al.*, *Microbiological Reviews* 56 (4): 577-592 (1992); see also Vaughan *et al.*, *Nature Biotechnology* 14:309-314 (1996). Thus, finding the desired clone using conventional phage screening methods is labor-intensive and unpredictable.

[0034] The present disclosure provides methods for integrating a heterologous nucleic acid sequence into a linear bacteriophage DNA genome, and isolating recombinant bacteriophages that express the heterologous nucleic acid sequence. The methods disclosed herein permit higher recovery of recombinant bacteriophage genomes that express the phenotypic properties associated with the heterologous nucleic acid sequence relative to that observed with other phage engineering methods, such as Break and Recombine 3.0 (BAR 3.0). For example, the overall yield of recombinant bacteriophage genomes obtained using the methods of the present technology was about 81% (13 out of 16 isolates). In contrast, no recombinant bacteriophages were generated using BAR 3.0 (*i.e.*, 0% recovery of recombinant bacteriophage genomes).

[0035] In practicing the present methods, many conventional techniques in molecular biology, protein biochemistry, cell biology, microbiology and recombinant DNA are used.

See, e.g., Sambrook and Russell eds. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edition; the series Ausubel *et al.* eds. (2007) *Current Protocols in Molecular Biology*; the series *Methods in Enzymology* (Academic Press, Inc., N.Y.); MacPherson *et al.* (1991) *PCR 1: A Practical Approach* (IRL Press at Oxford University Press); MacPherson *et al.* (1995) *PCR 2: A Practical Approach*; Harlow and Lane eds. (1999) *Antibodies, A Laboratory Manual*; Freshney (2005) *Culture of Animal Cells: A Manual of Basic Technique*, 5th edition; Gait ed. (1984) *Oligonucleotide Synthesis*; U.S. Patent No. 4,683,195; Hames and Higgins eds. (1984) *Nucleic Acid Hybridization*; Anderson (1999) *Nucleic Acid Hybridization*; Hames and Higgins eds. (1984) *Transcription and Translation; Immobilized Cells and Enzymes* (IRL Press (1986)); Perbal (1984) *A Practical Guide to Molecular Cloning*; Miller and Calos eds. (1987) *Gene Transfer Vectors for Mammalian Cells* (Cold Spring Harbor Laboratory); Makrides ed. (2003) *Gene Transfer and Expression in Mammalian Cells*; Mayer and Walker eds. (1987) *Immunochemical Methods in Cell and Molecular Biology* (Academic Press, London); and Herzenberg *et al.* eds (1996) *Weir's Handbook of Experimental Immunology*.

Definitions

[0036] Unless defined otherwise, all technical and scientific terms used herein generally have the same meaning as commonly understood by one of ordinary skill in the art to which this technology belongs. As used in this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the content clearly dictates otherwise. For example, reference to "a cell" includes a combination of two or more cells, and the like. Generally, the nomenclature used herein and the laboratory procedures in cell culture, molecular genetics, organic chemistry, analytical chemistry and nucleic acid chemistry and hybridization described below are those well-known and commonly employed in the art.

[0037] As used herein, the term "about" in reference to a number is generally taken to include numbers that fall within a range of 1%, 5%, or 10% in either direction (greater than or less than) of the number unless otherwise stated or otherwise evident from the context (except where such number would be less than 0% or exceed 100% of a possible value).

[0038] As used herein, "bacteriophage" or "phage" refers to a virus that infects bacteria. Bacteriophages are obligate intracellular parasites that multiply inside bacteria by co-opting some or all of the host biosynthetic machinery (*i.e.*, viruses that infect bacteria). Though different bacteriophages may contain different materials, they all contain nucleic acid and protein, and can under certain circumstances be encapsulated in a lipid membrane.

Depending upon the phage, the nucleic acid can be either DNA or RNA (but not both) and can exist in various forms.

[0039] As used herein, "expression" includes one or more of the following: transcription of the gene into precursor mRNA; splicing and other processing of the precursor mRNA to produce mature mRNA; mRNA stability; translation of the mature mRNA into protein (including codon usage and tRNA availability); and glycosylation and/or other modifications of the translation product, if required for proper expression and function.

[0040] As used herein, an "expression control sequence" refers to polynucleotide sequences which are necessary to affect the expression of coding sequences to which they are operably linked. Expression control sequences are sequences which control the transcription, post-transcriptional events and translation of nucleic acid sequences. Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (*e.g.*, ribosome binding sites); sequences that enhance protein stability; and when desired, sequences that enhance protein secretion. The nature of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include promoter, ribosomal binding site, and transcription termination sequence. The term "control sequences" is intended to encompass, at a minimum, any component whose presence is essential for expression, and can also encompass an additional component whose presence is advantageous, for example, leader sequences.

[0041] As used herein, "heterologous nucleic acid sequence" is any sequence placed at a location in the genome where it does not normally occur. A heterologous nucleic acid sequence may comprise a sequence that does not naturally occur in a bacteriophage, or it may comprise only sequences naturally found in the bacteriophage, but placed at a non-normally occurring location in the genome. In some embodiments, the heterologous nucleic acid sequence is not a natural phage sequence. In certain embodiments, the heterologous nucleic acid sequence is a natural phage sequence that is derived from a different phage. In other embodiments, the heterologous nucleic acid sequence is a sequence that occurs naturally in the genome of a wild-type phage but is then relocated to another site where it does not naturally occur, rendering it a heterologous sequence at that new site.

[0042] "Homology" or "identity" or "similarity" refers to sequence similarity between two peptides or between two nucleic acid molecules. Homology can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same nucleobase or amino acid, then the molecules are homologous at that position. A degree of homology between sequences is a function of the number of matching or homologous positions shared by the sequences. A polynucleotide or polynucleotide region (or a polypeptide or polypeptide region) has a certain percentage (for example, at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99%) of "sequence identity" to another sequence means that, when aligned, that percentage of bases (or amino acids) are the same in comparing the two sequences. This alignment and the percent homology or sequence identity can be determined using software programs known in the art. In some embodiments, default parameters are used for alignment. One alignment program is BLAST, using default parameters. In particular, programs are BLASTN and BLASTP, using the following default parameters: Genetic code=standard; filter=none; strand=both; cutoff=60; expect=10; Matrix=BLOSUM62; Descriptions=50 sequences; sort by =HIGH SCORE; Databases=non-redundant, GenBank+EMBL+DDBJ+PDB+GenBank CDS translations+SwissProtein+SPupdate+PIR. Details of these programs can be found at the National Center for Biotechnology Information. Biologically equivalent polynucleotides are those having the specified percent homology and encoding a polypeptide having the same or similar biological activity. Two sequences are deemed "unrelated" or "non-homologous" if they share less than 40% identity, or less than 25% identity, with each other.

[0043] As used herein, a "host cell" is a bacterial cell that can be infected by a phage to yield progeny phage particles. A host cell can form phage particles from a particular type of phage genomic DNA. In some embodiments, the phage genomic DNA is introduced into the host cell by infecting the host cell with a phage. In some embodiments, the phage genomic DNA is introduced into the host cell using transformation, electroporation, or any other suitable technique. In some embodiments, the phage genomic DNA is substantially pure when introduced into the host cell. In some embodiments, the phage genomic DNA is present in a vector when introduced into the host cell. The definition of host cell can vary from one phage to another. For example, *E. coli* may be the natural host cell for a particular type of phage, but *Klebsiella pneumoniae* is not.

[0044] As used herein, the term "isolated" refers to a substance or entity that has been separated from at least some of the components with which it was associated when initially produced (whether in nature or in an experimental setting). Isolated substances and/or entities may be separated from at least about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, or more of the other components with which they were initially associated. In some embodiments, isolated substances and/or entities are more than about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more than about 99% pure. As used herein, a substance is "pure" if it is substantially free of other components.

[0045] As used herein, "operably linked" means that expression control sequences are positioned relative to the nucleic acid of interest to initiate, regulate or otherwise control transcription of the nucleic acid of interest.

[0046] As used herein, a "phage genome" includes naturally occurring phage genomes and derivatives thereof. Generally, the derivatives possess the ability to propagate in the same hosts as the naturally occurring phage. In some embodiments, the only difference between a naturally occurring phage genome and a derivative phage genome is at least one of a deletion and an addition of nucleotides from at least one end of the phage genome (if the genome is linear) or at least one point in the genome (if the genome is circular).

[0047] As used herein, the term "polynucleotide" or "nucleic acid" means any RNA or DNA, which may be unmodified or modified RNA or DNA. Polynucleotides include, without limitation, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, RNA that is mixture of single- and double-stranded regions, and hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, polynucleotide refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term polynucleotide also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons.

[0048] As used herein, the term "recombinant" when used with reference, *e.g.*, to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration

of a native nucleic acid or protein, or that the material is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

[0049] As used herein, an endogenous nucleic acid sequence in the genome of an organism (or the encoded protein product of that sequence) is deemed "recombinant" herein if a heterologous sequence is placed adjacent to the endogenous nucleic acid sequence, such that the expression of this endogenous nucleic acid sequence is altered. In this context, a heterologous sequence is a sequence that is not naturally adjacent to the endogenous nucleic acid sequence, whether or not the heterologous sequence is itself endogenous to the organism (originating from the same organism or progeny thereof) or exogenous (originating from a different organism or progeny thereof). By way of example, a promoter sequence can be substituted (*e.g.*, by homologous recombination) for the native promoter of a gene in the genome of an organism, such that this gene has an altered expression pattern. This gene would be "recombinant" because it is separated from at least some of the sequences that naturally flank it. A nucleic acid is also considered "recombinant" if it contains any modifications that do not naturally occur in the corresponding nucleic acid in a genome. For instance, an endogenous coding sequence is considered "recombinant" if it contains an insertion, deletion or a point mutation introduced artificially, *e.g.*, by human intervention. A "recombinant nucleic acid" also includes a nucleic acid integrated into a host cell chromosome at a heterologous site and a nucleic acid construct present as an episome.

[0050] As used herein, a "recombinant bacteriophage genome" is a bacteriophage genome that has been genetically modified by the insertion of a heterologous nucleic acid sequence into the bacteriophage genome. A "recombinant bacteriophage" means a bacteriophage that comprises a recombinant bacteriophage genome. In some embodiments, the bacteriophage genome is modified by recombinant DNA technology to introduce a heterologous nucleic acid sequence into the genome at a defined site. In some embodiments, the heterologous nucleic acid sequence is introduced with no corresponding loss of endogenous phage genomic nucleotides. In other words, if bases N1 and N2 are adjacent in the wild-type bacteriophage genome, the heterologous nucleic acid sequence is inserted between N1 and N2. Thus, in the resulting recombinant bacteriophage genome, the heterologous nucleic acid sequence is flanked by nucleotides N1 and N2. In some embodiments, endogenous phage nucleotides are removed or replaced during the insertion of

the heterologous nucleic acid sequence. For example, in some embodiments, the heterologous nucleic acid sequence is inserted in place of some or all of the endogenous phage sequence which is removed. In some embodiments, endogenous phage sequences are removed from a position in the phage genome distant from the site(s) of insertion of the heterologous nucleic acid sequences.

[0051] As used herein, the term "sample" refers to clinical samples obtained from a subject or isolated microorganisms. In certain embodiments, a sample is obtained from a biological source (*i.e.*, a "biological sample"), such as tissue, bodily fluid, or microorganisms collected from a subject. Sample sources include, but are not limited to, mucus, sputum, bronchial alveolar lavage (BAL), bronchial wash (BW), whole blood, bodily fluids, cerebrospinal fluid (CSF), urine, plasma, serum, or tissue.

Bacteriophage

[0052] Bacteriophage are obligate intracellular parasites that multiply inside bacteria by co-opting some or all of the host biosynthetic machinery. Phages contain nucleic acid and protein, and may be enveloped by a lipid membrane. Depending upon the phage, the nucleic acid genome can be either DNA or RNA but not both, and can exist in either circular or linear forms. The size of the phage genome varies depending upon the phage. The simplest phages have genomes that are only a few thousand nucleotides in size, while the more complex phages may contain more than 100,000 nucleotides in their genome, and in rare instances no more than 500,000 bp. The number and amount of individual types of protein in phage particles will vary depending upon the phage. The proteins function in infection and to protect the nucleic acid genome from environmental nucleases.

[0053] Phage genomes come in a variety of sizes and shapes (*e.g.*, linear or circular). Most phages range in size from 24-200 nm in diameter. The capsid is composed of many copies of one or more phage proteins, and acts as a protective envelope around the phage genome. Many phages have tails attached to the phage capsid. The tail is a hollow tube through which the phage nucleic acid passes during infection. The size of the tail can vary and some phages do not even have a tail structure. In the more complex phages, the tail is surrounded by a contractile sheath which contracts during infection of the bacterial host cell. At the end of the tail, phages have a base plate and one or more tail fibers attached to it. The base plate and tail fibers are involved in the binding of the phage to the host cell.

[0054] Lytic or virulent phages are phages which can only multiply in bacteria and lyse the bacterial host cell at the end of the life cycle of the phage. The lifecycle of a lytic phage begins with an eclipse period. During the eclipse phase, no infectious phage particles can be found either inside or outside the host cell. The phage nucleic acid takes over the host biosynthetic machinery and phage specific mRNAs and proteins are produced. Early phage mRNAs code for early proteins that are needed for phage DNA synthesis and for shutting off host DNA, RNA and protein biosynthesis. In some cases, the early proteins actually degrade the host chromosome. After phage DNA is made late mRNAs and late proteins are made. The late proteins are the structural proteins that comprise the phage as well as the proteins needed for lysis of the bacterial cell. In the next phase, the phage nucleic acid and structural proteins are assembled and infectious phage particles accumulate within the cell. The bacteria begin to lyse due to the accumulation of the phage lysis protein, leading to the release of intracellular phage particles. The number of particles released per infected cell can be as high as 1000 or more. Lytic phage may be enumerated by a plaque assay. The assay is performed at a low enough concentration of phage such that each plaque arises from a single infectious phage. The infectious particle that gives rise to a plaque is called a PFU (plaque forming unit).

[0055] Lysogenic phages are those that can either multiply *via* the lytic cycle or enter a quiescent state in the host cell. In the quiescent state, the phage genome exists as a prophage (*i.e.*, it has the potential to produce phage). In most cases, the phage DNA actually integrates into the host chromosome and is replicated along with the host chromosome and passed on to the daughter cells. The host cell harboring a prophage is not adversely affected by the presence of the prophage and the lysogenic state may persist indefinitely. The lysogenic state can be terminated upon exposure to adverse conditions. Conditions which favor the termination of the lysogenic state include: desiccation, exposure to UV or ionizing radiation, exposure to mutagenic chemicals, *etc*. Adverse conditions lead to the production of proteases (rec A protein), the expression of the phage genes, reversal of the integration process, and lytic multiplication.

[0056] In some embodiments, a phage genome comprises at least 5 kilobases (kb), at least 10 kb, at least 15 kb, at least 20 kb, at least 25 kb, at least 30 kb, at least 35 kb, at least 40 kb, at least 45 kb, at least 50 kb, at least 55 kb, at least 60 kb, at least 65 kb, at least 70 kb, at least 75 kb, at least 80 kb, at least 85 kb, at least 90 kb, at least 95 kb, at least 100 kb, at least 105 kb, at least 110 kb, at least 115 kb, at least 120 kb, at least 125 kb, at least 130 kb, at

least 135 kb, at least 140 kb, at least 145 kb, at least 150 kb, at least 175 kb, at least 200 kb, at least 225 kb, at least 250 kb, at least 275 kb, at least 300 kb, at least 325 kb, at least 350 kb, at least 375 kb, at least 400 kb, at least 425 kb, at least 450 kb, at least 475 kb, or at least 500 kb of nucleic acids.

[0057] In certain embodiments, the linear bacteriophage DNA genomes correspond to a bacteriophage family or order selected from the group consisting of *Myoviridae, Styloviridae, Siphoviridae, Pedoviridae, Tectiviridae, Leviviridae, Podoviridae,* and *Plasmaviridae*. In some embodiments, the plurality of linear bacteriophage DNA genomes are derived from one or more bacteriophage genuses (or genera) selected from the group consisting of T7-like phage, phiKMV-like phage, LUZ24-like phage, phiKZ-like phage, PB1-like phage, Felix-O1-like phage, T4-like phage, phi92-like phage, rV5-like phage, SP6-like phage, N4-like phage, phiEco32-like phage, T5-like phage, KP34-like phage, KP15-like phage, GAP227-like phage, AP22-like phage, phiFel-like phage, Sap6-like phage, Silvia-like phage, Kay-like phage, Twort-like phage, P68-like phage, and phiETA-like phage.

[0058] Examples of bacteriophages with linear genomes useful in the methods of the present technology include *Klebsiella* phage K11, lambda phage, *Enterobacteria* phage T2, *Enterobacteria* phage T1, *Enterobacteria* phage T7, *Enterobacteria* phage T5, *Enterobacteria* phage P1, *Enterobacteria* phage PRD1, K1E phage, K1-5 phage, RB49 phage, RB16 phage, KP15 phage, KP27 phage, Miro phage, Matisse phage, phiEap-3 phage, ECP3 phage, EFDG1 phage, EFLK1 phage, vB_Efae230P-4 phage, vB_EfaP_IME195 phage, SA11 phage, Stau2 phage, K phage, G1 phage, SA12 phage, 812 phage, P68 phage, SAP-2 phage, 44AHJD phage, or SA97 phage.

Phage Engineering Methods of the Present Technology

[0059] In one aspect, the present disclosure provides a method for integrating a heterologous nucleic acid into a linear bacteriophage DNA genome comprising: (a) cleaving a first site at the 5' end of a plurality of linear bacteriophage DNA genomes and a second site at the 3' end of the plurality of linear bacteriophage DNA genomes with a CRISPR enzyme *in vitro*, wherein the plurality of linear bacteriophage DNA genomes is present in an isolated sample; (b) recombining *in vitro* the cleaved plurality of linear bacteriophage DNA genomes with the heterologous nucleic acid in the presence of a recombination system, wherein the heterologous nucleic acid comprises a 3' flanking region and a 5' flanking region that are homologous to the 5' and 3' ends of the cleaved plurality of linear bacteriophage DNA genomes respectively, thereby generating a plurality of circularized bacteriophage DNA

genomes; and (c) enriching the plurality of circularized bacteriophage DNA genomes by incubating the sample with at least one exonuclease. The at least one exonuclease may be bidirectional or unidirectional. Additionally or alternatively, in some embodiments, the at least one exonuclease is selected from the group consisting of Lambda exonuclease, Exonuclease III, RecBCD, Exonuclease VIII truncated, T5 exonuclease, and T7 exonuclease. In some embodiments, the at least one exonuclease is RecBCD. In certain embodiments, the method further comprises propagating the plurality of circularized bacteriophage DNA genomes in a non-natural bacterial host. The plurality of linear bacteriophage DNA genomes may be recombinant or non-recombinant. In some embodiments, the linear bacteriophage DNA genome is a terminally redundant linear bacteriophage DNA genome.

[0060] In another aspect, the present disclosure provides a method for making an enriched plurality of recombinant circularized bacteriophage DNA genomes from a plurality of linear bacteriophage DNA genomes comprising: (a) contacting a sample comprising the plurality of linear bacteriophage DNA genomes with a first sgRNA-CRISPR enzyme complex and a second sgRNA-CRISPR enzyme complex *in vitro* under conditions where (i) the first sgRNA-CRISPR enzyme complex cleaves a first site at the 5' end of the plurality of linear bacteriophage DNA genomes; and (ii) the second sgRNA-CRISPR enzyme complex cleaves a second site at the 3' end of the plurality of linear bacteriophage DNA genomes; (b) recombining *in vitro* the cleaved plurality of linear bacteriophage DNA genomes with a heterologous nucleic acid in the presence of a recombination system under conditions to produce recombinant circularized bacteriophage DNA genomes; and (c) contacting the sample with at least one exonuclease to obtain an enriched plurality of recombinant circularized bacteriophage DNA genomes. The at least one exonuclease may be bidirectional or unidirectional. Additionally or alternatively, in some embodiments, the at least one exonuclease is selected from the group consisting of Lambda exonuclease, Exonuclease III, RecBCD, Exonuclease VIII truncated, T5 exonuclease, and T7 exonuclease. In some embodiments, the at least one exonuclease is RecBCD. The plurality of linear bacteriophage DNA genomes may be recombinant or non-recombinant.

[0061] The heterologous nucleic acid comprises a 3' flanking region and a 5' flanking region that are homologous to the 5' and 3' ends of the cleaved plurality of linear bacteriophage DNA genomes respectively. In certain embodiments, the first sgRNA-CRISPR enzyme complex comprises the sequence 5' GCCACCUGAGGUUAGACCAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA

GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:2), and the second gRNA-CRISPR enzyme complex comprises the sequence 5' GGCUACUUGGCGAUUUCUGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:3).

[0062] In some embodiments of the methods disclosed herein, the homologous 5' flanking region of the heterologous nucleic acid sequence has a length of about 20-30 base pairs (bps), 30-40 bps, 40-50 bps, 50-60 bps, 60-70 bps, 70-80 bps, 80-90 bps, 90-100 bps, 100-110 bps, 110-120 bps, 120-130 bps, 130-140 bps, 140-150 bps, 150-160 bps, 160-170 bps, 170-180 bps, 180-190 bps, 190-200 bps, 200-210 bps, 210-220 bps, 220-230 bps, 230-240 bps, 240-250 bps, 250-260 bps, 260-270 bps, 270-280 bps, 280-290 bps, 290-300 bps, 300-310 bps, 310-320 bps, 320-330 bps, 330-340 bps, 340-350 bps, 350-360 bps, 360-370 bps, 370-380 bps, 380-390 bps, 390-400 bps, 400-410 bps, 410-420 bps, 420-430 bps, 430-440 bps, 440-450 bps, 450-460 bps, 460-470 bps, 470-480 bps, 480-490 bps, 490-500 bps, 500-510 bps, 510-520 bps, 520-530 bps, 530-540 bps, 540-550 bps, 550-560 bps, 560-570 bps, 570-580 bps, 580-590 bps, or 590-600 bps.

[0063] Additionally or alternatively, in some embodiments of the methods disclosed herein, the homologous 3' flanking region of the heterologous nucleic acid sequence has a length of about 20-30 base pairs (bps), 30-40 bps, 40-50 bps, 50-60 bps, 60-70 bps, 70-80 bps, 80-90 bps, 90-100 bps, 100-110 bps, 110-120 bps, 120-130 bps, 130-140 bps, 140-150 bps, 150-160 bps, 160-170 bps, 170-180 bps, 180-190 bps, 190-200 bps, 200-210 bps, 210-220 bps, 220-230 bps, 230-240 bps, 240-250 bps, 250-260 bps, 260-270 bps, 270-280 bps, 280-290 bps, 290-300 bps, 300-310 bps, 310-320 bps, 320-330 bps, 330-340 bps, 340-350 bps, 350-360 bps, 360-370 bps, 370-380 bps, 380-390 bps, 390-400 bps, 400-410 bps, 410-420 bps, 420-430 bps, 430-440 bps, 440-450 bps, 450-460 bps, 460-470 bps, 470-480 bps, 480-490 bps, 490-500 bps, 500-510 bps, 510-520 bps, 520-530 bps, 530-540 bps, 540-550 bps, 550-560 bps, 560-570 bps, 570-580 bps, 580-590 bps, or 590-600 bps.

[0064] In certain embodiments, the plurality of linear bacteriophage DNA genomes correspond to a bacteriophage family or order selected from the group consisting of *Myoviridae, Styloviridae, Siphoviridae, Pedoviridae, Tectiviridae, Leviviridae, Podoviridae,* and *Plasmaviridae*. In some embodiments, the plurality of linear bacteriophage DNA genomes are derived from one or more bacteriophage genuses (or genera) selected from the group consisting of T7-like phage, phiKMV-like phage, LUZ24-like phage, phiKZ-like

phage, PB1-like phage, Felix-O1-like phage, T4-like phage, phi92-like phage, rV5-like phage, SP6-like phage, N4-like phage, phiEco32-like phage, T5-like phage, KP34-like phage, KP15-like phage, GAP227-like phage, AP22-like phage, phiFel-like phage, Sap6-like phage, Silvia-like phage, Kay-like phage, Twort-like phage, P68-like phage, and phiETA-like phage.

[0065] Additionally or alternatively, in some embodiments, the plurality of linear bacteriophage DNA genomes correspond to *Klebsiella* phage K11, lambda phage, *Enterobacteria* phage T2, *Enterobacteria* phage T1, *Enterobacteria* phage T7, *Enterobacteria* phage T5, *Enterobacteria* phage P1, *Enterobacteria* phage PRD1, K1E phage, K1-5 phage, RB49 phage, RB16 phage, KP15 phage, KP27 phage, Miro phage, Matisse phage, phiEap-3 phage, ECP3 phage, EFDG1 phage, EFLK1 phage, vB_Efae230P-4 phage, vB_EfaP_IME195 phage, SA11 phage, Stau2 phage, K phage, G1 phage, SA12 phage, 812 phage, P68 phage, SAP-2 phage, 44AHJD phage, or SA97 phage.

[0066] In some embodiments of the methods disclosed herein, the plurality of linear bacteriophage DNA genomes correspond to *Klebsiella* phage K11, wherein the first site at the 5' end of the plurality of linear bacteriophage DNA genomes is after the nucleotide located at position 232 of SEQ ID NO:1 and the second site at the 3' end of the plurality of linear bacteriophage DNA genomes is after the nucleotide located at position 40,715 of SEQ ID NO:1. In certain embodiments, the CRISPR enzyme is Cas9 and is coupled to a sgRNA having the sequence 5' GCCACCUGAGGUUAGACCAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:2) or 5' GGCUACUUGGCGAUUUCUGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:3).

[0067] In another aspect, the present disclosure provides a method for integrating a heterologous nucleic acid into a non-terminally redundant linear bacteriophage DNA genome. In some embodiments, the method comprises: (a) contacting a sample comprising a linear bacteriophage DNA genome with a sgRNA-CRISPR enzyme complex *in vitro* under conditions where the sgRNA-CRISPR enzyme complex cleaves a first site at the 3' end of the linear bacteriophage DNA genome to produce a cleaved bacteriophage genome, wherein the cleaved bacteriophage genome comprises genomic sequence that is upstream of the first site; (b) recombining *in vitro* the cleaved bacteriophage genome with a heterologous nucleic acid

in the presence of a recombination system under conditions to produce a circularized recombinant bacteriophage genome; and (c) contacting the sample with at least one exonuclease to enrich for the circularized recombinant bacteriophage genome. In other embodiments, the method comprises: (a) contacting a sample comprising a linear bacteriophage DNA genome with a sgRNA-CRISPR enzyme complex *in vitro* under conditions where the sgRNA-CRISPR enzyme complex cleaves a first site at the 5' end of the linear bacteriophage DNA genome to produce a cleaved bacteriophage genome, wherein the cleaved bacteriophage genome comprises genomic sequence that is downstream of the first site; (b) recombining *in vitro* the cleaved bacteriophage genome with a heterologous nucleic acid in the presence of a recombination system under conditions to produce a circularized recombinant bacteriophage genome; and (c) contacting the sample with at least one exonuclease to enrich for the circularized recombinant bacteriophage genome. In any of the above embodiments of the method, the heterologous nucleic acid comprises a 5' flanking region that is homologous to the 3' end of the cleaved bacteriophage genome, and a 3' flanking region that is homologous to the 5' end of the cleaved bacteriophage genome. In any of the above embodiments, the linear bacteriophage DNA genome may be recombinant or non-recombinant.

[0068] In some embodiments of the methods disclosed herein, the homologous 5' flanking region of the heterologous nucleic acid sequence has a length of about 20-30 base pairs (bps), 30-40 bps, 40-50 bps, 50-60 bps, 60-70 bps, 70-80 bps, 80-90 bps, 90-100 bps, 100-110 bps, 110-120 bps, 120-130 bps, 130-140 bps, 140-150 bps, 150-160 bps, 160-170 bps, 170-180 bps, 180-190 bps, 190-200 bps, 200-210 bps, 210-220 bps, 220-230 bps, 230-240 bps, 240-250 bps, 250-260 bps, 260-270 bps, 270-280 bps, 280-290 bps, 290-300 bps, 300-310 bps, 310-320 bps, 320-330 bps, 330-340 bps, 340-350 bps, 350-360 bps, 360-370 bps, 370-380 bps, 380-390 bps, 390-400 bps, 400-410 bps, 410-420 bps, 420-430 bps, 430-440 bps, 440-450 bps, 450-460 bps, 460-470 bps, 470-480 bps, 480-490 bps, 490-500 bps, 500-510 bps, 510-520 bps, 520-530 bps, 530-540 bps, 540-550 bps, 550-560 bps, 560-570 bps, 570-580 bps, 580-590 bps, or 590-600 bps.

[0069] Additionally or alternatively, in some embodiments of the methods disclosed herein, the homologous 3' flanking region of the heterologous nucleic acid sequence has a length of about 20-30 base pairs (bps), 30-40 bps, 40-50 bps, 50-60 bps, 60-70 bps, 70-80 bps, 80-90 bps, 90-100 bps, 100-110 bps, 110-120 bps, 120-130 bps, 130-140 bps, 140-150 bps, 150-160 bps, 160-170 bps, 170-180 bps, 180-190 bps, 190-200 bps, 200-210 bps, 210-

220 bps, 220-230 bps, 230-240 bps, 240-250 bps, 250-260 bps, 260-270 bps, 270-280 bps, 280-290 bps, 290-300 bps, 300-310 bps, 310-320 bps, 320-330 bps, 330-340 bps, 340-350 bps, 350-360 bps, 360-370 bps, 370-380 bps, 380-390 bps, 390-400 bps, 400-410 bps, 410-420 bps, 420-430 bps, 430-440 bps, 440-450 bps, 450-460 bps, 460-470 bps, 470-480 bps, 480-490 bps, 490-500 bps, 500-510 bps, 510-520 bps, 520-530 bps, 530-540 bps, 540-550 bps, 550-560 bps, 560-570 bps, 570-580 bps, 580-590 bps, or 590-600 bps.

[0070] Additionally or alternatively, in some embodiments of the methods of the present technology, the at least one exonuclease is selected from the group consisting of Lambda exonuclease, Exonuclease III, RecBCD, Exonuclease VIII truncated, T5 exonuclease, and T7 exonuclease. The at least one exonuclease may be bidirectional or unidirectional. In some embodiments, the at least one exonuclease is RecBCD. In certain embodiments, the methods disclosed herein further comprise propagating the circularized recombinant bacteriophage genome in a bacterial host. The non-terminally redundant linear bacteriophage DNA genome may correspond to a bacteriophage family or order selected from the group consisting of *Myoviridae, Styloviridae, Siphoviridae, Pedoviridae, Tectiviridae, Leviviridae, Podoviridae,* and *Plasmaviridae*.

[0071] Additionally or alternatively, in some embodiments of any of the methods disclosed herein, the recombination system comprises a 5'-3' exonuclease, a DNA polymerase, and a DNA ligase. In one embodiment, the 5'-3' exonuclease is T5 exonuclease, the DNA polymerase is Phusion® DNA polymerase (Thermo Fisher Scientific, Waltham, MA), and the DNA ligase is *Taq* ligase. In other embodiments, the recombination system comprises a 3'-5' exonuclease, a DNA polymerase, and a DNA ligase.

[0072] Accurate identification of bacterial species within a biological sample informs the selection of suitable therapies for treating bacterial infections. Recombinant bacteriophage generated using the methods disclosed herein, may be used to identify bacteria present within a biological sample (*e.g.*, whole blood, plasma, serum). Such methods entail contacting the biological sample with a recombinant bacteriophage generated using the methods disclosed herein, and detecting the presence of bacterial host cells infected by the recombinant phage, wherein the recombinant phage comprises a heterologous nucleic acid that encodes a detectable gene product, thereby leading to the identification of bacteria present within the biological sample.

[0073] Additionally or alternatively, recombinant bacteriophage generated using the methods disclosed herein, may be used in methods for profiling antibiotic susceptibility of bacteria present within a biological sample (*e.g.*, whole blood, plasma, serum). These methods include (a) contacting the biological sample with an antibiotic and a recombinant bacteriophage generated using the methods disclosed herein, (b) detecting the presence of bacterial host cells infected by the recombinant phage, wherein the recombinant phage comprises a heterologous nucleic acid that encodes a detectable gene product, and (c) determining that the antibiotic is effective in inhibiting the bacteria present in the biological sample when the number of recombinant phage infected bacterial host cells is reduced relative to that observed in an untreated control sample.

CRISPR Enzymes

[0074] A variety of CRISPR enzymes are available for use in conjunction with any of the methods of the present disclosure. In some embodiments, the CRISPR enzyme is a Type II CRISPR enzyme. In some embodiments, the CRISPR enzyme catalyzes DNA cleavage. In some embodiments, the CRISPR enzyme catalyzes RNA cleavage. In some embodiments, the CRISPR enzyme is any Cas9 protein, for instance any naturally-occurring bacterial Cas9 as well as any variants, homologs or orthologs thereof. Non-limiting examples of Cas proteins include Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (also known as Csn1 and Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, homologues thereof, or variants thereof. In some embodiments, the CRISPR enzyme cleaves both strands of the target nucleic acid at the Protospacer Adjacent Motif (PAM) site.

Heterologous Nucleic Acids

[0075] In any of the above embodiments of the methods disclosed herein, the heterologous nucleic acid comprises an open reading frame that encodes a bioluminescent protein, a fluorescent protein, a chemiluminescent protein, or any combination thereof. In some embodiments, the encoded gene product(s) produces a detectable signal upon exposure to the appropriate stimuli, and the resulting signal permits detection of bacterial host cells infected by the recombinant phage. In certain embodiments, the open reading frame encodes a protein that serves as a marker that can be identified by screening bacterial host cells infected by a recombinant phage comprising a heterologous nucleic acid sequence comprising the open reading frame. Examples of such markers include by way of example

and without limitation: a fluorescent label, a luminescent label, a chemiluminescence label, or an enzymatic label. In some embodiments, the heterologous nucleic acid sequence further comprises sequences naturally found in the bacteriophage, but placed at a non-normally occurring location in the genome.

[0076] In some embodiments of the methods disclosed herein, the length of the heterologous nucleic acid sequence is at least 100 bases, at least 200 bases, at least 300 bases, at least 400 bases, at least 500 bases, at least 600 bases, at least 700 bases, at least 800 bases, at least 900 bases, at least 1 kilobase (kb), at least 1.1 kb, at least 1.2 kb, at least 1.3 kb, at least 1.4 kb, at least 1.5 kb, at least 1.6 kb, at least 1.7 kb, at least 1.8 kb, at least 1.9 kb, at least 2.0 kb, at least 2.1 kb, at least 2.2 kb, at least 2.3 kb, at least 2.4 kb, at least 2.5 kb, at least 2.6 kb, at least 2.7 kb, at least 2.8 kb, at least 2.9 kb, at least 3.0 kb, at least 3.1 kb, at least 3.2 kb, at least 3.3 kb, at least 3.4 kb, at least 3.5 kb, at least 3.6 kb, at least 3.7 kb, at least 3.8 kb, at least 3.9 kb, at least 4.0 kb, at least 4.5 kb, at least 5.0 kb, at least 5.5 kb, at least 6.0 kb, at least 6.5 kb, at least 7.0 kb, at least 7.5 kb, at least 8.0 kb, at least 8.5 kb, at least 9.0 kb, at least 9.5 kb, at least 10 kb, or more. In certain embodiments, the heterologous nucleic acid sequence comprises a length that is less than or equal to a length selected from the group consisting of 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb, 7 kb, 8 kb, 9 kb, and 10 kb. In some embodiments, the heterologous nucleic acid sequence comprises a length that is less than or equal to the maximum length of heterologous nucleic acid sequence that can be packaged into a phage particle comprising the linear phage genome.

[0077] In some embodiments, the length of the heterologous nucleic acid sequence is from 100 to 500 bases, from 200 to 1,000 bases, from 500 to 1,000 bases, from 500 to 1,500 bases, from 1 kb to 2 kb, from 1.5 kb to 2.5 kb, from 2.0 kb to 3.0 kb, from 2.5 kb to 3.5 kb, from 3.0 kb to 4.0 kb, from 3.5 kb to 4.5 kb, from 4.0 kb to 5.0 kb, from 4.5 kb to 5.5 kb, from 5.0 kb to 6.0 kb, from 5.5 kb to 6.5 kb, from 6.0 kb to 7.0 kb, from 6.5 kb to 7.5 kb, from 7.0 kb to 8.0 kb, from 7.5 kb to 8.5 kb, from 8.0 kb to 9.0 kb, from 8.5 kb to 9.5 kb, or from 9.0 kb to 10.0 kb.

[0078] In some embodiments, the heterologous nucleic acid sequence is inserted into the linear phage genome with no loss of endogenous phage genomic sequence. In some embodiments, the heterologous nucleic acid sequence replaces an endogenous linear phage genomic sequence. In some embodiments, the heterologous nucleic acid sequence includes an endogenous linear phage genomic sequence that was previously excised from the linear phage genome.

[0079] In certain embodiments, the heterologous nucleic acid sequence replaces an endogenous linear phage genomic sequence that is less than the length of the heterologous nucleic acid sequence. Accordingly, in some embodiments, the length of the recombinant linear phage genome is longer than the length of the wild-type linear phage genome. In some embodiments, the heterologous nucleic acid sequence replaces an endogenous linear phage genomic sequence that is greater than the length of the heterologous nucleic acid sequence. Thus, in some embodiments, the length of the recombinant linear phage genome is shorter than the length of the wild-type linear phage genome. In certain embodiments, the heterologous nucleic acid sequence replaces an endogenous linear phage genomic sequence that is equal to the length of the heterologous nucleic acid sequence.

[0080] In certain embodiments, the open reading frame of the heterologous nucleic acid encodes a protein that confers a phenotype of interest on a host cell infected by a recombinant phage expressing the heterologous nucleic acid. In some embodiments, the phenotype of interest is the expression of the gene product encoded by the open reading frame of the heterologous nucleic acid.

[0081] In certain embodiments, the open reading frame of the heterologous nucleic acid is operably linked to an expression control sequence that is capable of directing expression of the open reading frame, wherein the open reading frame encodes a bioluminescent protein, a fluorescent protein, a chemiluminescent protein, or any combination thereof. In some embodiments, the expression control sequence is located within the heterologous nucleic acid sequence. In other embodiments, the expression control sequence is located in the endogenous linear phage genome sequence. For example, the open reading frame may be inserted into the linear phage genome downstream of or in the place of an endogenous phage open reading frame sequence. In some embodiments, the expression control sequence is an inducible promoter or a constitutive promoter (*e.g.*, sarA promoter or lpp promoter). *See e.g.*, Djordjevic & Klaenhammer, *Methods in Cell Science* 20(1):119-126 (1998). The inducible promoter or constitutive promoter may be an endogenous phage promoter sequence, a non-endogenous phage promoter sequence, or a bacterial host promoter sequence. Additionally or alternatively, in some embodiments, the inducible promoter is a pH-sensitive promoter, or a temperature sensitive promoter.

[0082] In some embodiments, the heterologous nucleic acid sequence comprises a first open reading frame and at least one supplemental open reading frame. In certain embodiments, the first and the at least one supplemental open reading frames are operably

linked to the same expression control sequences. In some embodiments, the first and the at least one supplemental open reading frames are operably linked to different expression control sequences.

[0083] Fluorescent proteins include but are not limited to blue/UV fluorescent proteins (for example, TagBFP, Azurite, EBFP2, mKalama1, Sirius, Sapphire, and T-Sapphire), cyan fluorescent proteins (for example, ECFP, Cerulean, SCFP3A, mTurquoise, monomeric Midoriishi-Cyan, TagCFP, and mTFP1), green fluorescent proteins (for example, EGFP, Emerald, Superfolder GFP, Monomeric Azami Green, TagGFP2, mUKG, and mWasabi), yellow fluorescent proteins (for example, EYFP, Citrine, Venus, SYFP2, and TagYFP), orange fluorescent proteins (for example, Monomeric Kusabira-Orange, mKOκ, mKO2, mOrange, and mOrange2), red fluorescent proteins (for example, mRaspberry, mCherry, dsRed, mStrawberry, mTangerine, tdTomato, TagRFP, TagRFP-T, mApple, and mRuby), far-red fluorescent proteins (for example, mPlum, HcRed-Tandem, mKate2, mNeptune, and NirFP), near-IR fluorescent proteins (for example, TagRFP657, IFP1.4, and iRFP), long stokes-shift proteins (for example, mKeima Red, LSS-mKate1, and LSS-mKate2), photoactivatable fluorescent proteins (for example, PA-GFP, PAmCherry1, and PATagRFP), photoconvertible fluorescent proteins (for example, Kaede (green), Kaede (red), KikGR1 (green), KikGR1 (red), PS-CFP2, PS-CFP2, mEos2 (green), mEos2 (red), PSmOrange, and PSmOrange), fluorescein, rhodamine, and photoswitchable fluorescent proteins (for example, Dronpa).

[0084] Examples of bioluminescent proteins are aequorin (derived from the jellyfish *Aequorea victoria*) and luciferases (including luciferases derived from firefly and *Renilla,* nanoluciferase, red luciferase, luxAB, and the like). These proteins have also been genetically separated into two distinct functional domains that will generate light only when the protein domains are closely co-localized. A variety of emission spectrum-shifted mutant derivatives of both of these proteins have been generated over the past decade and have been used for multi-color imaging and co-localization within a living cell.

[0085] Examples of chemiluminescent protein include β-galactosidase, horseradish peroxidase (HRP), and alkaline phosphatase. Peroxidases generate peroxide that oxidizes luminol in a reaction that generates light, whereas alkaline phosphatases remove a phosphate from a substrate molecule, destabilizing it and initiating a cascade that results in the emission of light.

[0086] In some embodiments, the open reading frame of the heterologous nucleic acid comprises an epitope that can be detected with an antibody or other binding molecule. For example, an antibody that recognizes the epitope may be directly linked to a signal generating moiety (such as by covalent attachment of a chemiluminescent or fluorescent protein), or can be detected using at least one additional binding reagent such as a secondary antibody, directly linked to a signal generating moiety. In some embodiments, the epitope is absent in wild-type bacteriophage and the bacterial host cell. Accordingly, detection of the epitope in a sample demonstrates the presence of a bacterial host cell infected by a recombinant phage comprising a heterologous nucleic acid, wherein the open reading frame of the heterologous nucleic acid comprises the epitope.

[0087] In other embodiments, the open reading frame of the heterologous nucleic acid comprises a polypeptide tag sequence, such that the expression product of the open reading frame comprises the tag fused to a polypeptide or protein encoded by the open reading frame (*e.g.*, poly-histidine, FLAG, Glutathione S-transferase (GST) *etc.*).

[0088] In some embodiments, the open reading frame of the heterologous nucleic acid sequence comprises a biotin binding protein such as avidin, streptavidin, or neutrAvidin that can be detected with a biotin molecule conjugated to an enzyme (*e.g.*, β-galactosidase, horseradish peroxidase (HRP), and alkaline phosphatase) or an antibody. In some embodiments, the antibody conjugated to a biotin molecule may be directly linked to a signal generating moiety (such as by covalent attachment of a chemiluminescent or fluorescent protein), or can be detected using at least one additional binding reagent such as a secondary antibody, directly linked to a signal generating moiety.

Kits

[0089] The present technology provides kits for integrating a heterologous nucleic acid sequence into a linear bacteriophage DNA genome.

[0090] In one aspect, the kits of the present technology comprise (a) one or more coded/labeled vials that contain a plurality of linear bacteriophage DNA genomes, (b) at least one CRISPR enzyme, (c) a recombination system, and (d) at least one exonuclease. The at least one exonuclease may be bidirectional or unidirectional. Additionally or alternatively, in some embodiments, the at least one exonuclease is selected from the group consisting of Lambda exonuclease, Exonuclease III, RecBCD, Exonuclease VIII truncated, T5 exonuclease, and T7 exonuclease. In certain embodiments, the at least one exonuclease is

RecBCD. The plurality of linear bacteriophage DNA genomes may be recombinant or non-recombinant.

[0091] In some embodiments, each coded/labeled vial containing a plurality of linear bacteriophage DNA genomes corresponds to a different bacteriophage type. In other embodiments, each coded/labeled vial containing a plurality of linear bacteriophage DNA genomes corresponds to the same bacteriophage type. In some embodiments, each phage vial is assigned a unique code that identifies the bacteriophage in the phage vial, or the types of bacteria that the bacteriophage strain infects. The unique code can be encoded by a machine discernible pattern, such as a bar code, a QR code, an alphanumeric string, or any other pattern that can be discerned by a reader. Each unique code may be shown as, for example, a bar code sticker on a vial or container storing a corresponding phage sample. In some embodiments, the kit is stored under conditions that permit the preservation of the linear bacteriophage DNA genomes for extended periods, such as under bacteriophage-specific, controlled temperature, moisture, and pH conditions.

[0092] In some embodiments, the kits comprise a recombination system that includes a 5'-3' exonuclease, a DNA polymerase, and a DNA ligase. For example, in one embodiment, the 5'-3' exonuclease is T5 exonuclease, the DNA polymerase is Phusion® DNA polymerase (Thermo Fisher Scientific, Waltham, MA), and the DNA ligase is Taq ligase. In other embodiments, the kits comprise a recombination system that includes a 3'-5' exonuclease, a DNA polymerase, and a DNA ligase.

[0093] Additionally or alternatively, in some embodiments, the kits comprise one or more CRISPR enzymes selected from the group consisting of Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, and Csf4. The one or more CRISPR enzymes may be coupled to a sgRNA. In certain embodiments, the sgRNA has the sequence 5' GCCACCUGAGGUUAGACCAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:2) or 5' GGCUACUUGGCGAUUUCUGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:3).

[0094] Additionally or alternatively, in some embodiments, the kits further comprise vials containing natural or non-natural bacterial host cells. In some embodiments, the bacterial host cells are *E. coli*. In certain embodiments, the bacterial host cells are *E. coli* strain DH10β.

[0095] In some embodiments, the kits further comprise positive control heterologous nucleic acid sequences to correct for any variability in the recombination systems between experimental runs. The kits may also comprise instructions for use, software for automated analysis, containers, packages such as packaging intended for commercial sale and the like.

[0096] The kit may further comprise one or more of: wash buffers and/or reagents, hybridization buffers and/or reagents, labeling buffers and/or reagents, and detection means. The buffers and/or reagents are usually optimized for the particular detection technique for which the kit is intended. Protocols for using these buffers and reagents for performing different steps of the procedure may also be included in the kit. Further optional components of the kits may include expression media for gene products encoded by the heterologous nucleic acids disclosed herein, such as a medium containing nutrients and cofactors for bioluminescence, devices such as a lamp configured to illuminate at specific wavelengths of light to detect biofluorescence, and devices for measuring the extent of heterologous nucleic acid expression, such as a photometer or photodetector.

[0097] Additionally or alternatively, the kits disclosed herein may also include coded and labeled vials that contain a plurality of antibiotics.

EXAMPLES

Example 1: Phage Engineering Methods of the Present Technology in Klebsiella bacteriophage K11

[0098] This Example demonstrates that the methods of the present technology are useful for integrating a heterologous nucleic acid into a linear bacteriophage DNA genome (*e.g.*, *Klebsiella* bacteriophage K11) and for isolating recombinant bacteriophages that express the heterologous nucleic acid sequence.

[0099] The *Klebsiella* phage K11 has a 41,181 base pair linear double-stranded DNA genome that is shown in Figure 3 (GenBank Accession No. NC_011043, *a.k.a*, SEQ ID NO: 1). Figure 1 shows the schematic representation of the circular payload integration method disclosed herein. SgRNAs complexed with *S. pyogenes* Cas9 endonuclease were used to generate double-stranded breaks after the nucleotide located at position 232, and after the nucleotide located at position 40,715 of phage K11 genomic DNA. The sgRNA sequences

were 5' GCCACCUGAGGUUAGACCAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:2) and 5' GGCUACUUGGCGAUUUCUGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:3). The first 20 base pairs of each sgRNA represents the K11 phage sequence that was targeted. These cleavage sites were selected because they excised the very beginning and very end of the K11 genome. About 2.288 µg of K11 genomic DNA was digested with the Cas9:sgRNAs.

[00100] A heterologous nucleic acid sequence containing the nanoluciferase reporter gene was constructed. The heterologous nucleic acid sequence also included nucleotides 1-232, and 40,716-41,181 of the wild-type K11 phage genomic sequence. The 3' flanking region of the heterologous nucleic acid was homologous to the 5' end of the cleaved linear K11 phage genome, whereas the 5' flanking region of the heterologous nucleic acid was homologous to the 3' end of the cleaved linear K11 phage genome, thus facilitating recombination between the heterologous nucleic acid and the truncated linear K11 phage genome.

[00101] The cleaved linear K11 phage DNA and the heterologous nucleic acid sequence were subsequently joined together in a Gibson Assembly® (New England Biolabs, Ipswich, MA) reaction to form a circular recombinant phage genome. About 150 ng of the heterologous nucleic acid sequence (~2.3-fold molar excess) was added to the cleaved linear K11 phage DNA. The reaction was purified and about 1 µg total DNA was recovered. After completion of the Gibson Assembly® (New England Biolabs, Ipswich, MA) reaction, the K11 phage sample (which contained a mixture of circular recombinant K11 phage genome sequences and wild-type linear K11 phage genome sequences) was subjected to treatment with RecBCD, a bidirectional DNA exonuclease, according to the manufacturer's protocol. RecBCD degraded any unassembled wild-type linear phage genomic DNA, but did not affect the circularized recombinant phage genomic DNA. Next, 2 µL of the RecBCD treated reaction (about 67 ng total DNA) was transformed into the *E. coli* strain DH10β. Although not a natural host for the K11 phage, DH10β cells replicated the K11 phage genome and packaged fully infective recombinant phage particles that could be used to subsequently infect the native *Klebsiella pneumoniae* host.

[00102] Because most phages have mechanisms to resolve linear concatamers or circular replication intermediates, the recombinant circularized K11 phage genomes were converted into functional linear double-stranded DNA genomes that retained the heterologous nucleic acid sequence. Accordingly, the recombinant K11 phages contained nanoluciferase, a bioluminescent reporter that can be expressed in infected *Klebsiella pneumoniae* host cells. The phage engineering methods of the present technology yielded a total of 226 plaques.

[00103] A total of sixteen plaques were screened by PCR, out of which thirteen (81%) yielded the 907 bp amplicon that corresponded to the Nanoluciferase insert. See Figure 2. Four of the thirteen plaques containing the recombinant K11 bacteriophage were tested on a host population of *Klebsiella pneumoniae*. The infected bacterial hosts displayed luminescence that was nearly five orders of magnitude above the background level. See Figure 4.

[00104] These results demonstrate that the methods of the present technology permit the efficient recovery of recombinant linear bacteriophage DNA genomes that (a) contain a heterologous nucleic acid sequence of interest, and (b) express the phenotypic properties associated with the heterologous nucleic acid sequence of interest. Accordingly, the methods disclosed herein are useful for integrating heterologous nucleic acids into linear bacteriophage DNA genomes.

Example 2: Comparison Against BAR 3.0 Phage Engineering Method

[00105] This Example demonstrates that the methods of the present technology are useful for integrating a heterologous nucleic acid into a linear bacteriophage DNA genome (*e.g.*, *Klebsiella* bacteriophage K11) and for isolating recombinant bacteriophages that express the heterologous nucleic acid sequence. Moreover, this Example demonstrates that the methods disclosed herein show superior efficiency with respect to recovering recombinant phage genomes compared to other phage engineering techniques, such as BAR 3.0.

[00106] The CRISPR/Cas system was used to cleave the K11 phage genome after gene 4.5 to create an insertion site for a nanoluciferase reporter sequence into the phage genome. The desired chimeric guide sequence was placed under the control of a T7 promoter (Figure 5) and was transcribed using the NEB HiScribe T7 High Yield RNA Synthesis Kit (NEB E2040, Ipswich, MA). The resulting RNA product was purified using the Qiagen RNeasy Mini Kit (Qiagen 74104, Hilden Germany).

[00107] The Cas9 endonuclease from New England Biolabs (NEB M0386, Ipswich, MA) was complexed with the sgRNA using a modified protocol. Briefly, in a 27 µL volume, 30nM of sgRNA was incubated with 30nM of Cas9 endonuclease. NEBuffer 3.1 was used instead of the included Cas9 Nuclease Reaction buffer. After a 10 minute preincubation at 25 °C, 2.3 µg of K11 genomic DNA was added to the reaction to achieve a final concentration of 3nM target DNA. The reaction was then incubated at 37 °C for 1 hr before adding an additional 3nM Cas9 endonuclease. After incubation for another hour, 10 Units of RNAse A (ThermoFisher EN0351, Waltham, MA) was added to the reaction mixture to degrade any remaining RNA.

[00108] After cleavage, the DNA was purified using phenol/chloforom precipitation. Cleavage of the K11 genomic DNA was verified using gel imaging. See Figure 6.

[00109] A synthetic DNA construct containing 60 bp of homology to the K11 genome around the gene 4.5 cleavage site surrounding a nanoluciferase gene (Figure 7) was introduced into the cleaved K11 phage genome using NEBuilder HiFi DNA assembly mix (NEB E5520, Ipswich, MA) according to the manufacturer's protocol. Briefly, 4 µg of cleaved K11 DNA was mixed with 90 ng of the nanoluc/K11 homology construct. The reaction was incubated at 50 °C for 60 minutes.

[00110] The reaction was subsequently purified using phenol/chloforom precipitation and 1 µg of the reaction product was electroporated into competent *Klebsiella pneumoniae* Sp 390 using the following electroporation settings: 200Ω resistance, 25 µF capacitance, and 2.4 kV. After electroporation, 400 µl of SOC broth was added to the cultures and cells were allowed to recover for one hour at 37 °C with shaking. Cells were then plated on a 0.65% soft agar overlay on an LB plate and incubated overnight at 37 °C. Plaque formation was evaluated the following day. As shown in the Table below, no recombinant K11 bacteriophage were recovered using the BAR 3.0 protocol.

Electroplaquing Results for Cas9 cleaved K11 and recombined K11/nanoluciferase			
	60 ng K11 DNA (uncleaved control)	60 ng cleaved K11 DNA	1 µg recombinant K11 phage with nanoluciferase insert
pfu/ml	2.00E+08	0	0

[00111] These results demonstrate that the methods disclosed herein show superior efficiency with respect to recovering recombinant phage genomes compared to other phage engineering techniques, such as BAR 3.0. These results demonstrate that the methods of the

present technology permit the efficient recovery of recombinant linear bacteriophage DNA genomes that (a) contain a heterologous nucleic acid sequence of interest, and (b) express the phenotypic properties associated with the heterologous nucleic acid sequence of interest. Accordingly, the methods disclosed herein are useful for integrating heterologous nucleic acids into linear bacteriophage DNA genomes.

EQUIVALENTS

[00112] The present technology is not to be limited in terms of the particular embodiments described in this application, which are intended as single illustrations of individual aspects of the present technology. Many modifications and variations of this present technology can be made without departing from its spirit and scope, as will be apparent to those skilled in the art. Functionally equivalent methods and apparatuses within the scope of the present technology, in addition to those enumerated herein, will be apparent to those skilled in the art from the foregoing descriptions. Such modifications and variations are intended to fall within the scope of the present technology. It is to be understood that this present technology is not limited to particular methods, reagents, compounds compositions or biological systems, which can, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting.

[00113] In addition, where features or aspects of the disclosure are described in terms of Markush groups, those skilled in the art will recognize that the disclosure is also thereby described in terms of any individual member or subgroup of members of the Markush group.

[00114] As will be understood by one skilled in the art, for any and all purposes, particularly in terms of providing a written description, all ranges disclosed herein also encompass any and all possible subranges and combinations of subranges thereof. Any listed range can be easily recognized as sufficiently describing and enabling the same range being broken down into at least equal halves, thirds, quarters, fifths, tenths, *etc.* As a non-limiting example, each range discussed herein can be readily broken down into a lower third, middle third and upper third, *etc.* As will also be understood by one skilled in the art all language such as “up to,” “at least,” “greater than,” “less than,” and the like, include the number recited and refer to ranges which can be subsequently broken down into subranges as discussed above. Finally, as will be understood by one skilled in the art, a range includes each individual member. Thus, for example, a group having 1-3 cells refers to groups having 1, 2, or 3 cells. Similarly, a group having 1-5 cells refers to groups having 1, 2, 3, 4, or 5 cells, and so forth.

[00115] All patents, patent applications, provisional applications, and publications referred to or cited herein are incorporated by reference in their entirety, including all figures and tables, to the extent they are not inconsistent with the explicit teachings of this specification.

CLAIMS

1. A method for integrating a heterologous nucleic acid into a linear bacteriophage DNA genome comprising:

(a) cleaving a first site at the 5' end of a plurality of linear bacteriophage DNA genomes and a second site at the 3' end of the plurality of linear bacteriophage DNA genomes with a CRISPR enzyme *in vitro*, wherein the plurality of linear bacteriophage DNA genomes is present in an isolated sample;

(b) recombining *in vitro* the cleaved plurality of linear bacteriophage DNA genomes with the heterologous nucleic acid in the presence of a recombination system, wherein the heterologous nucleic acid comprises a 3' flanking region and a 5' flanking region that are homologous to the 5' and 3' ends of the cleaved plurality of linear bacteriophage DNA genomes respectively, thereby generating a plurality of circularized bacteriophage DNA genomes; and

(c) enriching the plurality of circularized bacteriophage DNA genomes by incubating the sample with at least one exonuclease.

2. The method of claim 1, further comprising propagating the plurality of circularized bacteriophage DNA genomes in a non-natural bacterial host.

3. The method of claim 1 or 2, wherein the plurality of linear bacteriophage DNA genomes correspond to a bacteriophage family or order selected from the group consisting of *Myoviridae, Styloviridae, Siphoviridae, Pedoviridae, Tectiviridae, Leviviridae, Podoviridae,* and *Plasmaviridae*.

4. The method of any one of claims 1-3, wherein the plurality of linear bacteriophage DNA genomes are derived from one or more bacteriophage genuses (or genera) selected from the group consisting of T7-like phage, phiKMV-like phage, LUZ24-like phage, phiKZ-like phage, PB1-like phage, Felix-O1-like phage, T4-like phage, phi92-like phage, rV5-like phage, SP6-like phage, N4-like phage, phiEco32-like phage, T5-like phage, KP34-like phage, KP15-like phage, GAP227-like phage, AP22-like phage, phiFel-like phage, Sap6-like phage, Silvia-like phage, Kay-like phage, Twort-like phage, P68-like phage, and phiETA-like phage.

5. The method of any one of claims 1-4, wherein the plurality of linear bacteriophage DNA genomes correspond to *Klebsiella* phage K11, lambda phage, *Enterobacteria* phage T2,

Enterobacteria phage T1, *Enterobacteria* phage T7, *Enterobacteria* phage T5, *Enterobacteria* phage P1, *Enterobacteria* phage PRD1, K1E phage, K1-5 phage, RB49 phage, RB16 phage, KP15 phage, KP27 phage, Miro phage, Matisse phage, phiEap-3 phage, ECP3 phage, EFDG1 phage, EFLK1 phage, vB_Efae230P-4 phage, vB_EfaP_IME195 phage, SA11 phage, Stau2 phage, K phage, G1 phage, SA12 phage, 812 phage, P68 phage, SAP-2 phage, 44AHJD phage, or SA97 phage.

6. The method of any one of claims 1-5, wherein the at least one exonuclease is selected from the group consisting of Lambda exonuclease, Exonuclease III, RecBCD, Exonuclease VIII truncated, T5 exonuclease, and T7 exonuclease.

7. The method of claim 6, wherein the at least one exonuclease is RecBCD.

8. The method of any one of claims 1-7, wherein the recombination system comprises a 5'-3' exonuclease, a DNA polymerase, and a DNA ligase.

9. The method of claim 8, wherein the 5'-3' exonuclease is T5 exonuclease, the DNA polymerase is Phusion® DNA polymerase, and the DNA ligase is *Taq* ligase.

10. The method of any one of claims 1-7, wherein the recombination system comprises a 3'-5' exonuclease, a DNA polymerase, and a DNA ligase.

11. The method of any one of claims 1-10, wherein the CRISPR enzyme is a Cas protein selected from the group consisting of Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, and Csf4.

12. The method of claim 11, wherein the CRISPR enzyme is coupled to a sgRNA.

13. The method of any one of claims 1-12, wherein the heterologous nucleic acid comprises an open reading frame that encodes a bioluminescent protein, a fluorescent protein, a chemiluminescent protein, or any combination thereof.

14. The method of claim 13, wherein the open reading frame of the heterologous nucleic acid is operably linked to an expression control sequence that is capable of directing

expression of the bioluminescent protein, the fluorescent protein, the chemiluminescent protein, or any combination thereof.

15. The method of claim 13 or 14, wherein the bioluminescent protein is Aequorin, firefly luciferase, *Renilla* luciferase, red luciferase, luxAB, or nanoluciferase.

16. The method of claim 13 or 14, wherein the chemiluminescent protein is β-galactosidase, horseradish peroxidase (HRP), or alkaline phosphatase.

17. The method of claim 13 or 14, wherein the fluorescent protein is TagBFP, Azurite, EBFP2, mKalama1, Sirius, Sapphire, T-Sapphire, ECFP, Cerulean, SCFP3A, mTurquoise, monomeric Midoriishi-Cyan, TagCFP, mTFP1, EGFP, Emerald, Superfolder GFP, Monomeric Azami Green, TagGFP2, mUKG, mWasabi, EYFP, Citrine, Venus, SYFP2, TagYFP, Monomeric Kusabira-Orange, mKOκ, mKO2, mOrange, mOrange2, mRaspberry, mCherry, dsRed, mStrawberry, mTangerine, tdTomato, TagRFP, TagRFP-T, mApple, mRuby, mPlum, HcRed-Tandem, mKate2, mNeptune, NirFP, TagRFP657, IFP1.4, iRFP, mKeima Red, LSS-mKate1, LSS-mKate2, PA-GFP, PAmCherry1, PATagRFP, Kaede (green), Kaede (red), KikGR1 (green), KikGR1 (red), PS-CFP2, PS-CFP2, mEos2 (green), mEos2 (red), PSmOrange, or Dronpa.

18. The method of any one of claims 14-17, wherein the expression control sequence is an inducible promoter or a constitutive promoter.

19. The method of any one of claims 1-18, wherein the heterologous nucleic acid is about 100-500 base pairs in length.

20. The method of any one of claims 1-18, wherein the heterologous nucleic acid is about 500-1500 base pairs in length.

21. The method of claim 5, wherein the plurality of linear bacteriophage DNA genomes correspond to *Klebsiella* phage K11.

22. The method of claim 21, wherein the first site at the 5' end of the plurality of linear bacteriophage DNA genomes is after the nucleotide located at position 232 of SEQ ID NO:1.

23. The method of claim 22, wherein the second site at the 3' end of the plurality of linear bacteriophage DNA genomes is after the nucleotide located at position 40,715 of SEQ ID NO:1.

24. The method of claim 23, wherein the CRISPR enzyme is Cas9.

25. The method of claim 24, wherein Cas9 is coupled to a sgRNA having the sequence 5' GCCACCUGAGGUUAGACCAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:2).

26. The method of claim 24, wherein Cas9 is coupled to a sgRNA having the sequence 5' GGCUACUUGGCGAUUUCUGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:3).

27. A method for making an enriched plurality of recombinant circularized bacteriophage DNA genomes from a plurality of linear bacteriophage DNA genomes comprising:

(a) contacting a sample comprising the plurality of linear bacteriophage DNA genomes with a first sgRNA-CRISPR enzyme complex and a second sgRNA-CRISPR enzyme complex *in vitro* under conditions where

(i) the first sgRNA-CRISPR enzyme complex cleaves a first site at the 5' end of the plurality of linear bacteriophage DNA genomes; and

(ii) the second sgRNA-CRISPR enzyme complex cleaves a second site at the 3' end of the plurality of linear bacteriophage DNA genomes;

(b) recombining *in vitro* the cleaved plurality of linear bacteriophage DNA genomes with a heterologous nucleic acid in the presence of a recombination system under conditions to produce recombinant circularized bacteriophage DNA genomes; and

(c) contacting the sample with at least one exonuclease to obtain an enriched plurality of recombinant circularized bacteriophage DNA genomes.

28. The method of claim 27, wherein the heterologous nucleic acid comprises a 3' flanking region and a 5' flanking region that are homologous to the 5' and 3' ends of the cleaved plurality of linear bacteriophage DNA genomes respectively.

29. The method of claim 27 or 28, wherein the plurality of linear bacteriophage DNA genomes correspond to *Klebsiella* phage K11.

30. The method of any one of claims 27-29, wherein

the first sgRNA-CRISPR enzyme complex comprises the sequence 5' GCCACCUGAGGUUAGACCAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:2), and

the second gRNA-CRISPR enzyme complex comprises the sequence 5' GGCUACUUGGCGAUUUCUGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU 3' (SEQ ID NO:3).

31. A method for integrating a heterologous nucleic acid into a non-terminally redundant linear bacteriophage DNA genome comprising:

(a) contacting a sample comprising a linear bacteriophage DNA genome with a sgRNA-CRISPR enzyme complex *in vitro* under conditions where the sgRNA-CRISPR enzyme complex cleaves a first site at the 3' end of the linear bacteriophage DNA genome to produce a cleaved bacteriophage genome, wherein the cleaved bacteriophage genome comprises genomic sequence that is upstream of the first site;

(b) recombining *in vitro* the cleaved bacteriophage genome with a heterologous nucleic acid in the presence of a recombination system under conditions to produce a circularized recombinant bacteriophage genome; and

(c) contacting the sample with at least one exonuclease to enrich for the circularized recombinant bacteriophage genome.

32. A method for integrating a heterologous nucleic acid into a non-terminally redundant linear bacteriophage DNA genome comprising:

(a) contacting a sample comprising a linear bacteriophage DNA genome with a sgRNA-CRISPR enzyme complex *in vitro* under conditions where the sgRNA-CRISPR enzyme complex cleaves a first site at the 5' end of the linear bacteriophage DNA genome to

produce a cleaved bacteriophage genome, wherein the cleaved bacteriophage genome comprises genomic sequence that is downstream of the first site;

(b) recombining *in vitro* the cleaved bacteriophage genome with a heterologous nucleic acid in the presence of a recombination system under conditions to produce a circularized recombinant bacteriophage genome; and

(c) contacting the sample with at least one exonuclease to enrich for the circularized recombinant bacteriophage genome.

33. The method of claim 31 or 32, wherein the heterologous nucleic acid comprises a 5' flanking region that is homologous to the 3' end of the cleaved bacteriophage genome, and a 3' flanking region that is homologous to the 5' end of the cleaved bacteriophage genome.

34. The method of any one of claims 31-33, further comprising propagating the circularized recombinant bacteriophage genome in a bacterial host.

35. The method of any one of claims 31-34, wherein the non-terminally redundant linear bacteriophage DNA genome corresponds to a bacteriophage family or order selected from the group consisting of *Myoviridae, Styloviridae, Siphoviridae, Pedoviridae, Tectiviridae, Leviviridae, Podoviridae,* and *Plasmaviridae*.

36. The method of any one of claims 31-35, wherein the at least one exonuclease is selected from the group consisting of Lambda exonuclease, Exonuclease III, RecBCD, Exonuclease VIII truncated, T5 exonuclease, and T7 exonuclease.

37. The method of claim 36, wherein the at least one exonuclease is RecBCD.

38. The method of any one of claims 31-37, wherein the recombination system comprises a 5'-3' exonuclease, a DNA polymerase, and a DNA ligase.

39. The method of claim 38, wherein the 5'-3' exonuclease is T5 exonuclease, the DNA polymerase is Phusion® DNA polymerase, and the DNA ligase is *Taq* ligase.

40. The method of any one of claims 31-37, wherein the recombination system comprises a 3'-5' exonuclease, a DNA polymerase, and a DNA ligase.

41. The method of any one of claims 31-40, wherein the CRISPR enzyme of the sgRNA-CRISPR enzyme complex is a Cas protein selected from the group consisting of Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, and Csf4.

42. The method of any one of claims 31-41, wherein the heterologous nucleic acid comprises an open reading frame that encodes a bioluminescent protein, a fluorescent protein, a chemiluminescent protein, or any combination thereof.

43. The method of claim 42, wherein the open reading frame of the heterologous nucleic acid is operably linked to an expression control sequence that is capable of directing expression of the bioluminescent protein, the fluorescent protein, the chemiluminescent protein, or any combination thereof.

44. The method of claim 42 or 43, wherein the bioluminescent protein is Aequorin, firefly luciferase, *Renilla* luciferase, red luciferase, luxAB, or nanoluciferase.

45. The method of claim 42 or 43, wherein the chemiluminescent protein is β-galactosidase, horseradish peroxidase (HRP), or alkaline phosphatase.

46. The method of claim 42 or 43, wherein the fluorescent protein is TagBFP, Azurite, EBFP2, mKalama1, Sirius, Sapphire, T-Sapphire, ECFP, Cerulean, SCFP3A, mTurquoise, monomeric Midoriishi-Cyan, TagCFP, mTFP1, EGFP, Emerald, Superfolder GFP, Monomeric Azami Green, TagGFP2, mUKG, mWasabi, EYFP, Citrine, Venus, SYFP2, TagYFP, Monomeric Kusabira-Orange, mKOκ, mKO2, mOrange, mOrange2, mRaspberry, mCherry, dsRed, mStrawberry, mTangerine, tdTomato, TagRFP, TagRFP-T, mApple, mRuby, mPlum, HcRed-Tandem, mKate2, mNeptune, NirFP, TagRFP657, IFP1.4, iRFP, mKeima Red, LSS-mKate1, LSS-mKate2, PA-GFP, PAmCherry1, PATagRFP, Kaede (green), Kaede (red), KikGR1 (green), KikGR1 (red), PS-CFP2, PS-CFP2, mEos2 (green), mEos2 (red), PSmOrange, or Dronpa.

47. The method of any one of claims 43-46, wherein the expression control sequence is an inducible promoter or a constitutive promoter.

48. The method of any one of claims 31-47, wherein the heterologous nucleic acid is about 100-500 base pairs in length, or about 500-1500 base pairs in length.

49. A kit comprising (a) one or more labeled vials containing a plurality of linear bacteriophage DNA genomes, (b) at least one CRISPR enzyme, (c) a recombination system, and (d) at least one exonuclease, and optionally instructions for use.

50. The kit of claim 49, wherein the at least one exonuclease is selected from the group consisting of Lambda exonuclease, Exonuclease III, RecBCD, Exonuclease VIII truncated, T5 exonuclease, and T7 exonuclease.

51. The kit of claim 49 or 50, wherein the recombination system comprises a 5'-3' exonuclease, a DNA polymerase, and a DNA ligase.

52. The kit of claim 49 or 50, wherein the recombination system comprises a 3'-5' exonuclease, a DNA polymerase, and a DNA ligase.

53. The kit of any one of claims 49-52, wherein the at least one CRISPR enzyme is a Cas protein selected from the group consisting of Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, and Csf4.

54. The kit of any one of claims 49-53, wherein the at least one exonuclease is bidirectional or unidirectional.

Figure 1

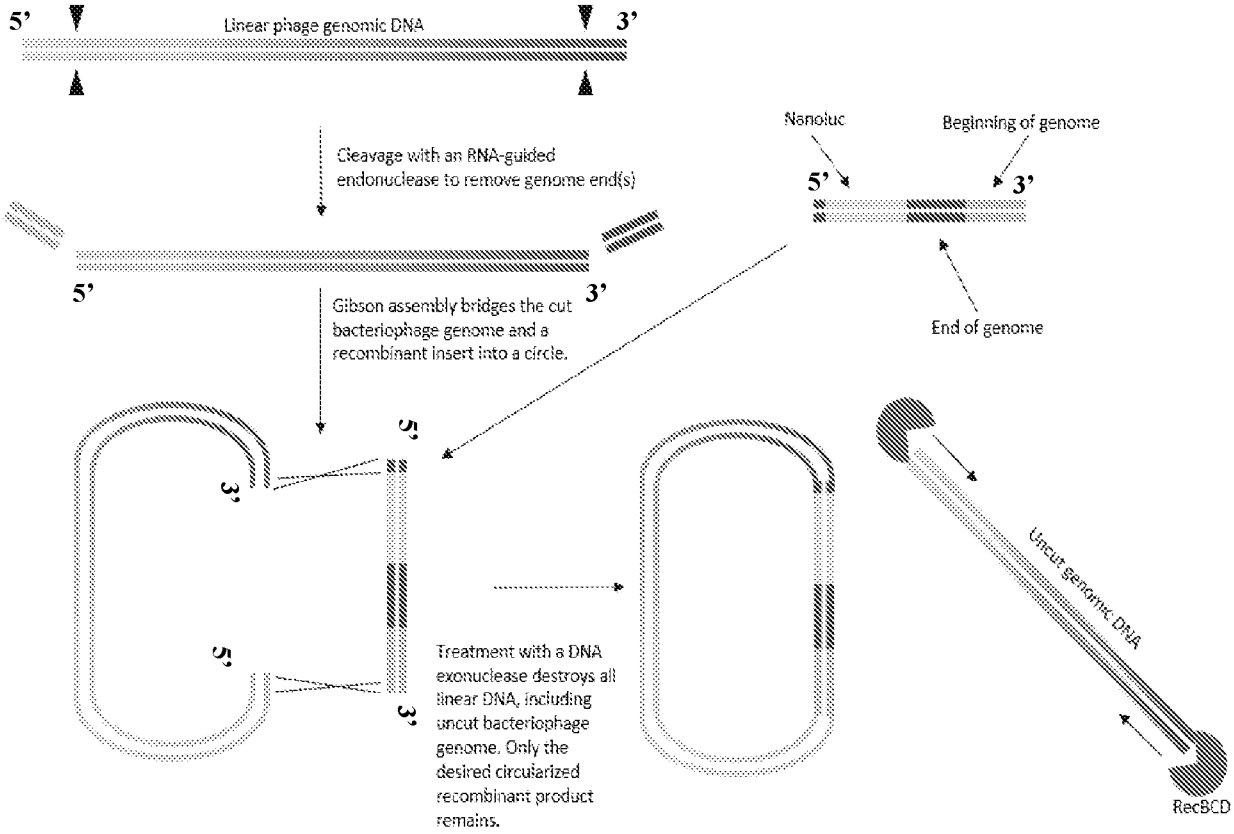

5'
Linear phage genomic DNA
3'
Cleavage with an RNA-guided endonuclease to remove genome end(s)
5'
3'
Gibson assembly bridges the cut bacteriophage genome and a recombinant insert into a circle.
Nanoluc
Beginning of genome
5'
3'
End of genome
5'
3'
5'
3'
Treatment with a DNA exonuclease destroys all linear DNA, including uncut bacteriophage genome. Only the desired circularized recombinant product remains.
Uncut genomic DNA
RecBCD

Figure 2

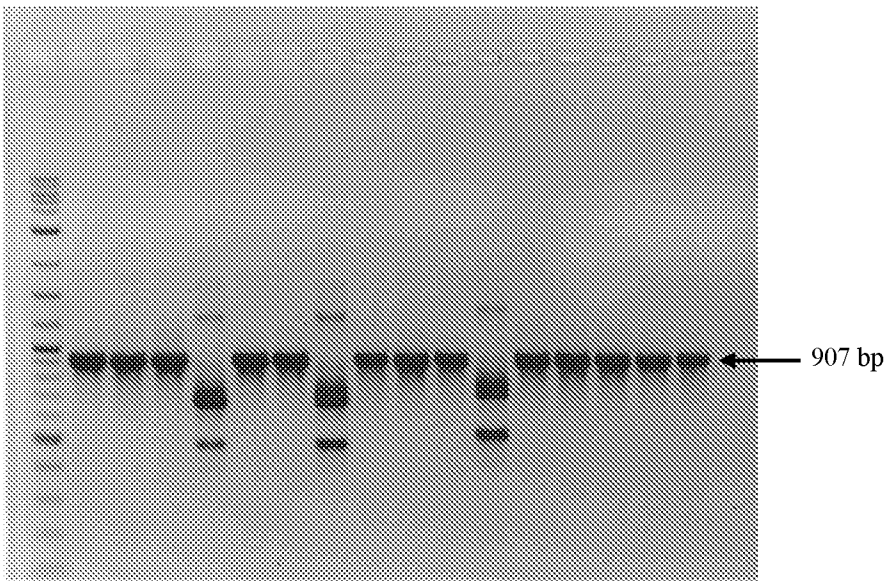


Figure 3

>NC_011043.1 Klebsiella phage K11, complete genome (SEQ ID NO: 1)
TCTCACAGTTTACACTTTTGGTTATCCCCCCGGTACCCTCCAGTTCACCCAAAGTAACCTAGGGTACCCC
TCTTTACCTTTGGTTTAACCTTGGGTGGTACCTTGGGAATCCCTTAGGTGATACCATATGTTGGGGTAAT
GGTGACCTGAGGACACTATATGTTGATGTCTCTGTGTCCCTATCTGTTGGTACTCATTAAGTCACACCTC
AAGTCGCCACCTGAGGTTAGACCAGAGGTAACCACCTGAGGTTATACCTGAGACCATATACCTAAGGTGA
GCTGACTGCTCACGAGGTTCACCGTTTGACTAACGTTTAGCAGTGACTGTTAGTAGGTCACATTAAGAGA
GTCGGTGCTATTAGTAATAGCGGTAAGTATCTCGTTTAGCAGTCCCTGAGACACTGAGAGCGGGACAAGA
GGGTATCGGTGAGTCATCACTATAAGGGCTATTGGTGGTCAGTGTCAACACCATAATCAATTAGGACACA
CTATAGGGAGACACTTAAAGTATTACTATGAGACCATCACCATAAAGATCACTATCACTATAGGTCTAAC
TAAAAGTTTAACTTTAAGTGTTGACATTCAGATTCCTTTATGAGACATTAGCAACCGTTGAGAGACACAA
CGTCACCAACGACCAGACAATACCACGAGTTATCTGGTTAGACTGAGGGTCTCAAGTAGTCATCAACCGG
ACATACGAAAGTGGTTGACTCAACGATGAACAAGTAGTAAGATGTACCACAGATTCACGAAGCACCGCTC
TTTAACAATATGGATTAGTCGCTGATATGTACACCATGACATTAGTGTTTAACTAGTGGTTACATTCAGG
TCTCTGGCAAGGTACGTCCTGTCACCCTGAGAGTAGCCACGATGATAACCACTAACATCGAGGATACACA
GCATGGAAATCGTAATGCAGGCACTGAACCACGGGGTCATTATGACGACAGCACGGGACTACACCGGGGC
CACCAAATACATGGTGCAATACGGCTTACAGTTCACGGTGTTTGACTCGTTCCGTGAGGCACTGCAAGAT
TACACAGATTTGCGTCACCCATTTCGCAAGAGTGTGGGGACTAGCGGTTAACGACAGGTCATCCAAGCGG
TGGCCTGAAAGATAACCACTAACTGAAGGATATACACGATGATTTTCACTAAAGAGCCAGCAAATAAAGC
CTTCGTATTCGTAACCGCTTACCGTGGCTATGAGTCGCTCGAAGTTAACGAGAAGGTCCTCAAGGGTCTC
ATCCGCACCATTAAGACCTATCCGGGTGCTTACGGTAACATCCGCGATGAGAATGTTGTGGGATGCTTCA
AAGAGGCTGGCATGGAGTACGCAACGGAAGAGCGCACGCTCAAGGTTGAATGCACCGTTAAACAAGCGGC
TGAACTGGCGTGGCTGGCATGTAAGACCTACCATCAAGACGCTGTACTAGTGGTTAACTCACAGACCCAC
ACAGCCTCCTTATGGTCTATTGAGAACGTAGGGGAGTATCCTCAGGTATACCCACGCTTGAAAGAGGTGT
CTTTAGGTGGTACGCTGCAACAAGTTGATGCACCTAAGGGTGAATGCTATTCAGTCATCGACGGGCAATA
CTGGGAGGTGGCGTGATGGTTGACTATGGTCTCACACAAGAACACTTGAAGTTATACCGCACGGCCATGG
CATATGGTGCATCGTTCGGTTACTGTATGGCCCAACTGGCCCAGACCTACCGCACACGCAAGGTGATGTA
TGGTAACCCTGTTCGTAATTAGTGTGTACGCCCTGATTGTCCTGTACTTTGTGCGGGACTTTCGCAAGGG
CCTCAAGGTGCACAAAGCATCATTCAGTTACATGAAGTGGGGCGTGTTACCTCGCTTTACTGTACGGCTA
CCTAATGGCCGCTTTAAGGCTAACAAGGTAGGTATTTTCTATATCGCAACCCATTAACACATCGCACATA
AGGAAACAACCAAATGAACTACACCGACATGCAAGAGCGCTTAGACGTCGTCCGTAACCTGCCAATCTGT
GAACTCGACAAGCGCCAGCCGCTGCTGGTAGCACTCATGGCGGACATTGTGAACGCTGAGACGTCCGATG
GTGACGATACGGATAGCGGTTGGGGTCTGGAACGTCAGGACTACTGGCAAACCCTGAAGATTAAGGCCAA
AGATGCTGGGTTTAACCTGCTGGGCAACGGTCACTTCAGCGCAGCGTTTAAGCACGAGCTGCTACCGGGT
AGGGCCATTAAGGTTGGCTTTAAGAAAGAGGACTCAGGGGCCGCATACGTGGCTTTCTGCCGGATGCACC
AAGGACGGGTAGGGATACCTAACGTCTATCACGTAGCGCGTCACGCTGGGTGCTACACGGTGGTACTTGA
TGAGCTGGAACCGTGCCAGCGCAGTGGGAACGATGAGCACGAGCACTACGCAGACCTAGCGTATTACTTT
GTCGAAGGTGAATCGGACCCAGCGGACTACTCGGAGGGCGACCAGCCGTTTATTGAGACGTGCCAAATGA
TTCGCAAGTTCTTCTACGGGATTGCGTCCTTTGATATGCACAGCGGTAACATCATGTTCACCAAGGACGG
CAAGCCAGTGATTACCGACCCGGTGTCATTCTCAGCGGACCGGGACCGGGAGCCTTTCTCACTGGAACCT
GAGGACCTGCTCGCAGAGATTGAGCAGATAGCGCACGACAAGATGATCGAACGCTGTAAGCGCAACAAGG
CTAAGCGTGACCCGAACGGAGAGCTGCGCATCGCACGCCGTAAGGCCAATAAGGAACGTCGAGCACGCCG
TAAGGCACACGCTCGGTGGCGTAAGGAGCGCGAGCGTATTAACGCTGATGCCTTAAAGTTTGACCTTGCT
AAAATCGAGGAGCGGGTACTAGCGTGGCAAATGGGACCAGGCCTGGCGATACAAATGGGCAAGCCGTTAC
CACTCGACAACTACCTTCAGGGTAGACTTATGGGTTAACGAGGTGTATCTTAGGTGTCTCCGAACGGTGA
GGCACCCATAGATAAACTTTATCCACAAAGAGGCACACAATGAACGCATTAAACATTGCACGTAATGACT
TCTCCGAGATTGAACTTGCTGCTATTCCGTACAACATCCTCAGCGAGCACTACGGGGACAAGCTGGCACG
TGAGCAGTTAGCACTGGAGCATGAAGCGTACGAGCTTGGCGAACAACGTTTCCTGAAGATGTTAGAACGT
CAGGTGAAAGCTGGTGAGTTCGCTGACAACGCGGCCGCTAAGCCGCTGGTCTTAACGTTGCACCCACAGC
TGACCAAGCGCATTGACGACTGGAAGGAGGAGCAAGCAAACGCTCGCGGTAAGAAGCCTCGCGCATACTA
CCCGATTAAGCACGGCGTCGCCTCAAAGTTAGCTGTTAGCATGGGCGCTGAGGTGCTAAAAGAGAAGCGC
GGAGTGTCCAGTGAGGCAATCGCACTGCTGACCATTAAGGTCGTCTTGGGGACGCTCACAGACGCCTCAA
AGGCCACAATCCAGCAGGTATCCTCTCAGTTAGGCAAGGCTCTTGAGGATGAGGCCCGCTTCGGTCGTAT
CCGTGAGCAGGAAGCCGCATACTTCAAGAAGAACGTAGCGGACCAGCTGGACAAGCGAGTAGGCCACGTG
TACAAGAAGGCTTTCATGCAGGTAGTCGAGGCCGATATGATATCCAAAGGGATGCTGGGCGGCGACAACT

GGGCGAGCTGGAAAACTGACGAGCAGATGCACGTAGGGACCAAGCTGCTGGAGCTACTCATTGAGGGAAC
TGGTCTGGTGGAAATGACCAAGAACAAGATGGCCGATGGCTCCGATGATGTAACCAGTATGCAGATGGTC
CAGCTGGCTCCGGCCTTTGTGGAACTCCTGAGCAAACGGGCAGGCGCACTCGCGGGTATCAGCCCGATGC
ACCAGCCGTGCGTAGTCCCTCCGAAACCTTGGGTGGAGACCGTAGGCGGTGGCTACTGGTCAGTCGGTCG
CCGTCCGCTGGCACTGGTGCGTACCCACTCCAAGAAGGCGCTGCGCCGCTACGCTGACGTGCACATGCCA
GAGGTATACAAGGCGGTAAACCTCGCGCAAAACACGCCGTGGAAGGTGAACAAGAAGGTGCTGGCGGTAG
TCAACGAGATTGTCAACTGGAAGCACTGCCCGGTAGGTGACGTCCCAGCGATTGAACGCGAAGAGTTACC
GCCGCGCCCGGACGATATTGACACCAACGAGGTGGCACGTAAGGCATGGCGCAAGGAGGCCGCAGCGGTC
TACCGTAAGGACAAGGCCCGCCAGTCTCGCCGTTTGTCGATGGAGTTCATGGTCGCACAGGCTAACAAGT
TCGCTAACCACAAGGCCATTTGGTTCCCGTACAACATGGACTGGCGCGGGCGTGTGTACGCTGTGAGCAT
GTTCAACCCACAGGGTAACGATATGACCAAGGGGATGCTGACGCTGGCCAAGGGTAAGCCAATTGGTCTC
GACGGGTTCTACTGGCTGAAGATTCACGGCGCAAACTGTGCAGGTGTCGACAAGGTTCCCTTCCCTGAGC
GCATCAAGTTCATCGAAGAGAACGAGGGCAACATTCTGGCGAGCGCAGCGGACCCGCTGAATAACACTTG
GTGGACCCAGCAAGATTCGCCGTTCTGTTTCTTAGCGTTCTGCTTCGAGTACGCAGGTGTTAAGCATCAC
GGCCTGAATTACAACTGCTCGCTGCCGCTGGCGTTCGATGGGTCCTGCTCTGGGATTCAGCACTTCAGCG
CGATGCTCCGAGATTCCATCGGTGGTCGTGCGGTTAACCTGCTGCCTTCTGATACCGTGCAGGATATCTA
CAAGATTGTGGCCGACAAGGTGAACGAAGTGCTCCACCAGCACGCCGTCAACGGGTCTCAGACCGTGGTC
GAGCAGATTGCTGACAAAGAGACTGGCGAGTTTCACGAGAAGGTGACTCTGGGCGAGTCCGTACTGGCTG
CGCAGTGGTTGCAATATGGTGTGACCCGCAAGGTGACTAAGCGTTCGGTCATGACGCTGGCATACGGTTC
CAAAGAGTTTGGCTTCCGCCAGCAGGTTCTTGAGGACACCATTCAGCCTGCTATTGACAACGGCGAGGGC
CTGATGTTTACGCACCCTAACCAAGCAGCTGGCTACATGGCTAAGCTGATTTGGGACGCTGTGACCGTGA
CCGTAGTGGCCGCTGTCGAGGCAATGAACTGGCTGAAGTCTGCCGCTAAGCTGCTGGCTGCTGAAGTCAA
GGACAAGAAGACCAAAGAGGTGCTGCGTAAGCGCTGCGCAATCCACTGGGTAACACCCGATGGCTTCCCG
GTGTGGCAGGAGTACCGCAAGCAGAACCAAGCGCGCCTGAAGCTGGTCTTCCTCGGGCAGGCCAACGTCA
AGATGACGTATAACACTGGGAAGGACTCCGAGATTGATGCCCACAAGCAGGAATCCGGCATCGCTCCTAA
CTTTGTTCACTCACAGGATGGCAGTCACCTGCGCATGACTGTAGTACACGCCAACGAGGTCTACGGGATT
GACTCCTTCGCACTCATTCACGACTCCTTTGGGACCATTCCGGCTGACGCTGGGAATCTCTTTAAGGCAG
TCCGCGAGACGATGGTCAAGACCTACGAGGACAACGATGTAATTGCAGACTTCTACGACCAGTTTGCCGA
CCAGCTGCACGAGTCTCAACTGGACAAGATGCCTGCGGTTCCGGCCAAAGGTGACCTGAATCTGCGCGAT
ATCTTGGAGTCTGACTTCGCGTTTGCGTAAGGTCTCAGGCAATTAGGGCACACTATAGGGAACCTTCGAA
TGACCGAGGGTTCCATTACTTAAAGTCTTAACTTAAAGAATACTTAAAGAGGCACGCTATGACTTACTCA
ATCGTTGTAACCATCTTGTTAATCATCACCCTTACGCTCCTCATTAACACCATACGCAATTCACTACGCA
GCGAGGAGCGGCTGGGGCGCAAGGTCCAAGAGGCCAACTCCGCGTTTAGCAGTGAGTCCTGCAAGGTCCT
GCGTCTGGCAGACAGGGCTGACTCGCTCAGTAGACAGGTTCGTTACTTAGAGGGTGAGCTTGAGAGCGAG
AAACAGAAGGTGCGCGATGTGAACGAACTTCGAGAGCACCAGCGGGAACGCATGAAGTTTCTTCGTAAGT
CCCTGAAGGAAGCACAAGACGAGCTGATGATGGTCTCCGACCTGATTCACGTTAAGTTCACCGCAGTGTT
GCCAGACGGTACCCACTCTAAGACGATCTTTAAGTTAGGACTCGGGCCGTGTGGTCTGCACGTTAAGTCC
CTGCGCTGGACCGAGCTGGACGACCGCTATCTGATAGACCAGCTGTGCACCAACGGTGAGCGCAAGCAGT
TCGTCTACTACAAGAGCGAAGTAGTAGGGCGCATCGAGTTCCGCCACGGTAAGCTGTAATTAGGACCCAC
TATCAGGAACATACTCAAGGTCATCATTCGGTGGCCTTCATGAATGTCCCTTACTATCACAATCAGGAGC
AACACCATGTATCAGAACACAATCAATTTCGAGCGCAACCGTGAACGTCAGCAGACTGAGGGTTATATCC
CTAAGGGCCGCAAGCTGAACAAGACGAAGCGCGGCGGTGGCGTGAAGGGTTCCTTCCGTAACGCTAAGGG
TGACAGCGTTGTTAACCAAGAGAAATACTTCGTAGGAGCGTAACAAATGGCTACGGAAAAAAGATGGCTC
TTCGATGGAAGCACCTCACAATGGTCTCGTTTAGGAGCAGCGGAGCGTAGACTACTAGATACGACAGGCC
TGCACGTGGTCATGCTTGACGACCCATTCACTAACACCGTGCTGTTCAACGTATTCGAGCCACGCGGGTC
ACTTCTAATAAGTAAGCGGTTCAGCCACTGGTCGATTGACTCAGCGTCAGACTGGCTGGCAAAACTCACC
GCAGACTACTCGAGCTGGAAGTAATTAGGACACACTATAGGCAGACTCAAGGTCATCGGATTCCGGCGGC
CTTTATGATTGCTTATTGCACACTAAATGAACACTACACTTCGGAGACATCATCATGATGAACATTAAGA
CTAATCCATTTAAGGCCGTATCGTTCGTTCGCTCTGCTATCGAGAAGGCGCTGGAGACTTCCGGTTACCT
CATCGCAGACACTAAGCATGATGGTGTACGCGGGAACATTTGCGTAGACAACACGGCTAACTCATCGTGG
CTCAGCCGGGTCTCCAAGACCATTCCGGCCCTTGAGCACCTCAACGGTTTCGACCAGCGCTGGCAGAAGT
TACTGAAAGATGACCGCTGGATTTTCCCGGATGGCTTCATGCTTGATGGTGAACTCATGGTCAAAGGCGT
GGACTTCAACACCGGGTCTGGCCTGCTGCGCACCAAGTGGCTCAAAGAGACCAACTGGATGTACTCCAGC
AAGGATGGAGTGGTGAAGGGCAAGAAGGAACCTTTCGAGCTGGATACCAAGCAACTAAAAGTTGTCCTCT
ATGATATCATTCCGCTTGACATTATCGAGTCCGGTGATGACTACAACGTGATGACCCTCCTCCGCCTTGA
GCATGTCAAGGTAGCCTTACCAGTCCTGCAAGACCACTTCCCTGAAGTCGAGTGGTGCCTCTCGGAGTCC
CATGAAGTTTACGACATGGACGAACTCGAAGCGCTGTACCGACAGAAACGTGAAGAAGGTCACGAAGGTC

Figure 3 contd.

TGGTGGTCAAGGACCCTCAGGGCATCTACAAGCGTGGTAAGAAGTCCGGCTGGTGGAAGATGAAGCCAGA
GAATGAAGCTGACGGTGTAGTTGTGGGACTCAACTGGGGAACTCCCGGTCTTGCCAACGAGGGCAAGGTG
ATTGGCTTCGAGGTCCTCCTTGAGTCTGGTCGCGTGGTATCCGCCAACAACATCTCTCAGGCACTTATGG
AGGAGTTCACAGCCAAAGTTAAGGCCCACACCATGTGCGCCAATGGTTGCCGGATGTCTAAGGATGTCGG
TATGGATAATCACTCCTGCGCTGGCAAGTGTGCTTACGACCAACACCCGTCGAATAACCCTTATGAGGGC
TGGGCGTGCCAAATCAAGTACATGGAGGAAACTCCAGACGGCTCCCTGCGTCACCCGACCGTTCGACAAA
TGGCGTGGCACTGAGGCTGACCCGACCATCAAGATGTAATTAGGACCCACTATAGGAGACACCAAATGTC
TATCAACCTGATTCTAATCATCGTGCTCATCCTCGCGGCTATCGTGTGGTCAATGAATGACGAGCCACCT
AAAGGAGCATAAACCATGCGCTTACACTTCAATAAATCCAACGGTATCTTCTCGGTTCGCCGGGAGGACC
GCAGCACTGTAGCGGCCACCGAGCGCCACGGTAAGATTCCACGTATCGGCGACACCTTCGAGCTGGCACC
TAGCGTTCACATCTTGGTTACTCGCGGTCTCTACGAATTGGCTCAGACCAAGAGCCGTCCTTTCGTACCC
GTTGTGGTAACCAAGTGGCCACGCCTTCGTCTGTTCTGGGAGCGCATCAAGGAGGTGGTCAATGACTGAA
CGTGAAATTCAAGTTGTGGACCTTCTGGTTGGGCAAAACACTGACCGCCCAGACTCCACAACGTGCGCTG
ATGGCGTCATATGCTACAAGGTATCGTGTAGCGAGTGTCCGCTAAACGTCAAAGGTACGACCATTGGGGA
GGTCCGTACAATGAAGGACAGCAAAGGCTCCGCCCACTTCCCGGAGTGCAAGATATGGAACGGCGCTGGT
CAGTGTACCTGCGAGCCGACCCGAGACGACGGTGTTAAGCAGCCGAGCCACTACCAGCTGTTCGACGGTG
TCGAGGCCATCGAGGTGATTGCTCGCAGCATGACCCAAGAGATGTTTAAGGGGTACTGCCTCGGGAACAT
CCTCAAGTACCGCCTTCGGGCCGGGAAGAAGTCCGAGCTGGCTACCTTAGAGAAAGACATGGCGAAGGCC
GCTTTCTATCTGGAGCTGTACACCAAGCACAAGGGTCTGTGTTATGACGCCAAGTGAGTGGGCAAGAAAG
ATGTACGAGAAGACGCTCGACCCTGCGTACATCACCCTGTATAACATGTGGAAGGAGCGAGAAGATGCAA
AAGTTCGTCGTAACGGTCGAGACAGCTAACGCATCGTACGAACTCCCGGTACACGCTGGGTCTCTTGATG
AGGCCCTCGAAGTTGCCGAGGCGGAGTACGAAGAGTTAGGCCAAGTGACTCGGGTACGCCCGGATAGTCA
TTAGGACACACTATAGGGACACAGGTTGTCCCTCTTTCTGTTATAAACCAAAGGAGATTCACCATGGCAT
TCGCTAAGAAGAAAATTTACACCACCAAGATTGGTACCTGTGAGCCGTACGCTTACTTCAACAAGCCGGA
CTATGGCGGTGAGGGTTTTGAGAACCCACGTGGTACCTACAAAGGTTACGTAACGTTCAAGAACGAAGAC
TGTCAGGAGCTGGTAGACCTCATCGTTAAGACCCATGAGGAAAACTACGCCGCTCGTCTGGAAGCGCACG
AAGCGAACCCGCCTAAGGTTCAGAAGGGTAAGAAACCTCTGAAGCCGTATGAAGGCGACATGCCGTTCTT
CGATAACGGTGACGGCACCACCACGTTCAACTTCAAGTGCTACGGTTCGTACGAGGACAAGAAGACTGGC
GAGACCAAGAAGATTGTTCTGGGCGTAGTAGACGCGAAGGGCAAGCGCATTCAGGACGTTCCGATTATCG
GTGGCGGCTCCAAAGTGAAGATTCGCTTCTCGCTGGTACCGTACGGCTGGTCTGCGGTAGCTGGCGCTTC
CGTTAAGTTGCAGCTGGAAGGCGTGATGCTGGTCGAACTGGCTACCTTTGGTGGTGGCGAAGACGACTGG
GCTGACGAAGCCGTAGAAGGCGGTTACGAAGCGGACGAATCTCGCAGCCGTAAACCTCAGGAAGACCCGG
AAGACTGGTCTGGTGAGGAAGCTGACGAGGGCGAAGCCGAAGAAGACGATGACTTCTAATGGCGGGCTAT
GGGGCCAAAGGGATTCGGAAGGTGGGTGCCTTCCGGTCTGGCCTTGAGGACAAGGTGTCCAAGCAGTTAG
AAGCAAAGGGCGTCACGTTCGATTACGAATTGTGGCGCATCCCTTACGTTATTCCTGCGAGTGACCACCT
TTACACTCCAGACTTCTTGTTACCCAACGGTATCTTCGTGGAGACTAAGGGTCTCTGGGAAGCCGAGGAC
CGCAAGAAGCACCTACTGATTCGTGAGCAGCACCCGGAGTTAGACATCCGGTTAGTGTTCTCTTCGAGTC
GCACTAAGATTTACAAAGGGTCACCAACCAGTTACGCTGAGTGGTGTGAGAAGCATAACATCTTGTTTGC
CGACAAACTGATTCCCGTAGACTGGCTGAAGGAGCCGAAGCGTGATGTACCGTTCGGCAAGTTCAAGCAG
AAGAAAGGAGCAAAGTAAGTATGGCCAAGGTTCAATTCACTAAGCGACAGGAGACCTCTCAGATTTTCGT
TCACTGTTCCGCCACCAAGGCAAACATGGACGTAGGCGTCCGTGAGATTCGCCAGTGGCACAAAGAGCAG
GGCTGGCTGGATGTAGGGTATCACTTCATCATCCGTCGTGACGGTACCGTTGAGGCGGGCCGCGACCAAG
ACGCTGTGGGTTCACACGTCAAGGGATACAACTCGACCTCTGTCGGTGTGTGTCTGGTAGGTGGTATCGA
CGCCAAGGGTAACCCCGAGGCAAACTTCACGCCTCAGCAGATGAGCGCACTGAATGGGTTGCTGCACGAG
CTGAGGGGGACCTACCCCAAGGCTGTCATTATGGCGCACCACGATGTAGCGCCGAAGGCTTGTCCTAGCT
TCGACCTGCAACGTTGGGTAAAGACTGGCGAGCTGGTCACTTCTGACCGTGGGTAAACATTAGGGCACAC
TACAGGGAGACAATTACGTTTCCCTGTTGTCACACATTCTGTACAAATTATGGTCAGGCTAAGGTGCACT
TGGCGTAGCGCTGCGTTTCATTCGGGTTCGATTCCCGGACTGACCACACCAACGGAGATTACTTTATGAA
CAAGTTCAAAGAACACTTTGCTGACTCATGGCCACTGTATGTGTACGCATCGGCATTCATCATTGGCGCA
CTGCGAGTGTTGCTCCCATGAGTTACGGGGACAGTCGAGAAGACGGTCAGGAAAGTATCTTCCTGTTCCA
CGCTCCGTGCGAAAACTGTGGTTCTTCTGATGGTAACTCAGTGTACTCTGACGGGCATGAGTATTGCTTC
GTGTGTCAACACCGGGTTCCCGGCTCAGAGGAACGTACCGAAAAGTTATCATCGAGAAGACCCAAAGGAG
GGAATTACGGGATGAATACACAAGGCTCAGGACTACTGGTATTCGGCGAGAGTGACGGTCGGTACACTGA
CCTGACTGCTCGTGGTATCTCAAAGGCGACATGCCAGAAGGCTGGCTATTGGGTCGCCAAGGTCAGAGGA
ACCGCCTATCAGGTGGCCGACTATCGTGACCAGAATGGCTCCATCGTCTCCCAGAAGCTGAGGGACAAGG
AGAAGAACTTCTCTACCCGAGGGTCCCACAAAGGGGATGCACTGTTTGGTAAGCACCTATGGAATGGTGG
TAAGAAGATTGTCATCACCGAGGGTGAAATCGACATGCTAACCGTGATGCAACTACAGGACTGTAAGTGG

Figure 3 contd.

CCTGTGGTTTCTCTCGGTCACGGTGCGTCAGCCGCTAAGAAAACTTGTAGTGCAAACTACGAGTATTTTG
ATAGCTTCGACCAGATTATCCTGATGTTCGACATGGATGACCCCGGTCGGGCAGCTGTAGAGGAAGCCGC
TCAGGTTCTCCCTCCCGGTAAGGTGCACGTAGCTGTGCTGACCGAGAAGGATGCCAACGAGTGTTTACTC
AAAGGTAAGGGAAAGGAGGTTCTCGACCAGATATGGAACGCGGCACCTTGGGTACCTGATGGTGTCATCG
GTGCGATGTCCATGAAGGACCGAGTGCGTGAGGCCATGACCAGCGAACAAAGCGTAGGATACCTTTTCTC
GGGATGCCCGGGACTGAATGACCGAACCTTGGGTGCACGTGGTGGCGAAGTCATCATGGTCACTTCTGGG
TCAGGAATGGGTAAGTCTACGTTCGTTCGTCAGCAGGCTCTAGGGTTCGCCAGAGGGCAAGGACTGAGGG
TAGGCATGGCGATGCTTGAGGAGTCCGTAGAGGAGACCATGGAGGATGTCCTAGGGATTGCTAACGGAAT
CCGCTTACGGCAGCAGCCTCGGGAGTTCAAGCAGAAACTCATTGAGGATGGTACGTACGATGAGTGGTTC
GATGAGCTGTATGGCTCCGACCAGTTCCATCTCTACGACTCCTTTGCGGAAGCTGAGGTGGACCGCCTGC
TGGCCAAGCTGCACTACATGCGCACAGGGTTGAACTGTGACGTAATCATTCTGGACCACATCTCAATCGT
AGTGTCTGCCTCGGAGGAATCCGATGAGCGCAAGATGATTGACCGACTCATGACCAAGCTGAAAGGGTTC
GCTAAGTCAACCGGAGTGGTACTTATTGTTATTTGCCACCTGAAGAACCCGGAGAAAGGTAAAGCTCATG
AAGAAGGACGTGCTGTTTCCATTACTGACCTGCGTGGGTCTGGGTCTCTGCGCCAGCTCTCTGATACTAT
CATTGCACTTGAGCGTAATCAGCAAGGGGATATGCCTAATCTTGTCCTCCTTCGTATTCTCAAGTGCCGC
TTTAATGGTATTGGCGTTGGCATTGCGGGGTACATGGAGTACAACGAAAAGACAGGACTCCTTGAACCGT
CTAGCTACACTGGCGGAGAAGGAGAGGGAGATACTGGCTGGGAAGGCCACGAAGAAGACGATTACTAAAC
GTAAATGCAATGGGGCGTACTGCTGGTGCGCCTTTGACCCTGATTATCAATAACGGAAGGAGAGCCATCA
TGTTTAAACTTATCGAAGCATTAGGCCGTCTGGTCATCGCACTGTACGTACGTGAAGCCAAGGCACTGGA
CAAAGCGTCCAAGGTGGAAGCGAAAGCAGCCGCTAAGCTGGCTAAGGCAGCCGACAAGGCACGTCAGGCA
TCTCTGGATGCAACCGCAGAGGCAGCTAAAGTTGCCGCTAAAGCTCAGAAACTTAAGGAGTTCTTCTAAT
GACTACCAAAGTTAAATTCCCCGGCAATACCATTCAGCTGTCCGACACCGTTGACCAGTGGGGACGCAAG
GTTCACATCAACGTTCGCAACGACAAGGTCACTCTGGTCTACCGCTGGAAGGCCAAGAGCGATAATCGTG
CGCATACTCAGCGTGTGACCCTCGACGACACACAGGCAGCTCGGCTGCTGGCGTCCGTAGCTGTAGCCGC
TACTGTGGCCATAGGTGAGGACAAAGTGCGTGAGGCAATCCTGAGCAAAGAGGTTGGCGAAACGTCCGTG
CGTCTGGCCGAAGCGTCAGAAGTTAAGTGATAAACTCAAGGTCATTACTATATGTAGTGGCCTTTATGAT
TATACACACAACATATTGAGAGGACATTACCATGCGTAAACCTGAAGAGATTCGTAAAGAGATTGAAGCG
CTGAACAAAGAGCTGGCTGAGGCCAAGACCTATGAGGCTAAGCGTGACGCTGCTGTGCACATTCTGGAGA
ACTTAGGGTGGACCCACAGTGGCCACAAGGGCTGGCAGAAGCCTTCGCAAAAGTGGAGCGACTATAAGGC
TCCCCTGAAGGCTGGTGAGCTGGCAACTTGGGACGACAAGGTACTAGGTGGGATAGTGTACATACGCAGT
GTGGGCGATAAGTACGCTCAGGTGTCCCACGTTCGTGGTGTTAGTAGACTGGGAGCTGATGTACTGAACA
GTAGCTTTGCTGTCGAGAAGAGTAAGTTAACCGTGCGTCCTCGTGAGTATTTCATCGGGCGTCGTTAAGC
AACAGGAGACCACTATGTTAGTAACCGATATCGAGGCGAACAACCTCTTAGAGAAAGTCACTCAGTTCCA
CTGTGGTGTCATTTATGACTACAGTACGGACGAGTACGTATCGTATCGACCTTGGGACTTCTCAGCGTAT
CTCGATGCGTTGGAAGCTGAGGTGGCTCGTGGTGGTCTCATCGTATTCCACAACGGTCACAAGTACGATG
CCCCAGTGTTAACCAAGCTGGCCAAGCTCCAGTTAAACCGAGAGTTCCACCTGCCGCGTGAGAACGTAGT
GGACACGTTGGTGCTCAGTCGTTTACTGTTTGCGAACATTAAGGACTCCGACATGGCCCTGCTGCGTTCC
GGTAAGTTACCCGGTAAGCGCTATGGGTCTCACGCTCTGGAGGCGTGGGGTTACCGCTTGGGCGAGATGA
AGGGTGAGTACAAGGACGACTTCAAGAAGCTACTTGAGGAACAGGGAGAGGACTATGTTGACGGTGCTGA
GTGGATTAGCTTCAACGAGCCGATGATGGCGTATAACGTTCAGGACGTTGTGGTGACCAAGGCTCTCTTA
GAGAAGCTGCTGAGCGACAAGCACTACTTCCCACTGTTTGGTAGTAACACCATAGAGTTCTACACCTCAG
CGTACTGCTTGAGGTTCTGGGAGGAGGCTTGTGAGGCCGTCTGGTTGGAACATCGGGCCGCTTGGTTACT
CGCTAAGCAGGAGCGCAACGGATTCCCGTTCAACACCAAGGCCATTGAGGAGTTGTACGTTGAACTCGCT
GGTCGTCGTTCTGAACTCCTTCAGACACTTACCGACACTTTCGGAACTTGGTACCAACCTAAAGGCGGCA
CTGAGTTATTCCTGCACCCGCGCACTGGTAAACCTCTGGGTAAATACCCACGAGTGAAGTACCCGAAACA
GGGTGGTATCTACAAGAAACCCAAGAACAAAGCTCAACGAGAGGGTCGTGAACCCTGTGAGCTGGACACT
CGGGATTACGTAGAGGGTGCTCCATACACACCAGTAGAGCACGTTGTGTTCAACCCAAGTAGCCGAGACC
ACATTGCGCTCAAGCTGAAGGAAGCCGGATGGGTACCCACAGAGTTCACCGAAAAGGGTGCACCTAAGGT
AGACGACGAGGTCCTTGAGCATGTTCGTGTGGGGGACCCTGAGAAGCAGCGCTGTATCGACCTCATCAAA
GAGTACCTGATGATACAGAAGCGTATCGGTCAGGCGGCTGAGGGCGACAAAGCGTGGCTACGTTACGTTC
AAGAGGATGGTAAAATCCATGGAAGTGTTAACCCTAATGGTGCAGTTACAGGGCGAGCAACGCATAGCTT
CCCTAACCTTGGTCAAGTTCCGGGCGTTCGTTCGCCGTATGGTGAGCCTTGTCGAGCAGCGTTCGGCGCA
GAGCATCACTTGGACGGACTTACCGGACAGCCTTGGGTTCAAGCAGGCATCGACGCCAGCGGACTCGAAC
TCCGTTGTCTGGCACACTTCATGTCTAAGTACGACGACGGGGCATATGCGGATGTCATTCTCAACGGTGA
TATACACACAGTCAACCAAACGGCGGCTGAGTTGCCAACACGTGATAACGCCAAGACATTCATCTACGGT
TTCCTCTATGGTGCTGGAGACGAAAAGATTGGACAGATTGTGGGCGCAGGTAAGGAACGCGGAAAGGAAC
TCAAGAAGAAATTCCTTGAGAACACCCCAGCAATCGCAGCCCTGCGTGAAGGAATCCAGCAGACCCTCGT

Figure 3 contd.

CGAGTCATCCCGATGGGTTGCCGGAGAGCAGAAGGTCAAGTGGAAACGACGCTGGATTAAGGGACTGGAT
GGAAGAAAGGTACACGTTCGGTCACCACATGCCGCGCTCAACACGTTGCTTCAGTCAGCGGGTGCGCTCA
TTTGTAAGCTGTGGATTGTCGAGACTGAAGAGTTGCTTCTTAAGGCAGGATTGAAGCACGGATGGGATGG
CGACTTCGCCTACATGGCGTGGGTTCACGATGAAATACAAGTGGCCTGCCGGACCTCAGAGATTGCACAG
CAGGTGATTGACATAGCGCAGCAAGCTATGCGTAACGTGGGAGACCACTTTAAGTTCCGTTGCCGTCTGG
ACACAGAAGGTAAGATGGGTCCGAACTGGGCCGTATGTCACTAATAATACAGGAGATTTATCATGGGTAT
TAACAAACAGTTTCGCGTAACGTTCGATGTAACGGCTACTATGAGTGATGACCAAGAGCGGGAGTTCCTT
GAGGACCTACTATCTCTTGCGTATGGCGTGGACGACAAACGTCAGGCGCACATTGTAACCGAAGCAATCA
CCAAAGGTCATGAGGCGGCACTGGCATTCGTCATGCAGAGTGGTCTGCGGGAAGCTATTAAGGACATCGG
TAAGGAGCTGAGCTGCTCCGCTGTGACAGTACGCTTCTCTCCGGCAACCGTGAGGGTGACTAAGTGAGCG
AGTACCTCAAAGTTCTGGCGGCCCTCAAGGGCTGCCCTAAGTCCTTCCAGTCGAACTACGTGCGGAACAA
CGCCGCGTTAGTCGCTGAGGCTGCGAGCCGTGGTCACATTTCATGTCTGACCATGAGTGGTCGTAACGGT
GGCGCTTGGGAAATTACCAGTGCCGGAGTGAAATTCCTTAAGGCCCATGGAGGTTGTCTATGAAAGACTT
TTTAGGTAACGATATCGAGATTGGCGACACCATTGTGTATGCTGACGCTGGTGGCCGTGGAGGCTCTTCG
GGTCTTAACAAGACAGTAGTTACCCGAATGACTGATAAACAGGTCATGGTGTACGAATCAACGTGGTCAA
AACTGTGGCGTCCGTTTGACCGTGTTGTGGTTGTTGCTAAGGGAGGTTCCCAATGAAGCACACATTGTTA
TCCTTCAGTGACTACCGGGCAACCCAGAAGATTGCCAAGGGTGTCCTTGTGATGGATGGTGACTGGTTGG
TATTCCAAGCCATGAGTGCCGCTGAGTTCGATGCCTCGTGGGAGGAGGAGATTTGGCACCGTTGCTGTGA
CCACGCTAAGGCCCGAGAGATTCTGGAGAACTCCATCGAGTCCTACAAGGGCCGCAAGAAGGCGTGGAAG
AATGCAGACGTTGTCCTAGCGTTCACTGACCGTGTCAACTGGCGCAAGCTGCTTGTGGACCCGACGTACA
AAGAGAACCGCGCAGTCGTCAAGAAACCTGTGGGTTACTTTGAGTTCCTTGAGTACGTCTTTGAGTCCTA
CACATGTGTCCTTGAGCCTCAGCTCGAAGGTGATGACGTGATGGGTATCATCGGGTCTAACCCTCTCGTG
TACAACTACGAGAAGGCCGTGCTGGTCTCCTGCGACAAGGACTTTAAGACCATCCCGGATTGTGATTTCC
TGTGGTGCACGACTGGTAACATCCTCGTTCAGACTCAGGAGACAGCCGACTACTGGCACCTCTTCCAGAC
TATCAAGGGTGACATCACCGATGGTTACGGTGGGATTCCCGGATGGGGAGATACCGCTGAGGACTTCCTC
AAGGAACCCTTCATTGTGGAGCCTGTAACGTCCGTGCTGAAGTCCGGTAAGAACAAGGGCCAAGAGGTAA
CCAAGTGGGTGAAACGCGCTCCTGAGCCGGGAGAGACGCTCTGGGACTGCATTAAGTCCATTGGTGCCAA
AGCAGGGATGACCGAAGCGGAAGTAATTAAGCAGGGCCAGATGGCTCGCATCCTCCGTTCTGATGAGTAC
AACATCGAGACTGGGGAGATTACTCTATGGCAACCGGGCAGCTGATTCTCATCGTCCTGACCATGGGCTT
AGTTGCTCGTGGTCTCTGGATGTTGGCCTTGATTATCAAGCAGATAGTCGAGCATAAAGCAGAGTGATAA
ACTCATGGGCACAATTAGGACCCACTATAGGGAAGTGCCCATTATGATTATTACTTAAAGATTACTTAGA
GAGGAGACTCAAATGTTAAAACCTATAGAGCACATCCTTAACAATCCTAATGACCTTCCTGACGTACCGC
GAGCTGTCAAGGAGTACCTACAGTCTCGCTTCAATGCTGACTTCCTGTATCAGTCAGAGGTCCGTAAGCT
GCGTGAGGCTGGCCACAGTGAGGAGTTCATCTCCGGTGTACTGTATGGTCACTACATGGCTTCTCGTGTC
CTTGACGAGATGGAGGGCCGCCAGCGTGCACTCAAAGAAGGAGATTGATTATGTGTTTCTCACCTAAGAT
GAAAGCACCTAAGGTCGACACAACGACTGTCCCTGAGCCAGCTCCGCTCACTGAGGAACCTAAGGGTATC
CAGTACGGTGGCGACGAAGACTCAAACAGCACCACTCCTGAGGTGTCAGGGCGTAAGTCACTCAAGGTGA
CCAAGACGACCGAGCCCACAGGGTCAGTCAGTAAAATCCGTAAGTCAGCTTTAGGAGGCTAACATGGGAC
TGTTCAAGAAAATCAAGAAGGCTATCTCCAAGGTAGTCAAGGCACCACTCAAGGCCGTGGGTCTAGCAGC
AGATGCGCCTAACGTGCAGACAGCCGCTGAGACACCTGTGGCAGCACCTCAGGAAGCACCGAAAGAGGTC
GTGGAGGACGTTGAGTCTTCAGCAGACACCGAGTCTGGTAAGAAGAAATCCCGAGCGTCTGGTAAGAAGT
CCCTCTCAGTTTCCCGCAGCTCAGGCGGTGGGATTAACTTATGATTGGTTACGGGGAGGGCTAACAAATG
GCAGAAGTTAAACTCGAAGGCTTCGCAGAGGAGGGAGCCAAGGCGGTGTATGACCGTCTGAAGAACGACC
GACAACCTTACGAGACACGAGCAGAGTCCTGTGCGCAGTACACGATTCCATCACTGTTCCCTAAGGACTC
CGATAACGCATCAACAGATTACACGACTCCGTGGCAATCCGTAGGTGCTCGCGGCCTGAACAACCTAGCG
TCCAAGCTGATGTTGGCCCTGTTCCCGATGCAGTCATGGATGAAGTTGACCATTAGTGAATACGAAGCGA
AGAACCTTCTGGGTGACGCTGAGGGTCTCGCTAAGGTCGATGAGGGCCTATCAATGGTAGAGCGAATCAT
CATGAACTACATCGAGTCCAACAGTTACCGAGTGACTCTCTTCGAGTGCTTGAAGCAACTGTGTGTGGCC
GGGAACGCATTGCTGTACTTACCGGAGCCTGAGGGTTACACCCCGATGAAGCTCTATCGCCTGAACTCGT
ATGTGGTCCAGCGAGACGCTTTCGGTAACGTACTCCAGATTGTCACTCTCGACAAGATTGCGTTCAACGC
TCTCCCTGAGGATGTCCGCAGCCAAGTGGAAGCAGCCCAAGGTGAGCAGAAGGAAGACGCTGAGGTTGAC
GTCTACACCCACGTGTACCTGAACGAATCCGGGGATGGCTACTCGAAGTACGAAGAGGTTGCCGAAGCAG
TAGTACCGGGCAGCGAGGCTGAATACCCGCTCGAAGAGTGTCCGTACATTCCGGTCCGCATGGTCCGCAT
CGACGGTGAATCCTACGGTCGTTCCTACGTGGAAGAGTATCTGGGTGACCTCAAGTCCCTAGAGAACCTC
CAAGAGTCCATCGTGAAGATGGCGATGATTACCGCGAAGGTCATCGGTCTGGTAGACCCGGCAGGTATCA
CTCAGGTCCGCCGACTCACGGCAGCACAGTCTGGTGCGTTCGTACCGGGCCGTAAGCAGGACATTGAGTT
CCTCCAGCTGGAGAAGTCCGGTGACTTTACCGTAGCGAAGAACGTAAGCGACACCATTGAGGCTCGCCTA

Figure 3 contd.

TCGTATGCCTTTATGCTCAACAGTGCGGTACAACGTACAGGCGAGCGAGTCACAGCCGAAGAGATTCGGT
ACGTGGCGTCAGAGCTGGAAGATACCCTAGGCGGTGTCTACTCGATTCTATCGCAGGAACTCCAGCTGCC
TCTGGTAAGAGTGCTCTTGAAGCAACTACAAGCCACGCAGCAAATCCCGGAGTTACCTAAAGAGGCCGTC
GAGCCAACTATCAGCACTGGCCTTGAGGCTATCGGACGTGGTCAGGACCTTGACAAGCTGGAGCGGTGCA
TTGCCGCATGGTCAGCCCTTAAGGCCCTCGAAGGTGATGACGACCTCAACTTGGCTAACCTCAAGTTACG
TATCGCTAACGCTATTGGACTCGACACTGCTGGTATGCTTCTCACTCAGGAGCAGAAGAACGCCCTTATG
GCACAGCAAGGTGCTCAGATTGCCACACAGCAAGGGGCCGCAGCGCTGGGTCAAGGGATGGCCGCACAGG
CTACTGCAAGTCCTGAAGCGATGGCCGCAGCAGCTGATTCAGTAGGTATGCAACCGGGCATGTAATTAGG
GCACACTATAGGGAGACCGATTGGTTTCCCTCTTAGTCTTAACTTTAAGGAGATTGAAATGGCTGGCGAA
TCTAACGCAGACGTATACGCATCCTTCGGTGTTAACAGTGCTGTACTGACTGGTAGTACACCTGAGGAGC
ACCAAGAAAACATGTTGGCTCTTGATGTTGCTGCCCGTGATGGCGATGATGCAATCGAGCTGAACACAAA
CAGTGATGACCCGTATGGTTCCGATGTGGACCCGTTCGGTGAACCTGAAGAGGGCCGTATGCAGGTCCGT
ATCTCCGCTGACGGTTCAGACGAACAGGACGGCGAAGAGGGTCAGGGTGACGAAGAACAGCAGGGCGACG
AAGAGAGTCAGCCGGAGGAAGTAACCGATGAAGGTGAACCTGAAGAGTTCAAACCTATTGGTGAAACTCC
GGCTGACATCAACGAAGCCTCTCAGCAGCTGGAAGAACACGAAGCTGGCTTTAACGACATGGTTGCTACT
GCAATCGAACGCGGTCTCTCACAGGATGCTGTGACCCGTATTCAGCAGGAGTACCAGAACGAGGACAGTT
TGTCCGACGAGTCTTACCGAGAGCTGGCCGAGGCGGGCTACAGTAAGGCGTTCGTCGATGCGTACATTCG
CGGTCAGGAGGCTCTGGTCAACCAGTACGTTGAGAAAGTGATGGACTTCGTGGGAGGCCGTGAGCGATTC
CAGCAGGTCTACAGTCACATGCAGACCAATAACCCTGAGGGTGCCGAGGCGCTCATCAAGGCTTTTGAGT
CTCGTGATGTAGCCACCATGAAGACGATTCTGAACCTAGCGGGACAGTCTCGTGATAAAACCTTTGGTAA
GAAAGCTGAGCGCTCTATTGCCAAGCGTGCAACCCCAGCGAAACCTGCTCCCCGCAAGGCTGTAGGCTTC
GAGTCTCAAGCTGAGATGATTAAGGCGATGTCCGACCCGCGCTACCGCACCGACTCTAAGTATCGTCGTG
AAGTAGAGCAAAAGGTAATCGACTCAACGTTCTAATGAATTAGGGCACACTATAGGGAGACCATCAGACT
GAACACGGTGACGTCCACTGGCTCCCTTCGAGTTACACAATGAGTATCACCTCGTTTCAAGTAGTAACTG
ACGCGACCTTAGGGCAAGACCTTATGATAGGCGCGGAGAATCACCCCAAGAGCTTGGCAACGATAGGCCC
GTTTGGTCAGCGTAATGACTAATTCTATTCGTAAACAACATAAGGAGATTCAACATGGCTAACATGCAAG
GTGGACAGCAGCTCGGTACTAACCAAGGTAAAGGTCAATCCGCAGCAGACAAGCTGGCGCTATTCCTGAA
AGTATTCGGCGGTGAAGTCCTGACCGCATTCGCTCGTACCTCTGTGACCACCAACCGTCACATGCAGCGT
CAAATCAGCTCCGGTAAGTCCGCACAGTTCCCTGTGATTGGCCGCACCAAGGCTGCTTACCTGCAACCGG
GCGAGTCTCTGGATGACAAACGTAAAGACATCAAGCACACCGAGAAGACCATTAACATTGATGGCCTGCT
GACCGCTGACGTGCTGATTTACGACATCGAAGACGCGATGAACCACTATGACGTGCGCTCCGAGTACACC
TCTCAGATTGGTGAATCTCTGGCGATGGCGGCGGATGGTGCGGTTCTGGCTGAGCTGGCTGGTCTGGTTA
ACCTCGCTGATTCCGTCAACGAGAACATCGCGGGTCTGGGCAAACCGTCCCTGCTGGAAGTTGGTGCTAA
GGCTGACCTGACCGACCCGGTTAAACTGGGCCAAGCGGTTATCGCACAGCTGACCATTGCTCGTGCGGCC
CTGACCAAGAACTACGTCCCGGCGAACGACCGTACGTTCTACACCACCCCGGACGTGTACTCTGCGATTC
TGGCGGCTCTGATGCCTAACGCTGCGAACTATGCGGCTCTGATTGACCCTGAGCGTGGCTCTATCCGTAA
CGTGATGGGCTTCGAAGTCGTAGAGGTTCCGCACCTGACCGCTGGTGGTGCTGGTGATGACCGCCCGGAC
GAAGGCGCAGAAGCGACCAACCAGAAGCACGCCTTCCCGGCAACTGGCGGTAAAGTCAACAAAGAGAACG
TTGTGGGCCTGTTCCAGCACCGTTCCGCTGTCGGCACCGTTAAACTGAAAGACCTGGCTCTGGAGCGTGC
TCGTCGTACTGAGTATCAGGCTGACCAGATTGTTGCCAAGTACGCGATGGGTCATGGTGGTCTGCGTCCA
GAATCTGCTGGTGCGCTGGTTTTCACAGCAGCCTAAGCGTAAATACCTTTAGTGCTCGGACGGTAACTCC
GTCTGAGTATGAGGTACAGACTGTGGCCATTACTGGTGATTCACTTAAGGTGACACTTGGTGGGCTGGAG
GGAGTAACGGACTGGTCAACACTTGAGGTAACTTATGGTACTTCCGGGATTGCCAGCCACACTCGCCGGA
CCAACACGCTGTACTTCAAAGGAATCGCTGTGGGCGAAACTCTAGTGACTGTCAGCTTTGACGGGTCTGA
AAGGAAGTCCTTTAAGCTGGTCGTGACTAATTAAAACTAAGCCAAACCCCTTGGGGACCACTCACGGTCT
CTGAGGGGTTTTTTCGTTAGGAGCTTACATTATGAACATGCAAGATGCTTACTTTGGGTCTGCCGCTGAG
CTGGATGCTATCAACGAGATGCTCGCAGCTATCGGTGAATCCCCGGTGACCACCCTTGACGAAGATGGTA
GCGCAGACGTAGCTAACGCTCGTCGTATCCTCAACAGGATTAACCGCCAGATTCAGTCTAAAGGTTGGGC
CTTCAACATCAACGAGTCGGCCACGCTGACCCCTGACGCGGACACTGGGCTTATCCCGTTCCGTCCGGCC
TACCTGTCCATCCTTGGTGGCCAGTACGTCAACCGTGGTGGTTGGGTGTACGACAAGTCCACAGAGACGG
ATACCTTCTCTGGGGCAATCACAGTGACCCTAATCACACTTCAGGACTACGACGAGATGCCTGAGTGTTT
CCGCCAGTGGATTGTCACCAAGGCCAGCCGTCAGTTCAACTCTCGGTTCTTCGGAGCGGAGGACGTAGAG
AACTCTCTGGCACAGGAAGAGATGGAAGCGCGTATGGCATGCAACGAGTACGAGATGGACTTCGGTCAGT
ACAACATGCTTGACGGCGACGCATACGTGCAGGGTCTCATCGGTCGTTAATCAGAAACTTAAGGAGGACC
AAATGGCTCTCGTATCACAATCAATCAAGAACCTCAAGGGAGGCATTAGCCAGCAGCCTGAAATCCTACG
GTACCCAGAGCAGGGTACACTTCAGGTCAACGGTTGGTCCTCCGAGACTGAGGGTCTCCAGAAGCGACCA
CCTATGGTGTTCATCAAGTCCCTTGGACCTAGGGGCTACTTGGGGGAAGACCCGTACATTCACCTCATCA

Figure 3 contd.

ACCGAGATGAATACGAGCAGTATTACGCAGTGTTCACTGGGAACGATGTTCGGGTATTCGACCTGTCCGG
CTATGAGTACCAAGTAAGAGGTGACCGCTCGTATATCTCAGTAGTCAACCCTAAGGATAACTTGCGGATG
ATAACCGTGGCCGACTACACGTTCATCGTTAACCGTACCCGACAGGTCCGCGAGAACCAGAACGTGACCA
ACGGTGGTACCTTCAGGGACAACGTGGACGGTATTGTCAACGTCCGTGGTGGCCAGTATGGTCGTAAGCT
CGAAGTGAACATTAACGGTGTATGGGTCAGCCACCAGCTGCCTCCGGGTGACAACGCTAAGGATGACCCG
CCCAAGGTTGACGCACAGGCCATTGCGGCTGCACTCGCTGACCTACTTCGTGTGGCCCACCCAACGTGGA
CATTCAACGTGGGGACTGGTTATATCCACTGCATCGCACCAGCTGGGGTAACTCTTGATGAGTTCCAGAC
GAGGGACGGTTACGCGGACCAGCTGATTAACCCGGTGACCCACTACGTTCAGAGCTTCTCTAAGTTGCCA
CTTAACGCGCCTGACGGGTACATGGTGAAGATTGTCGGGGACACGTCCAAGACTGCTGACCAGTATTACG
TGAAGTATGACGCTTCTCAGAAGGTCTGGAAGGAAACCGTGGGCTGGAACATCTCGGTCGGCCTTGAGTA
TCACACGATGCCTTGGACTCTGGTGCGTGCAGCTGACGGTAACTTTGACCTCGGGTATCACGAGTGGAGG
GACCGCCGTGCTGGTGACGACGACACTAACCCTCAGCCGTCCTTTGTTAACTCAACGATAACCGATGTGT
TCTTCTTCAGGAACCGCTTAGGGTTCATCTCTGGGGAGAACATCGTGCTTTCCCGCACCAGTAAATACTT
TGAGTTCTACCCGCCGTCAGTGGCCAACTATACGGACGATGACCCGCTAGATGTTGCCGTGAGTCATAAC
CGTGTGTCGGTCCTTAAGTACGCTGTGAGCTTCGCTGAGGAGCTTCTGCTGTGGTCCGATGAGGCTCAGT
TCGTCCTGTCGGCCAACGGTGTGTTATCCGCTAAGACTGCACAGCTGGACCTGACCACTCAGTTCGATGT
GTCAGACCGTGCGCGTCCTTACGGTATCGGCAGGAACATCTACTATGCGTCTCCTCGAAGCTCCTTTACG
TCCATCATGCGCTACTACGCGGTACAGGATGTAAGCTCTGTGAAGAACGCAGAGGACATGACGGCCCACG
TCCCGAACTACATCCCGAACGGTGTGTACAGTATCAACGGGTCCGGTACTGAGAACTTCGCGTGTGTGCT
GACCAAGGGTGCTCCCAGCAAGGTGTTCATCTACAAGTTCCTCTACATGGACGAGAACATTCGACAGCAG
TCATGGTCCCACTGGGACTTCGGAGATGGTGTGGAGGTGATGGCTGCAAACTGCATCAACTCAACGATGT
ACCTGCTGATGCGGAACGCCTACAACGTGTGGATAGCTGCTGTGGACTTTAAGAAGGAGTCGACTGACTT
CCCGTTCGAGCCTTACAGGTTCCACGTGGATGCCAAGCGGTCATATCACATCTCAGAGACTGCGTACGAC
ATCGAGACCAACCAGACGGTAGTGAACGTCAAGGACATCTACGGTGCGTCGTTCTCCAATGGTACGGTGG
CAATCTGCGAGAGTGACGGCAAAATCACCGAGTATGAGCCGATGGGTGACTCTTGGGATTCAACCCCAGA
CATCCGCATTAGCGGTGACATCTCTGGCAAGGATATCGTCATTGGGTTCCTGTACGACTTCCAATATGTG
TTCAGTCGGTTCCTCATCAAGCAGGAGCAGAACGATGGCACAACGTCCACAGAGGACGCCGGACGCCTAC
AACTTCGGAGAGCGTGGGTGAACTATCAGGACACTGGTGCGTTCACTGTGAGTGTCGAGAATGGCAACCG
TGAGTTCAACTATCTGGTCAACGCCAGAGTAGGCTCCACGGGTCTACGTCTTGGCCAGAAGGCAACGACC
ACTGGTCAGTATCGCTTCCCGGTGACAGGTAACGCCTTGTACCAGAAGGTGTCCCTGAGTTCCTTCAACG
CTTCCCCGGTGTCAATCATTGGGTGCGGCTGGGAAGGTAACTACAGCAGACGAGCCAACGGCATTTAACT
GAAGGAATCCTTATGGTGTGCTCAATTAGGGCACACTATAGGGAGACCACACTAAGAGGGGACTTAAAGC
ATGTACATAAGACAATCCACTAAAACTGACCTATTTGTGTTCAAGCCGTCCCGTGACGATAGACTTGAGG
CAGCAGCCTTGGGTATAGCTCCGGGATTCCCACCGCATACCGAATGTGTCTCACTGGTTACCGATGGTGA
CATAGAGGGCACATACAACCTTCTGGCTATTGGAGGCAACGTGGGTGACCAAGTGTGGTTCGTAACGGAC
CAGAAGGTATCACGCTTGACCAGAGAGGAGCGTTTAGAGTTTCGTAAGAACATTATCGAATACCGCGACA
GGTTACACGAGAAGTACCCAATCCTCTGGAACTACGTGTGGGTAGGTAACAAGTCGCACATTCGGTTCCT
GAAGACAATTGGTGCTGTATTCGAGAATGATTTTACACTCAACGGCACCTTCCAACTGTTCACCATAACG
AGGAGGTAACTATGTGCTGGATGGCAGCTATTCCTATCGCAATGACGGCGGTGCAGGCCATCGGCCAGTC
ACGCAATGAAGCCAAGATGATTGGCCTTCAGAATGACCAGATGCGCCGACAGTCTGCCCAGATGATTAAA
GAGTCAAACATTCAGAACGCTAACGCCAGCCTTGAGCAGAAGCAGAAGCTGGAAGAAGCCAGTTCGGACC
TGACCGCTAAGAATCTCGATAAGGTTCAGGCCATGGGTACAATCCGTGCAGCAATCGGAGAGGGAAACCT
TGAGGGTGCCAGCATGGACCGTATCAGTCGAATCGAGGAGGGCAAGTTCATTCGGGAGGCCAACGCGGTC
ACCGATAACTACCGTCGAGACTATGCGTCACTGTTCGCTCAGCAGCTGGGCAACTCAGAGTCAACTATTG
ACCAAGTTAAGTCCATGCAGAAGGCTGAGGGCAAAGGTAAGTCTAAGCTGGAACAGGTGCTGGACCCGCT
GGCATTGATGACCTCACAAGGCGCATCCGCATATTCGTCAGGTGCGTTCGACAGTAAGGGAACCAAGGCA
CCAATTAGTCAGGCCCAAGGTACTAAGGTAGGAGGTAAGTAATGGCCAGTAAATTAGAACAAGCATTAAG
CCAACTGCCGCAGGCCGGGTCTACCCGCATCCGTGGTGGCTCAGCGTCCATGCAGTATCGCCCAGTAACC
ATCCAACAGGAAGGGTTCCGTCAGTCCAACCTCGTGCAGTCCTTGGCGAAGTTTGGTACTGCGGTGGGTG
AGGCAGCGGATGCCTACGACAAGCGCCAACGGGACAAGGCCGATGAGCGGTCCGACGAGATTATCCGCAA
GTTGACCCCAGAGCAGCGCCGGGAGGCAATCAAGAACGGGACCCTGCTGTATCAGGATGACCCGTACGCT
ATGGAGGCCCTACGGTTCAAGACTGGTCGTAACGCAGCGTTCCTCATTGACGACGAAGTGGCACAGCGTG
TTCAGAACGGTGAGTTCCGTACCCGTGCTGAGATGGAAGAGTACCGCCACAAACGGTTGACCGAAGGTGC
CAACGAGTTCGCTGAACAGTTCATGATTAACCCTGAGGACTCTGAGTTCCAGAGAGGGTTCAACGCGAAC
ATCACTGAGCGCAACATCTCGCTGTACGGTAAGCACGATACGTTCCTGAGCGAGCAGGCCCAGAAGGGTG
CCATACTGGCCTCGAAGGTGGAGCTGTCAGGTGTGCTCAAAGACCCTGCCGTTCTGGCCCGTCCAGAGTC
CGGTGAGTTCTTCCAGCGCTACATCGACAACGCACTTAAGACTGGGAGTATCCCTAGCGACGCTCAGGCA

Figure 3 contd.

CAGCAGGTCATCATCGGGTCCCTTAACGACGTCATTCAGCGTCCGGGTGCTACCAACTTCCTCCAGAGCC
TTGAGGGCCGCCCAGTCACCCTTAATGGGAAGACCACGACCTATAAGGAGCTGATGGGAGAGGAGCAATG
GAACGCCCTGATGGTCAAGGCCCAGTCAACTCAGTTCGACAATGACGCTAAGTTGTCTGAAGGTTTCCGC
CTTGGGATTACCAGCGCGTTGAACCAAGACGATACCAGCAAGGGCTGGGAGATGCTTCAGGGTGCCAAAG
CGGAACTTGACCGCCTGCAACCCGGTGAGCAGATGACCCCAGAGCGTGAGCGCTTGATTCAGGCTGAGGA
GCAGATGCAGGCCCGTTTCCGTCAGGAGGCCCAAGCCGCAGCCAAGGAGATGGACAAGCGTCAGAAGACC
ATCAACAAGAATCAGGTCATCGACCAGCAGTTCACCAAGCGTATCAACGGTCAGTACGTGTCCACCAGCT
ACAAGGACATGCCGACCAATGAGAACACCGGAGAGTTCACGCACAGTGACATGGTGAACTACGCTAACGG
TAAGCTGGCCGAGATTGACCAGATGCAGCTCACGGAGCAACAGAAGGACCGCATGAAGCTGAGCTACCTC
CGGGCAGACTCAGAGGGTGGAGCCTTCCGTACCGTTGTGGGCCAGTTGGTAACCGACGCCGGGTCTGAAT
GGTCTGCCGCTGTGATTAACGGTAAGTTACCGGAGGACACCACAGCGTTGAACAAACTGCGCACCATGCG
TAACACCGACCCGGACCTCTTCGCTGCACTGTACCCGGACAAGGCTGACTTGTTCCTGACGATGGACATG
ATGGATAAGCAGGGCATTGACCCGCAGATTCTCATCGACGCTGACCGTTCTCGCCGCAGTCTCACCAAGG
AGATGCAGTACGAGGACGATAAGGCGTGGGCGTCCCTGAAGAACAACTCAGAGTCCCCAGAGCTGTCCCG
CATTCCGGCTAGTCTGGATGGTATGGCCCGTAAGATTTACGACAGCGTCAAGTACCGTACAGGCAACAGC
GACATGGCGATGCAGCAGACCGACAAGTTCCTCAAGGAATCCACTGTGACCTTCAAGGGTGATGACGTGG
ATGGCGATACCATTGGTATTATCCCGAAGAACATCCTACAGGTCAGTGATGACCCTAAGAGCTGGGAGCA
GGGCCGAGACATCCTCGAAGAAGCCCGTAAGGGAATCATTGCGGCTAACCCTTGGGTGACCAACAAGCAG
CTGACGATGTACCAGCAGGGTGACTCTATCTACATGATGGACACCACTGGCACTGTGCGAATCCGCTACG
ACAAGGAGCTACTGACTCGCACTTATCAGGAACAGCAGCAGCGTCTGGCCAAGGAAGCCGAAGAGAAGGC
ACTGAAGGAAGCAACCAAGCGTGCACCTATCGCCGCAGCCACTCAGGCCCGTAAGGCCGCTGGTGAGCGT
GTCCGTGCGAAACGTAAAGCCACTCCGAAGTTCATCTATGGAGGTGGTGACCAATAATCATTAAGGAGAC
AACATGAGCTACGATAAGTCCAAACCTAGCGATTACGATGGCATCTTTCAGAAGGCAGCAGACTCTCATG
GGGTCTCCTATGACCTCCTGCGTAAGTTATCGTTTAACGAATCATCCTTCAACCCTAAGGCCGTCTCTAA
GACTGGCCCTAAGGGAATCATGCAGTTCACCCGCAACACGGCCCGAGCGATGGGCCTTAACGTGACAGAT
GGTGACGACGATGGGCGCTACAACCCTGAGTTAGCCATTGACGCTGGCGCTAAGCTGCTTGCGAGCCTCG
TTAAGAAGTACAATGGGGATGAGCTTAAAGCGGCCCTAGCGTACAACCAAGGGGAAGGCCCAGCAGGTGC
CCCTCAGCTCCAAGCGTACGACAAGGGAGACTTCGGGTCTATCTCGGAGGAAGGGCGTAACTACATGCGC
AAGCTGCTGGATGTGGCCAAGAGTCCGAACTCAGGCGCACTGGAGGCGTTCGGTGGCATCACCCCAAAGG
GTAAAGGGATTCCCGCAGAGGATGCCTTCAAGGGCATCGCTAAGGCTGGAAAGGTTGGTACCGAACTGCC
GGAGTCCCATGGGTTCGACATTGAGGGTGTAGCGCAGGAAGCACCAAACACTCCATACGCTAAGGACTTC
TGGGAGAAGACCGGGACTACTCTCGATGAGTATAACTCTCGGTCAACCTTCTTCGGGTTCGGGGACGCTG
CTGAGGCTCAGATTCAGAACTCCACATTAGGTGTGGCCTTCCGTGCTGCGCGGGCTGACGATGGGTACGA
TGTGTTCAAGGACACGATGACCCCGACTCGCTGGAACTCTTATGTTCCCTCCAAGGAAGACCTACAGAAG
CTGCGCGACTCTGGGCTACCTCCGAGCTACTACGGTGTGGTGACTGGTGGTGACGGTGAGAACTGGGATG
CACTCATCAAGCTGGCCAAGGATAACTTCGAGGCTGACCAACGGGCCGCTGAGGCTGGTACTGGGGCGAA
ACTCGCTGCTGGTATCGTTGGGGCTGGTGTAGACCCACTCAGCTATGTACCTCTGGTCGGTGTGGCCGGG
AAGGGACTCAAGGTTGTCAATAAGGCCCTGCTAGTAGGCGCACAGGCTGGGGCACTCAGTGTTGCCTCTG
AGGGAATCCGTACGTCAGTGGCTGGTGGCGAGGCTCACTACGCTGATGCGGCACTCGGTGGGTTACTGTT
TGGCGCTGGTATGTCGGCTCTCAGTGATGCTGTGGCGGCTGGTATCCGTAAGGCCCGTGGAGTCGATTCT
GTGAATGAGTTCGCTGGACCAGCACTCCGTATGGAAGCGCGAGAGACTGCCATCAACACTGGTGGTCATG
ACACCTCGACGCTACCTCCAGAGAACTTCTCGTTCGAGCAGGACCACAGAGGCGTTCCGTTTGCTGACCA
CCCGACCGAAGAGGGAGCAGTGGTTCTGGCCAATGGTTCCATCCTGAGCGATACCAACCCGCTTAACCCA
AGGACTCAACGTGACTTCGCAGAGATTGACCCAGAGCGTGCAGCTCCCGGTATCAAACTCGGTGGGTTCA
CTGAGATTGGCCTGAAGACCTTAGGGTCCAAGGATGCTGGTGTACGTGCAATCGCTCAGGACCTCGTGCG
CTCTCCCACAGGGATGCAATCAGGGTCTAGTGGTAAGTTCGGTGCGACCGCTTCGGACATCCACGAGCGA
CTCCATGCGACTGACCAACGGATGTATAACCAACTGTATGACGCTGTTGACCGTGCCATGAAGGACCCAG
AGTTCTCCGTGGGCGAGCAGAAGATGTCACGCAGAGCCATCCGTCAGGAAGTCTACAAGCGTGCCTCATT
GGCGATTGAGCGCCCAGAGTTACAGGCTGATTTGACCAAAGGTGAACGTGAGGTGATGGACCTGCTGAAA
GAGCACTTCGACACCAAGCGTGAACTGATGGAACAGCCGGGTATCTTCGGTAACGCTAACGCCGTGAGCA
TCTTCCCCGGTAGTCGACACAAGGGTACTTACGTGCCTAACGTGTACGACAGGGGTGCCAAGGAACTGAT
GATGCAGAAGCTGGGCGGACCTGAAGGACTCCAACAGGCAATCGCTCAGAGCTGGCTTACCAGTTACCGA
GTGCGACCTGAGGTCAAGGCGCGTGTTGACGAGTACCTGATGGAACTCAACGGCTACAAGTCGGTAGACC
AAGTGACACCTGAGGTGGTCCAGAAGCACGCTATGGATAAGGCGTACGGTATCAGCCACACTGAGGACTT
CACAGCGTCCAGTGTCATTGACGACAACATCACAGGTCTGGTCGGTATCGAGAACAACTCGTTCCTTGAG
GCCCGTAACATGTTCGACAGCGACCTCCCGGTTACCTTACCGGATGGGTCAACCTTCAGCGTCAACGACC
TGAGGGACTTCGACATGGCACGGATTATCCCAGCGTACGACCGTCGAGTTAACGGTGATATCTCCATCAT

Figure 3 contd.

GGGCGGTAGCGGTAAGACCACGCAGCAGCTCAAGGACGAAATCATGGCGTTAGACAAGCGGGCTGAGCGT
AAGGGACAGCTGAAGGGCGAAGTGGAAGCACTGAAGGACACCGTTAAGATTCTCACGGGGCGTGCTCGTC
GTAACAACGATACAGCCTTTGAGACCGCCATGCGTACCCTGAACGACCTAGCGTTCTTCGCTAAGAACTT
CTACATGGGTCCGCAGAACCTCACAGAGATTGCTGGGATGTTGGCTAAGGGTAACGTTAAGGCGATGCTC
CACGGTATCCCGACGTTGCGTGACCTAGCCACCAGAACCTCTCCGGTGTCCGGTAGTGAACTCCGCGAAC
TCCATGGGGCGCTGTTCGGTAAGGAACTCGACCAGTTAATCCGTCCGGGGCGTGAGGATATCGTACAGCG
AATCCGCGAGGCTTCCGATACCAGTGGGGCCATGGCGTCAGTCATTGGCACCATCAAGTTCGGTACTCAG
GAGCTGTCGGCTCGTTCTCCTTGGACCAAGATGCTGAACGGTACGGCTAACTACATTCTGGACACTGCCC
GTCAGGGTGTGCTCGGTGATGTGGCTGGTGCGGCCCTAGGCGGTAAGGGTTCCAAGTTTGGCAAAGAGAA
CTTCCTCAAAGCTGCCTCTATCAGTCCTGAGCAGTGGAAGGGAATCAAGCAACTCTTTGTCGACCACGCA
ACTCGTGACGCTAACGGCCAGTTCACCATCAAGGACAAGAAGGCTTTCAGTCAGGACCCGAGAGCGATGG
ACCTGTGGCGTCTTGCCGATAAGGTTGCCGACGAGACCATGCTGCGCCCTCACAAGGTATCCCAGCAGGA
TTCCAAGGCGTACGGTGCTGGTGTCAAGATGGCTATGCAGTTCAAGAACTTCACCATCAAGTCACTCAAT
GCCAAGTTCATTCGGTCCTTCTACGAGGGCTACAAGAACAACCGCGCTATCGACATGGCGTTGACACACG
TGTTGTCTCTGGGTATCGCCGGGACTTACTTTGCGATGCAGGCCCACGTGAAGGCTTACGGCCTCCAAGA
GTCCCAACGTAAGGACTACCTGAAGAAAGCCCTGAACCCGACCATGCTGGGCTACGCAGCGTTGACTCGA
AGTTCCCACACTGGTGCCCCGCTGTCCATCGTTTCGATGATGGCTGGTGCCGCTGGGTTCCAAGACGCCA
ACATGCTGCGCTCCACCATCTTACCTAAGGAGGAACAATTCCAGAAGAAAGATGGAGCGTCCAAAGGTCG
AGCCGAGTCGAGCAACCTTGCGGGTAACTTAGGGTCTCAGGTCCCAGCTCTGGGTTACGTAGGGAACGTC
ATTGCTACCGCCAAGAACGCCTACGGTGTTGCTACAGCACCCAACAAGCCGACTGAGCGTGACTACATGA
CTGGCCTGATGAACTCCACCAAGGAGCTTGTTCCGAACGACCCACTGACCCAGCAGCTCATCATGAAAAT
CTATGAAGCCAACGGGGTCACCATCAAGCAGCAGCCGAAGCCTAACTAATTAGGACACACTATAGGGAGA
CCGATTGGTTTCCCCCCTTCTCATTCAACTAAAGGAGGTCACAATGGACCAAGACATTAAAACAGTCATT
CAGTACCCAGTAGGGGCCACTGAGTTCGACATCCCGTTCGACTACCTGTCCCGTAAGTTTGTCCGTGTGT
CGCTGGCAGCTGACGACAACCGCAGACTGCTGAGTAACATCACTGAGTACCGCTACGTGTCTAAGACCAG
AGTGAAGCTCCTTGTGGAAACTACCGGGTTCGACCGTGTGGAAATCCGCAGATTCACCTCAGCGTCTGAG
CGTATTGTTGACTTCAGCGACGGCTCCGTACTGCGGGCAACAGACCTTAACGTTTCTCAGATTCAGTCTG
CCCATATCGCAGAGGAAGCACGTGATTCAGCACTGTTGGCTATGCCGCAGGATGATGCTGGCAACCTTGA
TGCCCGTAACCGCAGAATCGTTCGGCTGGCTCCGGGTGTCGAAGGTACGGATGCAATCAACAAGAACCAG
CTGGACACCACCTTAGGTGAAGCTGGTGGCATCCTGTCGGAAATCAAACAGACCGAGAAGGACATTCAGG
ATTACATCGAGAACTTTGCAGATGACACCACGTCTCTCAAGGGAATCAACTGGGTGTATAACAATGGGTC
GGCCAATGGTGGCGAGACCTCCATCCTGATTACCCGCGAGGGGCCAGTGTTCGCTGTGCCTACCATTTAC
ATCAATGGGGACAGACAGTCTGTTGGTTACCACTACTCTTACGACTCCGGTGATAAGACCATTCACCTAG
TTAAGCCGCTAACTGCTGGAGACTTTGTGGAATGTGTTACCTCTGAGGGCGTACTGCCGCTGTCTAATCT
TCTGTCGACACCAGACGGGGCCAGTCAGATTGGCACTAAAAGCGGCCTGACTGTGCAAGACTACCTTAAC
GGCGTGAAGTCCGCTACCATCCTGCGCAACATTGAGCCAGTCATTGATGGACAGCGCATCGTCCTCTCTG
AGATTAGCCCTACTTTGGGGCCTAAGTCTGGAGGTACCTTGGTGTACGACCAGTCTGATACATCCTCTGT
GGACGACGGGTACACTGTTTTCGTGACAGCTGGCGGTAAACGGTGGAAGCGAGAAGAGTCCTACATTGAC
GTAGCGTGGTTCGGTCCTAACTTTGGCCTTGCCTTACAGACCGCTGTTAACCTCGTTGACAACTACGTGA
GAACTGTCGGTTTCTACAGTCGCAAGACCATCTACATTGCAGCTGGTACCTATACGACAGACCGTCAGGT
GGACATTCCATCTTATGTCTCTGTGGTGGCCATAGGTAACGTTAGCATCAATGGTTCTGGGCTTCCAGTA
AACTCCTACGTACTCCGCATAACGAACAAGGTTGGTGGCATTGTCACAACCCACCACTCAGGGTGGAACC
TCGGGGCCGTAGGTGGGACCCTTCGTCTTGTAGGAAACGGCAACACCGGGCAAGTGGATGGGCTTTATGT
GGGCGGTGCGACTTCTATGAGCGACGTACGGAACGTTAGCCTTTACGCTGTGTCAACTTCGGGTGTTCGC
TATGGGCTAACATTTGGTAGCACCAACACTTACCTCTTCACGGCAACCAAATGCCACTTTGAGACGTCTC
TTGTAAACCTGTACATTCCGGGCACCACAAGCTCTAACTCAGGGGAGAAGATGGTATTCAATGATACTGT
GTTCGGTGGCTCATCTAGGAACCATGTAGAGGTAAGCACCCCAGGCATGGACCTCACGTTCAATAACTGC
TCTTTCGACTTCACAAGCGGTAGCGTCCTGTACGGGACAGAGACTTGGGGCTATGCGAAAGTAGGCATGA
ATAATTGCCACTTCGAGGGGTTCAATAGTTTGTGGATAAAGGTGGATGCCCCGCAAGGTGGATTCATTGG
GTCAAACCGAGCGATAACCGTATCAAACGCCACAGTCCTTCCTAGGCTTCGCTCCAACACTGCTGGAACA
AACTCGGCGAGCCGTATGCACATTGATGCCAAGTCTACCCCGGTGTATATCAGTGGGCTGGACCTACGGC
ACGAGGTCGTACCATACACCGAGGAAATCTTCATGGCTTCAGCTGAAACTACCCTGTCTCTGCAAGGATA
TCTTAAGGACCCGCATTTCCAGATTCCAAGTGCTGCGCACATTCAGAACCGTGGGTGGAACATCGCTGAC
GAAACAACTGGAACTGTTGTGAACAGCCCCGCAACCTTGGATTCCCTTACGCGATTTACATGCACCGAGA
GGAACGCGATGTCTGCGGCTGTGGTCGATGGTGGAACTTCTGGTAAGCTCTTAGCAATGACTGGAGCGGG
TGGGTATTTCACTCTGGTCACTAAAGGATTCATTCCGGTGAGTACGTTCCAACGGATTGGCGGAGCAATG
TCGATTCAGGCAGCAGCAAGTACCGGAAACATCCAGTGCACGCTTGGTGTCCAGTGGTTCGACTACGATG

Figure 3 contd.

GTAACCTAATCGGGACAGACCAAGCCTTTGCGATTAACATGCGTGAGGTGTTCAACAACTCTTCTCTACC
TAACTTCGCCGAAGGCAACAACCGCTTCATCTCTACATCTGCGAGAACATTCCGTGCGCCAGCGGGGGCC
GCTAAGTGTAAACCATTGTGGCGAATCTCTGGTCATACTGGCGTTGTGAACATCTCGAGATTAGCATCAT
TTGTTTTATAAGGAGACAACATGCTGAATGATTTAAACCAACCACGAGGCTCGACGCTGGGCCTCTTTAC
TCCAAACCTTCCGTTGAAGAAGCGGTTGGACACCTTACCAAACATTTTAGATTTTGATTCAGACAGCCTT
AACGATGATAGCACTCGGTTTCAAAAGGCTATTACAGCTGGTGTGAAATCTTTATACGTCCCAGAACCTC
AGTTCTTTGGCAACAATAAGCCTCTTAAAATTGCTAACGTTGACATTGTGACCAATATGCACATCTACGG
GAACGGCTCAGCGGGATACCGTCAGGTTGGCGGGGCCATCACCATCCTAGATGGAGCGGACTATGGGTTT
AAACTGGCTGGTGTCGACTCTCAGACGCGAAACATTGGAGGCCGCATTGACGGTCTCTCGTTCCAAGGTG
AGTTCCCAACGACCGTGGCCGACGCCATCCGGTGCCAATCTGCCAGTAGCTTCGCGCTGGTCAACCTCTC
GTTCAGGAACCTCTCCGGGTCCGCTCTGGACCTGCGTGACTTCATGGAGAGCCACATTGAGCACTGCTAC
TTTAACTCAGTAGGTTCCGACACAAAGAACCCAATCAACATCGGGGACTTCGTCGGGTCGGCTCCTTGGA
ACGTCAACAACCTGCACATTGAGAACAACACCTTCGGGTCATGTAGTGGGAACATTATCAACATTAGTGA
CTCAGCTAACGCCGACCTCATTTGGATTCTCAACAATAAATTCGAATGGGACTCGACCCCAGTAAGCCCC
AACGTTTCCAACAAGGCGGTGGCATACATCGGGCGAGCCGAGCGTGTAAATGTGTCCGGTAACGGCTTCG
TGTACTACTACCCGGCCCACAACAAGTACGATGCCCTTATCCGAGTTTCCGATAAGTCGGCCTATGGTAA
CTTGTTCTCTGATAATACCGCTTGGGGCTGTACGCCTCCTTCAGGTAGTGACCTCACTCCAGCGTTCTAT
TGGGACATTGCTGGTGGGTCGTCTGCGGGGTCTAACAACAAGGCTAACACAAACCTCCCTACGCGCTGCA
CCAGTATCCACTCTCAGGATATCGACGAGCCGCTGGTAAGGACTACTCCGGGTAACCGACCAAACCTCCA
GAGCATCGGGGCAATGTCTCCCGGATATCTCTCTGCGCACTCCTTAGGTGGGGCTAACGCTTCCAACTTC
TTTGTGCCAGACACTGGTGCTACCAAGTACGGTACGGTGCTAGAGGCTCAAACTGGTGGTGAGGTTCGCC
GCTTGTTCATTCCTAAGGACATTGTTAGCCAGCGTGCTTGCGTTCGAGTTCAGGCCAGAGTGATGCCGTC
GCCGACAGCTGATGCCCTTGTGGGGCTGACCTGTGACGGCTCCATTGTTTCCACCACAATCCAAGGCGCA
ACCCAAGACTACCATACGGTGGCGGCTGGTGGCGGCTGGCAGATTGTCGAGTGGCTCATCCCGGCGTCTA
GTTACACTGCGGGCCAGTTAATCTTCACGAACCGTAGTGACACCGTCAAGTTCAAACTTGATGGCGTCCG
TGTGTCACGTGCAGACTTCGTAGATGTGACGATTGCATGGAGTCCGACCCCAATCTCCGCAGGGTCTGTG
GTAAACACCACTGCATCAATCACTCGCGTAAGTTCCCACGTGGTCGGCACTAGTGGTCTGAAGACAGACG
GTACGTTAGGTGGCGCTGTTAGTAGCTCTTATTTCAACCGTGGGGCCAATACCTTAGTGGTACAGCTGGC
AGCACTCACAGCAGCCACTCCGTCAATCACTCAGGTTACGGTTAGGCTGTTCCTTAACTAAGGAGGTAAC
ATGTTGTCCCTAGACTTCAACAACGAAGTTATCAAGGCGGCTCCCATTGCGGGGGTCGCTGGGGCCGATG
GTGTAGCGAGGCTCTTCTGGGGCCTCTCACTCAACGAGTGGTTCTACGTCGCGGCAATCGCCTACACAGT
GGTTCAGATTGGTGCCAAGGTAGTCGACAAAATCATTGACTGGAAGAAAGCAAATAAGGAGTAACATATG
GACCTGATTAAGTTCCTCGAAATGTTAGATACTGAGATGGCTCAGCAGATGCTCATGGACCTGAAGAATC
CCGAGAAGCGAACCCCTCAGCTGTACAACGCCATTGGTAAACTACTGGAGCGCCATAAGTTCCAAATCTC
TAAGCTGACCCCTGACGTTAACATCTTGGGCGGACTGGCTGAGGGTCTGGAGGCTTACAACTCCAAGGTG
GGCGCCGATGGTCTGACAGACGACGATACGTTCACCCTACAGTGATATACTCAAGGTACTACTATATGTA
GTGCCTTTATGGATGTCATTGCACTACGCTAGGCGTTCCTACGTGAAATCTGAGAAACAACGGGAGGCAT
TATGCTGGAGTTCACAAAGAGAATCGTCCCGTATCTTGTGGCTATCATGGTGTTTGCCTTCGGGTGGCAC
TTGGGGTCTCAATCTACGGACGCTAAATGGAAGGAGGTAGTACAGCATGAATACGTTAAGAAGCAAACGG
CTAGAGCTGAAACTCAGAAAGCGATTGACGCAATATCGGCTAAGTACCAAGCAGACCTTGAGGGGCTGGA
GGGCAGCACTGATAGGGTTATTGCTGATTTGCGTAGCGACAATAAGCGGCTGCGCGTCAGAGTCAAACCT
ACCAGTGTCGCCGCAGGACCAGACGGTCGATGCCTCGTTGATGGTTCCGTCGAACTACACGAAGCAACTG
CTCGAAGTCTTATCGCAATAACCCAGAAGGCCGACCTCAAAGAGAAGGCCCTACAGGACACTATTCGCAA
GCTACAGCGGAAAGGAGGTGAACATTGAGTAACTCTCAGCAAGCCAAGAACGCCTTAATCATTGCGCAAC
TGAAGGGTGACTTTGTCGCCTTTCTCTTCGTGCTCTGGAAGGCCCTGAACCTGCCGGAACCAACCAAGTG
TCAAATCGACATGGCCAAGTGTCTGGCGAACCCAAAGAACAAGAAGTTTATCCTTCAGGCTTTCCGTGGT
ATCGGGAAGTCATTCATCACGTGTGCGTTCGTAGTGTGGACCCTGTGGCGTGACCCTCAGTTAAAGATAC
TGATTGTCTCGGCCTCAAAGGAACGTGCGGACGCTAACTCCATCTTCATCAAGAACATCATCGACTTGTT
GCCTTTCCTGAGTGAGCTTAAGCCCCGCCCCGGTCAGCGTGACTCCGTGATTAGCTTTGATGTAGGCCCT
GCCAAGCCAGACCACAGCCCGTCAGTTAAGTCTGTGGGTATTACTGGTCAGCTTACTGGTAGCCGTGCTG
ATATCATCATTGCGGATGACGTGGAGATTCCCGGTAACTCTGCAACCCAAGGCGCTCGTGAGAAACTCTG
GACGCTGGTTCAGGAGTTCGCCGCACTGTTGAAACCTCTGCCGACTAGCCGTGTTATCTATCTGGGTACA
CCTCAGACCGAGATGACGCTCTACAAGGAACTTGAGGACAACCGTGGGTACTCCACCATTATCTGGCCTG
CACAGTATCCTCGCTCCAAAGAGGAGGACCTGTACTATGGCGACCGACTGGCCCCGATGCTCCGTAGTGA
GTACGATGAGGACAAAGAGGGCCTCAGCAGTCAGCCTACTGACCCGGTTCGATTCGACTCCATGGACCTT
CAGGAACGTGAGGTGGAATACGGCAAGGCTGGCTATACGCTTCAGTTCATGCTCAACCCGAACCTCAGTG
ACGCCGAGAAGTACCCGCTACGCCTCCGTGACGCTATCGTGTGCGGTCTACAGATGGACAAGGCCCCAAT

Figure 3 contd.

GCATTACCAGTGGTTGCCGAACCGTCAGAACCGCAATGAGGAGCTTCCTAACGTGGGCATGAAGGGTGAC
GAGATTTACTCCTTCCATACAGCCTCAAGTAACACTGGCGCGTATCAGGGTAAGATTCTGGTCATTGACC
CCAGCGGTCGCGGTAAGGATGAGACTGGCTGGTGCGTACTGTACACCCTCAACGGTTACATCTACTTGAT
GGACGCTGGCGGTACTCGTGGGTACGAAGAGAAGTCCCTTGAGTTCCTCGCTAAGAAAGCCAAACAGTGG
CAGGTTCAGACTGTGGTCTTCGAGAGCAACTTCGGTGACGGTATGTTCGGTAACGTGTTCCAGCCTGTGC
TCCTGAAGCATCACCCAGCGCAACTCGAAGAGATTCGTGCTCGTGGTATGAAAGAGGTCCGTATCTGCGA
TACCCTTGAGCCTGTACTGGCAAGTCACCGCTTGGTCATCCGTGATGAGGTTATCCGACAGGACTACCAG
ACGGCACGTGATGCAGACGGTAAGCACGCTCTGAAGTACAGTCTGTTCTACCAGATGACCCGTATGAGCC
GTGAGAAGGGCGCGGTGGCACACGATGACCGACTTGATGCGTTAGCATTGGGTGTCGAGTTCCTACGCTC
TACGATGCAGCAGGACGCTGTGAAGATAGAGGCTGAGGTACTTCAGGAGTTCTTGGAGCACCACATGGAG
AAGCCCCTGAGTAACATCTCCCAGTTCCGGGCCACCAGTAGCAACGGTGTGGACATCCGATGGGAAGACG
ATGGGGATGACACTATGTTCATCGCATGGTGATTATGCAGGGATTGTGCATAAGGATTCATTAGGCCACG
GAAGGCCACTTTGAGGAAACTCCATGTATAACAGACACTTGGAATTAGGACCCACTATAGGGAGAGACCC
TTGAAGACTTACTATAAGACAACTTAAAGATTCATTCATATAGTTATTCACTTTAAGTCTCCTTAAAGGC
AGAGGGTAGTGATGATAATATCACCCTCTCACTATAAGACACTAAGAGCCAACATAAGGAGGACCTATGC
GCTTATTGTTAACCTTACTGCGCCATAGGGCTACTTGGCGATTTCTGCTGGTACTTGCTGGTGCCCTTGG
GGCTTCACTGGTTACTCAGCAGCAACTCAGTGGACTGGAGACTCTCGTGTGCTCTCTACTCACTTGTAGC
GATTAGGGTCTTCCTGACGCGCTAGGGATTCCGTAGTGATGCTTATCAGCATACACCACTCCATCCCTCT
ACAGTCAATACTTAAAGTTAACCTTAGGTGATTCACTGGGTCTACCTACGGGTCTATGCAATGACCTGAG
GAGTACCTGAGGTTACCTTTAAGAATTTTACATAAAGTTCTGAGTGTACATCTCACAGTTTACACTTTTG
GTTATCCCCCCGGTACCCTCCAGTTCACCCAAAGTAACCTAGGGTACCCCTCTTTACCTTTGGTTTAACC
TTGGGTGGTACCTTGGGAATCCCTTAGGTGATACCATATGTTGGGGTAATGGTGACCTGAGGACACTATA
TGTTGATGTCTCTGTGTCCCT

Figure 3 contd.

Figure 4

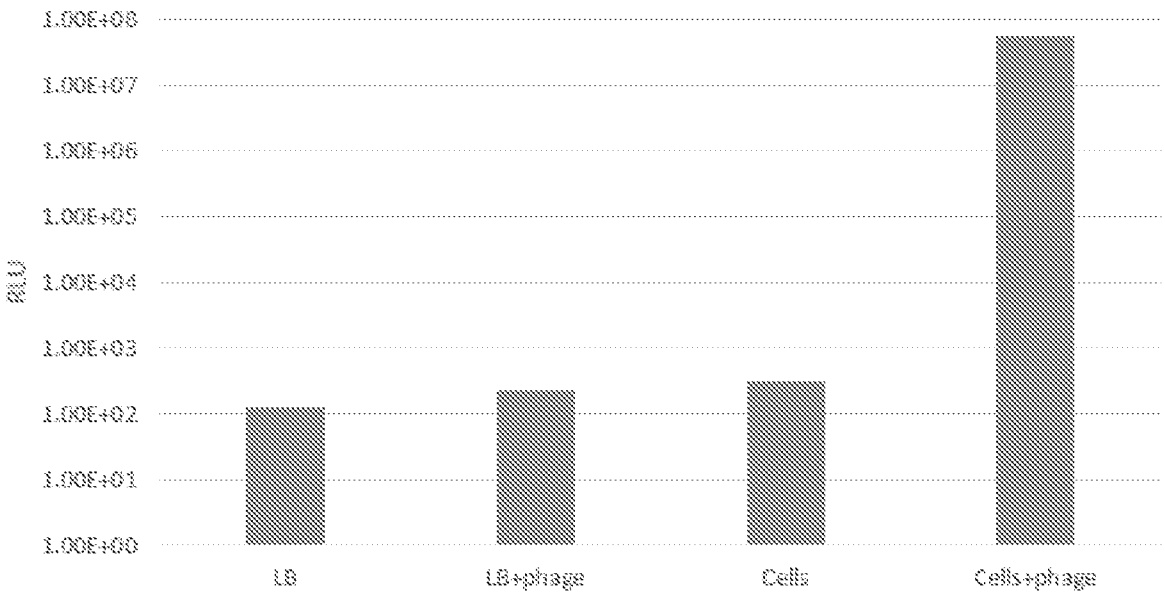

1.00E+08
1.00E+07
1.00E+06
1.00E+05
1.00E+04
1.00E+03
1.00E+02
1.00E+01
1.00E+00
RLU
LB
LB+phage
Cells
Cells+phage

Figure 5



Figure 6

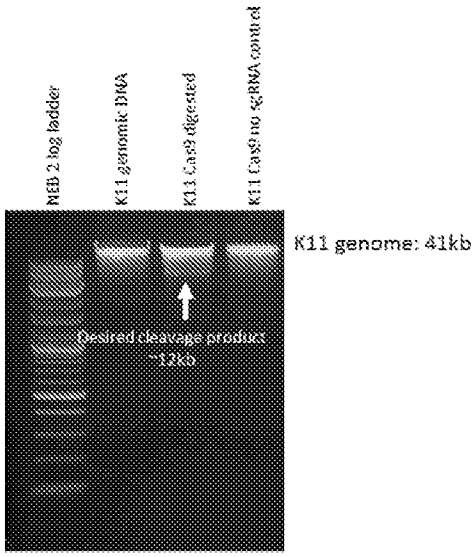
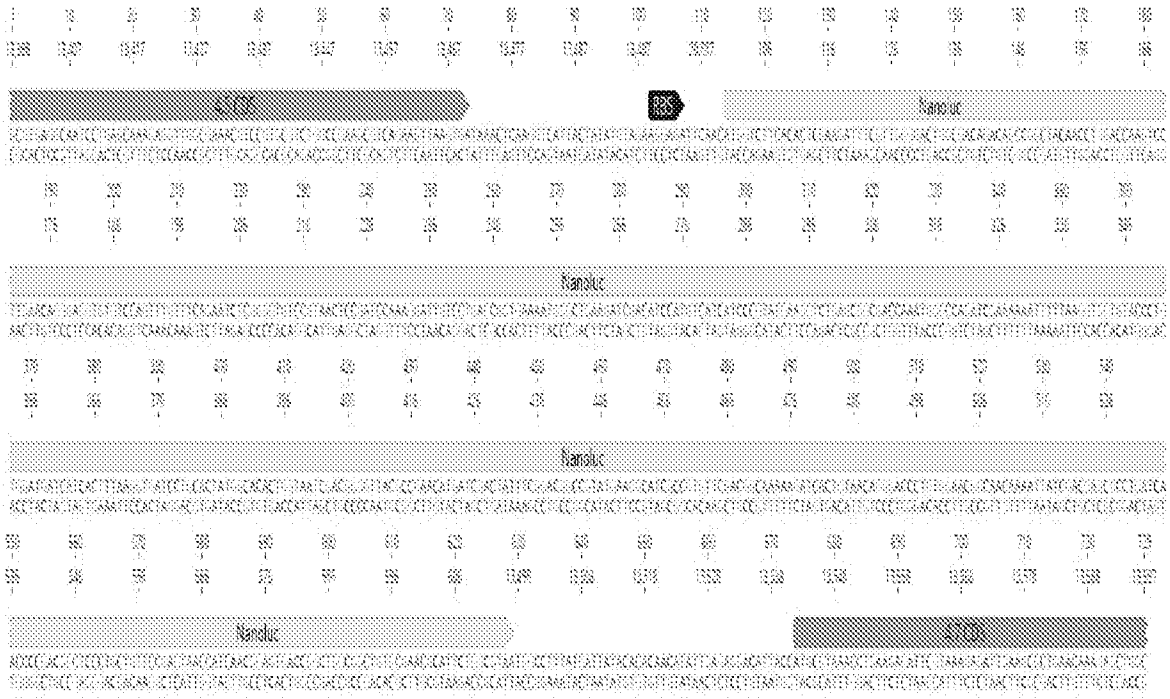

MEB 2 log ladder
K11 genomic DNA
K11 Cas9 digested
K11 Cas9 no gRNA control
K11 genome: 41kb
Desired cleavage product:
~12kb

Figure 7



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2018/035420

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. ☐ forming part of the international application as filed:

      ☐ in the form of an Annex C/ST.25 text file.

      ☐ on paper or in the form of an image file.

   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. ☒ furnished subsequent to the international filing date for the purposes of international search only:

      ☒ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2018/035420

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/86 C12N7/00 C12N15/90
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, BIOSIS, Sequence Search, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2016/100389 A1 (SYNTHETIC GENOMICS INC [US]) 23 June 2016 (2016-06-23) the whole document	1-54
A	JIA-WANG WANG ET AL: "CRISPR/Cas9 nuclease cleavage combined with Gibson assembly for seamless cloning", BIOTECHNIQUES RAPID DISPATCHES, vol. 58, no. 4, 1 April 2015 (2015-04-01), XP055385711, US ISSN: 0736-6205, DOI: 10.2144/000114261 the whole document	1-54

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 August 2018	11/09/2018

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Brenz Verca, Stefano

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2018/035420

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MARIE-LAURENCE LEMAY ET AL: "Genome Engineering of Virulent Lactococcal Phages Using CRISPR-Cas9", ACS SYNTHETIC BIOLOGY, vol. 6, no. 7, 30 March 2017 (2017-03-30), pages 1351-1358, XP055503259, Washington, DC,USA ISSN: 2161-5063, DOI: 10.1021/acssynbio.6b00388 the whole document -----	1-54
A,P	WO 2017/172645 A2 (CHARLES STARK DRAPER LABORATORY INC [US]) 5 October 2017 (2017-10-05) the whole document -----	1-54

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2018/035420

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2016100389 A1	23-06-2016	AU 2015362618 A1	13-07-2017
		CA 2971205 A1	23-06-2016
		CN 107278227 A	20-10-2017
		EA 201791343 A1	30-03-2018
		EP 3234136 A1	25-10-2017
		JP 2017537643 A	21-12-2017
		US 2016186147 A1	30-06-2016
		WO 2016100389 A1	23-06-2016
WO 2017172645 A2	05-10-2017	US 2017283779 A1	05-10-2017
		US 2017298456 A1	19-10-2017
		WO 2017172644 A2	05-10-2017
		WO 2017172645 A2	05-10-2017