

(57) Abstract: Described in certain example embodiments herein are engineered phages comprising one or more heterologous genes that encode one or more heterologous gene products and methods of using the engineered phages to deliver one or more heterologous gene products to a subject.

ENGINEERED PHAGES AND USES THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Application No. 63/472,620, filed June 13, 2023, and entitled "Engineered Phages and Uses Thereof", the entire contents of which are incorporated herein by reference.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH

[0002] This invention was made with government support under Grant No. 1R35GM147484 awarded by the National Institutes of Health. The government has certain rights in the invention.

SEQUENCE LISTING XML

[0003] The instant application contains a sequence listing, which has been submitted in XML file format by electronic submission and is hereby incorporated by reference in its entirety. The XML file, created on June 12, 2024, is named 94247-14-WOU1_Sequence Listing.xml and is 18,988 bytes in size.

TECHNICAL FIELD

[0004] The subject matter disclosed herein is generally directed to engineered phages.

BACKGROUND

[0005] Oral delivery is a primary administration route for pharmaceuticals and nutraceuticals. For small molecule therapeutics, there has been extensive development into coatings, such as enteric coatings, that provide extended release as well as responsive release (e.g., pH responsive) such that the therapeutic is protected from degradation until release in the small intestine. While this works well for non-biologic small molecule therapeutics, these approaches have not translated well to delivery of biologic molecules, such as proteins.

[0006] Oral administration of biologics is challenging due to the degradative nature of the upper gastrointestinal tract (GIT) and resident commensal bacteria that impairs the efficacy of engineered microbes for producing proteins in the lower GIT. Virulent bacteriophages (phages) have been of renewed interest because of their antibacterial activities against pathogenic bacteria, but emergent work indicates they are a major component of the gut microbiota that coexists with resident commensal bacteria. As such there remains an unmet need for improved systems and techniques for delivery of biologic agents, such as proteins.

[0007] Citation or identification of any document in this application is not an admission that such a document is available as prior art to the present disclosure.

SUMMARY

[0008] Described in certain example embodiments herein are engineered phages comprising one or more heterologous cargo genes, wherein the one or more exogenous genes encode one or more heterologous gene products (e.g., nucleic acids and/or polypeptides); and one or more regulatory

elements, wherein each of the one or more heterologous cargo genes is operatively coupled to at least one of the one or more regulatory elements.

[0009] In certain example embodiments, the phage is a lytic phage. In certain example embodiments, the phage is a T4 phage, T2 phage, MS2 phage, M13 phage, T3 phage, T5 phage, T7 phage, Lamda phage, P1 phage, PhiX174 phage, Mu phage, P22 phage, *Pseudomonas* phage phiKZ, *Bacillus* phage phi29, *Salmonella* phage P22, *Mycobacterium* phage D29, *Enterobacteria* phage T6, *Mycobacterium smegmatis* phage L5, *Escherichia* phage T5, *listeria* phage A511, *Vibrio* phage KVP40, *Staphylococcus* phage K, or *Staphylococcus* phage PhiETA.

[0010] In certain example embodiments, the phage is capable of infecting a bacterial cell or a yeast cell. In certain example embodiments, the phage is capable of infecting a cell of a microbiome. In certain example embodiments, the microbiome is a gastrointestinal microbiome, a skin microbiome, an airway microbiome, a vaginal microbiome, or any combination thereof. In certain example embodiments, the gastrointestinal microbiome is an oral cavity microbiome, a stomach microbiome, a small intestine microbiome, a large intestine microbiome, an appendix microbiome, a cecum microbiome, or any combination thereof.

[0011] In certain example embodiments, the one or more heterologous gene products are therapeutic proteins. In certain example embodiments, the one or more heterologous gene products are capable of (a) stimulating an immune response in a subject in which it is delivered; (b) treating a disease or a symptom thereof; (c) modulating a microbiome; or any combination of (a)-(c). In certain example embodiments, stimulating an immune response comprises stimulating a B-cell and/or T-cell mediated response.

[0012] In certain example embodiments, the one or more regulatory elements comprises a phage promoter, optionally wherein the phage promoter is an endogenous phage promoter. In certain example embodiments, the one or more regulatory elements comprises an early phage promoter, a late phage promoter, or both.

[0013] In certain example embodiments, the engineered phage is a T4 phage and the one or more regulatory elements comprise motB, gp18, gp22, or any combination thereof.

[0014] In certain example embodiments, the one or more heterologous gene products comprises one or more signaling domains and/or translocation domains or nucleic acid sequence encoding the same. In certain example embodiments, the one or more signaling domains and/or translocation domains are capable of translocating the one or more heterologous gene products (such as a protein), when expressed in a host cell, to a periplasm.

[0015] In certain example embodiments, the one or more heterologous gene products (e.g., proteins) comprises an N-terminal signal peptide for periplasmic secretion, optionally wherein the N-terminal signal peptide comprises (a) a signal peptide selected from Tat, appA, dsbD, degQ, cueO, pldA, ushA, pagP, fhuA, ptrA, lptA, bepA, dacA, agp, malS, spy, amiC, tolB, rseB, amtB, dsbC, amiA, torZ, FhuD, zinT, mglB, eco, nikA, mepH, fadL, ompL, ompN, thiB, MalE, PhoA, or any combination thereof; (b) a leader peptide from *spA*, *phoA*, *ribose binding protein*, *pelB*, *ompA*, *ompT*, *dsbA*, *torA*, *torT*, *tolT* or any combination thereof; (c) or any combination thereof.

[0016] Described in certain example embodiments herein are pharmaceutical formulations comprising an engineered phage described herein, and a pharmaceutically acceptable carrier.

[0017] Described in certain example embodiments herein are vaccines comprising an engineered phage described herein.

[0018] Described in certain example embodiments herein are methods of delivering a protein to a subject in need thereof and/or treating a disease or a symptom thereof in the subject in need thereof, the method comprising administering an amount of the engineered phage, pharmaceutical formulation, or vaccine as in any one of the preceding claims to the subject in need thereof, whereby after administration the phage infects a host cell of the subject in need thereof, stimulates production of the one or more heterologous gene products (e.g., nucleic acids and/or proteins) by the host cell, and lyses the host cell so as to release the one or more heterologous gene products produced by the host cell thereby delivering the one or more heterologous gene products to the subject in need thereof.

[0019] In certain example embodiments, administering comprises oral administration.

[0020] In certain example embodiments, the host cell is a bacterial cell or a yeast cell. In certain example embodiments, the host cell is part of a microbiome of the subject in need thereof. In certain example embodiments, the microbiome is a gastrointestinal microbiome, a skin microbiome, an airway microbiome, a vaginal microbiome, or any combination thereof. In certain example embodiments, the gastrointestinal microbiome is an oral cavity microbiome, a stomach microbiome, a small intestine microbiome, a large intestine microbiome, an appendix microbiome, a cecum microbiome, or any combination thereof.

[0021] These and other aspects, objects, features, and advantages of the example embodiments will become apparent to those having ordinary skill in the art upon consideration of the following detailed description of example embodiments.

BRIEF DESCRIPTION OF THE DRAWINGS

[0022] An understanding of the features and advantages of the present disclosure will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the disclosure may be utilized, and the accompanying drawings of which:

[0023] **FIG. 1A-1B** - Phage-induced lysis of bacteria releases intracellular proteins. (**FIG. 1A**) Scheme of the potential utility of this strategy in the gut, by introducing a recombinant phage to target an endogenous commensal gut bacterium. (**FIG. 1B**) Batch culture of *E. coli* and T4 phage demonstrates the presence of phage increases the extracellular quantity of sfGFP.

[0024] **FIG. 2A-2D** - Survey of T4 phage promoters for *in vitro* sfGFP production. (**FIG. 2A**) During T4 infection, promoters are expressed in a series of stages. Representative promoters were selected from early, middle, and late stages of expression. (**FIG. 2B**) To survey the protein production from these promoters, they were encoded upstream of *sfGFP* and inserted into the *ac* region of the T4 phage genome to produce a small library of recombinant T4 phages. (**FIG. 2C**) After 12 h of coculture of the T4 library with *E. coli*, the fluorescence and (**FIG. 2D**) phage concentration was measured in the supernatant. Statistical analyses were performed by one-way ANOVA compared to the Wildtype phage with Dunnett's multiple comparison test. *, $p < 0.05$; **, $p < 0.01$.

[0025] FIG. 3A-3O - *In vivo* sfGFP production by engineered phage. **FIG. 3A** shows a schematic representation of the mouse model. **FIG. 3B** shows graphs of fecal *E. coli* and phage concentrations immediately prior to phage administration and at experimental conclusion. **FIG. 3C** shows a graph of the mean fluorescence intensity of the murine mucosa, showing significantly higher for mice colonized with T4::sfgfp phage compared to wildtype T4. **FIG. 3D** shows fluorescence was visually evident within fluorescence microscopy images of unfixed colonic sections, with minimal autofluorescence in mice receiving wildtype T4 phage. Dashed lines represent the intestinal mucosa **FIG. 3E** shows that mice colonized with T4::sfgfp had substantially greater mucosal fluorescence as shown by the fluorescence microscopy images. Dashed lines represent the intestinal mucosa. **FIG. 3F** shows a fluorescence microscopy image of a colonic section fixed and fluorescently stained with nuclear and mucin stains. **FIG. 3G** shows a fluorescence microscopy image of a colonic section fixed and fluorescently stained with nuclear and mucin stains. **FIG. 3H** shows a fluorescence microscopy image of a colonic section fixed and fluorescently stained with nuclear and mucin stains. **FIG. 3I** shows a fluorescence microscopy image of a colonic section fixed and fluorescently stained with nuclear and mucin stains. **FIG. 3J** shows a fluorescence microscopy image of a colonic section fixed and fluorescently stained with nuclear and mucin stains. **FIG. 3K** shows a fluorescence microscopy image of a colonic section fixed and fluorescently stained with nuclear and mucin stains. **FIG. 3L** shows a fluorescence microscopy image of a colonic section fixed and fluorescently stained with nuclear and mucin stains. **FIG. 3M** shows a fluorescence microscopy image of a colonic section fixed and fluorescently stained with nuclear and mucin stains. **FIG. 3N** shows a fluorescence microscopy image of a colonic section fixed and fluorescently stained with nuclear and mucin stains. **FIG. 3O** shows a fluorescence microscopy image of a colonic section fixed and fluorescently stained with nuclear and mucin stains. Scale bars = 10 μ m. All bars represent mean values. Statistical analysis was performed by a two-tailed unpaired t-test. ****, $p < 0.0001$.

[0026] FIG. 4A-4G – The efficacy of T4::clpB to reduce food consumption and weight gain. **FIG. 4A** shows a schematic representation of the *in vivo* experiment using a murine model of diet-induced obesity with C57BL/6 mice. Mice were colonized with *E. coli* alone, wildtype T4 phage, or T4::clpB phage. **FIG. 4B** shows a graph of the concentration of Phage, by plaque assay, showing that wildtype T4 and T4::clpB are able to coexist with their *E. coli* host *in vivo* and that the presence of phage does not markedly alter *E. coli* colonization. **FIG. 4C** shows a graph of the concentration of bacteria, by selective culture, showing that wildtype T4 and T4::clpB are able to coexist with their *E. coli* host *in vivo* and that the presence of phage does not markedly alter *E. coli* colonization. **FIG. 4D** shows a graph of longitudinal weight measurements collected for *E. coli*, *E. coli* + T4, and *E. coli* + T4::clpB. **FIG. 4E** shows a graph of food consumption for *E. coli*, *E. coli* + T4, and *E. coli* + T4::clpB. **FIG. 4F** shows a graph of overall activity measured over a 24 h period through distance traveled (m) for *E. coli*, *E. coli* + T4, and *E. coli* + T4::clpB. **FIG. 4G** shows a cytokine panel of serum at sacrifice. Samples with undetectable concentrations were set to the limit of detection. (B,C,G) Lines represent median values and (D,E,F) bars represent mean values. Statistical analyses were performed by (D,F,G) two-way ANOVA or (E) one-way ANOVA with (D,E,F,G) Tukey's multiple comparison test. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

[0027] FIG. 5 - *E. coli* EC001 growth alone and infected with wildtype T4 phage, demonstrating reduced OD600 when co- colonized with phage and indicating proficient cell lysis.

[0028] FIG. 6A-6C - Quantitation of His6-tagged ClpB and supernatant protein of phage lysates. Total lysate concentration analyzed via Bradford reagent (FIG. 6A) was used to normalize His6-tag ELISA results (FIG. 6B), indicating significant production of His6-tagged protein via phage infection (FIG. 6C). Statistical analyses were performed by a two-tailed unpaired t-test. *, $p < 0.05$; **, $p < 0.01$.

[0029] FIG. 7 - T4 phage engineered to express a serine protease inhibitor. The engineered T4 phage delivered a functional serine protease inhibitor.

[0030] The figures herein are for illustrative purposes only and are not necessarily drawn to scale.

DETAILED DESCRIPTION

[0031] Before the present disclosure is described in greater detail, it is to be understood that this disclosure is not limited to particular embodiments described, and as such may, 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.

[0032] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. Although any methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present disclosure, the preferred methods and materials are now described.

[0033] All publications and patents cited in this specification are cited to disclose and describe the methods and/or materials in connection with which the publications are cited. All such publications and patents are herein incorporated by references as if each individual publication or patent were specifically and individually indicated to be incorporated by reference. Such incorporation by reference is expressly limited to the methods and/or materials described in the cited publications and patents and does not extend to any lexicographical definitions from the cited publications and patents. Any lexicographical definition in the publications and patents cited that is not also expressly repeated in the instant application should not be treated as such and should not be read as defining any terms appearing in the accompanying claims. The citation of any publication is for its disclosure prior to the filing date and should not be construed as an admission that the present disclosure is not entitled to antedate such publication by virtue of prior disclosure. Further, the dates of publication provided could be different from the actual publication dates that may need to be independently confirmed.

[0034] As will be apparent to those of skill in the art upon reading this disclosure, each of the individual embodiments described and illustrated herein has discrete components and features which may be readily separated from or combined with the features of any of the other several embodiments without departing from the scope or spirit of the present disclosure. Any recited method can be carried out in the order of events recited or in any other order that is logically possible.

[0035] Where a range is expressed, a further aspect includes from the one particular value and/or to the other particular value. Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between

the upper and lower limit of that range and any other stated or intervening value in that stated range, is encompassed within the disclosure. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges and are also encompassed within the disclosure, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the disclosure. For example, where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the disclosure, e.g., the phrase “x to y” includes the range from ‘x’ to ‘y’ as well as the range greater than ‘x’ and less than ‘y’. The range can also be expressed as an upper limit, e.g. ‘about x, y, z, or less’ and should be interpreted to include the specific ranges of ‘about x’, ‘about y’, and ‘about z’ as well as the ranges of ‘less than x’, less than y’, and ‘less than z’. Likewise, the phrase ‘about x, y, z, or greater’ should be interpreted to include the specific ranges of ‘about x’, ‘about y’, and ‘about z’ as well as the ranges of ‘greater than x’, greater than y’, and ‘greater than z’. In addition, the phrase “about ‘x’ to ‘y’”, where ‘x’ and ‘y’ are numerical values, includes “about ‘x’ to about ‘y’”.

[0036] It should be noted that ratios, concentrations, amounts, and other numerical data can be expressed herein in a range format. It will be further understood that the endpoints of each of the ranges are significant both in relation to the other endpoint, and independently of the other endpoint. It is also understood that there are a number of values disclosed herein, and that each value is also herein disclosed as “about” that particular value in addition to the value itself. For example, if the value “10” is disclosed, then “about 10” is also disclosed. Ranges can be expressed herein as from “about” one particular value, and/or to “about” another particular value. Similarly, when values are expressed as approximations, by use of the antecedent “about,” it will be understood that the particular value forms a further aspect. For example, if the value “about 10” is disclosed, then “10” is also disclosed.

[0037] It is to be understood that such a range format is used for convenience and brevity, and thus, should be interpreted in a flexible manner to include not only the numerical values explicitly recited as the limits of the range, but also to include all the individual numerical values or sub-ranges encompassed within that range as if each numerical value and sub-range is explicitly recited. To illustrate, a numerical range of “about 0.1% to 5%” should be interpreted to include not only the explicitly recited values of about 0.1% to about 5%, but also include individual values (e.g., about 1%, about 2%, about 3%, and about 4%) and the sub-ranges (e.g., about 0.5% to about 1.1%; about 5% to about 2.4%; about 0.5% to about 3.2%, and about 0.5% to about 4.4%, and other possible sub-ranges) within the indicated range.

General Definitions

[0038] Unless defined otherwise, technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure pertains. Definitions of common terms and techniques in molecular biology may be found in *Molecular Cloning: A Laboratory Manual*, 2nd edition (1989) (Sambrook, Fritsch, and Maniatis); *Molecular Cloning: A Laboratory Manual*, 4th edition (2012) (Green and Sambrook); *Current Protocols in Molecular Biology* (1987) (F.M. Ausubel et al. eds.); the series *Methods in Enzymology* (Academic Press, Inc.); *PCR 2: A Practical Approach* (1995) (M.J. MacPherson, B.D. Hames, and G.R. Taylor eds.); *Antibodies*, A

Laboratory Manual (1988) (Harlow and Lane, eds.): Antibodies A Laboratory Manual, 2nd edition 2013 (E.A. Greenfield ed.); Animal Cell Culture (1987) (R.I. Freshney, ed.); Benjamin Lewin, Genes IX, published by Jones and Bartlett, 2008 (ISBN 0763752223); Kendrew *et al.* (eds.), The Encyclopedia of Molecular Biology, published by Blackwell Science Ltd., 1994 (ISBN 0632021829); Robert A. Meyers (ed.), Molecular Biology and Biotechnology: a Comprehensive Desk Reference, published by VCH Publishers, Inc., 1995 (ISBN 9780471185710); Singleton *et al.*, Dictionary of Microbiology and Molecular Biology 2nd ed., J. Wiley & Sons (New York, N.Y. 1994), March, Advanced Organic Chemistry Reactions, Mechanisms and Structure 4th ed., John Wiley & Sons (New York, N.Y. 1992); and Marten H. Hofker and Jan van Deursen, Transgenic Mouse Methods and Protocols, 2nd edition (2011).

[0039] Definitions of common terms and techniques in chemistry and organic chemistry can be found in Smith. Organic Synthesis, published by Academic Press. 2016; Tinoco *et al.* Physical Chemistry, 5th edition (2013) published by Pearson; Brown *et al.*, Chemistry, The Central Science 14th ed. (2017), published by Pearson, Clayden *et al.*, Organic Chemistry, 2nd ed. 2012, published by Oxford University Press; Carey and Sunberg, Advanced Organic Chemistry, Part A: Structure and Mechanisms, 5th ed. 2008, published by Springer; Carey and Sunberg, Advanced Organic Chemistry, Part B: Reactions and Synthesis, 5th ed. 2010, published by Springer, and Vollhardt and Schore, Organic Chemistry, Structure and Function; 8th ed. (2018) published by W.H. Freeman.

[0040] Definitions of common terms, analysis, and techniques in genetics can be found in e.g., Hartl and Clark. Principles of Population Genetics. 4th Ed. 2006, published by Oxford University Press. Published by Booker. Genetics: Analysis and Principles, 7th Ed. 2021, published by McGraw Hill; Isik *et al.*, Genetic Data Analysis for Plant and Animal Breeding. First ed. 2017. published by Springer International Publishing AG; Green, E. L. Genetics and Probability in Animal Breeding Experiments. 2014, published by Palgrave; Bourdon, R. M. Understanding Animal Breeding. 2000 2nd Ed. published by Prentice Hall; Pal and Chakravarty. Genetics and Breeding for Disease Resistance of Livestock. First Ed. 2019, published by Academic Press; Fasso, D. Classification of Genetic Variance in Animals. First Ed. 2015, published by Callisto Reference; Megahed, M. Handbook of Animal Breeding and Genetics, 2013, published by Omniscryptum Gmbh & Co. Kg., LAP Lambert Academic Publishing; Reece. Analysis of Genes and Genomes. 2004, published by John Wiley & Sons. Inc; Deonier *et al.*, Computational Genome Analysis. 5th Ed. 2005, published by Springer-Verlag, New York; Meneely, P. Genetic Analysis: Genes, Genomes, and Networks in Eukaryotes. 3rd Ed. 2020, published by Oxford University Press.

[0041] As used herein, the singular forms "a", "an", and "the" include both singular and plural referents unless the context clearly dictates otherwise.

[0042] As used herein, "about," "approximately," "substantially," and the like, when used in connection with a measurable variable such as a parameter, an amount, a temporal duration, and the like, are meant to encompass variations of and from the specified value including those within experimental error (which can be determined by e.g. given data set, art accepted standard, and/or with e.g., a given confidence interval (e.g. 90%, 95%, or more confidence interval from the mean), such as variations of +/-10% or less, +/-5% or less, +/-1% or less, and +/-0.1% or less of and from the specified

value, insofar such variations are appropriate to perform in the disclosure. As used herein, the terms "about," "approximate," "at or about," and "substantially" can mean that the amount or value in question can be the exact value or a value that provides equivalent results or effects as recited in the claims or taught herein. That is, it is understood that amounts, sizes, formulations, parameters, and other quantities and characteristics are not and need not be exact, but may be approximate and/or larger or smaller, as desired, reflecting tolerances, conversion factors, rounding off, measurement error and the like, and other factors known to those of skill in the art such that equivalent results or effects are obtained. In some circumstances, the value that provides equivalent results or effects cannot be reasonably determined. In general, an amount, size, formulation, parameter or other quantity or characteristic is "about," "approximate," or "at or about" whether or not expressly stated to be such. It is understood that where "about," "approximate," or "at or about" is used before a quantitative value, the parameter also includes the specific quantitative value itself, unless specifically stated otherwise.

[0043] The term "optional" or "optionally" means that the subsequent described event, circumstance or substituent may or may not occur, and that the description includes instances where the event or circumstance occurs and instances where it does not.

[0044] The recitation of numerical ranges by endpoints includes all numbers and fractions subsumed within the respective ranges, as well as the recited endpoints.

[0045] As used herein, "adjuvant" refers to an additional compound, composition, or ingredient that can facilitate stimulation an immune response in addition to the main antigen of a composition, formulation, or vaccine. Generally, an adjuvant can increase the immune response of an antigen as compared to the antigen alone. This can improve and/or facilitate any protective immunity developed in the recipient subject in response to the antigen. "Adjuvant" as used herein can refer to a component that potentiates the immune responses to an antigen and/or modulates it towards the desired immune response(s).

[0046] As used herein, "antigen" refers to a molecule (biologic (e.g., nucleic acid, protein, fat, or any combination thereof) or non-biologic) with one or more epitopes that stimulate a host's immune system to make a secretory, humoral and/or cellular antigen-specific response, which can be B-cell mediated and/or T-cell mediated. The term can also be used herein to refer to a nucleic acid molecule that is capable of producing such an antigen in a subject. The term is also used interchangeably with "immunogen." For example, a specific antigen can be complete protein, portions of a protein, peptides, fusion proteins, glycosylated proteins and combinations thereof.

[0047] As used herein, "anti-infective" refers to compounds or molecules that can either kill an infectious agent or inhibit it from spreading. Anti-infectives include, but are not limited to, antibiotics, antibacterials, antifungals, antivirals, and antiprotozoals.

[0048] As used herein, a "biological sample" refers to a sample obtained from, made by, secreted by, excreted by, or otherwise containing part of or from a biologic entity. A biologic sample can contain whole cells and/or live cells and/or cell debris, and/or cell products, and/or virus particles. The biological sample can contain (or be derived from) a "bodily fluid". The biological sample can be obtained from an environment (e.g., water source, soil, air, and the like). Such samples are also referred to herein as environmental samples. As used herein "bodily fluid" refers to any non-solid excretion, secretion, or

other fluid present in an organism and includes, without limitation unless otherwise specified or is apparent from the description herein, amniotic fluid, aqueous humor, vitreous humor, bile, blood or component thereof (e.g. plasma, serum, etc.), breast milk, cerebrospinal fluid, cerumen (earwax), chyle, chyme, endolymph, perilymph, exudates, feces, female ejaculate, gastric acid, gastric juice, lymph, mucus (including nasal drainage and phlegm), pericardial fluid, peritoneal fluid, pleural fluid, pus, rheum, saliva, sebum (skin oil), semen, sputum, synovial fluid, sweat, tears, urine, vaginal secretion, vomit and mixtures of one or more thereof. Biological samples include cell cultures, bodily fluids, cell cultures from bodily fluids. Bodily fluids may be obtained from an organism, for example by puncture, or other collecting or sampling procedures.

[0049] As used herein, "dose," "unit dose," or "dosage" refers to physically discrete units suitable for use in a subject, each unit containing a predetermined quantity of a pharmaceutical formulation described herein calculated to produce the desired response or responses in association with its administration.

[0050] As used herein, "immunomodulator," refers to an agent, such as a therapeutic agent, which is capable of modulating or regulating one or more immune function or response.

[0051] As used herein, "immune response" refers to the reaction of the molecules, components, pathways, organs, fluids and/or cells of the body to the presence of a substance that is foreign or recognized by the body as foreign to the body. The immune response can include a cell mediated immune response and/or an adaptative or humoral immune response. The immune response can include a B-cell mediated and/or a T-cell mediated immune response.

[0052] As used herein, "modulate or modulation of the immune response" refers to change in the immune response that results from the introduction of a composition, vaccine, or other compound or formulation described herein in a recipient subject as compared to a suitable control.

[0053] The term "molecular weight", as used herein, can generally refer to the mass or average mass of a material. If a polymer or oligomer, the molecular weight can refer to the relative average chain length or relative chain mass of the bulk polymer. In practice, the molecular weight of polymers and oligomers can be estimated or characterized in various ways including gel permeation chromatography (GPC) or capillary viscometry. GPC molecular weights are reported as the weight-average molecular weight (M_w) as opposed to the number-average molecular weight (M_n). Capillary viscometry provides estimates of molecular weight as the inherent viscosity determined from a dilute polymer solution using a particular set of concentration, temperature, and solvent conditions.

[0054] As used herein, "pharmaceutical formulation" refers to the combination of an active agent, compound, or ingredient with a pharmaceutically acceptable carrier or excipient, making the composition suitable for diagnostic, therapeutic, or preventive use in vitro, in vivo, or ex vivo.

[0055] As used herein, "pharmaceutically acceptable carrier" refers to a carrier or excipient that is useful in preparing a pharmaceutical formulation that is generally safe, non-toxic, and is neither biologically or otherwise undesirable, and includes a carrier or excipient that is acceptable for veterinary use and/or human pharmaceutical use. A "pharmaceutically acceptable carrier or excipient" as used in the specification and claims includes both one and more than one such carrier or excipient.

[0056] As used herein, "pharmaceutically acceptable salt" refers to any acid or base addition salt whose counter-ions are non-toxic to the subject to which they are administered in pharmaceutical doses of the salts.

[0057] The terms "subject," "individual," and "patient" are used interchangeably herein to refer to a vertebrate, preferably a mammal, more preferably a human. Mammals include, but are not limited to, murines, simians, humans, farm animals, sport animals, and pets. Tissues, cells and their progeny of a biological entity obtained in vivo or cultured in vitro are also encompassed.

[0058] As used herein, the term "T cell antigen" trgrtd to any antigen that is recognized by and triggers an immune response in a T cell (also referred to herein as e.g., a T-cell immune response or T-cell mediated immune response) (e.g., an antigen that is specifically recognized by a T cell receptor on a T cell via presentation of the antigen or portion thereof bound to a major histocompatibility complex molecule (MHC). In some embodiments, an antigen that is a T cell antigen is also a B cell antigen. In other embodiments, the T cell antigen is not also a B cell antigen. T cells antigens generally are proteins or peptides. T cell antigens may be an antigen that stimulates a CD8+ T cell response, a CD4+ T cell response, or both. The nanocarriers, therefore, in some embodiments can effectively stimulate both types of responses.

[0059] As used herein, the terms "treating" and "treatment" refer generally to obtaining a desired pharmacological and/or physiological effect. The effect can be, but does not necessarily have to be, prophylactic in terms of preventing or partially preventing a disease, symptom or condition thereof. The term "treatment" as used herein covers any treatment of disease or symptom thereof, such as one that can be treated by the one or more gene products encoded by the one or more heterologous genes of the engineered phage of the present disclosure, in a subject, such as a in a mammal, particularly a human and includes any one or more of the following: (a) preventing the disease from occurring in a subject which may be predisposed to the disease but has not yet been diagnosed as having it; (b) inhibiting the disease, i.e., arresting its development; (c) relieving the disease, i.e., mitigating or ameliorating the disease and/or its symptoms or conditions (d) and/or curing the disease (i.e., eliminating the causative agent or genetic cause of the disease from the subject). The term "treatment" as used herein can refer to therapeutic treatment alone, prophylactic treatment alone, or both therapeutic and prophylactic treatment. Those in need of treatment (subjects in need thereof) can include those already with the disorder and/or those in which the disease or disorder is to be prevented, such as those where it is predicted they are to suffer from the disease or disorder and/or have been determined to be at risk (e.g., due to one or more factors, exposure or potential exposure, genetic make-up, lifestyle, etc.) of developing the disease and/or condition. Efficacy can be measured using objective or subjective techniques. For example, where the engineered phage is to deliver a gene product that is designed to stimulate an immune response efficacy can be measured via determining antibody titers and comparing them to a standard and/or control. Efficacy can be measured by measuring the occurrence of nicotine use and comparing the amount of use to a standard, control, and/or over a period of time. Likewise, efficacy for other gene products to be delivered can be objectively determined by measuring the effects expected to be observed (e.g., direct measurement of the gene product delivered, modulation in a biologic activity that is directly affected by the gene product, etc.) or an indirect effect

thereof in the subject (e.g., measuring bloodwork parameters, or other phenotypic characteristics of the subject treated). Other methods of determining efficacy will be appreciated by those of skill in the art.

[0060] As used herein, "vaccine" refers to a compound, molecule, compositions, and formulations or compounds, molecules, etc. delivered by a composition (e.g., an engineered phage) of the present disclosure that are capable of inducing an immune response in a subject. The term "vaccine" can also be used to refer to a compound, molecule, compositions, and formulations that are capable of providing protective immunity against an organism, or provide desensitization (i.e., a reduced immune response) to an antigen. The vaccine can be capable of stimulating a B-cell and/or T-cell immune response in the subject.

[0061] Various embodiments are described hereinafter. It should be noted that the specific embodiments are not intended as an exhaustive description or as a limitation to the broader aspects discussed herein. One aspect described in conjunction with a particular embodiment is not necessarily limited to that embodiment and can be practiced with any other embodiment(s). Reference throughout this specification to "one embodiment", "an embodiment," "an example embodiment," means that a particular feature, structure or characteristic described in connection with the embodiment is included in at least one embodiment of the present disclosure. Thus, appearances of the phrases "in one embodiment," "in an embodiment," or "an example embodiment" in various places throughout this specification are not necessarily all referring to the same embodiment, but may. Furthermore, the particular features, structures or characteristics may be combined in any suitable manner, as would be apparent to a person skilled in the art from this disclosure, in one or more embodiments. Furthermore, while some embodiments described herein include some but not other features included in other embodiments, combinations of features of different embodiments are meant to be within the scope of the disclosure. For example, in the appended claims, any of the claimed embodiments can be used in any combination.

[0062] As used herein the term "altered expression" may particularly denote altered production of the recited gene products by a cell. As used herein, the term "gene product(s)" includes RNA transcribed from a gene (e.g., mRNA), or a polypeptide encoded by a gene or translated from RNA.

[0063] Also, "altered expression" as intended herein may encompass modulating the activity of one or more endogenous gene products. Accordingly, "altered expression", "altering expression", "modulating expression", or "detecting expression" or similar may be used interchangeably with respectively "altered expression or activity", "altering expression or activity", "modulating expression or activity", or "detecting expression or activity" or similar terms. As used herein, "modulating" or "to modulate" generally means either reducing or inhibiting the activity of a target or antigen, or alternatively increasing the activity of the target or antigen, as measured using a suitable *in vitro*, cellular or *in vivo* assay. In particular, "modulating" or "to modulate" can mean either reducing or inhibiting the (relevant or intended) activity of, or alternatively increasing the (relevant or intended) biological activity of the target or antigen, as measured using a suitable *in vitro*, cellular or *in vivo* assay (which will usually depend on the target or antigen involved), by at least 5%, at least 10%, at least 25%, at least 50%, at least 60%, at least 70%, at least 80%, or 90% or more, compared to activity of the target or antigen in

the same assay under the same conditions but without the presence of the inhibitor/antagonist agents or activator/agonist agents described herein.

[0064] As will be clear to the skilled person, "modulating" can also involve effecting a change (which can either be an increase or a decrease) in affinity, avidity, specificity and/or selectivity of a target or antigen, for one or more of its targets compared to the same conditions but without the presence of a modulating agent. Again, this can be determined in any suitable manner and/or using any suitable assay known per se, depending on the target. In particular, an action as an inhibitor/antagonist or activator/agonist can be such that an intended biological or physiological activity is increased or decreased, respectively, by at least 5%, at least 10%, at least 25%, at least 50%, at least 60%, at least 70%, at least 80%, or 90% or more, compared to the biological or physiological activity in the same assay under the same conditions but without the presence of the inhibitor/antagonist agent or activator/agonist agent. Modulating can also involve activating the target or antigen or the mechanism or pathway in which it is involved.

[0065] All publications, published patent documents, and patent applications cited herein are hereby incorporated by reference to the same extent as though each individual publication, published patent document, or patent application was specifically and individually indicated as being incorporated by reference.

OVERVIEW

[0066] Delivery of therapeutic proteins has presented significant challenges compared to non-biologic therapeutics. Strategies employed to improve bioavailability of non-biologic therapeutics have not translated to biologics. For example, for small molecule therapeutics, there has been extensive development into coatings, such as enteric coatings, that provide extended release as well as responsive release (e.g., pH responsive) such that the therapeutic is protected from degradation until release in the small intestine. However, such approaches have not translated to protein delivery. More recent approaches include engineered microbial cells to express proteins, but such methods require extremely high doses (e.g., 1×10^{11} cells given three times a day). Further the engineered microbes suffer from longevity, due in part, to the metabolic burden placed from production of the therapeutic protein.

[0067] With that said, embodiments disclosed herein provide engineered phages to deliver exogenous proteins to a subject. The engineered phages can contain one or more genes encoding one or more heterologous gene products (e.g., nucleic acids and/or proteins). Without being bound by theory, when the engineered phages are administered to a subject in need thereof they infect a microbe in the subject in need thereof, such as in a microbiome. After infection the host microbial cell infected with the engineered phage produces the heterologous gene products, which then can be released by the host cells, such as by lysis that is part of the phage life cycle. Without being bound by theory, Applicant can leverage the microbiome to target delivery of gene product, such as a protein, to the microenvironment of the microbiome. Without being bound by theory, this can provide an increase in oral bioavailability of the gene products as compared to other strategies for oral delivery of the same gene product. Other compositions, compounds, methods, features, and advantages of the present disclosure will be or become apparent to one having ordinary skill in the art upon examination of the

following drawings, detailed description, and examples. It is intended that all such additional compositions, compounds, methods, features, and advantages be included within this description, and be within the scope of the present disclosure.

ENGINEERED PHAGES

[0068] Generally, phages are viruses that infect microbes such as bacteria and yeast. Bacterial phages, also known as bacteriophages, are viruses that specifically infect bacteria. They are incredibly diverse and abundant, with estimated numbers reaching trillions upon trillions in various environments around the world. Generally, phages are composed of genetic material, which can be either DNA or RNA, surrounded by a protein coat. They can be non-enveloped. They vary in size and structure, with some being relatively simple while others possess complex architectures. The genetic material of a virulent or lytic phage encodes all the necessary instructions to hijack the bacterial cell's machinery, leading to the production of new phage particles and ultimately the lysis, or destruction, of the infected bacterium or yeast. Yeast phages are members of the virus family Cystoviridae or Totiviridae and are classified based on their genetic material and structural characteristics. See also e.g., Dion et al., *Nat Rev Microbiol.* 2020 Mar;18(3):125-138.

[0069] Described in certain example embodiments herein are engineered phages containing one or more heterologous cargo genes, wherein the one or more exogenous genes encode one or more heterologous gene products (e.g., nucleic acids and/or polypeptides); and one or more regulatory elements, wherein each of the one or more heterologous cargo genes is operatively coupled to at least one of the one or more regulatory elements. In some embodiments, the one or more heterologous cargo genes are present in the genome of the engineered phage.

[0070] In certain example embodiments, the phage is a lytic phage. In certain example embodiments, the phage is a T4 phage, T2 phage, MS2 phage, M13 phage, T3 phage, T5 phage, T7 phage, Lamda phage, P1 phage, PhiX174 phage, Mu phage, P22 phage, *Pseudomonas* phage phiKZ, *Bacillus* phage phi29, *Salmonella* phage P22, *Mycobacterium* phage D29, *Enterobacteria* phage T6, *Mycobacterium smegmatis* phage L5, *Escherichia* phage T5, *listeria* phage A511, *Vibrio* phage KVP40, *Staphylococcus* phage K, or *Staphylococcus* phage PhiETA.

[0071] In certain example embodiments, the phage is capable of infecting a bacterial cell or a yeast cell. In certain example embodiments, the phage is capable of infecting a cell of a microbiome. In certain example embodiments, the microbiome is a gastrointestinal microbiome, a skin microbiome, an airway microbiome, a vaginal microbiome, or any combination thereof. In certain example embodiments, the gastrointestinal microbiome is an oral cavity microbiome, a stomach microbiome, a small intestine microbiome, a large intestine microbiome, an appendix microbiome, a cecum microbiome, or any combination thereof.

[0072] In some embodiments, the host cell is a bacterial cell. In some embodiments, the bacterial cell is an *Escherichia coli* (*E. coli*). In some embodiments, the *E. coli* is *E. coli* B, *E. coli* K-12, *E. coli* C, *E. coli* W3110, *E. coli* BL21. In some embodiments, the bacterial cell is pathogenic to the subject. In some embodiments, the bacterial cell is a commensal bacteria cell. In some embodiments, the bacterial cell is a *Salmonella*, *Shigella*, *Campylobacter*, *E. coli*, or *Yersinia* cell. In some embodiments, the bacterial cell is an *E. coli* or a *Bifidobacterium*. In some embodiments, the bacteria cell is *C. difficile*.

Exemplary microorganisms present in various microbiomes are generally known and will be appreciated by those of skill in the art. See e.g., Ruan et al., *Dig Dis Sci.* 2020 Mar;65(3):695-705 and Deng and Swanson. *Br J Nutr.* 2015 Jan;113 Suppl:S6-17 (gastrointestinal microbiome); Buchta, V., *Ceska Gynekol.* 2018 Winter;83(5):371-379 (vaginal microbiome); Loso et al., *Asia Pac Allergy.* 2022 Jul 29;12(3):e32 (airway microbiome); Byrd et al., *Nat Rev Microbiol.* 2018 Mar;16(3):143-155 (skin microbiome); Yagi et al., *Int J Mol Sci.* 2021 Oct 8;22(19) (lung microbiome). Any such cells can be host cells in accordance with the present disclosure.

[0073] In some embodiments, the host cell is a yeast cell. In some embodiments, the yeast cell is pathogenic to the subject. In some embodiments, the yeast cell is a commensal yeast cell. In some embodiments, the yeast cell is a yeast cell of the genus *Candida* or *Saccharomyces*. In some embodiments, the *Candida* is *Candida albicans*. In some embodiments, the *Saccharomyces* is *S. boulardii* or *S. cerevisiae*.

[0074] In certain example embodiments, the one or more heterologous gene products are therapeutic proteins. In certain example embodiments, the one or more heterologous gene products are capable of (a) stimulating an immune response in a subject in which it is delivered; (b) treating a disease or a symptom thereof; (c) modulating a microbiome; or any combination of (a)-(c). In certain example embodiments, stimulating an immune response comprises stimulating a B-cell and/or T-cell mediated response. Such gene products in this context can also be referred to as antigens or antigenic components.

[0075] In certain example embodiments, the one or more regulatory elements comprises a phage promoter, optionally wherein the phage promoter is an endogenous phage promoter. In certain example embodiments, the one or more regulatory elements comprises an early phage promoter, a late phage promoter, or both. Exemplary early promoters include, but are not limited to, T7 early promoters A1, A2, and A3; lambda (λ) phage early promoter; and T4 early promoters *motB*, *gp55*, *lpl*, *NAligase*, *ndd*. Exemplary late promoters include, but are not limited to, *gp18*, *gp22*, *gp23*, *gp67*, and *soc*. In certain example embodiments, the engineered phage is a T4 phage and the one or more regulatory elements comprise *motB*, *gp18*, *gp22*, or any combination thereof.

[0076] In certain example embodiments, the one or more heterologous gene products comprises one or more signaling domains and/or translocation domains or nucleic acid sequence encoding the same. In certain example embodiments, the one or more signaling domains and/or translocation domains are capable of translocating the one or more heterologous gene products (such as a protein), when expressed in a host cell, to a periplasm.

[0077] In certain example embodiments, the one or more heterologous gene products (e.g., proteins) comprises an N-terminal signal peptide for periplasmic secretion, optionally wherein the N-terminal signal peptide comprises (a) a signal peptide selected from *Tat*, *appA*, *dsbD*, *degQ*, *cueO*, *pldA*, *ushA*, *pagP*, *fhuA*, *ptrA*, *lptA*, *bepA*, *dacA*, *agp*, *malS*, *spy*, *amiC*, *tolB*, *rseB*, *amtB*, *dsbC*, *amiA*, *torZ*, *FhuD*, *zinT*, *mgIB*, *eco*, *nikA*, *mepH*, *fadL*, *ompL*, *ompN*, *thiB*, *MalE*, *PhoA*, or any combination thereof; (b) a leader peptide from *spA*, *phoA*, *ribose binding protein*, *pelB*, *ompA*, *ompT*, *dsbA*, *torA*, *torT*, *tolT* or any combination thereof; (c) or any combination thereof.

Phagemid Vectors and Encoding Polynucleotides

[0078] Also described herein are vectors, such as phagemid vectors, that encode the engineered phages of the present disclosure. Also described herein are polynucleotides that encode one or more components of the engineered phages of the present disclosure.

[0079] Phagemids, short for phage plasmids, are hybrid genetic constructs that combine the properties of both plasmids and bacteriophages. Phagemids have been used extensively in other contexts, such as for cloning and manipulating DNA in molecular biology research. Thus, one of ordinary skill in the art will be familiar with the principles, components, and design of phagemid vectors suitable for encoding and expressing the engineered phagemids of the present disclosure. Phagemids are generally plasmids that contain a phage origin of replication, allowing them to be packaged into viral particles and propagated as phages.

[0080] A basic phagemid vector contains a phage genome, such as a genome of an engineered phage of the present disclosure. In some embodiments, the phagemid can include a selectable marker. In some embodiments, the phage genome and/or phagemid contains, in addition to one or more engineered heterologous genes and/or regulatory elements driving expression of the one or more heterologous genes (i.e., a heterologous gene construct), genes encoding one or more structural proteins of the phage (e.g., capsid proteins, tail proteins, tail fibers, any other components necessary for phage assembly and/or attachment to host cells), gene(s) encoding the phage replication machinery (e.g., proteins involved in DNA replication (including but not limited to DNA polymerases, helicases, primases, and other replication related-factors), gene(s) encoding the packaging machinery (e.g., packaging signals), gene(s) encoding lytic proteins (e.g., proteins that will cause lysis of the host cell); regulatory elements (e.g., regulatory elements that drive/control expression of the different genes during the life cycle of the phage, which are in addition to those that may be part of the heterologous gene construct); integrase or repressor gene(s) (such as in the case of temperate phages). One of ordinary skill in the art will appreciate that the precise arrangement and organization of these elements can differ significantly between phages, reflecting the diverse strategies and adaptations that different phages have evolved to infect and replicate within their host cells. Such genomic arrangements of wild-type phages that can be engineered to include a heterologous construct described herein is known or is reasonably available to one of ordinary skill in the art. Standard recombinant and nucleic acid engineering techniques can be used to generate a heterologous gene construct and/or generate an engineered phage genome and/or phagemid for propagation of the engineered phages of the present disclosure. For example, the vectors described herein can be constructed using any suitable process or technique. In some embodiments, one or more suitable recombination and/or cloning methods or techniques can be used to the vector(s) described herein. Suitable recombination and/or cloning techniques and/or methods can include, but not limited to, those described in U.S. Patent Publication No. US 2004/0171156 A1.

Codon Optimization of Vector Polynucleotides

[0081] In some embodiments, one or more polynucleotides, such as engineered phage genome and/or component thereof are contained in a vector ("vector polynucleotides") and are codon optimized. Different vector polynucleotides (e.g., genes, constructs, etc. within the vector) can be codon optimized

separately. Thus, different genes or gene constructs can be codon optimized for expression in different cell types. In some embodiments, the vector polynucleotides corresponding to phage production or the phage itself (e.g., structural proteins etc.) are codon optimized for expression in the host cell (e.g., a bacterial or yeast cell) and the heterologous gene(s) are codon optimized for expression in a non-phage host and/or non-yeast eukaryotic cell of the subject (e.g., a mammal or non-mammal animal). In general, codon optimization refers to a process of modifying a nucleic acid sequence for enhanced expression in the host cells of interest by replacing at least one codon (e.g., about or more than about 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, or more codons) of the native sequence with codons that are more frequently or most frequently used in the genes of that host cell while maintaining the native amino acid sequence. Various species exhibit particular bias for certain codons of a particular amino acid. Codon bias (differences in codon usage between organisms) often correlates with the efficiency of translation of messenger RNA (mRNA), which is in turn believed to be dependent on, among other things, the properties of the codons being translated and the availability of particular transfer RNA (tRNA) molecules. The predominance of selected tRNAs in a cell is generally a reflection of the codons used most frequently in peptide synthesis. Accordingly, genes can be tailored for optimal gene expression in a given organism based on codon optimization. Codon usage tables are readily available, for example, at the "Codon Usage Database" available at www.kazusa.or.jp/codon/ and these tables can be adapted in a number of ways. See Nakamura, Y., et al. "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" *Nucl. Acids Res.* 28:292 (2000). Computer algorithms for codon optimizing a particular sequence for expression in a particular host cell are also available, such as Gene Forge (Aptagen; Jacobus, PA), are also available. In some embodiments, one or more codons (e.g., 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, or more, or all codons) in a sequence encoding a DNA/RNA-targeting Cas protein corresponds to the most frequently used codon for a particular amino acid. As to codon usage in yeast, reference is made to the online Yeast Genome database available at http://www.yeastgenome.org/community/codon_usage.shtml, or *Codon selection in yeast*, Bennetzen and Hall, *J Biol Chem.* 1982 Mar 25;257(6):3026-31. As to codon usage in plants including algae, reference is made to *Codon usage in higher plants, green algae, and cyanobacteria*, Campbell and Gowri, *Plant Physiol.* 1990 Jan; 92(1): 1-11.; as well as *Codon usage in plant genes*, Murray et al, *Nucleic Acids Res.* 1989 Jan 25;17(2):477-98; or *Selection on the codon bias of chloroplast and cyanelle genes in different plant and algal lineages*, Morton BR, *J Mol Evol.* 1998 Apr;46(4):449-59.

[0082] A vector polynucleotide can be codon optimized for expression in a specific cell-type, tissue type, organ type, and/or subject type. In some embodiments, a codon optimized sequence is a sequence optimized for expression in a eukaryote, e.g., humans (i.e., being optimized for expression in a human or human cell), or for another eukaryote, such as another animal (e.g. a mammal or avian) as is described elsewhere herein. Such codon optimized sequences are within the ambit of the ordinary skilled artisan in view of the description herein. In some embodiments, the polynucleotide is codon optimized for a specific cell type. Such cell types can include, but are not limited to, epithelial cells (including skin cells, cells lining the gastrointestinal tract, cells lining other hollow organs), nerve cells (nerves, brain cells, spinal column cells, nerve support cells (e.g. astrocytes, glial cells, Schwann cells etc.) , muscle cells (e.g. cardiac muscle, smooth muscle cells, and skeletal muscle cells), connective

tissue cells (fat and other soft tissue padding cells, bone cells, tendon cells, cartilage cells), blood cells, stem cells and other progenitor cells, immune system cells, germ cells, and combinations thereof. Such codon optimized sequences are within the ambit of the ordinary skilled artisan in view of the description herein. In some embodiments, the polynucleotide is codon optimized for a specific tissue type. Such tissue types can include, but are not limited to, muscle tissue, connective tissue, connective tissue, nervous tissue, and epithelial tissue. Such codon optimized sequences are within the ambit of the ordinary skilled artisan in view of the description herein. In some embodiments, the polynucleotide is codon optimized for a specific organ. Such organs include, but are not limited to, muscles, skin, intestines, liver, spleen, brain, lungs, stomach, heart, kidneys, gallbladder, pancreas, bladder, thyroid, bone, blood vessels, blood, and combinations thereof. Such codon optimized sequences are within the ambit of the ordinary skilled artisan in view of the description herein.

[0083] In some embodiments, a vector polynucleotide is codon optimized for expression in particular cells, such as prokaryotic or eukaryotic cells. The eukaryotic cells may be those of or derived from a particular organism, such as a plant or a mammal, including but not limited to human, or non-human eukaryote or animal or mammal as discussed herein, e.g., mouse, rat, rabbit, dog, livestock, or non-human mammal or primate.

[0084] Also described herein are cells and populations thereof that contain a phagemid described herein, such as those useful for propagation of the engineered phages described herein.

Regulatory Elements

[0085] In certain embodiments, the polynucleotides and/or vectors thereof described herein can include one or more regulatory elements that can be operatively linked a polynucleotide of the engineered phagemid, genome or component thereof (e.g., a heterologous gene or construct). The term "regulatory element" is intended to include promoters, enhancers, internal ribosomal entry sites (IRES), other expression control elements (e.g., transcription termination signals, such as polyadenylation signals and poly-U sequences) and cellular localization signals (e.g., nuclear localization or export signals). Exemplary regulatory elements are described, for example, in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory elements include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). A tissue-specific promoter can direct expression primarily in a desired tissue of interest, such as muscle, neuron, bone, skin, blood, specific organs (e.g., liver, pancreas), or particular cell types (e.g., lymphocytes). Regulatory elements may also direct expression in a temporal-dependent manner, such as in a cell-cycle dependent or developmental stage-dependent manner, which may or may not also be tissue or cell-type specific. In some embodiments, a vector comprises one or more pol III promoter (e.g., 1, 2, 3, 4, 5, or more pol III promoters), one or more pol II promoters (e.g., 1, 2, 3, 4, 5, or more pol II promoters), one or more pol I promoters (e.g., 1, 2, 3, 4, 5, or more pol I promoters), or combinations thereof. Examples of pol III promoters include, but are not limited to, U6 and H1 promoters. Examples of pol II promoters include, but are not limited to, the retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV

enhancer) (see, e.g., Boshart et al, Cell, 41:521-530 (1985)), the SV40 promoter, the dihydrofolate reductase promoter, the β -actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1 α promoter. Also encompassed by the term "regulatory element" are enhancer elements, such as WPRE; CMV enhancers; the R-U5' segment in LTR of HTLV-I (Mol. Cell. Biol., Vol. 8(1), p. 466-472, 1988); SV40 enhancer; and the intron sequence between exons 2 and 3 of rabbit β -globin (Proc. Natl. Acad. Sci. USA., Vol. 78(3), p. 1527-31, 1981). Exemplary promoters also include bovine U6 (bU6) and bovine 7SK (b7SK), and other bovine PolIII promoters (see e.g., Lambeth et al., Anim Genet. 2006 Aug;37(4):369-72), bovine papillomavirus-1 promoters (BPV-1) (Linz and Baker. J Virol. 1988 Aug;62(8):2537-43. doi: 10.1128/JVI.62.8.2537-2543.1988), the bovine SIX1 gene promoter (see e.g., Wei et al. Scientific Reports volume 7, Article number: 12599 (2017)), bovine growth hormone promoter (see e.g., Jiang et al., Nuc Acid Prot Syn Mol Gen. 1999. 274(12): 7893-7900), bovine pyruvate carboxylase (see e.g., Hazelton et al. J. Dairy Sci. 91:91-99), a bidirectional promoter (see e.g., Meersserman et al. DNA Research, Volume 24, Issue 3, June 2017, Pages 221-233), a bovine Akt3 promoter (see e.g., Farmanullah et al. Journal of Genetic Engineering and Biotechnology (2021) 19:164), bovine alpha-lactalbumin promoter (see e.g., FEBS Lett. 1991 Jun 17;284(1):19-22), bovine beta-casein promoter (see e.g., Cerdan et al., Mol Reprod Dev. 1998 Mar;49(3):236-45), any combination thereof.

[0086] In some embodiments, the regulatory sequence can be a regulatory sequence described in U.S. Pat. No. 7,776,321, U.S. Pat. Pub. No. 2011/0027239, or International Patent Publication No. WO 2011/028929, the contents of which are incorporated by reference herein in their entireties. In some embodiments, the vector can contain a minimal promoter. In some embodiments, the minimal promoter is the Mecp2 promoter, tRNA promoter, or U6. In a further embodiment, the minimal promoter is tissue specific. In some embodiments, the length of the vector polynucleotide the minimal promoters and polynucleotide sequences is less than 4.4Kb.

[0087] To express a polynucleotide, the vector can include one or more transcriptional and/or translational initiation regulatory sequences, e.g., promoters, that direct the transcription of a gene and/or translation of the encoded protein in a cell. In some embodiments a constitutive promoter may be employed. Exemplary constitutive promoters for mammalian cells are generally known in the art and include, but are not limited to SV40, CAG, CMV, EF-1 α , β -actin, RSV, and PGK. Suitable constitutive promoters for bacterial cells, yeast cells, and fungal cells are generally known in the art, such as a T-7 promoter for bacterial expression and an alcohol dehydrogenase promoter for expression in yeast.

[0088] In some embodiments, the regulatory element can be a regulated promoter. As used herein, "regulated promoter" refers to promoters that direct gene expression not constitutively, but in a temporally- and/or spatially-regulated manner, and includes tissue-specific, tissue-preferred and inducible promoters. Regulated promoters include conditional promoters and inducible promoters. In some embodiments, conditional promoters can be employed to direct expression of a polynucleotide in a specific cell type, under certain environmental conditions, and/or during a specific state of development. Suitable tissue specific promoters can include, but are not limited to, liver specific promoters (e.g. APOA2, SERPIN A1 (hAAT), CYP3A4, and MIR122), pancreatic cell promoters (e.g.,

INS, IRS2, Pdx1, Alx3, Ppy), cardiac specific promoters (e.g. Myh6 (alpha MHC), MYL2 (MLC-2v), TN13 (cTnl), NPPA (ANF), Slc8a1 (Ncx1)), central nervous system cell promoters (SYN1, GFAP, INA, NES, MOBP, MBP, TH, FOXA2 (HNF3 beta)), skin cell specific promoters (e.g., FLG, K14, TGM3), immune cell specific promoters, (e.g., ITGAM, CD43 promoter, CD14 promoter, CD45 promoter, CD68 promoter), urogenital cell specific promoters (e.g., Pbsn, Upk2, Sbp, Fer114), endothelial cell specific promoters (e.g., ENG), pluripotent and embryonic germ layer cell specific promoters (e.g., Oct4, NANOG, Synthetic Oct4, T brachyury, NES, SOX17, FOXA2, MIR122), and muscle cell specific promoter (e.g. myostatin, Desmin). Other tissue and/or cell specific promoters are generally known in the art and are within the scope of this disclosure.

[0089] Inducible/conditional promoters can be positively inducible/conditional promoters (e.g., a promoter that activates transcription of the polynucleotide upon appropriate interaction with an activated activator, or an inducer (compound, environmental condition, or other stimulus) or a negative/conditional inducible promoter (e.g., a promoter that is repressed (e.g., bound by a repressor) until the repressor condition of the promoter is removed (e.g., inducer binds a repressor bound to the promoter stimulating release of the promoter by the repressor or removal of a chemical repressor from the promoter environment). The inducer can be a compound, environmental condition, or other stimulus. Thus, inducible/conditional promoters can be responsive to any suitable stimuli such as chemical, biological, or other molecular agents, temperature, light, and/or pH. Suitable inducible/conditional promoters include, but are not limited to, Tet-On, Tet-Off, Lac promoter, pBad, AlcA, LexA, Hsp70 promoter, Hsp90 promoter, pDawn, XVE/OlexA, GVG, and pOp/LhGR.

[0090] Examples of promoters that are inducible and that can allow for spatiotemporal control of gene editing or gene expression may use a form of energy. The form of energy may include but is not limited to sound energy, electromagnetic radiation, chemical energy and/or thermal energy. Examples of inducible systems include tetracycline inducible promoters (Tet-On or Tet-Off), small molecule two-hybrid transcription activations systems (FKBP, ABA, etc.), or light inducible systems (Phytochrome, LOV domains, or cryptochrome), such as a Light Inducible Transcriptional Effector (LITE) that direct changes in transcriptional activity in a sequence-specific manner. The components of a light inducible system may include one or more elements of the CRISPR-Cas system described herein, a light-responsive cytochrome heterodimer (e.g., from *Arabidopsis thaliana*), and a transcriptional activation/repression domain. In some embodiments, the vector can include one or more of the inducible DNA binding proteins provided in International Patent Publication No. WO 2014/018423 and U.S. Patent Publication Nos., 2015/0291966, 2017/0166903, 2019/0203212, which describe e.g., embodiments of inducible DNA binding proteins and methods of use and can be adapted for use with the present disclosure.

[0091] In some embodiments, transient or inducible expression can be achieved by including, for example, chemical-regulated promoters, i.e., whereby the application of an exogenous chemical induces gene expression. Modulation of gene expression can also be obtained by including a chemical-repressible promoter, where application of the chemical represses gene expression. Chemical-inducible promoters include, but are not limited to, the maize In2-2 promoter, activated by benzene sulfonamide herbicide safeners (De Veylder et al., (1997) *Plant Cell Physiol* 38:568-77), the maize GST promoter

(GST-II-27, WO93/01294), activated by hydrophobic electrophilic compounds used as pre-emergent herbicides, and the tobacco PR-1 a promoter (Ono et al., (2004) Biosci Biotechnol Biochem 68:803-7) activated by salicylic acid. Promoters that are regulated by antibiotics, such as tetracycline-inducible and tetracycline-repressible promoters (Gatz et al., (1991) Mol Gen Genet 227:229-37; U.S. Patent Nos. 5,814,618 and 5,789,156) can also be used herein.

[0092] In some embodiments where multiple elements are to be expressed from the same vector or within the same vector system, different promoters or regulatory elements can be used for each element to be expressed to avoid or limit loss of expression due to competition between promoters and/or other regulatory elements.

[0093] In some embodiments, the polynucleotide, vector or system thereof can include one or more elements capable of translocating and/or expressing a polynucleotide to/in a specific cell component or organelle. Such organelles can include, but are not limited to, nucleus, ribosome, endoplasmic reticulum, Golgi apparatus, chloroplast, mitochondria, vacuole, lysosome, cytoskeleton, plasma membrane, cell wall, peroxisome, centrioles, etc. Such regulatory elements can include, but are not limited to, nuclear localization signals (examples of which are described in greater detail elsewhere herein), any such as those that are annotated in the LocSigDB database (see e.g., genome.unmc.edu/LocSigDB/ and Negi et al., 2015. Database. 2015: bav003; doi: 10.1093/database/bav003), nuclear export signals (e.g., LXXXLXXLXL and others described elsewhere herein), endoplasmic reticulum localization/retention signals (e.g., KDEL, KDXX, KKXX, KXX, and others described elsewhere herein; and see e.g., Liu et al. 2007 Mol. Biol. Cell. 18(3):1073-1082 and Gorleku et al., 2011. J. Biol. Chem. 286:39573-39584), mitochondria targeting signals (see e.g., Chin, R.M., et al, 2018, Cell Reports. 22:2818-2826, particularly at Fig. 2; Doyle et al. 2013. PLoS ONE 8, e67938; Funes et al. 2002. J. Biol. Chem. 277:6051-6058; Matouschek et al. 1997. PNAS USA 85:2091-2095; Oca-Cossio et al., 2003. 165:707-720; Waltner et al., 1996. J. Biol. Chem. 271:21226-21230; Wilcox et al., 2005. PNAS USA 102:15435-15440; Galanis et al., 1991. FEBS Lett 282:425-430), and peroxisome targeting signals (e.g., (S/A/C)-(K/R/H)-(L/A), SLK, (R/K)-(L/V/I)-XXXXX-(H/Q)-(L/A/F)). Suitable protein targeting motifs can also be designed or identified using any suitable database or prediction tool, including but not limited to Minimotif Miner (minimotifminer.org, mitominer.mrc-mbu.cam.ac.uk/release-4.0/embodiment.do?name=Protein%20MTS), LocDB (see above), PTSs predictor, TargetP-2.0 (www.cbs.dtu.dk/services/TargetP/), ChloroP (www.cbs.dtu.dk/services/ChloroP/), NetNES (www.cbs.dtu.dk/services/NetNES/), Predotar (urgi.versailles.inra.fr/predotar/), and SignalP (www.cbs.dtu.dk/services/SignalP/).

Selectable Markers and Tags

[0094] One or more of the polynucleotides described herein, such as those of or encoding the can be operably linked, fused to, or otherwise modified to include a polynucleotide that encodes or is a selectable marker or tag, which can be a polynucleotide or polypeptide. In some embodiments, the polypeptide encoding a polypeptide selectable marker is incorporated in the genetic modifying system polynucleotide or other polynucleotide of the present disclosure such that the selectable marker polypeptide, when translated, is inserted between two amino acids between the N- and C- terminus of the genetic modifying system polypeptide (or other polypeptide of the present disclosure) or at the N-

and/or C-terminus of the genetic modifying system polypeptide (or other polypeptide of the present disclosure). In some embodiments, the selectable marker or tag is a polynucleotide barcode or unique molecular identifier (UMI).

[0095] It will be appreciated that the polynucleotide encoding such selectable markers or tags can be incorporated into a polynucleotide encoding one or more components of the genetic modifying system (or other polynucleotide) described herein in an appropriate manner to allow expression of the selectable marker or tag. Such techniques and methods are described elsewhere herein and will be instantly appreciated by one of ordinary skill in the art in view of this disclosure. Many such selectable markers and tags are generally known in the art and are intended to be within the scope of this disclosure.

[0096] Suitable selectable markers and tags include, but are not limited to, affinity tags, such as chitin binding protein (CBP), maltose binding protein (MBP), glutathione-S-transferase (GST), poly(His) tag; solubilization tags such as thioredoxin (TRX) and poly(NANP), MBP, and GST; chromatography tags such as those consisting of polyanionic amino acids, such as FLAG-tag; epitope tags such as V5-tag, Myc-tag, HA-tag and NE-tag; protein tags that can allow specific enzymatic modification (such as biotinylation by biotin ligase) or chemical modification (such as reaction with FIAsh-EDT2 for fluorescence imaging), DNA and/or RNA segments that contain restriction enzyme or other enzyme cleavage sites; DNA segments that encode products that provide resistance against otherwise toxic compounds including antibiotics, such as, spectinomycin, ampicillin, kanamycin, tetracycline, Basta, neomycin phosphotransferase II (NEO), hygromycin phosphotransferase (HPT)) and the like; DNA and/or RNA segments that encode products that are otherwise lacking in the recipient cell (e.g., tRNA genes, auxotrophic markers); DNA and/or RNA segments that encode products which can be readily identified (e.g., phenotypic markers such as β -galactosidase, GUS; fluorescent proteins such as green fluorescent protein (GFP), cyan (CFP), yellow (YFP), red (RFP), luciferase, and cell surface proteins); polynucleotides that can generate one or more new primer sites for PCR (e.g., the juxtaposition of two DNA sequences not previously juxtaposed), DNA sequences not acted upon or acted upon by a restriction endonuclease or other DNA modifying enzyme, chemical, etc.; epitope tags (e.g., GFP, FLAG- and His-tags), and, DNA sequences that make a molecular barcode or unique molecular identifier (UMI), DNA sequences required for a specific modification (e.g., methylation) that allows its identification. Other suitable markers will be appreciated by those of skill in the art.

[0097] Selectable markers and tags can be operably linked to one or more components of the genetic modifying system (or other polypeptide) described herein via suitable linker, such as a glycine or glycine serine linkers as short as GS or GG up to (GGGGG)₃ or (GGGGS)₃. Other suitable linkers are described elsewhere herein.

Exemplary Heterologous Genes and/or Gene Products

[0098] As described elsewhere herein the engineered phages contain one or more heterologous genes that encode one or more heterologous gene products (e.g., nucleic acids (e.g., RNA) or proteins) (collectively cargo or cargos) to be produced by the host cell after infection of a host cell by the engineered phage. Cargos include, but are not necessarily limited to, biologically active agents, including, but not limited to, therapeutic gene products, imaging gene products, and monitoring gene

products. A cargo may be an exogenous (i.e., not native to a cell of the subject to which the cargo is delivered after being produced by a host cell) or an endogenous, (i.e., native to a cell of the subject to which the cargo is delivered after being produced by a host cell). One or more or two or more different cargoes may be delivered by the engineered phages described herein.

Polynucleotides

[0099] In some embodiments, the cargo is a cargo polynucleotide. As used herein, “nucleic acid,” “nucleotide sequence,” and “polynucleotide” can be used interchangeably herein and can generally refer to a string of at least two base-sugar-phosphate combinations and refers to, among others, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, 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 as used herein can refer to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions can be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide. “Polynucleotide” and “nucleic acids” also encompasses such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including simple and complex cells, inter alia. For instance, the term polynucleotide as used herein can include DNAs or RNAs as described herein that contain one or more modified bases. Thus, DNAs or RNAs including unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein. “Polynucleotide”, “nucleotide sequences” and “nucleic acids” also includes PNAs (peptide nucleic acids), phosphorothioates, and other variants of the phosphate backbone of native nucleic acids. Natural nucleic acids have a phosphate backbone, artificial nucleic acids can contain other types of backbones, but contain the same bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are “nucleic acids” or “polynucleotides” as that term is intended herein. As used herein, “nucleic acid sequence” and “oligonucleotide” also encompasses a nucleic acid and polynucleotide as defined elsewhere herein.

[0100] As used herein, “deoxyribonucleic acid (DNA)” and “ribonucleic acid (RNA)” can generally refer to any polyribonucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. RNA can be in the form of non-coding RNA, including but not limited to, tRNA (transfer RNA), snRNA (small nuclear RNA), rRNA (ribosomal RNA), anti-sense RNA, RNAi (RNA interference construct), siRNA (short interfering RNA), microRNA (miRNA), or ribozymes, aptamers, guide RNA (gRNA), or coding mRNA (messenger RNA).

[0101] In some embodiments, the cargo polynucleotide is DNA. In some embodiments, the cargo polynucleotide is RNA. In some embodiments, the cargo polynucleotide is a polynucleotide (a DNA or an RNA) that encodes an RNA and/or a polypeptide. As used herein with reference to the relationship between DNA, cDNA, cRNA, RNA, protein/peptides, and the like “corresponding to” or “encoding” (used interchangeably herein) refers to the underlying biological relationship between these different molecules. As such, one of skill in the art would understand that operatively “corresponding to” can

direct them to determine the possible underlying and/or resulting sequences of other molecules given the sequence of any other molecule which has a similar biological relationship with these molecules. For example, from a DNA sequence an RNA sequence can be determined and from an RNA sequence a cDNA sequence can be determined.

Polynucleotide Modifications

[0102] In some embodiments, the cargo polynucleotides include one or more modifications capable of modifying the e.g., functionality, packaging ability, stability, degradation localization, increase expression lifetime, resistance to degradation, or any combination thereof, of the at least one or more cargo polynucleotides. Modifications can be sequence modifications (e.g., mutations), chemical modifications, or other modifications, such as complexing to a lipid, polymer, etc. In some embodiments, the cargo polynucleotide is modified to protect it against degradation, by e.g., nucleases or otherwise prevent its degradation.

[0103] In some embodiments, one or more polynucleotides in the engineered polynucleotide are modified. In some embodiments, the engineered polynucleotide includes one or more non-naturally occurring nucleotides, which can be the result of modifying a naturally occurring nucleotide. In some embodiments, the modification is selected independently for each polynucleotide modified. In some embodiments, the modification(s) increase or decrease the stability of the polynucleotide, reduce the immunogenicity of the polynucleotide, increase or decrease the rate of transcription and/or translation, or any combination thereof. Non-naturally occurring nucleotides and/or nucleotide analogs may be modified at the ribose, phosphate, and/or base moiety.

[0104] Suitable modifications include, without limitation, methylpseudouridine, a phosphorothioate linkage, a locked nucleic acid (LNA) nucleotides comprising a methylene bridge between the 2' and 4' carbons of the ribose ring, or bridged nucleic acids (BNA), 2'-O-methyl analogs, 2'-deoxy analogs, or 2'-fluoro analogs, 2-aminopurine, 5-bromo-uridine, pseudouridine, (Ψ), N1-methylpseudouridine (me1 Ψ), 5-methoxyuridine(5moU), inosine, 7-methylguanosine, inosine, 7-methylguanosine. Examples of RNA, including but not limited to guide RNA, chemical modifications include, without limitation, incorporation of 2'-O-methyl (M), 2'-O-methyl 3'phosphorothioate (MS), S-constrained ethyl(cEt), or 2'-O-methyl 3'thioPACE (MSP) at one or more terminal nucleotides.

[0105] In some embodiments, the polynucleotide (DNA and/or RNA) is modified with a 5'- and/or 3'-cap structure. In some embodiments, the 5' cap structure is cap0, cap1, ARCA, inosine, N1-methyl-guanosine, 2'-fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, or 2-azido-guanosine. In some embodiments, the 5'terminal cap is 7mG(5')ppp(5')NmpNp, m7GpppG cap, N⁷-methylguanine. In some embodiments, the 3'terminal cap is a 3'-O-methyl-m7GpppG, 2'Fluoro bases, inverted dT and dTTs, phosphorylation of the 3' end nucleotide, a C3 spacer. Exemplary 5'- and/or 3' that protect against degradation are described in e.g., Gagliardi and Dziembowski. Philosophical transactions of the Royal Society B. 2018. 313(1762). <https://doi.org/10.1098/rstb.2018.0160>; Boo and Kim. Experimental & Molecular Medicine volume 52, pages 400–408 (2020); and Adachai et al., 2021. Biomedicines 2021, 9, 550. <https://doi.org/10.3390/biomedicines9050550>.

[0106] In some embodiments, the 5'-UTR comprises a Kozak sequence.

[0107] In some embodiments, the polynucleotide can be modified with a tailing sequence may range from absent to 500 nucleotides in length (e.g., at least 60, 70, 80, 90, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, or 500 nucleotides). In some embodiments, the tailing region is or includes a polyA tail. Where the tailing region is a polyA tail, the length may be determined in units of or as a function of polyA Binding Protein binding. In this embodiment, the polyA tail is long enough to bind at least 4 monomers of PolyA Binding Protein. PolyA Binding Protein monomers bind to stretches of approximately 38 nucleotides. As such, it has been observed that polyA tails of about 80 nucleotides and 160 nucleotides are functional. In some embodiments, the poly-A tail is at least 160 nucleotides in length.

[0108] In some embodiments, about 10%, 15%, 20%, 24%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, to/or about 100% of the uracils in of a polynucleotide of the present disclosure have a chemical modification, In some embodiments, about 10%, 15%, 20%, 24%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, to/or about 100% of the uracils of a polynucleotide of the present disclosure have a N1-methyl pseudouridine in the 5-position of the uracil.

[0109] In some embodiments, the polynucleotide, optionally an RNA (e.g., an mRNA) includes a stabilization element. In some embodiments, the stabilization element is a histone stem-loop. In some embodiments, the stabilization element is a nucleic acid sequence having increased GC content relative to wild type sequence.

[0110] In some embodiments, a polynucleotide of the present disclosure includes a sequence encoding a self-cleaving peptide. The self-cleaving peptide may be, but is not limited to, a 2A peptide. In one embodiment, this sequence may be used to separate the coding regions of two or more polypeptides.

[0111] In some embodiments, the polynucleotides (e.g., mRNAs) are linear. In yet another embodiment, the polynucleotides of the present disclosure that are circular are known as "circular polynucleotides" or "circP." As used herein, "circular polynucleotides" or "circP" means a single stranded circular polynucleotide which acts substantially like, and has the properties of, an R A. The term "circular" is also meant to encompass any secondary or tertiary configuration of the circP.

[0112] Other RNA modifications, such as mRNA modifications, that can be incorporated into a polynucleotide of the present disclosure include, but are not limited to, any one or more of those described e.g., U.S. Pat. 8,278,036, 8,691,966, 8,748,089, 9,750,824, 10,232,055, 10,703,789, 10,702,600, 10,577,403, 10,442,756, 10,266,485, 10,064,959, 9,868,692, 10,064,959, 10,272,150; U.S. Publications, US20130197068, US20170043037, US20130261172, US20200030460, US20150038558, US20190274968, US20180303925, US20200276300; International Patent Application Publication Nos. WO/2018/081638A1, WO/2016/176330A1, which are incorporated herein by reference and can be adapted for use with the present disclosure.

Signaling and Localization Molecules

[0113] In some embodiments, the cargo polynucleotide includes a signaling and/or localization molecule (e.g., a polynucleotide that is a signaling or localization molecule or a polynucleotide that encodes a signaling or localization peptide or polypeptide).

[0114] In some embodiments, the signaling or localization molecule directs a function (e.g., secretion, folding, etc.) and/or trafficking to a particular location within a cell (e.g., nucleus, Golgi, lysosome, peroxisome, cytoplasm, membrane, chloroplast, vacuole, mitochondria, etc.). In some embodiments, the signaling or localization molecule(s) is/are positioned at the 3' and/or 5' end of a polynucleotide of the present disclosure, such as a cargo polynucleotide. In some embodiments, the signaling or localization molecule(s) is/are located at one or more positions between the 3' and 5' end of a polynucleotide of the present disclosure. In some embodiments, the signaling or localization molecule(s) are located at the 3' and/or 5' end of a polynucleotide of the present disclosure and at one or more positions between the 3' and 5' end of a polynucleotide of the present disclosure. In some embodiments, the signaling and/or localization molecule(s) is/are incorporated in a polynucleotide, such as a cargo polynucleotide, such that it is at the C-terminus, N-terminus, or one or more positions between the C-terminus and N-terminus of a polypeptide encoded by the polynucleotide.

[0115] In some embodiments, a polynucleotide of the present disclosure includes a polynucleotide sequence that is or encodes one or more signal peptides, leucine rich repeat (LRR) sequences, nuclear localization signals, a Type IX secretion system (T9SS) substrate, secretion signal peptide, an amino acid sequence capable of directing clearance from a cell or organism, an Fc receptor directing binding to a dendritic cell, and/or directing antigen processing, an F-box domain or polypeptide, a subcellular localization sequence, a TOM70, TOM20, or TOM22 binding polypeptide, a stromal import sequence, a thylakoid targeting sequence, a peroxisome targeting signal 1 sequence, a peroxisome targeting signal 2 sequence, an endoplasmic reticulum signaling sequence.

[0116] Exemplary nuclear localization molecules are described in e.g., Lu et al., *Cell Communication and Signaling*. 2021. 19(60): 1-10 (particularly at Table 1 therein), which can be adapted for use with the present disclosure. Exemplary signal peptides are described in e.g., Owji et al., *European J Cell Biol*. 2018. 97(6):422-441, which can be adapted for use with the present disclosure. Exemplary peroxisome targeting sequences are described in e.g., Baerends et al., 2000. *FEMS Microbiol Rev*. 24(3): 291-301, which can be adapted for use with the present disclosure. Exemplary endoplasmic reticulum signaling molecules are described in e.g., Walter et al., *J Cell Biol*. 1981. 91(2 Pt. 1):545-50 doi:10.1083/jcb.91.2.545, which can be adapted for use with the present disclosure. Exemplary lysosomal and endosomal signaling molecules are described in e.g., Bonifacino and Traub. 2003. *Ann. Rev. Biochem*. 72:395-447, which can be adapted for use with the present disclosure. Exemplary endoplasmic reticulum signaling sequences are described in e.g., *J Cell Biol*. 1996 Jul 2; 134(2): 269-278, which can be adapted for use with the present disclosure. Exemplary Golgi signaling sequences are described in e.g., Gleeson et al., 1994. *Glycoconjugat J*. 11:381-394, which can be adapted for use with the present disclosure.

Interference RNAs

[0117] In certain example embodiments, the one or more cargo polynucleotides are or encode one or more interference RNAs. Interference RNAs are RNA molecules capable of suppressing gene expressions. Example types of interference RNAs include small interfering RNA (siRNA), microRNA (miRNA), and short hairpin RNA (shRNA).

[0118] In certain example embodiments, the interference RNA may be a siRNAs. Small interfering RNA (siRNA) molecules are capable of inhibiting target gene expression by interfering RNA. siRNAs may be chemically synthesized, or may be obtained by in vitro transcription, or may be synthesized in vivo in a target cell. siRNAs may comprise double-stranded RNA from 15 to 40 nucleotides in length and can contain a protuberant region 3' and/or 5' from 1 to 6 nucleotides in length. Length of protuberant region is independent from total length of siRNA molecule. siRNAs may act by post-transcriptional degradation or silencing of target messenger. In some cases, the exogenous polynucleotides encode shRNAs. In shRNAs, the antiparallel strands that form siRNA are connected by a loop or hairpin region.

[0119] The interference RNA (e.g., siRNA) may suppress expression of genes to promote long term survival and functionality of cells after transplanted to a subject. In some examples, the interference RNAs suppress genes in TGF β pathway, e.g., TGF β , TGF β receptors, and SMAD proteins. In some examples, the interference RNAs suppress genes in colony-stimulating factor 1 (CSF1) pathway, e.g., CSF1 and CSF1 receptors. In certain embodiments, the one or more interference RNAs suppress genes in both the CSF1 pathway and the TGF β pathway. TGF β pathway genes may comprise one or more of ACVR1, ACVR1C, ACVR2A, ACVR2B, ACVRL1, AMH, AMHR2, BMP2, BMP4, BMP5, BMP6, BMP7, BMP8A, BMP8B, BMPR1A, BMPR1B, BMPR2, CDKN2B, CHRD, COMP, CREBBP, CUL1, DCN, E2F4, E2F5, EP300, FST, GDF5, GDF6, GDF7, ID1, ID2, ID3, ID4, IFNG, INHBA, INHBB, INHBC, INHBE, LEFTY1, LEFTY2, LOC728622, LTBP1, MAPK1, MAPK3, MYC, NODAL, NOG, PITX2, PPP2CA, PPP2CB, PPP2R1A, PPP2R1B, RBL1, RBL2, RBX1, RHOA, ROCK1, ROCK2, RPS6KB1, RPS6KB2, SKP1, SMAD1, SMAD2, SMAD3, SMAD4, SMAD5, SMAD6, SMAD7, SMAD9, SMURF1, SMURF2, SP1, TFDP1, TGFB1, TGFB2, TGFB3, TGFB1, TGFB2, THBS1, THBS2, THBS3, THBS4, TNF, ZFYVE16, and/or ZFYVE9.

[0120] In some embodiments, the cargo polynucleotide is an RNAi molecule, antisense molecule, and/or a gene silencing oligonucleotide or a polynucleotide that encodes an RNAi molecule, antisense molecule, and/or gene silencing oligonucleotide.

[0121] As used herein, "gene silencing oligonucleotide" refers to any oligonucleotide that can alone or with other gene silencing oligonucleotides utilize a cell's endogenous mechanisms, molecules, proteins, enzymes, and/or other cell machinery or exogenous molecule, agent, protein, enzyme, and/or polynucleotide to cause a global or specific reduction or elimination in gene expression, RNA level(s), RNA translation, RNA transcription, that can lead to a reduction or effective loss of a protein expression and/or function of a non-coding RNA as compared to wild-type or a suitable control. This is synonymous with the phrase "gene knockdown" Reduction in gene expression, RNA level(s), RNA translation, RNA transcription, and/or protein expression can range from about 100, 99, 98, 97, 96, 95, 94, 93, 92, 91, 90, 89, 88, 87, 86, 85, 84, 83, 82, 81, 80, 79, 78, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, to 1% or less reduction. "Gene silencing oligonucleotides" include, but are not limited to, any antisense oligonucleotide, ribozyme, any oligonucleotide (single or double stranded) used to stimulate the RNA interference (RNAi) pathway in a cell (collectively RNAi oligonucleotides), small interfering RNA (siRNA), microRNA, and short-hairpin RNA (shRNA). Commercially available

programs and tools are available to design the nucleotide sequence of gene silencing oligonucleotides for a desired gene, based on the gene sequence and other information available to one of ordinary skill in the art.

Therapeutic Polynucleotides

[0122] In some embodiments, the cargo molecule is a therapeutic polynucleotide. Therapeutic polynucleotides are those that provide a therapeutic effect when delivered to a recipient cell. The polynucleotide can be a toxic polynucleotide (a polynucleotide that when transcribed or translated results in the death of the cell) or polynucleotide that encodes a lytic peptide or protein. In embodiments, delivery vesicles having a toxic polynucleotide as a cargo molecule can act as an antimicrobial or antibiotic. This is discussed in greater detail elsewhere herein. In some embodiments, the cargo molecule can be exogenous to the producer cell and/or a first cell. In some embodiments, the cargo molecule can be endogenous to the producer cell and/or a first cell. In some embodiments, the cargo molecule can be exogenous to the recipient cell and/or a second cell. In some embodiments, the cargo molecule can be endogenous to the recipient cell and/or second cell.

[0123] As described herein the cargo polynucleotide can be any polynucleotide endogenous or exogenous to the eukaryotic cell. For example, the cargo polynucleotide can be a polynucleotide residing in the nucleus of the eukaryotic cell. The cargo polynucleotide can be a sequence coding a gene product (e.g., a protein) or a non-coding sequence (e.g., a regulatory polynucleotide).

[0124] In some embodiments, the cargo polynucleotide is a DNA or RNA (e.g., a mRNA) vaccine.

Aptamers

[0125] In certain example embodiments, the cargo polynucleotide is an aptamer. In certain embodiments, the one or more agents is an aptamer. Nucleic acid aptamers are nucleic acid species that have been engineered through repeated rounds of in vitro selection or equivalently, SELEX (systematic evolution of ligands by exponential enrichment) to bind to various molecular targets such as small molecules, proteins, nucleic acids, cells, tissues and organisms. Nucleic acid aptamers have specific binding affinity to molecules through interactions other than classic Watson-Crick base pairing. Aptamers are useful in biotechnological and therapeutic applications as they offer molecular recognition properties similar to antibodies. In addition to their discriminate recognition, aptamers offer advantages over antibodies as they can be engineered completely in a test tube, are readily produced by chemical synthesis, possess desirable storage properties, and elicit little or no immunogenicity in therapeutic applications. In certain embodiments, RNA aptamers may be expressed from a DNA construct. In other embodiments, a nucleic acid aptamer may be linked to another polynucleotide sequence. The polynucleotide sequence may be a double stranded DNA polynucleotide sequence. The aptamer may be covalently linked to one strand of the polynucleotide sequence. The aptamer may be ligated to the polynucleotide sequence. The polynucleotide sequence may be configured, such that the polynucleotide sequence may be or capable of being linked to a solid support or ligated to another polynucleotide sequence.

[0126] Aptamers, like peptides generated by phage display or monoclonal antibodies ("mAbs"), are capable of specifically binding to selected targets and modulating the target's activity, e.g., through binding, aptamers may block their target's ability to function. A typical aptamer is 10-15 kDa in size (30-

45 nucleotides), binds its target with sub-nanomolar affinity, and discriminates against closely related targets (e.g., aptamers will typically not bind other proteins from the same gene family). Structural studies have shown that aptamers are capable of using the same types of binding interactions (e.g., hydrogen bonding, electrostatic complementarity, hydrophobic contacts, steric exclusion) that drives affinity and specificity in antibody-antigen complexes.

[0127] Aptamers have a number of desirable characteristics for use in research and as therapeutics and diagnostics including high specificity and affinity, biological efficacy, and excellent pharmacokinetic properties. In addition, they offer specific competitive advantages over antibodies and other protein biologics. Aptamers are chemically synthesized and are readily scaled as needed to meet production demand for research, diagnostic or therapeutic applications. Aptamers are chemically robust. They are intrinsically adapted to regain activity following exposure to factors such as heat and denaturants and can be stored for extended periods (>1 yr) at room temperature as lyophilized powders. Not being bound by a theory, aptamers bound to a solid support or beads may be stored for extended periods.

[0128] Oligonucleotides in their phosphodiester form may be quickly degraded by intracellular and extracellular enzymes such as endonucleases and exonucleases. Aptamers can include modified nucleotides conferring improved characteristics on the ligand, such as improved in vivo stability or improved delivery characteristics. Examples of such modifications include chemical substitutions at the ribose and/or phosphate and/or base positions. SELEX identified nucleic acid ligands containing modified nucleotides are described, e.g., in U.S. Pat. No. 5,660,985, which describes oligonucleotides containing nucleotide derivatives chemically modified at the 2' position of ribose, 3' of pyrimidines, and 8 position of purines, U.S. Pat. No. 5,756,703 which describes oligonucleotides containing various 2' - modified pyrimidines, and U.S. Pat. No. 5,580,737 which describes highly specific nucleic acid ligands containing one or more nucleotides modified with 2'-amino (2'-NH₂), 2'-fluoro (2'-F), and/or 2'-O-methyl (2'-OMe) substituents. Modifications of aptamers may also include modifications at exocyclic amines, substitution of 4- thiouridine, substitution of 5-bromo or 5-iodo-uracil; backbone modifications, phosphonothioate or allyl phosphate modifications, methylations, and unusual base-pairing combinations such as the isobases isocytidine and isoguanosine. Modifications can also include 3' and 5' modifications such as capping. As used herein, the term phosphonothioate encompasses one or more non-bridging oxygen atoms in a phosphodiester bond replaced by one or more sulfur atoms. In further embodiments, the oligonucleotides comprise modified sugar groups, for example, one or more of the hydroxyl groups is replaced with halogen, aliphatic groups, or functionalized as ethers or amines. In one embodiment, the 2'-position of the furanose residue is substituted by any of an O-methyl, O-alkyl, O-allyl, S-alkyl, S-allyl, or halo of synthesis of 2'-modified sugars are described, e.g., in Sproat, et al., Nucl. Acid Res. 19:733-738 (1991); Cotten, et al, Nucl. Acid Res. 19:2629-2635 (1991); and Hobbs, et al, Biochemistry 12:5138-5145 (1973). Other modifications are known to one of ordinary skill in the art. In certain embodiments, aptamers include aptamers with improved off-rates as described in International Patent Publication No. WO 2009012418, "Method for generating aptamers with improved off-rates," incorporated herein by reference in its entirety. In certain embodiments aptamers are chosen from a library of aptamers. Such libraries include but are not limited to those described in Rohloff et al.,

"Nucleic Acid Ligands with Protein-like Side Chains: Modified Aptamers and Their Use as Diagnostic and Therapeutic Agents," *Molecular Therapy Nucleic Acids* (2014) 3, e201. Aptamers are also commercially available (see, e.g., SomaLogic, Inc., Boulder, Colorado). In certain embodiments, the present disclosure may utilize any aptamer containing any modification as described herein.

[0129] In certain other example embodiments, the polynucleotide may be a ribozyme or other enzymatically active polynucleotide.

Biologically active agents

[0130] In some embodiments, the cargo is a biologically active agent. Biologically active agents include any molecule that induces, directly or indirectly, an effect in a cell. Biologically active agents may be a protein, a nucleic acid, a small molecule, a carbohydrate, a lipid or any combination thereof. Biologically active agents can be therapeutic agents. Exemplary Therapeutic agents include, without limitation, chemotherapeutic agents, anti-oncogenic agents, anti-angiogenic agents, tumor suppressor agents, anti-microbial agents, enzyme replacement agents, gene expression modulating agents and expression constructs comprising a nucleic acid encoding a therapeutic protein or nucleic acid, and vaccines. Therapeutic agents may be peptides, proteins (including enzymes, antibodies and peptidic hormones), ligands of cytoskeleton, nucleic acid, small molecules, non-peptidic hormones and the like. To increase affinity for the nucleus, agents may be conjugated to a nuclear localization sequence. Nucleic acids that may be delivered by the method of the disclosure include synthetic and natural nucleic acid material, including DNA, RNA, transposon DNA, antisense nucleic acids, dsRNA, siRNAs, transcription RNA, messenger RNA, ribosomal RNA, small nucleolar RNA, microRNA, ribozymes, plasmids, expression constructs, etc.

Imaging Agents

[0131] In some embodiments, the cargo is an imaging agent. Such agents when delivered to a cell or cells in the subject can facilitate the imaging of those cell or cells. In some embodiments, the imaging agent is a monitoring agents. Monitoring agents can allow for monitoring of a cells condition or state over time, such as by imaging. The monitoring agent can change or otherwise be responsive to cell state or status (e.g., such as pH, presence of certain cell products, etc.) such that it changes and produces a different signal to be imaged. Imaging and monitoring agents include reporter probes, biosensors, optically active protein (e.g., a fluorescent protein) and the like. In some embodiments, the imaging and/or monitoring agent are biosensors are molecules that detect and transmit information regarding a physiological change or process, for instance, by detecting the presence or change in the presence of a chemical or environment (e.g., pH, osmolarity, etc.). The information obtained by the biosensor typically activates a signal that is detected with a transducer. The transducer typically converts the biological response into an electrical signal. Examples of biosensors include enzymes, antibodies, DNA, receptors and regulator proteins used as recognition elements, which can be used either in whole cells or isolated and used independently (D'Souza, 2001, *Biosensors and Bioelectronics* 16:337-353).

Polynucleotide Modifying Systems

[0132] In some embodiments, the cargo is a polynucleotide modifying system or component(s) thereof. In some embodiments the polynucleotide modifying system is a gene modifying system. In

some embodiments, the gene modifying system is or is composed of a gene (or genetic) modulating agent. In some embodiments, the genetic modulating agent may comprise one or more polypeptide and/or polynucleotide components of a polynucleotide modification system (e.g., a gene editing system) and/or polynucleotides encoding thereof.

[0133] In some embodiments, the gene editing system may be an RNA-guided system or other programmable nuclease system. In some embodiments, the gene editing system is an IscB system. In some embodiments, the gene editing system may be a CRISPR-Cas system. In some embodiments, the polynucleotide modifying system is a recombinase, zinc finger nuclease, or TALEN. Other suitable polynucleotide modifying systems and/or components thereof that can be included as cargos will be appreciated by one of ordinary skill in the art in view of the description herein.

CRISPR-Cas Systems

[0134] In general, a CRISPR-Cas or CRISPR system as used in herein and in documents, such as WO 2014/093622 (PCT/US2013/074667), refers collectively to transcripts and other elements involved in the expression of or directing the activity of CRISPR-associated ("Cas") genes, including sequences encoding a Cas gene, a tracr (trans-activating CRISPR) sequence (e.g. tracrRNA or an active partial tracrRNA), a tracr-mate sequence (encompassing a "direct repeat" and a tracrRNA-processed partial direct repeat in the context of an endogenous CRISPR system), a guide sequence (also referred to as a "spacer" in the context of an endogenous CRISPR system), or "RNA(s)" as that term is herein used (e.g., RNA(s) to guide Cas, such as Cas9, e.g. CRISPR RNA and transactivating (tracr) RNA or a single guide RNA (sgRNA) (chimeric RNA)) or other sequences and transcripts from a CRISPR locus. In general, a CRISPR system is characterized by elements that promote the formation of a CRISPR complex at the site of a target sequence (also referred to as a protospacer in the context of an endogenous CRISPR system). See, e.g., Shmakov et al. (2015) "Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems", *Molecular Cell*, DOI: dx.doi.org/10.1016/j.molcel.2015.10.008.

Specialized Cas-based Systems

[0135] In some embodiments, the system is a Cas-based system that is capable of performing a specialized function or activity. For example, the Cas protein may be fused, operably coupled to, or otherwise associated with one or more functional domains. In certain example embodiments, the Cas protein may be a catalytically dead Cas protein ("dCas") and/or have nickase activity. A nickase is a Cas protein that cuts only one strand of a double stranded target. In such embodiments, the dCas or nickase provide a sequence specific targeting functionality that delivers the functional domain to or proximate a target sequence. Example functional domains that may be fused to, operably coupled to, or otherwise associated with a Cas protein can be or include, but are not limited to a nuclear localization signal (NLS) domain, a nuclear export signal (NES) domain, a translational activation domain, a transcriptional activation domain (e.g. VP64, p65, MyoD1, HSF1, RTA, and SET7/9), a translation initiation domain, a transcriptional repression domain (e.g., a KRAB domain, NuE domain, NcoR domain, and a SID domain such as a SID4X domain), a nuclease domain (e.g., FokI), a histone modification domain (e.g., a histone acetyltransferase), a light inducible/controllable domain, a chemically inducible/controllable domain, a transposase domain, a homologous recombination

machinery domain, a recombinase domain, an integrase domain, and combinations thereof. Methods for generating catalytically dead Cas9 or a nickase Cas9 (WO 2014/204725, Ran et al. Cell. 2013 Sept 12; 154(6):1380-1389), Cas12 (Liu et al. Nature Communications, 8, 2095 (2017), and Cas13 (International Patent Publication Nos. WO 2019/005884 and WO2019/060746) are known in the art and incorporated herein by reference.

[0136] In some embodiments, the functional domains can have one or more of the following activities: methylase activity, demethylase activity, translation activation activity, translation initiation activity, translation repression activity, transcription activation activity, transcription repression activity, transcription release factor activity, histone modification activity, nuclease activity, single-strand RNA cleavage activity, double-strand RNA cleavage activity, single-strand DNA cleavage activity, double-strand DNA cleavage activity, molecular switch activity, chemical inducibility, light inducibility, and nucleic acid binding activity. In some embodiments, the one or more functional domains may comprise epitope tags or reporters. Non-limiting examples of epitope tags include histidine (His) tags, V5 tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Examples of reporters include, but are not limited to, glutathione-S-transferase (GST), horseradish peroxidase (HRP), chloramphenicol acetyltransferase (CAT) beta-galactosidase, beta-glucuronidase, luciferase, green fluorescent protein (GFP), HcRed, DsRed, cyan fluorescent protein (CFP), yellow fluorescent protein (YFP), and auto-fluorescent proteins including blue fluorescent protein (BFP).

[0137] The one or more functional domain(s) may be positioned at, near, and/or in proximity to a terminus of the effector protein (e.g., a Cas protein). In embodiments having two or more functional domains, each of the two can be positioned at or near or in proximity to a terminus of the effector protein (e.g., a Cas protein). In some embodiments, such as those where the functional domain is operably coupled to the effector protein, the one or more functional domains can be tethered or linked via a suitable linker (including, but not limited to, GlySer linkers) to the effector protein (e.g., a Cas protein). When there is more than one functional domain, the functional domains can be same or different. In some embodiments, all the functional domains are the same. In some embodiments, all of the functional domains are different from each other. In some embodiments, at least two of the functional domains are different from each other. In some embodiments, at least two of the functional domains are the same as each other.

[0138] Other suitable functional domains can be found, for example, in International Patent Publication No. WO 2019/018423.

Split CRISPR-Cas systems

[0139] In some embodiments, the CRISPR-Cas system is a split CRISPR-Cas system. See e.g., Zetche et al., 2015. Nat. Biotechnol. 33(2): 139-142 and International Patent Publication WO 2019/018423, the compositions and techniques of which can be used in and/or adapted for use with the present disclosure. Split CRISPR-Cas proteins are set forth herein and in documents incorporated herein by reference in further detail herein. In certain embodiments, each part of a split CRISPR protein is attached to a member of a specific binding pair, and when bound with each other, the members of the specific binding pair maintain the parts of the CRISPR protein in proximity. In certain embodiments, each part of a split CRISPR protein is associated with an inducible binding pair. An inducible binding

pair is one which is capable of being switched “on” or “off” by a protein or small molecule that binds to both members of the inducible binding pair. In some embodiments, CRISPR proteins may preferably split between domains, leaving domains intact. In particular embodiments, said Cas split domains (e.g., RuvC and HNH domains in the case of Cas9) can be simultaneously or sequentially introduced into the cell such that said split Cas domain(s) process the target nucleic acid sequence in the algae cell. The reduced size of the split Cas compared to the wild-type Cas allows other methods of delivery of the systems to the cells, such as the use of cell penetrating peptides as described herein.

DNA and RNA Base Editing

[0140] In some embodiments, the cargo is a base editor system or component thereof, such as Cas-deaminase. In some embodiments, a Cas protein is connected or fused to a nucleotide deaminase. Thus, in some embodiments the Cas-based system can be a base editing system. As used herein, “base editing” refers generally to the process of polynucleotide modification via a CRISPR-Cas-based or Cas-based system that does not include excising nucleotides to make the modification. Base editing can convert base pairs at precise locations without generating excess undesired editing byproducts that can be made using traditional CRISPR-Cas systems.

[0141] In certain example embodiments, the nucleotide deaminase may be a DNA base editor used in combination with a DNA binding Cas protein such as, but not limited to, Class 2 Type II and Type V systems. Two classes of DNA base editors are generally known: cytosine base editors (CBEs) and adenine base editors (ABEs). CBEs convert a C•G base pair into a T•A base pair (Komor et al. 2016. *Nature*. 533:420-424; Nishida et al. 2016. *Science*. 353; and Li et al. *Nat. Biotech.* 36:324-327) and ABEs convert an A•T base pair to a G•C base pair. Collectively, CBEs and ABEs can mediate all four possible transition mutations (C to T, A to G, T to C, and G to A). Rees and Liu. 2018. *Nat. Rev. Genet.* 19(12): 770-788, particularly at Figures 1b, 2a-2c, 3a-3f, and Table 1. In some embodiments, the base editing system includes a CBE and/or an ABE. In some embodiments, a polynucleotide of the present disclosure described elsewhere herein can be modified using a base editing system. Rees and Liu. 2018. *Nat. Rev. Genet.* 19(12):770-788. Base editors also generally do not need a DNA donor template and/or rely on homology-directed repair. Komor et al. 2016. *Nature*. 533:420-424; Nishida et al. 2016. *Science*. 353; and Gaudeli et al. 2017. *Nature*. 551:464-471. Upon binding to a target locus in the DNA, base pairing between the guide RNA of the system and the target DNA strand leads to displacement of a small segment of ssDNA in an “R-loop”. Nishimasu et al. *Cell*. 156:935-949. DNA bases within the ssDNA bubble are modified by the enzyme component, such as a deaminase. In some systems, the catalytically disabled Cas protein can be a variant or modified Cas can have nickase functionality and can generate a nick in the non-edited DNA strand to induce cells to repair the non-edited strand using the edited strand as a template. Komor et al. 2016. *Nature*. 533:420-424; Nishida et al. 2016. *Science*. 353; and Gaudeli et al. 2017. *Nature*. 551:464-471.

[0142] Other Example Type V base editing systems are described in International Patent Publication Nos. WO 2018/213708, WO 2018/213726, and International Patent Applications No. PCT/US2018/067207, PCT/US2018/067225, and PCT/US2018/067307, each of which is incorporated herein by reference.

[0143] In certain example embodiments, the base editing system may be an RNA base editing system. As with DNA base editors, a nucleotide deaminase capable of converting nucleotide bases may be fused to a Cas protein. However, in these embodiments, the Cas protein will need to be capable of binding RNA. Example RNA binding Cas proteins include, but are not limited to, RNA-binding Cas9s such as *Francisella novicida* Cas9 ("FnCas9"), and Class 2 Type VI Cas systems. The nucleotide deaminase may be a cytidine deaminase or an adenosine deaminase, or an adenosine deaminase engineered to have cytidine deaminase activity. In certain example embodiments, the RNA base editor may be used to delete or introduce a post-translation modification site in the expressed mRNA. In contrast to DNA base editors, whose edits are permanent in the modified cell, RNA base editors can provide edits where finer, temporal control may be needed, for example in modulating a particular immune response. Example Type VI RNA-base editing systems are described in Cox et al. 2017. Science 358: 1019-1027, International Patent Publication Nos. WO 2019/005884, WO 2019/005886, and WO 2019/071048, and International Patent Application Nos. PCT/US20018/05179 and PCT/US2018/067207, which are incorporated herein by reference. An example FnCas9 system that may be adapted for RNA base editing purposes is described in International Patent Publication No. WO 2016/106236, which is incorporated herein by reference.

[0144] An example method for delivery of base-editing systems, including use of a split-intein approach to divide CBE and ABE into reconstitutable halves, is described in Levy et al. Nature Biomedical Engineering doi.org/10.1038/s41441-019-0505-5 (2019), which is incorporated herein by reference.

Prime Editors

[0145] In one example embodiment, the cargo may be prime editing systems or one or more polynucleotides encoding one or more components thereof. Prime editing systems comprise a programmable nuclease (e.g., Cas), most often a nickase, linked to a reverse transcriptase domain and a guide molecule (prime editing guide pegRNA), which comprises a target-specific spacer, a primer binding site, and RT template. See e.g., Anzalone et al. 2019. Nature. 576: 149-157; and International Patent Application Publication No. WO2022150790A2. In some embodiments, the prime editing guide molecule can specify both the target polynucleotide information (e.g., sequence) and contain a new polynucleotide cargo that replaces target polynucleotides. To initiate transfer from the guide molecule to the target polynucleotide, the PE system can nick the target polynucleotide at a target site to expose a 3'hydroxyl group, which can prime reverse transcription of an edit-encoding extension region of the guide molecule (e.g., a prime editing guide molecule or peg guide molecule) directly into the target site in the target polynucleotide. See e.g., Anzalone et al. 2019. Nature. 576: 149-157, particularly at Figures 1b, 1c, related discussion, and Supplementary discussion.

[0146] Prime editing systems can also be used in tandem such that, the two pegRNAs template the synthesis of complementary DNA flaps on opposing strands of genomic DNA, which replace the endogenous DNA sequence between the PE-induced nick sites. See, e.g., WO 2021/138469; Anzalone AV, Gao XD, Podracky CJ, et al. Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing. *Nat Biotechnol.* 2022;40(5):731-740. Thus, use of two pegRNAs allows for larger insertions or deletions because of the two overlapping 3' flaps created by

the two nicked sites. In one example embodiment, the system can be used to insert or replace a sequence into one or more target genes. In example embodiments, the insertion or replacement results in an inactive target gene or less active form of the target gene. In one example embodiment, the system is used to replace all or a portion of the entire target gene. In one example embodiment, the system is used to replace all or a portion of an enhancer controlling the target gene expression.

Recombinase-mediated Modifications

[0147] Prime editing and twinPE systems can also be further combined with site-specific recombinases, such as integrases, to facilitate even larger insertions, substitutions and deletions. See e.g., WO 2021/138469; Anzalone AV, Gao XD, Podracky CJ, et al. Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing. *Nat Biotechnol.* 2022;40(5):731-740; Yarnall et al., *Nat Biotechnol* (2022). doi.org/10.1038/s41587-022-01527-4, which is incorporated by reference as if expressed in its entirety herein. Such systems or component(s) thereof can be included as cargo of the engineered phages of the present disclosure. The prime editing system is used to insert a recombinase recognition site at the desire site of modification and an integrase facilitates the insertion of a donor sequence from a donor template. "Uni-directional recombinases" or "integrases" refer to recombinase enzymes whose recognition sites are destroyed after the recombination has taken place. The term "integrase" refers to a type of recombinase. In other words, the sequence recognized by the recombinase is changed into one that is not recognized by the recombinase upon recombination. As a result, once a sequence is subjected to recombination by the uni-directional recombinase, the continued presence of the recombinase cannot reverse the previous recombination event.

[0148] Typically, two different sites are involved (in regard to recombination termed "complementary sites"), one present in the target nucleic acid (e.g., a chromosome or episome of a eukaryote) and another on the nucleic acid that is to be integrated at the target recombination site. The terms "attB" and "attP," which refer to attachment (or recombination) sites originally from a bacterial target (attachment site of bacteria) and a phage donor (attachment site of phage), respectively, are used herein although recombination sites for particular enzymes may have different names. The two attachment sites can share as little sequence identity as a few base pairs. The recombination sites typically include left and right arms separated by a core or spacer region. Thus, an attB recombination site consists of BOB', where B and B' are the left and right arms, respectively, and O is the core region. Similarly, attP is POP', where P and P' are the arms and O is again the core region. Upon recombination between the attB and attP sites, and concomitant integration of a nucleic acid at the target, the recombination sites that flank the integrated DNA are referred to as "attL" and "attR." The attL and attR sites, using the terminology above, thus consist of BOP' and POB', respectively. In some representations herein, the "O" is omitted and attB and attP, for example, are designated as BB' and PP', respectively.

[0149] In example embodiments, the recombinase of the present disclosure is a serine integrase. In example embodiments, serine integrases specifically recombine when recognizing the two attachment sites specific for the integrase. In example embodiments, the heterologous sites are referred to as attP and attB, however, these terms refer to the specific sequences recognized by the specific

integrase and do not refer to a single consensus sequence. Serine integrases mediate site-specific recombination between short recognition sites located in phage genomes and bacterial chromosomes, respectively, the attachment site of phage (attP) and attachment site of bacteria (attB) (i.e., the target sites of the integrase), to form the hybrid attachment sites attL and attR. Unlike Cre and Flp recombinases that catalyze reversible site-specific recombination reactions, serine integrases are unidirectional and catalyze only attP and attB recombination without RDF or Xis accessory proteins. Thus, in the absence of any accessory factors integrase is unidirectional. In addition, DNA substrates identified by serine integrases (attP and attB) are relatively short (30–50 bp) and have a minimal length of approximately 34–40 base pairs (bp) (Groth AC et al., *Proc. Natl. Acad. Sci. USA* 97, 5995–6000 (2000)). The compatibility of distinct DNA topological structures is also quite different from recognition of DNA by Hin recombinase or Tn3 resolvase. Serine integrases recognize DNA substrates specifically, not at random, but can facilitate recombination at sequences with partial identity with wild-type recombination sites, termed pseudo attachment sites (either pseudo attP or pseudo attB). A "pseudo-recombination site" is a DNA sequence recognized by a recombinase enzyme such that the recognition site differs in one or more base pairs from the wild-type recombinase recognition sequence and/or is present as an endogenous sequence in a genome that differs from the genome where the wild-type recognition sequence for the recombinase resides. "Pseudo attP site" or "pseudo attB site" refer to pseudo sites that are similar to wild-type phage or bacterial attachment site sequences, respectively, for phage integrase enzymes. "Pseudo att site" is a more general term that can refer to either a pseudo attP site or a pseudo attB site. Specific attB and attP sequences for use in the present disclosure include all wildtype sequences as well as pseudo attB and attP sequences.

[0150] Recombination sites used in the present methods include those recognized by unidirectional, site-directed recombinases (e.g., integrases). Non-limiting examples of serine integrases and recombination sites applicable to the present disclosure include ϕ C31 integrase, Bxb1, ϕ BT1 integrase, A118, TP901-1, and R4 and the corresponding recombination sites for each (see, e.g., Groth, A. C. and Calos, M. P. (2004) *J. Mol. Biol.* 335, 667–678; Lei, et al., *FEBS Lett.* 2018 Apr;592(8):1389–1399; Singh, et al., *Attachment Site Selection and Identity in Bxb1 Serine Integrase-Mediated Site-Specific Recombination*, *PLoS Genet.* 2013 May;9(5):e1003490; and Gupta, et al., *Nucleic Acids Res.* 2007 May; 35(10): 3407–3419). Additional serine recombinases and recombination sites may be any of those disclosed in US 20180346934A1 and US 2010/0190178. In certain embodiments, a functional domain of the serine integrase is used.

[0151] In one example embodiment, the system can be used to insert or replace a sequence into one or more target genes. In example embodiments, the insertion or replacement results in an inactive target gene or less active form of the target gene. In one example embodiment, the system is used to replace all or a portion of the entire target gene. In one example embodiment, the system is used to replace all or a portion of an enhancer controlling the target gene expression.

[0152] The peg guide molecule can be about 10 to about 200 or more nucleotides in length, such as 10 to/or 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88,

89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, or 200 or more nucleotides in length. Optimization of the peg guide molecule can be accomplished as described in Anzalone et al. 2019. *Nature*. 576: 149-157, particularly at pg. 3, Fig. 2a-2b, and Extended Data Figs. 5a-c.

CRISPR Associated Transposase (CAST) Systems

[0153] In one embodiment, the cargo may be a CAST system or component thereof and/or one or more polynucleotides encoding one or more components thereof. CAST systems are able to insert DNA sequences at a target site in a DNA molecule without relying on host cell repair machinery. CAST systems can be Class 1 or Class 2 CAST systems. For example, a Class 1 system is described in Klompe et al. *Nature*, doi:10.1038/s41586-019-1323, which is incorporated herein by reference. An example Class 2 system is described in Strecker et al. *Science*. 10/1126/science. aax9181 (2019), and PCT/US2019/066835 which are incorporated herein by reference. Suitable hybrid systems have also been described such as those described in Tou et al. *bioRxiv* 2022.01.07.475005, doi.org/10.1101/2022.01.07.475005, which is incorporated herein by reference. Additional CAST systems are disclosed, e.g., in WO 2020/131862 to Zhang et al and WO/2021/257997 to Zhang et al. (CAST-12k); WO 2021/087394 to Zhang et al. and WO/2022/147321 to Zhang et al. (CAST-1b); WO 2022/076820 to Zhang et al. (CAST-1f); WO/2022/150651 to Zhang et al. (minimal Tn7-like CAST systems), all of which are incorporated herein by reference.

[0154] The CAST system may comprise a Cas linked to a transposase subunit to achieve RNA-guided DNA-transposition, optionally linked to a guide molecule. In some embodiments, the Cas may be catalytically inactive (e.g., Type I, IV, or Type V systems). Transposases suitable for the CAST system may be of any variety generally derived Tn7-like transposons (e.g. non-limiting examples, including TnsA, TnsB, TnsC, or TniQ). Guide molecules can guide the catalytically inactive Cas and a Tn7 or Tn7-like subunit to a target site to direct insertion of a donor at the target site.

[0155] CAST systems may require combinatorial transposases for efficient deposition. For example, TnsA is an endonuclease that cleaves the 5'-ends of the transposon and interacts with TnsB, TnsC, and DNA. TnsB is a recombinase capable of cleaving the 3'-end of the transposon. In some CAST systems with these components, the interaction between TnsA and TnsB achieves catalysis. In another instance, TnsC can direct TnsA and TnsB to the insertion site. In another instance, TniQ and DNA can be recognized by TnsC and enable the Cas complex to achieve insertion at the site.

[0156] In one example embodiment, the CAST system can be used to insert or replace a sequence into one or more target genes. In example embodiments, the insertion or replacement results in an inactive target gene or less active form of the target gene. In one example embodiment, a CAST system is used to replace all or a portion of an enhancer controlling the target gene expression.

[0157] CAST systems may be used to introduce one or more modifications (insertions, deletions, substitutions) that modify expression of the one or more genes in a cell in the subject to which the gene

products produced by the host cell infected by the engineered phage of the present disclosure are delivered. The modifications may be made in a non-coding region that controls expression of the one or more target genes, in a coding region encoding a gene expression product (e.g., a polypeptide), or both. Example modifications are described in further detail below.

Non-LTR Systems

[0158] In one embodiment, the cargo may comprise a Non-LTR Retrotransposon system and/or one or more polynucleotide encoding to either decrease expression of one or more target genes in a cell of the the subject to which the gene products produced by the host cell infected by the engineered phage of the present disclosure are delivered.

[0159] A Non-LTR retrotransposon system may comprise one or more components of a retrotransposon, e.g., a non-LTR retrotransposon. Native or wild-type non-LTR retrotransposons encode the protein machinery necessary for their self-mobilization. The non-LTR retrotransposon element comprises a DNA element integrated into a host genome. The DNA element may encode one or two open reading frames (ORFs). For example, the R2 element of *Bombyx mori* encodes a single ORF containing reverse transcriptase (RT) activity and a restriction enzyme-like (REL) domain. L1 elements encode two ORFs, ORF1 and ORF2. ORF1 contains a leucine zipper domain involved in protein-protein interactions and a C-terminal nucleic acid binding domain. ORF2 has a N-terminal apurinic/apyrimidinic endonuclease (APE), a central RT domain, and a C-terminal cysteine histidine rich domain. An example replicative cycle of a non-LTR retrotransposon may comprise transcription of the full-length retrotransposon element to generate an mRNA active element (retrotransposon RNA). The active element mRNA is translated to generate the encoded retrotransposon proteins or polypeptides. A ribonucleoprotein complex comprising the active element and retrotransposon protein or polypeptide is formed and this RNP facilitates integration of the active element into the genome. In an example embodiment, the RNA-transposase complex nicks the genome and the 3' end of the nicked DNA serves as a primer to allow the reverse transcription of the transposon RNA into cDNA. The transposase proteins may then integrate the cDNA into the genome.

[0160] Elements of these systems may be engineered to work within the context of the disclosure. For example, a non-LTR retrotransposon polypeptide may be fused to a programmable nuclease. The binding elements that allow a non-LTR retrotransposon polypeptide to bind to the native retrotransposon DNA element, may be engineered into a donor construct to facilitate entry of a donor polynucleotide sequence into a target polypeptide.

[0161] In certain embodiments, the protein component of the non-LTR retrotransposon may be connected to or otherwise engineered to form a complex with a programmable nuclease, e.g., a Cas polypeptide. The retrotransposon RNA may be engineered to encode a donor polynucleotide sequence. Thus, in certain example embodiments, the Cas polypeptide, via formation of a CRISPR-Cas complex with a guide sequence, directs the retrotransposon complex (i.e., the retrotransposon polypeptide(s) and retrotransposon RNA) to a target sequence in a target polynucleotide, where the retrotransposon RNP complex facilitates integration of the donor polynucleotide sequence into the target polynucleotide. Accordingly, the one or more non-LTR retrotransposon components may comprise retrotransposon polypeptides, or function domains thereof, that facilitate binding of the retrotransposon RNA, reverse

transcription of the retrotransposon RNA into cDNA, and/or integration of the donor polynucleotide into the target polynucleotide, as well as retrotransposon RNA elements modified to encode the donor polynucleotide sequence. Example non-LTR retrotransposon systems are disclosed in WO 2021/102042, WO 2022/173830, which are incorporated herein by reference.

[0162] Examples of non-LTR retrotransposons may include those described in Christensen SM et al., RNA from the 5' end of the R2 retrotransposon controls R2 protein binding to and cleavage of its DNA target site, *Proc Natl Acad Sci U S A.* 2006 Nov 21;103(47):17602-7; Eickbush TH et al, Integration, Regulation, and Long-Term Stability of R2 Retrotransposons, *Microbiol Spectr.* 2015 Apr;3(2):MDNA3-0011-2014. doi: 10.1128/microbiolspec.MDNA3-0011-2014; Han JS, Non-long terminal repeat (non-LTR) retrotransposons: mechanisms, recent developments, and unanswered questions, *Mob DNA.* 2010 May 12;1(1):15. doi: 10.1186/1759-8753-1-15; Malik HS et al., The age and evolution of non-LTR retrotransposable elements, *Mol Biol Evol.* 1999 Jun;16(6):793-805, which are incorporated by reference herein in their entireties.

[0163] A non-LTR retrotransposon may comprise multiple retrotransposon polypeptides or polynucleotides encoding same. In some embodiments, the retrotransposon polypeptides may form a complex. For example, a non-LTR retrotransposon is a dimer, e.g., comprising two retrotransposon polypeptides forming a dimer. The dimer subunits may be connected or form a tandem fusion. A Cas protein or polypeptide may be associated with (e.g., connected to) one or more subunits of such complex. In some examples, the non-LTR retrotransposon is a dimer of two retrotransposon polypeptides; one of the retrotransposon polypeptides comprises nuclease or nickase activity and is connected with a Cas protein or polypeptide.

[0164] The retrotransposon polypeptides may be enzymes or variants thereof. In some examples, a retrotransposon polypeptide may be a reverse transcriptase, a nuclease, a nickase, a transposase, nucleic acid polymerase, ligase, or a combination thereof. In one example, a retrotransposon polypeptide is a reverse transcriptase. In another example, a retrotransposon polypeptide is a nuclease. In another example, a retrotransposon polypeptide is nickase. In a particular example, a non-LTR retrotransposon comprises a first retrotransposon polypeptide and a second retrotransposon polypeptide, wherein the second retrotransposon polypeptide comprises nuclease or nickase activity. In certain cases, a retrotransposon polypeptide may comprise an inactive enzyme. For example, a retrotransposon polypeptide may comprise a nuclease domain that is inactivated. Such inactivated domain may serve as a nucleic acid binding domain.

[0165] The retrotransposon polypeptides may comprise one or more modifications to, for example, enhance specificity or efficiency of donor polynucleotide recognition, target-primed template recognition (TPTR), and/or reduce or eliminate homing function. The retrotransposon polypeptides may also comprise one or more truncations or excisions to remove domains or regions of wild-type protein to arrive at a minimal polypeptide that retain donor polynucleotide recognition and TPTR. In some example embodiments, the native endonuclease activity may be mutated to eliminate endonuclease activity.

[0166] In certain example embodiments, the modifications or truncations of the non-LTR retrotransposon peptide may be in a zinc finger region, a Myb region, a basic region, a reverse transcriptase domain, a cysteine-histidine rich motif, or an endonuclease domain.

[0167] A non-LTR retrotransposon may comprise polynucleotide encoding one or more retrotransposon RNA molecules. The polynucleotide may comprise one or more regulatory elements. The regulatory elements may be promoters. The regulatory elements and promoters on the polynucleotides include those described throughout this application. For example, the polynucleotide may comprise a pol2 promoter, a pol3 promoter, or a T7 promoter.

[0168] In some cases, the polynucleotide encodes a retrotransposon RNA with at least a portion of its sequence complementary to a target sequence. For example, the 3' end of the retrotransposon RNA may be complementary to a target sequence. The RNA may be complementary to a portion of a nicked target sequence. In some embodiments, a retrotransposon RNA may comprise one or more donor polynucleotides. In certain cases, a retrotransposon RNA may encode one or more donor polynucleotides.

[0169] A retrotransposon RNA may be capable of binding to a retrotransposon polypeptide. Such retrotransposon RNA may comprise one or more elements for binding to the retrotransposon polypeptide. Examples of binding elements include hairpin structures, pseudoknots (e.g., a nucleic acid secondary structure containing at least two stem-loop structures in which half of one stem is intercalated between the two halves of another stem), stem loops, and bulges (e.g., unpaired stretches of nucleotides located within one strand of a nucleic acid duplex). In certain examples, the retrotransposon RNA comprises one or more hairpin structures. In some examples, the retrotransposon RNA comprises one or more pseudoknots. In certain examples, a retrotransposon RNA comprises a sequence encoding a donor polynucleotide and one or more binding elements for forming a complex with the retrotransposon polypeptide. The binding elements may be located on the 5' end, the 3' end, or a location in between.

[0170] In some embodiments, a retrotransposon RNA comprises a region capable of hybridizing with an overhang of a target polynucleotide at the target site. The overhang may be a stretch of single-stranded DNA. The overhang may function as a primer for reverse transcription of at least a portion of the retrotransposon RNA to a cDNA. In some cases, a region of the cDNA may be capable of hybridizing a second overhang of the target polynucleotide. The second overhang may function as a primer for the synthesis of a second strand to generate a double-stranded cDNA. The cDNA may comprise a donor polynucleotide sequence. The two overhangs may be from different strands of the target polynucleotide.

Donor Constructs

[0171] The systems may comprise one or more donor constructs comprising one or more donor polynucleotide sequences for insertion into a target polynucleotide. The donor construct comprises one or more binding elements. Examples of binding elements include hairpin structures, pseudoknots (e.g., a nucleic acid secondary structure containing at least two stem-loop structures in which half of one stem is intercalated between the two halves of another stem), stem loops, and bulges (e.g., unpaired stretches of nucleotides located within one strand of a nucleic acid duplex). In certain examples, the retrotransposon RNA comprises one or more hairpin structures. In some examples, the retrotransposon RNA comprises one or more pseudoknots. In certain examples, a retrotransposon RNA comprises a sequence encoding a donor polynucleotide and one or more binding elements for interacting to the retrotransposon polypeptide.

[0172] In certain example embodiments, the donor construct comprises a 5' binding element and a 3' binding element with a donor polynucleotide sequence located between the 5' and 3' prime binding element.

[0173] A donor polynucleotide may be any type of polynucleotides, including, but not limited to, a gene, a gene fragment, a non-coding polynucleotide, a regulatory polynucleotide, a synthetic polynucleotide, etc.

[0174] A target polynucleotide may comprise a protospacer adjacent motif (PAM) sequence. An example of the PAM sequence is AT.

[0175] The donor construct may further comprise one or more processing element. The processing element is an element that may be added to ensure accurate processing and incorporation of the donor polynucleotide sequence by the fusion proteins disclosed herein. Example processing elements include, but are not limited to, LRNA processing elements (e.g. GGCTCGTTGGGAGGTCCCGGTTGAAATCCCGACGAGCCCG), human 28s processing elements (e.g. TAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAATGGATGAACGAGATTCCCACTGTCCCTACCTACTATCCAGCGAAACACAGCCAAGGGAA), and natural retrotransposon processing elements such as R2 processing elements from *Bombyx mori* (e.g. tagccaaatgcctcgtcatctaattagtgacgcatgaatggattaacgagattcccactgtccctatctactatctagcgaaccacagccaa gggaacgggcttgggagaatcagcggggaa).

[0176] The donor construct may comprise one or more homology sequence. A homology sequence is a sequence that shares or complete or partial homology with a target sequence at the site the targeted site of insertion. The homology sequence may be located on the 5' end, 3' end, or on both the 5' and 3' end of the donor construct. In certain example embodiments, the homology sequence is only located on the 5' end of the donor construct. In certain example embodiments, the homology sequence is located only on the 3' end of the donor construct. In certain example embodiments, the location of the homology sequence may depend on whether the site-specific nuclease is being directed to create a nick or cut 5' or 3' of the targeted insertion site, e.g. a 5' homology sequence on the donor construct may be used when the site specific nuclease creates a nick or cut 5' of the targeted insertion site and a 3' homology sequence may be used when the site-specific nuclease is configured to create a nick or cut 3' of the targeted insertion site. In certain example embodiments, the homology sequence is included on both the 5' and 3' ends of the donor construct regardless of whether the site-specific nuclease creates a nick or cut 5' or 3' of the targeted insertion site. In certain example embodiments the donor construct may comprise in a 5' to 3', a binding element, and the donor sequence. In certain example embodiments the donor construct may comprise in a 5' to 3' direction a homology sequence, a binding element, and the donor sequence. In certain example embodiments the donor construct may comprise in a 5' to 3' direction a homology sequence, a first binding element, the donor sequence, and second binding element. In certain example embodiments, the donor construct may comprise in a 5' to 3' direction a first homology sequence, a first binding element, the donor sequence, and a second homology sequence. In certain example embodiments the donor construct may comprise, in a 5' to 3' direction, a first homology sequence, a first binding element, the donor sequence, a second binding

element, and a second homology sequence. In certain example embodiments, the donor construct may comprise, in a 5' to 3' direction, the donor sequence and a binding element. In certain example embodiments, the donor construct may comprise, in a 5' to 3' direction, the donor sequence, a binding element, and a homology sequence. A processing element may be further incorporated 3' of the donor sequence in any of the above donor construct configurations.

[0177] The homology sequence may have at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 175, 200 bases of homology to the target DNA. In certain example embodiments, the homology sequence may have between 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 base pairs of homology to the target sequence. In embodiments, with a homology sequence on both the 5' and 3' end of the donor construct, the size of the homology may be the same or different on each end. In some examples, the homology sequence comprises from 1 to 30, from 4 to 10, or from 10 to 25 nucleotides. For example, the homology sequence comprises from 4 to 10 nucleotides. For example, the homology sequence comprises from 10 to 25 nucleotides. For example, the homology sequence comprises 1 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides.

[0178] The donor polynucleotides may be inserted to the upstream or downstream of the PAM sequence of a target polynucleotide. For example, the donor polynucleotide may be inserted at a position between 10 bases and 200 bases, e.g., between 20 bases and 150 bases, between 30 bases and 100 bases, between 45 bases and 70 bases, between 45 bases and 60 bases, between 55 bases and 70 bases, between 49 bases and 56 bases or between 60 bases and 66 bases, from a PAM sequence on the target polynucleotide. In some cases, the insertion is at a position upstream of the PAM sequence. In some cases, the insertion is at a position downstream of the PAM sequence. In some cases, the insertion is at a position from 49 to 56 bases or base pairs downstream from a PAM sequence. In some cases, the insertion is at a position from 60 to 66 bases or base pairs downstream from a PAM sequence.

[0179] In a strand of a polynucleotide, anything towards the 5' end of a reference point is "upstream" of that point, and anything towards the 3' end of a reference point is "downstream" of that point. A location upstream of a PAM sequence refers to a location at the 5' side of the PAM sequence on the PAM-containing strand of the target sequence. A location downstream of a PAM sequence refers to a location at the 3' side of the PAM sequence on the PAM-containing strand of the target sequence.

[0180] The compositions and systems herein may be used to insert a donor polynucleotide with desired orientation. For example, appropriate homology sequence may be selected to control the orientation of insertion on the 5' or 3' strand of the target sequence.

[0181] The donor polynucleotide comprises a homology sequence of a region of the target sequence. The homology sequence may share at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or 100% sequence identity with the region of the target sequence. In an example, the homology sequence shares 100% sequence identity with the region of the target sequence.

[0182] In some embodiments, the donor polynucleotide may be inserted to the strand on the target sequence that contains the PAM (e.g., the PAM sequence of the site-specific nuclease such as Cas).

In such cases, the donor polynucleotide may comprise a homology sequence of a region on the PAM containing strand of the target sequence. Such region may comprise the PAM sequence. The region may be at the 3' side of the cleavage site of the site-specific nuclease. In some examples, the homology sequence may comprise from 4 to 10, or from 10 to 25 nucleotides in length.

[0183] In some embodiments, the donor polynucleotide may be inserted to the strand on the target sequence that binds to the guide, e.g., the strand that contains a guide-binding sequence. In such cases, the donor polynucleotide may comprise a homology sequence of a region that comprises at least a portion of the guide-binding sequence. In some cases, the region may comprise the entire guide-binding sequence. Such region may further comprise a sequence at the 3' side of the guide-binding sequence. For example, the region may comprise from 5 to 15 nucleotides, e.g., 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 nucleotides from the 3' side of the guide-binding sequence. In some cases, the region may be adjacent to the R-loop of the guide. For example, in the cases where the guide forms a RNA-DNA duplex with the guide-binding sequence, the region comprises a sequence at the 3' side from the RNA-DNA duplex, e.g., from 5 to from 5 to 15 nucleotides, e.g., 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 nucleotides from the 3' side from the RNA-DNA duplex.

[0184] In some examples, the homology sequence is of a region on the target sequence at 3' side of a PAM-containing strand. In certain examples, the homology sequence is of a region on the target sequence 10 nucleotides from 3' side of a RNA-DNA duplex formed by a guide molecule and a target sequence. For example, the guide molecule forms an RNA-DNA duplex with the target sequence, and the homology sequence is of a region on the target sequence 5 to 15 nucleotides from 3' side of the RNA-DNA duplex. In some embodiments, the donor polynucleotide is inserted to a region on the target sequence that is 3' side of a PAM-containing strand. In some cases, the donor polynucleotide is inserted to a region on the target sequence that is 3' side of a sequence complementary to the guide molecule.

[0185] The donor polynucleotide may be used for editing the target polynucleotide. In some cases, the donor polynucleotide comprises one or more mutations to be introduced into the target polynucleotide. Examples of such mutations include substitutions, deletions, insertions, or a combination thereof. The mutations may cause a shift in an open reading frame on the target polynucleotide. In some cases, the donor polynucleotide alters a stop codon in the target polynucleotide. For example, the donor polynucleotide may correct a premature stop codon. The correction may be achieved by deleting the stop codon or introduces one or more mutations to the stop codon. In other example embodiments, the donor polynucleotide addresses loss of function mutations, deletions, or translocations that may occur, for example, in certain disease contexts by inserting or restoring a functional copy of a gene, or functional fragment thereof, or a functional regulatory sequence or functional fragment of a regulatory sequence. A functional fragment refers to less than the entire copy of a gene by providing sufficient nucleotide sequence to restore the functionality of a wild type gene or non-coding regulatory sequence (e.g., sequences encoding long non-coding RNA). In certain example embodiments, the systems disclosed herein may be used to replace a single allele of a defective gene or defective fragment thereof. In another example embodiment, the systems disclosed herein may be used to replace both alleles of a defective gene or defective gene fragment. A "defective gene" or "defective gene fragment" is a gene or portion of a gene that when expressed fails to generate a

functioning protein or non-coding RNA with functionality of the corresponding wild-type gene. In certain example embodiments, these defective genes may be associated with one or more disease phenotypes. In certain example embodiments, the defective gene or gene fragment is not replaced but the systems described herein are used to insert donor polynucleotides that encode gene or gene fragments that compensate for or override defective gene expression such that cell phenotypes associated with defective gene expression are eliminated or changed to a different or desired cellular phenotype.

[0186] In certain embodiments, the donor may include, but not be limited to, genes or gene fragments, encoding proteins or RNA transcripts to be expressed, regulatory elements, repair templates, and the like. According to the disclosure, the donor polynucleotides may comprise left end and right end sequence elements that function with transposition components that mediate insertion.

[0187] In certain cases, the donor polynucleotide manipulates a splicing site on the target polynucleotide. In some examples, the donor polynucleotide disrupts a splicing site. The disruption may be achieved by inserting the polynucleotide to a splicing site and/or introducing one or more mutations to the splicing site. In certain examples, the donor polynucleotide may restore a splicing site. For example, the polynucleotide may comprise a splicing site sequence.

[0188] The donor polynucleotide to be inserted may have a size from 5 bases to 50 kb in length, e.g., from 50 to 40kb, from 100 and 30 kb, from 100 bases to 300 bases, from 200 bases to 400 bases, from 300 bases to 500 bases, from 400 bases to 600 bases, from 500 bases to 700 bases, from 600 bases to 800 bases, from 700 bases to 900 bases, from 800 bases to 1000 bases, from 900 bases to 1100 bases, from 1000 bases to 1200 bases, from 1100 bases to 1300 bases, from 1200 bases to 1400 bases, from 1300 bases to 1500 bases, from 1400 bases to 1600 bases, from 1500 bases to 1700 bases, from 600 bases to 1800 bases, from 1700 bases to 1900 bases, from 1800 bases to 2000 bases, from 1900 bases to 2100 bases, from 2000 bases to 2200 bases, from 2100 bases to 2300 bases, from 2200 bases to 2400 bases, from 2300 bases to 2500 bases, from 2400 bases to 2600 bases, from 2500 bases to 2700 bases, from 2600 bases to 2800 bases, from 2700 bases to 2900 bases, from 2800 bases to 3000 bases, from 2900 bases to 3100 bases, from 3000 bases to 3200 bases, from 3100 bases to 3300 bases, from 3200 bases to 3400 bases, from 3300 bases to 3500 bases, from 3400 bases to 3600 bases, from 3500 bases to 3700 bases, from 3600 bases to 3800 bases, from 3700 bases to 3900 bases, from 3800 bases to 4000 bases, from 3900 bases to 4100 bases, from 4000 bases to 4200 bases, from 4100 bases to 4300 bases, from 4200 bases to 4400 bases, from 4300 bases to 4500 bases, from 4400 bases to 4600 bases, from 4500 bases to 4700 bases, from 4600 bases to 4800 bases, from 4700 bases to 4900 bases, or from 4800 bases to 5000 bases in length.

OMEGA System

[0189] In one example embodiment, the cargo may be oligonucleotides encoding one or more component of a transposon-encoded RNA-guided nuclease system, referred to herein as OMEGA (obligate mobile element-guided activity). See, e.g., Altae-Tran H, Kannan S, Demircioglu FE, et al. The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided

endonucleases. *Science*. 2021;374(6563):57-65. OMEGA systems include, but are not limited to IscB, IsrB, TnpB systems.

[0190] In some embodiments, the nucleic acid-guided nucleases herein may be an IscB protein (see, e.g., International patent application publication No. WO2022087494A1; and Altae-Tran H, et al. 2021). An IscB protein may comprise an X domain and a Y domain as described herein. In some examples, the IscB proteins may form a complex with one or more guide molecules. In some cases, the IscB proteins may form a complex with one or more hRNA molecules which serve as a scaffold molecule and comprise guide sequences. In some examples, the IscB proteins are CRISPR-associated proteins, e.g., the loci of the nucleases are associated with an CRISPR array. In some examples, the IscB proteins are not CRISPR-associated. In some examples, the IscB protein may be homolog or ortholog of IscB proteins described in Kapitonov VV et al., ISC, a Novel Group of Bacterial and Archaeal DNA Transposons That Encode Cas9 Homologs, *J Bacteriol.* 2015 Dec 28;198(5):797-807. doi: 10.1128/JB.00783-15, which is incorporated by reference herein in its entirety.

[0191] In some embodiments, the nucleic acid-guided nucleases herein may be an IsrB (Insertion sequence RuvC-like OrfB) protein (see, e.g., International patent application publication No. WO2022087494A1; and Altae-Tran H, et al. 2021). IsrB refers to a group of shorter, ~350 aa IscB homologs that are also encoded in IS200/605 superfamily transposons. These proteins contain a PLMP domain and split RuvC but lack the HNH domain.

[0192] In some embodiments, the nucleic acid-guided nucleases herein may be a TnpB protein (see, e.g., International patent application publication No. WO2022159892A1; and Altae-Tran H, et al. 2021). TnpB is a putative endonuclease distantly related to IscB and thought to be the ancestor of Cas12, the type V CRISPR effector. The TnpB system comprises a TnpB polypeptide and a nucleic acid component capable of forming a complex with the TnpB polypeptide and directing the complex to a target polynucleotide. The TnpB systems and TnpB/nucleic acid component complexes may also be referred to herein as OMEGA (Obligate Mobile Element Guided Activity) systems or complexes, or W systems or complexes for short. TnpB systems are a distinct type of W system, which further include IscB, IsrB, and IshB systems. The nucleic acid component of W systems is structurally distinct from other RNA-guided nucleases, such as CRISPR-Cas systems, and may also be referred to as a wRNA. In certain example embodiments, the TnpB systems are RNA-predominate, that is the nucleic acid component makes a larger contribution to the overall size of the TnpB complex relative to other RNA-guided nuclease systems such as CRISPR-Cas. Also, given the more minimal structural features of TnpB relative other known programmable nucleases such as CRISPR-Cas, the polynucleotide binding pocket is open and more accessible, which can facilitate greater access to and ability to manipulate, modify, edit, remove, or delete nucleotides at a target region on the bound polynucleotide.

TALE Nucleases

[0193] In some embodiments, the cargo encodes a TALE nuclease or TALE nuclease system that can be used to modify a polynucleotide, such as in the cell of a subject to which the gene products are delivered after production in a host cell. In some embodiments, the methods provided herein use isolated, non-naturally occurring, recombinant or engineered DNA binding proteins that comprise TALE

monomers or TALE monomers or half monomers as a part of their organizational structure that enable the targeting of nucleic acid sequences with improved efficiency and expanded specificity.

[0194] Naturally occurring TALEs or “wild type TALEs” are nucleic acid binding proteins secreted by numerous species of proteobacteria. TALE polypeptides contain a nucleic acid binding domain composed of tandem repeats of highly conserved monomer polypeptides that are predominantly 33, 34 or 35 amino acids in length and that differ from each other mainly in amino acid positions 12 and 13. In advantageous embodiments the nucleic acid is DNA. As used herein, the term “polypeptide monomers”, “TALE monomers” or “monomers” will be used to refer to the highly conserved repetitive polypeptide sequences within the TALE nucleic acid binding domain and the term “repeat variable di-residues” or “RVD” will be used to refer to the highly variable amino acids at positions 12 and 13 of the polypeptide monomers. As provided throughout the disclosure, the amino acid residues of the RVD are depicted using the IUPAC single letter code for amino acids. A general representation of a TALE monomer which is comprised within the DNA binding domain is $X_{1-11}-(X_{12}X_{13})-X_{14-33}$ or 34 or 35 , where the subscript indicates the amino acid position and X represents any amino acid. $X_{12}X_{13}$ indicate the RVDs. In some polypeptide monomers, the variable amino acid at position 13 is missing or absent and in such monomers, the RVD consists of a single amino acid. In such cases the RVD may be alternatively represented as X^* , where X represents X_{12} and (*) indicates that X_{13} is absent. The DNA binding domain comprises several repeats of TALE monomers and this may be represented as $(X_{1-11}-(X_{12}X_{13})-X_{14-33}$ or 34 or $35)_z$, where in an advantageous embodiment, z is at least 5 to 40. In a further advantageous embodiment, z is at least 10 to 26.

[0195] The TALE monomers can have a nucleotide binding affinity that is determined by the identity of the amino acids in its RVD. For example, polypeptide monomers with an RVD of NI can preferentially bind to adenine (A), monomers with an RVD of NG can preferentially bind to thymine (T), monomers with an RVD of HD can preferentially bind to cytosine (C) and monomers with an RVD of NN can preferentially bind to both adenine (A) and guanine (G). In some embodiments, monomers with an RVD of IG can preferentially bind to T. Thus, the number and order of the polypeptide monomer repeats in the nucleic acid binding domain of a TALE determines its nucleic acid target specificity. In some embodiments, monomers with an RVD of NS can recognize all four base pairs and can bind to A, T, G or C. The structure and function of TALEs is further described in, for example, Moscou et al., Science 326:1501 (2009); Boch et al., Science 326:1509-1512 (2009); and Zhang et al., Nature Biotechnology 29:149-153 (2011).

[0196] The polypeptides used in methods of the disclosure can be isolated, non-naturally occurring, recombinant or engineered nucleic acid-binding proteins that have nucleic acid or DNA binding regions containing polypeptide monomer repeats that are designed to target specific nucleic acid sequences.

[0197] As described herein, polypeptide monomers having an RVD of HN or NH preferentially bind to guanine and thereby allow the generation of TALE polypeptides with high binding specificity for guanine containing target nucleic acid sequences. In some embodiments, polypeptide monomers having RVDs RN, NN, NK, SN, NH, KN, HN, NQ, HH, RG, KH, RH and SS can preferentially bind to guanine. In some embodiments, polypeptide monomers having RVDs RN, NK, NQ, HH, KH, RH, SS

and SN can preferentially bind to guanine and can thus allow the generation of TALE polypeptides with high binding specificity for guanine containing target nucleic acid sequences. In some embodiments, polypeptide monomers having RVDs HH, KH, NH, NK, NQ, RH, RN and SS can preferentially bind to guanine and thereby allow the generation of TALE polypeptides with high binding specificity for guanine containing target nucleic acid sequences. In some embodiments, the RVDs that have high binding specificity for guanine are RN, NH RH and KH. Furthermore, polypeptide monomers having an RVD of NV can preferentially bind to adenine and guanine. In some embodiments, monomers having RVDs of H*, HA, KA, N*, NA, NC, NS, RA, and S* bind to adenine, guanine, cytosine and thymine with comparable affinity.

[0198] The predetermined N-terminal to C-terminal order of the one or more polypeptide monomers of the nucleic acid or DNA binding domain determines the corresponding predetermined target nucleic acid sequence to which the polypeptides of the disclosure will bind. As used herein the monomers and at least one or more half monomers are "specifically ordered to target" the genomic locus or gene of interest. In plant genomes, the natural TALE-binding sites always begin with a thymine (T), which may be specified by a cryptic signal within the non-repetitive N-terminus of the TALE polypeptide; in some cases, this region may be referred to as repeat 0. In animal genomes, TALE binding sites do not necessarily have to begin with a thymine (T) and polypeptides of the disclosure may target DNA sequences that begin with T, A, G or C. The tandem repeat of TALE monomers always ends with a half-length repeat or a stretch of sequence that may share identity with only the first 20 amino acids of a repetitive full-length TALE monomer and this half repeat may be referred to as a half-monomer. Therefore, it follows that the length of the nucleic acid or DNA being targeted is equal to the number of full monomers plus two.

[0199] As described in Zhang et al., Nature Biotechnology 29:149-153 (2011), TALE polypeptide binding efficiency may be increased by including amino acid sequences from the "capping regions" that are directly N-terminal or C-terminal of the DNA binding region of naturally occurring TALEs into the engineered TALEs at positions N-terminal or C-terminal of the engineered TALE DNA binding region. Thus, in certain embodiments, the TALE polypeptides described herein further comprise an N-terminal capping region and/or a C-terminal capping region.

[0200] An exemplary amino acid sequence of a N-terminal capping region is:

MDPIRSRTPSPARELLSGPQPDGVQPTADRGVSPAGGPLDGLPARR
TMSRTRLPSPPAPSPAFSADSFSDLLRQFDPSTLNTSLFDSLPPFGAH
HTEAATGEWDEVQSGRLAADAPPTMRVAVTAARPPRAKPAPRRRAA
QPSDASPAAQVDLRTLGYSSQQQKEKIKPKVRSTVAQHHEALVGHGFT
HAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARA
LEALLTVAGELRGPPPLQLDTGQLLKIAGRGGVTAVEAVHAWRNALTGA
PLN

[0201] An exemplary amino acid sequence of a C-terminal capping region is:

RPALESIVAQLSRPDPALAAALTNDHLVALACLGGRPALDAVKKGLPHA
PALIKRTNRRIPERTSHRVADHAQVVRVLGFFQCHSHPAQAFDDAMTQ

FGMSRHGLLQLFRRVGVTELEARSGTLPPASQRWDRILQASGMKRAK
PSPTSTQTPDQASLHAFADSLERDLDPSPMHEGDQTRAS

[0202] As used herein the predetermined "N-terminus" to "C terminus" orientation of the N-terminal capping region, the DNA binding domain comprising the repeat TALE monomers and the C-terminal capping region provide structural basis for the organization of different domains in the d-TALEs or polypeptides of the disclosure.

[0203] The entire N-terminal and/or C-terminal capping regions are not necessary to enhance the binding activity of the DNA binding region. Therefore, in certain embodiments, fragments of the N-terminal and/or C-terminal capping regions are included in the TALE polypeptides described herein.

[0204] In certain embodiments, the TALE polypeptides described herein contain a N-terminal capping region fragment that included at least 10, 20, 30, 40, 50, 54, 60, 70, 80, 87, 90, 94, 100, 102, 110, 117, 120, 130, 140, 147, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260 or 270 amino acids of an N-terminal capping region. In certain embodiments, the N-terminal capping region fragment amino acids are of the C-terminus (the DNA-binding region proximal end) of an N-terminal capping region. As described in Zhang et al., *Nature Biotechnology* 29:149-153 (2011), N-terminal capping region fragments that include the C-terminal 240 amino acids enhance binding activity equal to the full-length capping region, while fragments that include the C-terminal 147 amino acids retain greater than 80% of the efficacy of the full length capping region, and fragments that include the C-terminal 117 amino acids retain greater than 50% of the activity of the full-length capping region.

[0205] In some embodiments, the TALE polypeptides described herein contain a C-terminal capping region fragment that included at least 6, 10, 20, 30, 37, 40, 50, 60, 68, 70, 80, 90, 100, 110, 120, 127, 130, 140, 150, 155, 160, 170, 180 amino acids of a C-terminal capping region. In certain embodiments, the C-terminal capping region fragment amino acids are of the N-terminus (the DNA-binding region proximal end) of a C-terminal capping region. As described in Zhang et al., *Nature Biotechnology* 29:149-153 (2011), C-terminal capping region fragments that include the C-terminal 68 amino acids enhance binding activity equal to the full-length capping region, while fragments that include the C-terminal 20 amino acids retain greater than 50% of the efficacy of the full-length capping region.

[0206] In certain embodiments, the capping regions of the TALE polypeptides described herein do not need to have identical sequences to the capping region sequences provided herein. Thus, in some embodiments, the capping region of the TALE polypeptides described herein have sequences that are at least 50%, 60%, 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical or share identity to the capping region amino acid sequences provided herein. Sequence identity is related to sequence homology. Homology comparisons may be conducted by eye, or more usually, with the aid of readily available sequence comparison programs. These commercially available computer programs may calculate percent (%) homology between two or more sequences and may also calculate the sequence identity shared by two or more amino acid or nucleic acid sequences. In some preferred embodiments, the capping region of the TALE polypeptides described herein have sequences that are at least 95% identical or share identity to the capping region amino acid sequences provided herein.

[0207] Sequence homologies can be generated by any of a number of computer programs known in the art, which include but are not limited to BLAST or FASTA. Suitable computer programs for carrying out alignments like the GCG Wisconsin Bestfit package may also be used. Once the software has produced an optimal alignment, it is possible to calculate % homology, preferably % sequence identity. The software typically does this as part of the sequence comparison and generates a numerical result.

[0208] In some embodiments described herein, the TALE polypeptides of the disclosure include a nucleic acid binding domain linked to the one or more effector domains. The terms “effector domain” or “regulatory and functional domain” refer to a polypeptide sequence that has an activity other than binding to the nucleic acid sequence recognized by the nucleic acid binding domain. By combining a nucleic acid binding domain with one or more effector domains, the polypeptides of the disclosure may be used to target the one or more functions or activities mediated by the effector domain to a particular target DNA sequence to which the nucleic acid binding domain specifically binds.

[0209] In some embodiments of the TALE polypeptides described herein, the activity mediated by the effector domain is a biological activity. For example, in some embodiments the effector domain is a transcriptional inhibitor (i.e., a repressor domain), such as an mSin interaction domain (SID). SID4X domain or a Krüppel-associated box (KRAB) or fragments of the KRAB domain. In some embodiments, the effector domain is an enhancer of transcription (i.e., an activation domain), such as the VP16, VP64 or p65 activation domain. In some embodiments, the nucleic acid binding is linked, for example, with an effector domain that includes but is not limited to a transposase, integrase, recombinase, resolvase, invertase, protease, DNA methyltransferase, DNA demethylase, histone acetylase, histone deacetylase, nuclease, transcriptional repressor, transcriptional activator, transcription factor recruiting, protein nuclear-localization signal or cellular uptake signal.

[0210] In some embodiments, the effector domain of the TALE nucleicase is a protein domain which exhibits activities which include but are not limited to transposase activity, integrase activity, recombinase activity, resolvase activity, invertase activity, protease activity, DNA methyltransferase activity, DNA demethylase activity, histone acetylase activity, histone deacetylase activity, nuclease activity, nuclear-localization signaling activity, transcriptional repressor activity, transcriptional activator activity, transcription factor recruiting activity, or cellular uptake signaling activity.

Zinc Finger Nucleases

[0211] In some embodiments, the cargo is or encodes a zinc finger nuclease and can be used to modify a polynucleotide in a cell of subject to which the cargo is delivered after being produced in a host cell. The first synthetic zinc finger nucleases (ZFNs) were developed by fusing a ZF protein to the catalytic domain of the Type IIS restriction enzyme FokI. (Kim, Y. G. et al., 1994, Chimeric restriction endonuclease, Proc. Natl. Acad. Sci. U.S.A. 91, 883–887; Kim, Y. G. et al., 1996, Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc. Natl. Acad. Sci. U.S.A. 93, 1156–1160). Increased cleavage specificity can be attained with decreased off target activity by use of paired ZFN heterodimers, each targeting different nucleotide sequences separated by a short spacer. (Doyon, Y. et al., 2011, Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat. Methods 8, 74–79). ZFPs can also be designed as transcription activators and repressors and have been used to target many genes in a wide variety of organisms. Exemplary methods of genome

editing using ZFNs can be found for example in U.S. Patent Nos. 6,534,261, 6,607,882, 6,746,838, 6,794,136, 6,824,978, 6,866,997, 6,933,113, 6,979,539, 7,013,219, 7,030,215, 7,220,719, 7,241,573, 7,241,574, 7,585,849, 7,595,376, 6,903,185, and 6,479,626, all of which are specifically incorporated by reference.

Meganucleases

[0212] In some embodiments, the cargo is or encodes a meganuclease or system thereof and can be used to modify a polynucleotide in a cell of subject to which the cargo is delivered after being produced in a host cell. Meganucleases, which are endodeoxyribonucleases characterized by a large recognition site (double-stranded DNA sequences of 12 to 40 base pairs). Exemplary methods for using meganucleases can be found in US Patent Nos. 8,163,514, 8,133,697, 8,021,867, 8,119,361, 8,119,381, 8,124,369, and 8,129,134, which are specifically incorporated herein by reference.

RNAi

[0213] In certain embodiments, the genetic modifying agent cargo is an RNAi molecule (e.g., shRNA). RNAi can reduce the transcription and/or translation of a target RNA molecule, by e.g., blocking RNA transcription and/or translation and/or reducing the amount of target RNA molecules by inducing their degradation.

[0214] As used herein, the term "RNAi" refers to any type of interfering RNA, including but not limited to, siRNAi, shRNAi, endogenous microRNA and artificial microRNA. For instance, it includes sequences previously identified as siRNA, regardless of the mechanism of down-stream processing of the RNA (i.e., although siRNAs are believed to have a specific method of *in vivo* processing resulting in the cleavage of mRNA, such sequences can be incorporated into the vectors in the context of the flanking sequences described herein). The term "RNAi" can include both gene silencing RNAi molecules, and also RNAi effector molecules which activate the expression of a gene.

[0215] As used herein, a "siRNA" refers to a nucleic acid that forms a double stranded RNA, which double stranded RNA has the ability to reduce or inhibit expression of a gene or target gene when the siRNA is present or expressed in the same cell as the target gene. The double stranded RNA siRNA can be formed by the complementary strands. In one embodiment, a siRNA refers to a nucleic acid that can form a double stranded siRNA. The sequence of the siRNA can correspond to the full-length target gene, or a subsequence thereof. Typically, the siRNA is at least about 15-50 nucleotides in length (e.g., each complementary sequence of the double stranded siRNA is about 15-50 nucleotides in length, and the double stranded siRNA is about 15-50 base pairs in length, preferably about 19-30 base nucleotides, preferably about 20-25 nucleotides in length, e.g., 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides in length).

[0216] As used herein "shRNA" or "small hairpin RNA" (also called stem loop) is a type of siRNA. In one embodiment, these shRNAs are composed of a short, e.g., about 19 to about 25 nucleotide, antisense strand, followed by a nucleotide loop of about 5 to about 9 nucleotides, and the analogous sense strand. Alternatively, the sense strand can precede the nucleotide loop structure and the antisense strand can follow.

[0217] The terms "microRNA" or "miRNA" are used interchangeably herein are endogenous RNAs, some of which are known to regulate the expression of protein-coding genes at the

posttranscriptional level. Endogenous microRNAs are small RNAs naturally present in the genome that are capable of modulating the productive utilization of mRNA. The term artificial microRNA includes any type of RNA sequence, other than endogenous microRNA, which is capable of modulating the productive utilization of mRNA. MicroRNA sequences have been described in publications such as Lim, et al., *Genes & Development*, 17, p. 991 - 1008 (2003), Lim et al *Science* 299, 1540 (2003), Lee and Ambros *Science*, 294, 862 (2001), Lau et al., *Science* 294, 858-861 (2001), Lagos-Quintana et al, *Current Biology*, 12, 735-739 (2002), Lagos Quintana et al, *Science* 294, 853- 857 (2001), and Lagos-Quintana et al, *RNA*, 9, 175- 179 (2003), which are incorporated herein by reference. Multiple microRNAs can also be incorporated into a precursor molecule. Furthermore, miRNA-like stem-loops can be expressed in cells as a vehicle to deliver artificial miRNAs and short interfering RNAs (siRNAs) for the purpose of modulating the expression of endogenous genes through the miRNA and or RNAi pathways.

[0218] As used herein, "double stranded RNA" or "dsRNA" refers to RNA molecules that are comprised of two strands. Double-stranded molecules include those comprised of a single RNA molecule that doubles back on itself to form a two-stranded structure. For example, the stem loop structure of the progenitor molecules from which the single-stranded miRNA is derived, called the pre-miRNA (Bartel et al. 2004. *Cell* 116:281 -297), comprises a dsRNA molecule.

Polypeptides

[0219] In certain example embodiments, the cargo is or encodes one or more polypeptides. The polypeptide may be a full-length protein or a functional fragment or functional domain thereof, that is a fragment or domain that maintains the desired functionality of the full-length protein. As used within this section "protein" is meant to refer to full-length proteins and functional fragments and domains thereof. A wide array of polypeptides may be delivered using the engineered phages described herein, including but not limited to, secretory proteins, immunomodulatory proteins, anti-fibrotic proteins, proteins that promote tissue regeneration and/or transplant survival functions, hormones, anti-satiety, anti-cancer, anti-microbial proteins, anti-fibrillating polypeptides, and antibodies (or fragments thereof), and/or the like). The one or more polypeptides may also comprise combinations of the aforementioned example classes of polypeptides.

Secretory Proteins

[0220] In certain example embodiments, the one or more polypeptides may comprise one or more secretory proteins. A secretory is a protein that is actively transported out of the cell, for example, the protein, whether it be endocrine or exocrine, is secreted by a cell. Secretory pathways have been shown conserved from yeast to mammals, and both conventional and unconventional protein secretion pathways have been demonstrated in plants. Chung et al., "An Overview of Protein Secretion in Plant Cells," *MIMB*, 1662:19-32, September 1, 2017. Accordingly, identification of secretory proteins in which one or more polynucleotides may be inserted can be identified for particular cells and applications. In embodiments, one of skill in the art can identify secretory proteins based on the presence of a signal peptide, which consists of a short hydrophobic N-terminal sequence.

[0221] In embodiments, the protein is secreted by the secretory pathway. In embodiments, the proteins are exocrine secretion proteins or peptides, comprising enzymes in the digestive tract. In

embodiments the protein is endocrine secretion protein or peptide, for example, insulin and other hormones released into the blood stream. In other embodiments, the protein is involved in signaling between or within cells via secreted signaling molecules, for example, paracrine, autocrine, endocrine or neuroendocrine. In embodiments, the secretory protein is selected from the group of cytokines, kinases, hormones and growth factors that bind to receptors on the surface of target cells.

[0222] As described, secretory proteins include hormones, enzymes, toxins, and antimicrobial peptides. Examples of secretory proteins include serine proteases (e.g., pepsins, trypsin, chymotrypsin, elastase and plasminogen activators), amylases, lipases, nucleases (e.g. deoxyribonucleases and ribonucleases), peptidases enzyme inhibitors such as serpins (e.g., α 1-antitrypsin and plasminogen activator inhibitors), cell attachment proteins such as collagen, fibronectin and laminin, hormones and growth factors such as insulin, growth hormone, prolactin platelet-derived growth factor, epidermal growth factor, fibroblast growth factors, interleukins, interferons, apolipoproteins, and carrier proteins such as transferrin and albumins. In some examples, the secretory protein is insulin or a fragment thereof. In one example, the secretory protein is a precursor of insulin or a fragment thereof. In certain examples, the secretory protein is c-peptide. In a preferred embodiment, the one or more polynucleotides is inserted in the middle of the c-peptide. In some embodiments, the secretory protein is GLP-1, glucagon, betatrophin, pancreatic amylase, pancreatic lipase, carboxypeptidase, secretin, CCK, a PPAR (e.g., PPAR-alpha, PPAR-gamma, PPAR-delta or a precursor thereof (e.g., preprotein or preproprotein). In aspects, the secretory protein is fibronectin, a clotting factor protein (e.g., Factor VII, VIII, IX, etc.), α 2-macroglobulin, α 1-antitrypsin, antithrombin III, protein S, protein C, plasminogen, α 2-antiplasmin, complement components (e.g., complement component C1-9), albumin, ceruloplasmin, transcortin, haptoglobin, hemopexin, IGF binding protein, retinol binding protein, transferrin, vitamin-D binding protein, transthyretin, IGF-1, thrombopoietin, hepcidin, angiotensinogen, or a precursor protein thereof. In aspects, the secretory protein is pepsinogen, gastric lipase, sucrase, gastrin, lactase, maltase, peptidase, or a precursor thereof. In aspects, the secretory protein is renin, erythropoietin, angiotensin, adrenocorticotrophic hormone (ACTH), amylin, atrial natriuretic peptide (ANP), calcitonin, ghrelin, growth hormone (GH), leptin, melanocyte-stimulating hormone (MSH), oxytocin, prolactin, follicle-stimulating hormone (FSH), thyroid stimulating hormone (TSH), thyrotropin-releasing hormone (TRH), vasopressin, vasoactive intestinal peptide, or a precursor thereof.

Immunomodulatory Polypeptides

[0223] In certain example embodiments, the one or more polypeptides may comprise one or more immunomodulatory protein. In certain embodiments, the present disclosure provides for modulating immune states. The immune state can be modulated by modulating T cell function or dysfunction. In particular embodiments, the immune state is modulated by expression and secretion of IL-10 and/or other cytokines as described elsewhere herein. In certain embodiments, T cells can affect the overall immune state, such as other immune cells in proximity.

[0224] In some embodiments, the immunomodulatory protein is an immunosuppressive protein or an immunostimulant protein. In some embodiments, the cargo is polynucleotide(s) that encode one or more immunomodulatory proteins. The term "immunosuppressive" means that immune response in an organism is reduced or depressed. An immunosuppressive protein may suppress, reduce, or mask the

immune system or degree of response of the subject being treated. For example, an immunosuppressive protein may suppress cytokine production, downregulate or suppress self-antigen expression, or mask the MHC antigens. The term "immunostimulant" means that immune response in an organism is increased or activated.

[0225] As used herein, the term "immune response" refers to a response by a cell of the immune system, such as a B cell, T cell (CD4+ or CD8+), regulatory T cell, antigen-presenting cell, dendritic cell, monocyte, macrophage, NKT cell, NK cell, basophil, eosinophil, or neutrophil, to a stimulus. In some embodiments, the response is specific for a particular antigen (an "antigen-specific response") and refers to a response by a CD4 T cell, CD8 T cell, or B cell via their antigen-specific receptor. In some embodiments, an immune response is a T cell response, such as a CD4+ response or a CD8+ response. Such responses by these cells can include, for example, cytotoxicity, proliferation, cytokine or chemokine production, trafficking, or phagocytosis, and can be dependent on the nature of the immune cell undergoing the response. In some cases, the immunosuppressive proteins may exert pleiotropic functions. In some cases, the immunomodulatory proteins may maintain proper regulatory T cells versus effector T cells (Treg/Teff) balance. For examples, the immunomodulatory proteins may expand and/or activate the Tregs and blocks the actions of Teffs, thus providing immunoregulation without global immunosuppression. Target genes associated with immune suppression include, for example, checkpoint inhibitors such PD1, Tim3, Lag3, TIGIT, CTLA-4, and combinations thereof.

[0226] The term "immune cell" as used throughout this specification generally encompasses any cell derived from a hematopoietic stem cell that plays a role in the immune response. The term is intended to encompass immune cells both of the innate or adaptive immune system. The immune cell as referred to herein may be a leukocyte, at any stage of differentiation (e.g., a stem cell, a progenitor cell, a mature cell) or any activation stage. Immune cells include lymphocytes (such as natural killer cells, T-cells (including, e.g., thymocytes, Th or Tc; Th1, Th2, Th17, Th $\alpha\beta$, CD4+, CD8+, effector Th, memory Th, regulatory Th, CD4+/CD8+ thymocytes, CD4-/CD8- thymocytes, $\gamma\delta$ T cells, etc.) or B-cells (including, e.g., pro-B cells, early pro-B cells, late pro-B cells, pre-B cells, large pre-B cells, small pre-B cells, immature or mature B-cells, producing antibodies of any isotype, T1 B-cells, T2, B-cells, naïve B-cells, GC B-cells, plasmablasts, memory B-cells, plasma cells, follicular B-cells, marginal zone B-cells, B-1 cells, B-2 cells, regulatory B cells, etc.), such as for instance, monocytes (including, e.g., classical, non-classical, or intermediate monocytes), (segmented or banded) neutrophils, eosinophils, basophils, mast cells, histiocytes, microglia, including various subtypes, maturation, differentiation, or activation stages, such as for instance hematopoietic stem cells, myeloid progenitors, lymphoid progenitors, myeloblasts, promyelocytes, myelocytes, metamyelocytes, monoblasts, promonocytes, lymphoblasts, prolymphocytes, small lymphocytes, macrophages (including, e.g., Kupffer cells, stellate macrophages, M1 or M2 macrophages), (myeloid or lymphoid) dendritic cells (including, e.g., Langerhans cells, conventional or myeloid dendritic cells, plasmacytoid dendritic cells, mDC-1, mDC-2, Mo-DC, HP-DC, veiled cells), granulocytes, polymorphonuclear cells, antigen-presenting cells (APC), etc.

[0227] "T cell response" refers more specifically to an immune response in which T cells directly or indirectly mediate or otherwise contribute to an immune response in a subject. T cell-mediated response may be associated with cell mediated effects, cytokine mediated effects, and even effects

associated with B cells if the B cells are stimulated, for example, by cytokines secreted by T cells. By means of an example but without limitation, effector functions of MHC class I restricted Cytotoxic T lymphocytes (CTLs), may include cytokine and/or cytolytic capabilities, such as lysis of target cells presenting an antigen peptide recognized by the T cell receptor (naturally-occurring TCR or genetically engineered TCR, e.g., chimeric antigen receptor, CAR), secretion of cytokines, preferably IFN gamma, TNF alpha and/or or more immunostimulatory cytokines, such as IL-2, and/or antigen peptide-induced secretion of cytotoxic effector molecules, such as granzymes, perforins or granulysin. By means of example but without limitation, for MHC class II restricted T helper (Th) cells, effector functions may be antigen peptide-induced secretion of cytokines, preferably, IFN gamma, TNF alpha, IL-4, IL5, IL-10, and/or IL-2. By means of example but without limitation, for T regulatory (Treg) cells, effector functions may be antigen peptide-induced secretion of cytokines, preferably, IL-10, IL-35, and/or TGF-beta. B cell response refers more specifically to an immune response in which B cells directly or indirectly mediate or otherwise contribute to an immune response in a subject. Effector functions of B cells may include in particular production and secretion of antigen-specific antibodies by B cells (e.g., polyclonal B cell response to a plurality of the epitopes of an antigen (antigen-specific antibody response)), antigen presentation, and/or cytokine secretion.

[0228] During persistent immune activation, such as during uncontrolled tumor growth or chronic infections, subpopulations of immune cells, particularly of CD8+ or CD4+ T cells, become compromised to different extents with respect to their cytokine and/or cytolytic capabilities. Such immune cells, particularly CD8+ or CD4+ T cells, are commonly referred to as “dysfunctional” or as “functionally exhausted” or “exhausted”. As used herein, the term “dysfunctional” or “functional exhaustion” refer to a state of a cell where the cell does not perform its usual function or activity in response to normal input signals, and includes refractivity of immune cells to stimulation, such as stimulation via an activating receptor or a cytokine. Such a function or activity includes, but is not limited to, proliferation (e.g., in response to a cytokine, such as IFN-gamma) or cell division, entrance into the cell cycle, cytokine production, cytotoxicity, migration and trafficking, phagocytotic activity, or any combination thereof. Normal input signals can include, but are not limited to, stimulation via a receptor (e.g., T cell receptor, B cell receptor, co-stimulatory receptor). Unresponsive immune cells can have a reduction of at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or even 100% in cytotoxic activity, cytokine production, proliferation, trafficking, phagocytotic activity, or any combination thereof, relative to a corresponding control immune cell of the same type. In some particular embodiments of the aspects described herein, a cell that is dysfunctional is a CD8+ T cell that expresses the CD8+ cell surface marker. Such CD8+ cells normally proliferate and produce cell killing enzymes, e.g., they can release the cytotoxins perforin, granzymes, and granulysin. However, exhausted/dysfunctional T cells do not respond adequately to TCR stimulation, and display poor effector function, sustained expression of inhibitory receptors and a transcriptional state distinct from that of functional effector or memory T cells. Dysfunction/exhaustion of T cells thus prevents optimal control of infection and tumors. Exhausted/dysfunctional immune cells, such as T cells, such as CD8+ T cells, may produce reduced amounts of IFN-gamma, TNF-alpha and/or one or more immunostimulatory cytokines, such as IL-2, compared to functional immune cells. Exhausted/dysfunctional immune cells, such as T cells, such as

CD8+ T cells, may further produce (increased amounts of) one or more immunosuppressive transcription factors or cytokines, such as IL-10 and/or Foxp3, compared to functional immune cells, thereby contributing to local immunosuppression. Dysfunctional CD8+ T cells can be both protective and detrimental against disease control. As used herein, a “dysfunctional immune state” refers to an overall suppressive immune state in a subject or microenvironment of the subject (e.g., tumor microenvironment). For example, increased IL-10 production leads to suppression of other immune cells in a population of immune cells. In some embodiments, the cargo mitigates or prevents T cell exhaustion.

[0229] CD8+ T cell function is associated with their cytokine profiles. It has been reported that effector CD8+ T cells with the ability to simultaneously produce multiple cytokines (polyfunctional CD8+ T cells) are associated with protective immunity in patients with controlled chronic viral infections as well as cancer patients responsive to immune therapy (Spranger et al., 2014, J. Immunother. Cancer, vol. 2, 3). In the presence of persistent antigen CD8+ T cells were found to have lost cytolytic activity completely over time (Moskophidis et al., 1993, Nature, vol. 362, 758–761). It was subsequently found that dysfunctional T cells can differentially produce IL-2, TNFα and IFNγ in a hierarchical order (Wherry et al., 2003, J. Virol., vol. 77, 4911–4927). Decoupled dysfunctional and activated Cell states have also been described (see, e.g., Singer, et al. (2016). A Distinct Gene Module for Dysfunction Uncoupled from Activation in Tumor-Infiltrating T Cells. Cell 166, 1500-1511 e1509; WO/2017/075478; and WO/2018/049025).

[0230] In some embodiments, the cargo(s) modulate T cell balance. The disclosure provides T cell modulating agents that modulate T cell balance. For example, in some embodiments, the disclosure provides T cell modulating agents and methods of using these T cell modulating agents to regulate, influence or otherwise impact the level of and/or balance between T cell types, e.g., between Th17 and other T cell types, for example, Th1-like cells. For example, in some embodiments, the disclosure provides T cell modulating agents and methods of using these T cell modulating agents to regulate, influence or otherwise impact the level of and/or balance between Th17 activity and inflammatory potential. As used herein, terms such as “Th17 cell” and/or “Th17 phenotype” and all grammatical variations thereof refer to a differentiated T helper cell that expresses one or more cytokines selected from the group consisting of interleukin 17A (IL-17A), interleukin 17F (IL-17F), and interleukin 17A/F heterodimer (IL17-AF). As used herein, terms such as “Th1 cell” and/or “Th1 phenotype” and all grammatical variations thereof refer to a differentiated T helper cell that expresses interferon gamma (IFNγ). As used herein, terms such as “Th2 cell” and/or “Th2 phenotype” and all grammatical variations thereof refer to a differentiated T helper cell that expresses one or more cytokines selected from the group consisting of interleukin 4 (IL-4), interleukin 5 (IL-5) and interleukin 13 (IL-13). As used herein, terms such as “Treg cell” and/or “Treg phenotype” and all grammatical variations thereof refer to a differentiated T cell that expresses Foxp3.

[0231] In some examples, immunomodulatory proteins are immunosuppressive cytokines. In general, cytokines are small proteins and include interleukins, lymphokines and cell signal molecules, such as tumor necrosis factor and the interferons, which regulate inflammation, hematopoiesis, and response to infections. Examples of immunosuppressive cytokines include interleukin 10 (IL-10), TGF-

β , IL-Ra, IL-18Ra, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-19, IL-20, IL-21, IL-22, IL-23, IL-24, IL-25, IL-26, IL-27, IL-28, IL-29, IL-30, IL-31, IL-32, IL-33, IL-34, IL-35, IL-36, IL-37, PGE2, SCF, G-CSF, CSF-1R, M-CSF, GM-CSF, IFN- α , IFN- β , IFN- γ , IFN- λ , bFGF, CCL2, CXCL1, CXCL8, CXCL12, CX3CL1, CXCR4, TNF- α and VEGF. Examples of immunosuppressive proteins may further include FOXP3, AHR, TRP53, IKZF3, IRF4, IRF1, and SMAD3. In one example, the immunosuppressive protein is IL-10. In one example, the immunosuppressive protein is IL-6. In one example, the immunosuppressive protein is IL-2.

Anti-fibrotic proteins

[0232] In certain example embodiments, the one or more cargo polypeptides may comprise an anti-fibrotic protein. Examples of anti-fibrotic proteins include any protein that reduces or inhibits the production of extracellular matrix components, fibronectin, proteoglycan, collagen, elastin, TGIFs, and SMAD7. In embodiments, the anti-fibrotic protein is a peroxisome proliferator-activated receptor (PPAR) or may include one or more PPARs. In some embodiments, the protein is PPAR α , PPAR γ is a dual PPAR α/γ . Derosa et al., "The role of various peroxisome proliferator-activated receptors and their ligands in clinical practice" January 18, 2017 J. Cell. Phys. 223:1 153-161.

Proteins that promote tissue regeneration and/or transplant survival functions

[0233] In certain example embodiments, the one or more cargo polypeptides may comprise a proteins that proteins that promote tissue regeneration and/or transplant survival functions. In some cases, such proteins may induce and/or up-regulate the expression of genes for pancreatic β cell regeneration. In some cases, the proteins that promote transplant survival and functions include the products of genes for pancreatic β cell regeneration. Such genes may include proislet peptides that are proteins or peptides derived from such proteins that stimulate islet cell neogenesis. Examples of genes for pancreatic β cell regeneration include Reg1, Reg2, Reg3, Reg4, human proislet peptide, parathyroid hormone-related peptide (1-36), glucagon-like peptide-1 (GLP-1), extendin-4, prolactin, Hgf, Igf-1, Gip-1, adipisin, resistin, leptin, IL-6, IL-10, Pdx1, Ptfa1, Mafa, Pax6, Pax4, Nkx6.1, Nkx2.2, PDGF, vglycin, placental lactogens (somatomammotropins, e.g., CSH1, CHS2), isoforms thereof, homologs thereof, and orthologs thereof. In certain embodiments, the protein promoting pancreatic B cell regeneration is a cytokine, myokine, and/or adipokine.

Hormones

[0234] In certain embodiments, the one or more cargo polypeptides may comprise one or more hormones. The term "hormone" refers to polypeptide hormones, which are generally secreted by glandular organs with ducts. Hormones include proteins from natural sources or from recombinant cell culture and biologically active equivalents of the native sequence hormone, including synthetically produced small-molecule entities and pharmaceutically acceptable derivatives and salts thereof. Included among the hormones are, for example, growth hormone such as human growth hormone, N-methionyl human growth hormone, and bovine growth hormone; parathyroid hormone; thyroxine; insulin; proinsulin; relaxin; prorelaxin; glycoprotein hormones such as follicle stimulating hormone (FSH), thyroid stimulating hormone (TSH), and luteinizing hormone (LH); prolactin, placental lactogen, mouse gonadotropin-associated peptide, inhibin; activin; mullerian-inhibiting substance; and thrombopoietin, growth hormone (GH), adrenocorticotrophic hormone (ACTH), dehydroepiandrosterone

(DHEA), cortisol, epinephrine, thyroid hormone, estrogen, progesterone, placental lactogens (somatomammotropins, e.g., CSH1, CHS2), testosterone, and neuroendocrine hormones. In certain examples, the hormone is secreted from pancreas, e.g., insulin, glucagon, somatostatin, pancreatic polypeptide and ghrelin. In some examples, the hormone is insulin.

[0235] Hormones herein may also include growth factors, e.g., fibroblast growth factor (FGF) family, bone morphogenic protein (BMP) family, platelet derived growth factor (PDGF) family, transforming growth factor beta (TGFbeta) family, nerve growth factor (NGF) family, epidermal growth factor (EGF) family, insulin related growth factor (IGF) family, hepatocyte growth factor (HGF) family, hematopoietic growth factors (HeGFs), platelet-derived endothelial cell growth factor (PD-ECGF), angiopoietin, vascular endothelial growth factor (VEGF) family, and glucocorticoids. In a particular embodiment, the hormone is insulin or incretins such as exenatide, GLP-1.

Neurohormones

[0236] In embodiments, the cargo is a neurohormone, which is a hormone produced and released by neuroendocrine cells. In some embodiments, the neurohormone is a polypeptide. Example neurohormones include Thyrotropin-releasing hormone, Corticotropin-releasing hormone, Histamine, Growth hormone-releasing hormone, Somatostatin, Gonadotropin-releasing hormone, Serotonin, Dopamine, Neurotensin, Oxytocin, Vasopressin, Epinephrine, and Norepinephrine.

Anti-microbial Proteins

[0237] In some embodiments, the one or more polypeptides may comprise one or more anti-microbial proteins. In embodiments where the cell is mammalian cell, human host defense antimicrobial peptides and proteins (AMPs) play a critical role in warding off invading microbial pathogens. In certain embodiments, the anti-microbial is α -defensin HD-6, HNP-1 and β -defensin hBD-3, lysozyme, cathelicidin LL-37, C-type lectin RegIIIalpha, for example. See, e.g., Wang, "Human Antimicrobial Peptide and Proteins" *Pharma*, May 2014, 7(5): 545-594, incorporated herein by reference.

Anti-fibrillating Proteins

[0238] In certain example embodiments, the one or more polypeptides may comprise one or more anti-fibrillating polypeptides. The anti-fibrillating polypeptide can be the secreted polypeptide. In some embodiments, the anti-fibrillating polypeptide is co-expressed with one or more other polynucleotides and/or polypeptides described elsewhere herein. The anti-fibrillating agent can be secreted and act to inhibit the fibrillation and/or aggregation of endogenous proteins and/or exogenous proteins that it may be co-expressed with. In some embodiments, the anti-fibrillating agent is P4 (VITYF), P5 (VVVVV), KR7 (KPWWPRR), NK9 (NIVNVSLVK), iAb5p (Leu-Pro-Phe-Phe-Asp), KLVF and derivatives thereof, indolicidin, carnosine, a hexapeptide as set forth in Wang et al. 2014. *ACS Chem Neurosci*. 5:972-981, alpha sheet peptides having alternating D-amino acids and L-amino acids as set forth in Hopping et al. 2014. *Elife* 3:e01681, D-(PGKLVYA), RI-OR2-TAT, cyclo(17, 21)-(Lys17, Asp21)A₍₁₋₂₈₎, SEN304, SEN1576, D3, R8-A β (25-35), human γ D-crystallin (HGD), poly-lysine, heparin, poly-Asp, poly-GI, poly-L-lysine, poly-L-glutamic acid, LVEALYL, RGFFYT, a peptide set forth or as designed/generated by the method set forth in US Pat. No. 8,754,034, and combinations thereof. In aspects, the anti-fibrillating agent is a D-peptide. In aspects, the anti-fibrillating agent is an L-peptide. In aspects, the anti-fibrillating agent is a retro-inverso modified peptide. Retro-inverso modified peptides are derived from peptides by

substituting the L-amino acids for their D-counterparts and reversing the sequence to mimic the original peptide since they retain the same spatial positioning of the side chains and 3D structure. In aspects, the retro-inverso modified peptide is derived from a natural or synthetic A β peptide. In some embodiments, the polynucleotide encodes a fibrillation resistant protein. In some embodiments, the fibrillation resistant protein is a modified insulin, see e.g., U.S. Pat. No.: 8,343,914.

Antibodies

[0239] In some embodiments, the one or more cargo polypeptides may be or comprise one or more antibodies. The term "antibody" is used interchangeably with the term "immunoglobulin" throughout the specification herein, and includes intact antibodies, fragments of antibodies, e.g., Fab, F(ab')₂ fragments, and intact antibodies and fragments that have been mutated either in their constant and/or variable region (e.g., mutations to produce chimeric, partially humanized, or fully humanized antibodies, as well as to produce antibodies with a desired trait, e.g., enhanced binding and/or reduced FcR binding). The term "fragment" refers to a part or portion of an antibody or antibody chain comprising fewer amino acid residues than an intact or complete antibody or antibody chain. Fragments can be obtained via chemical or enzymatic treatment of an intact or complete antibody or antibody chain. Fragments can also be obtained by recombinant means. Exemplary fragments include Fab, Fab', F(ab')₂, Fabc, Fd, dAb, V_{HH} and scF "and/or Fv fragments." As used herein, a preparation of antibody protein "having less than about 50% of non-antibody protein (also referred to herein as a "contaminating protein"), or of chemical precursors, is considered to be "substantially free." 40%, 30%, 20%, 10% and more preferably 5% (by dry weight), of non-antibody protein, or of chemical precursors is considered to be substantially free. When the antibody protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 30%, preferably less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume or mass of the protein preparation.

[0240] The term "antigen-binding fragment" refers to a polypeptide fragment of an immunoglobulin or antibody that binds antigen or competes with intact antibody (i.e., with the intact antibody from which they were derived) for antigen binding (i.e., specific binding). As such these antibodies or fragments thereof are included in the scope of the disclosure, provided that the antibody or fragment binds specifically to a target molecule.

[0241] It is intended that the term "antibody" encompass any Ig class or any Ig subclass (e.g., the IgG1, IgG2, IgG3, and IgG4 subclasses of IgG obtained from any source (e.g., humans and non-human primates, and in rodents, lagomorphs, caprines, bovines, equines, ovines, etc.).

[0242] The term "Ig class" or "immunoglobulin class", as used herein, refers to the five classes of immunoglobulin that have been identified in humans and higher mammals, IgG, IgM, IgA, IgD, and IgE. The term "Ig subclass" refers to the two subclasses of IgM (H and L), three subclasses of IgA (IgA1, IgA2, and secretory IgA), and four subclasses of IgG (IgG1, IgG2, IgG3, and IgG4) that have been identified in humans and higher mammals. The antibodies can exist in monomeric or polymeric form; for example, IgM antibodies exist in pentameric form, and IgA antibodies exist in monomeric, dimeric or multimeric form.

[0243] The term "IgG subclass" refers to the four subclasses of immunoglobulin class IgG - IgG1, IgG2, IgG3, and IgG4 that have been identified in humans and higher mammals by the heavy chains of the immunoglobulins, $V1 - \gamma 4$, respectively. The term "single-chain immunoglobulin" or "single-chain antibody" (used interchangeably herein) refers to a protein having a two-polypeptide chain structure consisting of a heavy and a light chain, said chains being stabilized, for example, by interchain peptide linkers, which has the ability to specifically bind antigen. The term "domain" refers to a globular region of a heavy or light chain polypeptide comprising peptide loops (e.g., comprising 3 or 4 peptide loops) stabilized, for example, by β pleated sheet and/or intrachain disulfide bond. Domains are further referred to herein as "constant" or "variable", based on the relative lack of sequence variation within the domains of various class members in the case of a "constant" domain, or the significant variation within the domains of various class members in the case of a "variable" domain. Antibody or polypeptide "domains" are often referred to interchangeably in the art as antibody or polypeptide "regions". The "constant domains" of an antibody light chain are referred to interchangeably as "light chain constant regions", "light chain constant domains", "CL regions" or "CL domains". The constant domains of an antibody heavy chain are referred to interchangeably as "heavy chain constant region", "heavy chain constant domains", "CH regions" or "CH domains". The "variable domains" of an antibody light chain are referred to interchangeably as "light chain variable regions", "light chain variable domains", "VL" regions or "VL" domains." The variable domains of an "antibody heavy" chain are referred to interchangeably as "heavy chain constant regions", "heavy chain constant domains", "VH" regions or "VH" domains.

[0244] The term "region" can also refer to a part or portion of an antibody chain or antibody chain domain (e.g., a part or portion of a heavy or light chain or a part or portion of a constant or variable domain, as defined herein), as well as more discrete parts or portions of said chains or domains. For example, light and heavy chains or light and heavy chain variable domains include "complementarity determining regions" or "CDRs" interspersed among "framework regions" or "FRs", as defined herein.

[0245] The term "conformation" refers to the tertiary structure of a protein or polypeptide (e.g., an antibody, antibody chain, domain or region thereof). For example, the phrase "light (or heavy) chain conformation" refers to the tertiary structure of a light (or heavy) chain variable region, and the phrase "antibody conformation" or "antibody fragment conformation" refers to the tertiary structure of an antibody or fragment thereof.

[0246] The term "antibody-like protein scaffolds" or "engineered protein scaffolds" broadly encompasses proteinaceous non-immunoglobulin specific-binding agents, typically obtained by combinatorial engineering (such as site-directed random mutagenesis in combination with phage display or other molecular selection techniques). Usually, such scaffolds are derived from robust and small soluble monomeric proteins (such as Kunitz inhibitors or lipocalins) or from a stably folded extra-membrane domain of a cell surface receptor (such as protein A, fibronectin or the ankyrin repeat).

[0247] Such scaffolds have been extensively reviewed in Binz et al. (Engineering novel binding proteins from nonimmunoglobulin domains. *Nat Biotechnol* 2005, 23:1257-1268), Gebauer and Skerra (Engineered protein scaffolds as next-generation antibody therapeutics. *Curr Opin Chem Biol.* 2009, 13:245-55), Gill and Damle (Biopharmaceutical drug discovery using novel protein scaffolds. *Curr Opin*

Biotechnol 2006, 17:653-658), Skerra (Engineered protein scaffolds for molecular recognition. J Mol Recognit 2000, 13:167-187), and Skerra (Alternative non-antibody scaffolds for molecular recognition. Curr Opin Biotechnol 2007, 18:295-304), and include without limitation affibodies, based on the Z-domain of staphylococcal protein A, a three-helix bundle of 58 residues providing an interface on two of its alpha-helices (Nygren, Alternative binding proteins: Affibody binding proteins developed from a small three-helix bundle scaffold. FEBS J 2008, 275:2668-2676); engineered Kunitz domains based on a small (58 residues) and robust, disulphide-crosslinked serine protease inhibitor, typically of human origin (e.g., LACI-D1), which can be engineered for different protease specificities (Nixon and Wood, Engineered protein inhibitors of proteases. Curr Opin Drug Discov Dev 2006, 9:261-268); monoclonal antibodies or adnectins based on the 10th extracellular domain of human fibronectin III (10Fn3), which adopts an Ig-like beta-sandwich fold (94 residues) with 2–3 exposed loops but lacks the central disulphide bridge (Koide and Koide, Monoclonal antibodies: antibody mimics based on the scaffold of the fibronectin type III domain. Methods Mol Biol 2007, 352:95-109); anticalins derived from the lipocalins, a diverse family of eight-stranded beta-barrel proteins (ca. 180 residues) that naturally form binding sites for small ligands by means of four structurally variable loops at the open end, which are abundant in humans, insects, and many other organisms (Skerra, Alternative binding proteins: Anticalins—harnessing the structural plasticity of the lipocalin ligand pocket to engineer novel binding activities. FEBS J 2008, 275:2677-2683); DARPins, designed ankyrin repeat domains (166 residues), which provide a rigid interface arising from typically three repeated beta-turns (Stumpp et al., DARPins: a new generation of protein therapeutics. Drug Discov Today 2008, 13:695-701); avimers (multimerized LDLR-A module) (Silverman et al., Multivalent avimer proteins evolved by exon shuffling of a family of human receptor domains. Nat Biotechnol 2005, 23:1556-1561); and cysteine-rich knottin peptides (Kolmar, Alternative binding proteins: biological activity and therapeutic potential of cystine-knot miniproteins. FEBS J 2008, 275:2684-2690).

[0248] "Specific binding" of an antibody means that the antibody exhibits appreciable affinity for a particular antigen or epitope and, generally, does not exhibit significant cross reactivity. "Appreciable" binding includes binding with an affinity of at least 25 μM . Antibodies with affinities greater than $1 \times 10^7 \text{ M}^{-1}$ (or a dissociation coefficient of $1 \mu\text{M}$ or less or a dissociation coefficient of 1 nM or less) typically bind with correspondingly greater specificity. Values intermediate of those set forth herein are also intended to be within the scope of the present disclosure and antibodies of the disclosure bind with a range of affinities, for example, 100 nM or less, 75 nM or less, 50 nM or less, 25 nM or less, for example 10 nM or less, 5 nM or less, 1 nM or less, or in embodiments 500 pM or less, 100 pM or less, 50 pM or less or 25 pM or less. An antibody that "does not exhibit significant cross reactivity" is one that will not appreciably bind to an entity other than its target (e.g., a different epitope or a different molecule). For example, an antibody that specifically binds to a target molecule will appreciably bind the target molecule but will not significantly react with non-target molecules or peptides. An antibody specific for a particular epitope will, for example, not significantly cross-react with remote epitopes on the same protein or peptide. Specific binding can be determined according to any art-recognized means for determining such binding. Preferably, specific binding is determined according to Scatchard analysis and/or competitive binding assays.

[0249] As used herein, the term "affinity" refers to the strength of the binding of a single antigen-combining site with an antigenic determinant. Affinity depends on the closeness of stereochemical fit between antibody combining sites and antigen determinants, on the size of the area of contact between them, on the distribution of charged and hydrophobic groups, etc. Antibody affinity can be measured by equilibrium dialysis or by the kinetic BIACORE™ method. The dissociation constant, K_d , and the association constant, K_a , are quantitative measures of affinity.

[0250] As used herein, the term "monoclonal antibody" refers to an antibody derived from a clonal population of antibody-producing cells (e.g., B lymphocytes or B cells) which is homogeneous in structure and antigen specificity. The term "polyclonal antibody" refers to a plurality of antibodies originating from different clonal populations of antibody-producing cells which are heterogeneous in their structure and epitope specificity, but which recognize a common antigen. Monoclonal and polyclonal antibodies may exist within bodily fluids, as crude preparations, or may be purified, as described herein.

[0251] The term "binding portion" of an antibody (or "antibody portion") includes one or more complete domains, e.g., a pair of complete domains, as well as fragments of an antibody that retain the ability to specifically bind to a target molecule. It has been shown that the binding function of an antibody can be performed by fragments of a full-length antibody. Binding fragments are produced by recombinant DNA techniques, or by enzymatic or chemical cleavage of intact immunoglobulins. Binding fragments include Fab, Fab', F(ab')₂, Fabc, Fd, dAb, Fv, single chains, single-chain antibodies, e.g., scFv, and single domain antibodies.

[0252] "Humanized" forms of non-human (e.g., murine) antibodies are chimeric antibodies that contain minimal sequence derived from non-human immunoglobulin. For the most part, humanized antibodies are human immunoglobulins (recipient antibody) in which residues from a hypervariable region of the recipient are replaced by residues from a hypervariable region of a non-human species (donor antibody) such as mouse, rat, rabbit, or nonhuman primate having the desired specificity, affinity, and capacity. In some instances, FR residues of the human immunoglobulin are replaced by corresponding non-human residues. Furthermore, humanized antibodies may comprise residues that are not found in the recipient antibody or in the donor antibody. These modifications are made to further refine antibody performance. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the hypervariable regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin sequence. The humanized antibody optionally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin.

[0253] Examples of portions of antibodies or epitope-binding proteins encompassed by the present definition include: (i) the Fab fragment, having V_L , C_L , V_H and C_{H1} domains; (ii) the Fab' fragment, which is a Fab fragment having one or more cysteine residues at the C-terminus of the C_{H1} domain; (iii) the Fd fragment having V_H and C_{H1} domains; (iv) the Fd' fragment having V_H and C_{H1} domains and one or more cysteine residues at the C-terminus of the C_{H1} domain; (v) the Fv fragment having the V_L and V_H domains of a single arm of an antibody; (vi) the dAb fragment (Ward et al., 341

Nature 544 (1989)) which consists of a V_H domain or a V_L domain that binds antigen; (vii)'isolated CDR regions or isolated CDR regions presented in a functional framework; (viii) F(ab')₂ fragments which are bivalent fragments including two Fab' fragments linked by a disulphide bridge at the hinge region; (ix) single chain antibody molecules (e.g., single chain Fv; scFv) (Bird et al., 242 Science 423 (1988); and Huston et al., 85 PNAS 5879 (1988)); (x) "diabodies" with two antigen binding sites, comprising a heavy chain variable domain (V_H) connected to a light chain variable domain (V_L) in the same polypeptide chain (see, e.g., EP 404,097; WO 93/11161; Hollinger et al., 90 PNAS 6444 (1993)); (xi) "linear antibodies" comprising a pair of tandem Fd segments (V_H-C_H1-V_H-C_H1) which, together with complementary light chain polypeptides, form a pair of antigen binding regions (Zapata et al., Protein Eng. 8(10):1057-62 (1995); and U.S. Patent No. 5,641,870).

[0254] As used herein, a "blocking" antibody or an antibody "antagonist" is one which inhibits or reduces biological activity of the antigen(s) it binds. In certain embodiments, the blocking antibodies or antagonist antibodies or portions thereof described herein completely inhibit the biological activity of the antigen(s).

[0255] Antibodies may act as agonists or antagonists of the recognized polypeptides. For example, the present disclosure includes antibodies which disrupt receptor/ligand interactions either partially or fully. The disclosure features both receptor-specific antibodies and ligand-specific antibodies. The disclosure also features receptor-specific antibodies which do not prevent ligand binding but prevent receptor activation. Receptor activation (i.e., signaling) may be determined by techniques described herein or otherwise known in the art. For example, receptor activation can be determined by detecting the phosphorylation (e.g., tyrosine or serine/threonine) of the receptor or of one of its down-stream substrates by immunoprecipitation followed by western blot analysis. In specific embodiments, antibodies are provided that inhibit ligand activity or receptor activity by at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 60%, or at least 50% of the activity in absence of the antibody.

[0256] The disclosure also features receptor-specific antibodies which both prevent ligand binding and receptor activation as well as antibodies that recognize the receptor-ligand complex. Likewise, encompassed by the disclosure are neutralizing antibodies which bind the ligand and prevent binding of the ligand to the receptor, as well as antibodies which bind the ligand, thereby preventing receptor activation, but do not prevent the ligand from binding the receptor. Further included in the disclosure are antibodies which activate the receptor. These antibodies may act as receptor agonists, i.e., potentiate or activate either all or a subset of the biological activities of the ligand-mediated receptor activation, for example, by inducing dimerization of the receptor. The antibodies may be specified as agonists, antagonists or inverse agonists for biological activities comprising the specific biological activities of the peptides disclosed herein. The antibody agonists and antagonists can be made using methods known in the art.

[0257] The antibodies as defined for the present disclosure include derivatives that are modified, i.e., by the covalent attachment of any type of molecule to the antibody such that covalent attachment does not prevent the antibody from generating an anti-idiotypic response. For example, but not by way of limitation, the antibody derivatives include antibodies that have been modified, e.g., by glycosylation,

acetylation, pegylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other protein, etc. Any of numerous chemical modifications may be carried out by known techniques, including, but not limited to specific chemical cleavage, acetylation, formylation, metabolic synthesis of tunicamycin, etc. Additionally, the derivative may contain one or more non-classical amino acids.

Nanobodies and Monobodies

[0258] In some embodiments, the one or more cargo polypeptides are nanobodies. As used throughout the specification herein, "nanobody(sies)" refers to engineered antigen binding V_HH fragments, typically of an antibody, such as an IgG. They are also referred to in the art a single domain antibodies (sdAb). Methods of designing, engineering, and producing nanobodies for specific targets and uses is generally known in the art. See e.g., S. Muyldermans. *Annu Rev Anim Biosci.* 2021 Feb 16;9:401-421; S. Muyldermans. *Annu Rev Biochem.* 2013;82:775-97; and S. Muyldermans. *FEBS J.* 2021 Apr;288(7):2084-2102.

[0259] In some embodiments, the one or more cargo polypeptides are monobodies. As used throughout the specification herein, "monobody(ies)" refers to engineered or synthetic binding proteins constructed using a fibronectin type III domain or variant thereof (see e.g., Koide et al. 1998, *J. Mol. Biol.* 284(4):1141-1151, Koide et al. 2012. *Meth. Enzymol.* 503:135-156; and Koide et al., 2012. *J. Mol. Biol.* 415(2): 393-405). Methods of designing, engineering, and producing monobodies for specific targets and uses is generally known in the art. See e.g., Sha et al., *Protein Sci.* 2017 May;26(5):910-924; *Annu Rev Pharmacol Toxicol.* 2020 Jan 6;60:391-415; and Hantschel et al., *Curr Opin Struct Biol.* 2020 Feb;60:167-17.

Protease Cleavage Sites

[0260] The one or more cargo polypeptides, as exemplified above, may comprise one or more protease cleavage sites, i.e., amino acid sequences that can be recognized and cleaved by a protease. The protease cleavage sites may be used for generating desired gene products (e.g., intact gene products without any tags or portion of other proteins). The protease cleavage site may be one end or both ends of the protein. Examples of protease cleavage sites that can be used herein include an enterokinase cleavage site, a thrombin cleavage site, a Factor Xa cleavage site, a human rhinovirus 3C protease cleavage site, a tobacco etch virus (TEV) protease cleavage site, a dipeptidyl aminopeptidase cleavage site and a small ubiquitin-like modifier (SUMO)/ubiquitin-like protein-1(ULP-1) protease cleavage site. In certain examples, the protease cleavage site comprises Lys-Arg.

Signaling Peptides

[0261] In some embodiments and as previously described the cargo polypeptide has one or more C- and/or N-terminal signaling peptides so as to direct translocation to different parts of a cell, such as a host cell or cell of the subject to which it is delivered after production by the host cell. In some embodiments, the signaling peptide can increase or improve post-translational processing (e.g., folding and/or di-sulfide bonding) and/or post-translational modifications (e.g., glycosylation) of the cargo protein during its production by a host cell. Exemplary signaling peptides are described in greater detail elsewhere herein.

Targeting Moieties

[0262] The cargo, such as any cargo polypeptides or other cargo gene products, described herein can include a targeting moiety. In some embodiments, the targeting moiety encoding polynucleotides can be included in the heterologous gene construct included in the engineered phage of the present disclosure such that the cargo polynucleotide(s) and/or gene products expressed therefrom include the targeting moiety and can be targeted to specific cells, tissues, organs, etc. In some embodiments, such as non-viral carriers, the targeting moiety can be attached to the carrier (e.g., polymer, lipid, inorganic molecule etc.) and can be capable of targeting the carrier and any attached or associated genetic modifying system polynucleotide(s) to specific cells, tissues, organs, etc. In some embodiments, the targeting moieties can target integrins on cell surfaces. Optionally, the binding affinity of the targeting moiety is in the range of 1 nM to 1 μ M.

Targeted Delivery

[0263] In some embodiments, the targeting moiety targets the cargo to a specific cell, tissue, organ, or system. Such delivery of a cargo is also referred to herein as targeted delivery. In such embodiments, the cargo can include one or more targeting moieties that can direct targeted delivery of the cargo(s). Exemplary targeting moieties include, without limitation, small molecule, polypeptide, and/or polynucleotide ligands for cell surface molecules, antibodies, affibodies, aptamers, or any combination thereof. In some embodiments a multivalent approach can be employed. Multivalent presentation or inclusion of targeting moieties (e.g., antibodies) can also increase the uptake and signaling properties of targeting moiety fragments. In some embodiments, targeted delivery can be to one cell type or to multiple cell types. Methods of a targeting moiety with a cargo as in the context of the present disclosure are generally known in the art.

[0264] In some embodiments, the targeting moiety is an aptamer. Aptamers are ssDNA or RNA oligonucleotides that impart high affinity and specific recognition of the target molecules by electrostatic interactions, hydrogen bonding and hydrophobic interactions as opposed to the Watson–Crick base pairing, which is typical for the bonding interactions of oligonucleotides. Aptamers as a targeting moiety can have advantages over antibodies: aptamers can demonstrate higher target antigen recognition as compared with antibodies; aptamers can be more stable and smaller in size as compared with antibodies; aptamers can be easily synthesized and chemically modified for molecular conjugation; and aptamers can be changed in sequence for improved selectivity and can be developed to recognize poorly immunogenic targets.

[0265] In some embodiments, the targeting moiety is or includes a cell penetrating peptide or a modified CPP that can facilitate intracellular delivery via macropinocytosis followed by endosomal escape.

[0266] In some embodiments, targeted delivery is organelle-specific targeted delivery. A cargo can include a targeting moiety that can direct organelle specific delivery, such as a nuclear localization sequence, ribosomal entry sequence, mitochondria specific moiety, and/or the like.

[0267] In some embodiments, the targeted delivery is multifunctional targeted delivery that can be accomplished by coupling more than one targeting moiety to cargo. In some embodiments, such an enhances accumulation in a desired site and/or promotes organelle-specific delivery and/or target a

particular type of cell and/or respond to the local environmental stimuli such as temperature (e.g., elevated), pH (e.g., acidic or basic), respond to targeted or localized externally applied stimuli such as a magnetic field, light, energy, heat or ultrasound and/or promote intracellular delivery of the cargo.

[0268] Exemplary targeting moieties are generally known in the art, and include without limitation, those described in e.g., in e.g., Deshpande et al, "Current trends in the use of liposomes for tumor targeting," *Nanomedicine (Lond)*. 8(9), doi:10.2217/nnm.13.118 (2013), International Patent Publication No. WO 2016/027264, Lorenzer et al, "Going beyond the liver: Progress and challenges of targeted delivery of siRNA therapeutics," *Journal of Controlled Release*, 203: 1–15 (2015); Surace et al, "Lipoplexes targeting the CD44 hyaluronic acid receptor for efficient transfection of breast cancer cells," *J. Mol Pharm* 6(4):1062-73; doi: 10.1021/mp800215d (2009); Sonoke et al, "Galactose-modified cationic liposomes as a liver-targeting delivery system for small interfering RNA," *Biol Pharm Bull*. 34(8):1338-42 (2011); Torchilin, "Antibody-modified liposomes for cancer chemotherapy," *Expert Opin. Drug Deliv.* 5 (9), 1003-1025 (2008); Manjappa et al, "Antibody derivatization and conjugation strategies: application in preparation of stealth immunoliposome to target chemotherapeutics to tumor," *J. Control. Release* 150 (1), 2-22 (2011); Sofou S "Antibody-targeted liposomes in cancer therapy and imaging," *Expert Opin. Drug Deliv.* 5 (2): 189-204 (2008); Gao J et al, "Antibody-targeted immunoliposomes for cancer treatment," *Mini. Rev. Med. Chem.* 13(14): 2026-2035 (2013); Molavi et al, "Anti-CD30 antibody conjugated liposomal doxorubicin with significantly improved therapeutic efficacy against anaplastic large cell lymphoma," *Biomaterials* 34(34):8718-25 (2013), Zhao et al., 2020. *Cell* 181:151-167, particularly at tables 1-5; Liu et al., *Front. Bioeng. Biotechnol.* 2021. 9:701504. doi: 10.3389/fbioe.2021.701504; US20210379192 (describes exemplary skeletal muscle cell targeting moieties), Snow-Lisy et al., *Drug. Deliv. Transl. Res.* 1:351(2011); US20060263336 (describes exemplary progenitor cell targeting moieties) each of which and the documents cited therein are hereby incorporated herein by reference.

[0269] Other exemplary targeting moieties are described elsewhere herein, such as epitope tags, reporter and selectable markers, and/or the like which can be configured for and/or operate in some embodiments as targeting moieties.

Pharmaceutical Formulations

[0270] Also described herein are pharmaceutical formulations that can contain an amount, effective amount, and/or least effective amount, and/or therapeutically effective amount of one or more compounds, molecules, compositions, vectors, vector systems, cells, or a combination thereof (which are also referred to as the primary active agent or ingredient elsewhere herein) described in greater detail elsewhere herein and a pharmaceutically acceptable carrier or excipient. As used herein, "pharmaceutical formulation" refers to the combination of an active agent, compound, or ingredient with a pharmaceutically acceptable carrier or excipient, making the composition suitable for diagnostic, therapeutic, or preventive use in vitro, in vivo, or ex vivo. As used herein, "pharmaceutically acceptable carrier or excipient" refers to a carrier or excipient that is useful in preparing a pharmaceutical formulation that is generally safe, non-toxic, and is neither biologically or otherwise undesirable, and includes a carrier or excipient that is acceptable for veterinary use as well as human pharmaceutical

use. A “pharmaceutically acceptable carrier or excipient” as used in the specification and claims includes both one and more than one such carrier or excipient. When present, the compound can optionally be present in the pharmaceutical formulation as a pharmaceutically acceptable salt. In some embodiments, the pharmaceutical formulation can include, such as an active ingredient, one or more engineered phages of the present disclosure. In some embodiments, the one or more engineered phages contain the same one or more heterologous genes. In some embodiments, the one or more engineered phages each contain a different one or more heterologous genes. In some embodiments, the pharmaceutical formulation contains two or more engineered phages where at least two of the two or more engineered phages contain a different one or more heterologous genes. In some embodiments, the pharmaceutical formulation contains two or more engineered phages where at least two of the two or more engineered phages contain the same one or more heterologous genes.

[0271] The pharmaceutical formulations described herein can be administered to a subject in need thereof via any suitable method or route to a subject in need thereof. Suitable administration routes can include, but are not limited to auricular (otic), buccal, conjunctival, cutaneous, dental, electro-osmosis, endocervical, endosinusal, endotracheal, enteral, epidural, extra-amniotic, extracorporeal, hemodialysis, infiltration, interstitial, intra-abdominal, intra-amniotic, intra-arterial, intra-articular, intrabiliary, intrabronchial, intrabursal, intracardiac, intracartilaginous, intracaudal, intracavernous, intracavitary, intracerebral, intracisternal, intracorneal, intracoronar (dental), intracoronary, intracorporus cavernosum, intradermal, intradiscal, intraductal, intraduodenal, intradural, intraepidermal, intraesophageal, intragastric, intragingival, intraileal, intralesional, intraluminal, intralymphatic, intramedullary, intrameningeal, intramuscular, intraocular, intraovarian, intrapericardial, intraperitoneal, intrapleural, intraprostatic, intrapulmonary, intrasinal, intraspinal, intrasynovial, intratendinous, intratesticular, intrathecal, intrathoracic, intratubular, intratumor, intratympanic, intrauterine, intravascular, intravenous, intravenous bolus, intravenous drip, intraventricular, intravesical, intravitreal, iontophoresis, irrigation, laryngeal, nasal, nasogastric, occlusive dressing technique, ophthalmic, oral, oropharyngeal, other, parenteral, percutaneous, periarticular, peridural, perineural, periodontal, rectal, respiratory (inhalation), retrobulbar, soft tissue, subarachnoid, subconjunctival, subcutaneous, sublingual, submucosal, topical, transdermal, transmucosal, transplacental, transtracheal, transtympanic, ureteral, urethral, and/or vaginal administration, and/or any combination of the above administration routes, which typically depends on the disease to be treated and/or the active ingredient(s).

[0272] Where appropriate, compounds, molecules, compositions, vectors, vector systems, cells, or a combination thereof described in greater detail elsewhere herein can be provided to a subject in need thereof as an ingredient, such as an active ingredient or agent, in a pharmaceutical formulation.

[0273] As used herein, “agent” refers to any substance, compound, molecule, and the like, which can be biologically active or otherwise can induce a biological and/or physiological effect on a subject to which it is administered to. As used herein, “active agent” or “active ingredient” refers to a substance, compound, or molecule, which is biologically active or otherwise, induces a biological or physiological effect on a subject to which it is administered to. In other words, “active agent” or “active ingredient” refers to a component or components of a composition to which the whole or part of the effect of the

composition is attributed. An agent can be a primary active agent, or in other words, the component(s) of a composition to which the whole or part of the effect of the composition is attributed. An agent can be a secondary agent, or in other words, the component(s) of a composition to which an additional part and/or other effect of the composition is attributed.

[0274] In some embodiments, a secondary active ingredient is included in the pharmaceutical formulation as a pharmaceutically acceptable salt of the secondary active ingredient. As used herein, "pharmaceutically acceptable salt" refers to any acid or base addition salt whose counter-ions are non-toxic to the subject to which they are administered in pharmaceutical doses of the salts. Suitable salts include, hydrobromide, iodide, nitrate, bisulfate, phosphate, isonicotinate, lactate, salicylate, acid citrate, tartrate, oleate, tannate, pantothenate, bitartrate, ascorbate, succinate, maleate, gentisinate, fumarate, gluconate, glucuronate, saccharate, formate, benzoate, glutamate, methanesulfonate, ethanesulfonate, benzenesulfonate, p-toluenesulfonate, camphorsulfonate, naphthalenesulfonate, propionate, malonate, mandelate, malate, phthalate, and pamoate.

Pharmaceutically Acceptable Carriers and Secondary Ingredients and Agents

[0275] The pharmaceutical formulation can include a pharmaceutically acceptable carrier. Suitable pharmaceutically acceptable carriers include, but are not limited to water, salt solutions, alcohols, gum arabic, vegetable oils, benzyl alcohols, polyethylene glycols, gelatin, carbohydrates such as lactose, amylose or starch, magnesium stearate, talc, silicic acid, viscous paraffin, perfume oil, fatty acid esters, hydroxy methylcellulose, and polyvinyl pyrrolidone, which do not deleteriously react with the active agent(s) present in the formulation.

[0276] The pharmaceutical formulations, or components thereof, can be sterilized, and if desired, mixed with agents, such as lubricants, preservatives, stabilizers, wetting agents, emulsifiers, salts for influencing osmotic pressure, buffers, coloring, flavoring and/or aromatic substances, and the like which do not deleteriously react with active agent(s) present in the formulation.

[0277] In some embodiments, the pharmaceutical formulation can also include an effective amount of secondary active agents, including but not limited to, non-biologic compounds and compositions (e.g., chemical compounds also referred to in the art as small molecule agents or small molecule therapeutic agents, minerals, vitamins, etc.), biologic agents or molecules (e.g., including, but not limited to, polynucleotides, amino acids, peptides, polypeptides, antibodies, aptamers, ribozymes, hormones). In some embodiments, the secondary active agent is an immunomodulator, antipyretic, anxiolytic, antipsychotic, analgesic, antispasmodic, anti-inflammatory agent, anti-histamine, anti-infective, chemotherapeutic, probiotic, prebiotic, non-virulent microbial cell, or any combinations thereof.

Effective Amounts

[0278] In some embodiments, the amount of the primary active agent and/or optional secondary agent can be an effective amount, least effective amount, and/or therapeutically effective amount. As used herein, "effective amount", "effective concentration", and/or the like refers to the amount, concentration, etc. of the primary and/or optional secondary agent included in the pharmaceutical formulation that achieve one or more therapeutic effects or desired effect. As used herein, "least effective", "least effective concentration", and/or the like amount refers to the lowest amount,

concentration, etc. of the primary and/or optional secondary agent that achieves the one or more therapeutic or other desired effects. As used herein, "therapeutically effective amount", "therapeutically effective concentration" and/or the like refers to the amount, concentration, etc. of the primary and/or optional secondary agent included in the pharmaceutical formulation that achieves one or more therapeutic effects. In some embodiments, the one or more therapeutic effects are those attributed to the delivered protein.

[0279] The effective amount, least effective amount, and/or therapeutically effective amount of the primary and optional secondary active agent described elsewhere herein contained in the pharmaceutical formulation can be any non-zero amount ranging from about 0 to 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 pg, ng, µg, mg, or g or be any numerical value or subrange within any of these ranges.

[0280] In some embodiments, the effective amount, least effective amount, and/or therapeutically effective amount of the primary and/or optional secondary active agent is an effective concentration, least effective concentration, and/or therapeutically effective concentration, which can each be any non-zero amount ranging from about 0 to 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 pM, nM, µM, mM, or M or be any numerical value or subrange within any of these ranges.

[0281] In other embodiments, the effective amount, least effective amount, and/or therapeutically effective amount of the primary and/or optional secondary active agent be any non-zero amount ranging from about 0 to 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 IU or be any numerical value or subrange within any of these ranges.

[0282] In some embodiments, the primary and/or the optional secondary active agent present in the pharmaceutical formulation can be any non-zero amount ranging from about 0 to 0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19, 0.2, 0.21, 0.22, 0.23, 0.24, 0.25, 0.26, 0.27, 0.28, 0.29, 0.3, 0.31, 0.32, 0.33, 0.34, 0.35, 0.36, 0.37, 0.38, 0.39, 0.4, 0.41, 0.42, 0.43, 0.44, 0.45, 0.46, 0.47, 0.48, 0.49, 0.5, 0.51, 0.52, 0.53, 0.54, 0.55, 0.56, 0.57, 0.58, 0.59, 0.6, 0.61, 0.62, 0.63, 0.64, 0.65, 0.66, 0.67, 0.68, 0.69, 0.7, 0.71, 0.72, 0.73, 0.74, 0.75, 0.76, 0.77, 0.78, 0.79, 0.8, 0.81, 0.82,

0.83, 0.84, 0.85, 0.86, 0.87, 0.88, 0.89, 0.9, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.9, to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9 % w/w, v/v, or w/v of the pharmaceutical formulation or be any numerical value or subrange within any of these ranges.

[0283] In some embodiments where a cell or cell population is present in the pharmaceutical formulation (e.g., as a primary and/or or secondary active agent), the effective amount of cells is an amount ranging from about 1 or 2 cells to 1×10^1 cells /mL, 1×10^{20} cells /mL or more, such as about 1×10^1 cells /mL, 1×10^2 cells /mL, 1×10^3 cells /mL, 1×10^4 cells /mL, 1×10^5 cells /mL, 1×10^6 cells /mL, 1×10^7 cells /mL, 1×10^8 cells /mL, 1×10^9 cells /mL, 1×10^{10} cells /mL, 1×10^{11} cells /mL, 1×10^{12} cells /mL, 1×10^{13} cells /mL, 1×10^{14} cells /mL, 1×10^{15} cells /mL, 1×10^{16} cells /mL, 1×10^{17} cells /mL, 1×10^{18} cells /mL, 1×10^{19} cells /mL, to/or about 1×10^{20} cells/mL or any numerical value or subrange within any of these ranges.

[0284] In some embodiments, the amount or effective amount, particularly where an infective particle or virus or phage is being delivered (e.g., an engineered phage or phage particle of the present disclosure), the effective amount of virus particles can be expressed as a titer (plaque forming units per unit of volume) or as a MOI (multiplicity of infection). In some embodiments, the effective amount can be about 1×10^1 particles per pL, nL, μ L, mL, or L to 1×10^{20} particles per pL, nL, μ L, mL, or L or more, such as about 1×10^1 , 1×10^2 , 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^7 , 1×10^8 , 1×10^9 , 1×10^{10} , 1×10^{11} , 1×10^{12} , 1×10^{13} , 1×10^{14} , 1×10^{15} , 1×10^{16} , 1×10^{17} , 1×10^{18} , 1×10^{19} , to/or about 1×10^{20} particles per pL, nL, μ L, mL, or L. In some embodiments, the effective titer can be about 1×10^1 transforming units per pL, nL, μ L, mL, or L to 1×10^{20} transforming units per pL, nL, μ L, mL, or L or more, such as about 1×10^1 , 1×10^2 , 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^7 , 1×10^8 , 1×10^9 , 1×10^{10} , 1×10^{11} , 1×10^{12} , 1×10^{13} , 1×10^{14} , 1×10^{15} , 1×10^{16} , 1×10^{17} , 1×10^{18} , 1×10^{19} , to/or about 1×10^{20} transforming units per pL, nL, μ L, mL, or L or any numerical value or subrange within these ranges. In some embodiments, the MOI of the pharmaceutical formulation can range from about 0.1 to 10 or more, such as 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, 6, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, 7, 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 8, 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9, 9, 9.1, 9.2, 9.3, 9.4, 9.5, 9.6, 9.7, 9.8, 9.9, 10 or more or any numerical value or subrange within these ranges.

[0285] In some embodiments, the amount or effective amount of the one or more of the active agent(s) described herein contained in the pharmaceutical formulation can range from about 1 pg/kg to about 10 mg/kg based upon the bodyweight of the subject in need thereof or average bodyweight of the specific patient population to which the pharmaceutical formulation can be administered.

[0286] In embodiments where there is a secondary agent contained in the pharmaceutical formulation, the effective amount of the secondary active agent will vary depending on the secondary

agent, the primary agent, the administration route, subject age, disease, stage of disease, among other things, which will be one of ordinary skill in the art.

[0287] When optionally present in the pharmaceutical formulation, the secondary active agent can be included in the pharmaceutical formulation or can exist as a stand-alone compound or pharmaceutical formulation that can be administered contemporaneously or sequentially with the compound, derivative thereof, or pharmaceutical formulation thereof.

[0288] In some embodiments, the effective amount of the secondary active agent, when optionally present, is any non-zero amount ranging from about 0 to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9 % w/w, v/v, or w/v of the total active agents present in the pharmaceutical formulation or any numerical value or subrange within these ranges. In additional embodiments, the effective amount of the secondary active agent is any non-zero amount ranging from about 0 to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9 % w/w, v/v, or w/v of the total pharmaceutical formulation or any numerical value or subrange within these ranges.

Dosage Forms

[0289] In some embodiments, the pharmaceutical formulations described herein can be provided in a dosage form. The dosage form can be administered to a subject in need thereof. The dosage form can be effective generate specific concentration, such as an effective concentration, at a given site in the subject in need thereof. As used herein, "dose," "unit dose," or "dosage" can refer to physically discrete units suitable for use in a subject, each unit containing a predetermined quantity of the primary active agent, and optionally present secondary active ingredient, and/or a pharmaceutical formulation thereof calculated to produce the desired response or responses in association with its administration. In some embodiments, the given site is proximal to the administration site. In some embodiments, the given site is distal to the administration site. In some cases, the dosage form contains a greater amount of one or more of the active ingredients present in the pharmaceutical formulation than the final intended amount needed to reach a specific region or location within the subject to account for loss of the active components such as via first and second pass metabolism.

[0290] The dosage forms can be adapted for administration by any appropriate route. Appropriate routes include, but are not limited to, oral (including buccal or sublingual), rectal, intraocular, inhaled, intranasal, topical (including buccal, sublingual, or transdermal), vaginal, parenteral, subcutaneous, intramuscular, intravenous, internasal, and intradermal. Other appropriate routes are described elsewhere herein. Such formulations can be prepared by any method known in the art. In some embodiments, the dosage form is adapted for administration to an area or part of a subject having a microbiome.

[0291] Dosage forms adapted for oral administration can discrete dosage units such as capsules, pellets or tablets, powders or granules, solutions, or suspensions in aqueous or non-aqueous liquids; edible foams or whips, or in oil-in-water liquid emulsions or water-in-oil liquid emulsions. In some embodiments, the pharmaceutical formulations adapted for oral administration also include one or more agents which flavor, preserve, color, or help disperse the pharmaceutical formulation. Dosage forms prepared for oral administration can also be in the form of a liquid solution that can be delivered as a foam, spray, or liquid solution. The oral dosage form can be administered to a subject in need thereof. Where appropriate, the dosage forms described herein can be microencapsulated.

[0292] The dosage form can also be prepared to prolong or sustain the release of any ingredient. In some embodiments, compounds, molecules, compositions, vectors, vector systems, cells, or a combination thereof described herein can be the ingredient whose release is delayed. In some embodiments the primary active agent is the ingredient whose release is delayed. In some embodiments, an optional secondary agent can be the ingredient whose release is delayed. Suitable methods for delaying the release of an ingredient include, but are not limited to, coating or embedding the ingredients in material in polymers, wax, gels, and the like. Delayed release dosage formulations can be prepared as described in standard references such as "Pharmaceutical dosage form tablets," eds. Liberman et. al. (New York, Marcel Dekker, Inc., 1989), "Remington - The science and practice of pharmacy", 20th ed., Lippincott Williams & Wilkins, Baltimore, MD, 2000, and "Pharmaceutical dosage forms and drug delivery systems", 6th Edition, Ansel et al., (Media, PA: Williams and Wilkins, 1995). These references provide information on excipients, materials, equipment, and processes for preparing tablets and capsules and delayed release dosage forms of tablets and pellets, capsules, and granules. The delayed release can be anywhere from about an hour to about 3 months or more.

[0293] Examples of suitable coating materials include, but are not limited to, cellulose polymers such as cellulose acetate phthalate, hydroxypropyl cellulose, hydroxypropyl methylcellulose, hydroxypropyl methylcellulose phthalate, and hydroxypropyl methylcellulose acetate succinate; polyvinyl acetate phthalate, acrylic acid polymers and copolymers, and methacrylic resins that are commercially available under the trade name EUDRAGIT® (Roth Pharma, Westerstadt, Germany), zein, shellac, and polysaccharides.

[0294] Coatings may be formed with a different ratio of water-soluble polymer, water insoluble polymers, and/or pH dependent polymers, with or without water insoluble/water soluble non-polymeric excipient, to produce the desired release profile. The coating is either performed on the dosage form (matrix or simple) which includes, but is not limited to, tablets (compressed with or without coated beads), capsules (with or without coated beads), beads, particle compositions, "ingredient as is" formulated as, but not limited to, suspension form or as a sprinkle dosage form.

[0295] Where appropriate, the dosage forms described herein can be a liposome. In these embodiments, primary active ingredient(s), and/or optional secondary active ingredient(s), and/or pharmaceutically acceptable salt thereof where appropriate are incorporated into a liposome. In embodiments where the dosage form is a liposome, the pharmaceutical formulation is thus a liposomal formulation. The liposomal formulation can be administered to a subject in need thereof.

[0296] Dosage forms adapted for topical administration can be formulated as ointments, creams, suspensions, lotions, powders, solutions, pastes, gels, sprays, aerosols, or oils. In some embodiments for treatments of the eye or other external tissues, for example the mouth or the skin, the pharmaceutical formulations are applied as a topical ointment or cream. When formulated in an ointment, a primary active ingredient, optional secondary active ingredient, and/or pharmaceutically acceptable salt thereof where appropriate can be formulated with a paraffinic or water-miscible ointment base. In other embodiments, the primary and/or secondary active ingredient can be formulated in a cream with an oil-in-water cream base or a water-in-oil base. Dosage forms adapted for topical administration in the mouth include lozenges, pastilles, and mouth washes.

[0297] Dosage forms adapted for nasal or inhalation administration include aerosols, solutions, suspension drops, gels, or dry powders. In some embodiments, a primary active ingredient, optional secondary active ingredient, and/or pharmaceutically acceptable salt thereof where appropriate can be in a dosage form adapted for inhalation is in a particle-size-reduced form that is obtained or obtainable by micronization. In some embodiments, the particle size of the size reduced (e.g., micronized) compound or salt or solvate thereof, is defined by a D_{50} value of about 0.5 to about 10 microns as measured by an appropriate method known in the art. Dosage forms adapted for administration by inhalation also include particle dusts or mists. Suitable dosage forms wherein the carrier or excipient is a liquid for administration as a nasal spray or drops include aqueous or oil solutions/suspensions of an active (primary and/or secondary) ingredient, which may be generated by various types of metered dose pressurized aerosols, nebulizers, or insufflators. The nasal/inhalation formulations can be administered to a subject in need thereof.

[0298] In some embodiments, the dosage forms are aerosol formulations suitable for administration by inhalation. In some of these embodiments, the aerosol formulation contains a solution or fine suspension of a primary active ingredient, secondary active ingredient, and/or pharmaceutically acceptable salt thereof where appropriate and a pharmaceutically acceptable aqueous or non-aqueous solvent. Aerosol formulations can be presented in single or multi-dose quantities in sterile form in a sealed container. For some of these embodiments, the sealed container is a single dose or multi-dose nasal or an aerosol dispenser fitted with a metering valve (e.g., metered dose inhaler), which is intended for disposal once the contents of the container have been exhausted.

[0299] Where the aerosol dosage form is contained in an aerosol dispenser, the dispenser contains a suitable propellant under pressure, such as compressed air, carbon dioxide, or an organic propellant, including but not limited to a hydrofluorocarbon. The aerosol formulation dosage forms in other embodiments are contained in a pump-atomizer. The pressurized aerosol formulation can also contain a solution or a suspension of a primary active ingredient, optional secondary active ingredient, and/or pharmaceutically acceptable salt thereof. In further embodiments, the aerosol formulation also contains co-solvents and/or modifiers incorporated to improve, for example, the stability and/or taste and/or fine particle mass characteristics (amount and/or profile) of the formulation. Administration of the aerosol formulation can be once daily or several times daily, for example 2, 3, 4, or 8 times daily, in which 1, 2, 3 or more doses are delivered each time. The aerosol formulations can be administered to a subject in need thereof.

[0300] For some dosage forms suitable and/or adapted for inhaled administration, the pharmaceutical formulation is a dry powder inhalable-formulations. In addition to a primary active agent, optional secondary active ingredient, and/or pharmaceutically acceptable salt thereof where appropriate, such a dosage form can contain a powder base such as lactose, glucose, trehalose, mannitol, and/or starch. In some of these embodiments, a primary active agent, secondary active ingredient, and/or pharmaceutically acceptable salt thereof where appropriate is in a particle-size reduced form. In further embodiments, a performance modifier, such as L-leucine or another amino acid, cellobiose octaacetate, and/or metals salts of stearic acid, such as magnesium or calcium stearate. In some embodiments, the aerosol formulations are arranged so that each metered dose of aerosol contains a predetermined amount of an active ingredient, such as the one or more of the compositions, compounds, vector(s), molecules, cells, and combinations thereof described herein.

[0301] Dosage forms adapted for vaginal administration can be presented as pessaries, tampons, creams, gels, pastes, foams, or spray formulations. Dosage forms adapted for rectal administration include suppositories or enemas. The vaginal formulations can be administered to a subject in need thereof.

[0302] Dosage forms adapted for parenteral administration and/or adapted for injection can include aqueous and/or non-aqueous sterile injection solutions, which can contain antioxidants, buffers, bacteriostats, solutes that render the composition isotonic with the blood of the subject, and aqueous and non-aqueous sterile suspensions, which can include suspending agents and thickening agents. The dosage forms adapted for parenteral administration can be presented in a single-unit dose or multi-unit dose containers, including but not limited to sealed ampoules or vials. The doses can be lyophilized and re-suspended in a sterile carrier to reconstitute the dose prior to administration. Extemporaneous injection solutions and suspensions can be prepared in some embodiments, from sterile powders, granules, and tablets. The parenteral formulations can be administered to a subject in need thereof.

[0303] For some embodiments, the dosage form contains a predetermined amount of a primary active agent, secondary active ingredient, and/or pharmaceutically acceptable salt thereof where appropriate per unit dose. In an embodiment, the predetermined amount of primary active agent, secondary active ingredient, and/or pharmaceutically acceptable salt thereof where appropriate can be an effective amount, a least effect amount, and/or a therapeutically effective amount. In other embodiments, the predetermined amount of a primary active agent, secondary active agent, and/or pharmaceutically acceptable salt thereof where appropriate, can be an appropriate fraction of the effective amount of the active ingredient.

Co-Therapies and Combination Therapies

[0304] In some embodiments, the pharmaceutical formulation(s) described herein are part of a combination treatment or combination therapy. The combination treatment can include the pharmaceutical formulation described herein and an additional treatment modality. The additional treatment modality can be a chemotherapeutic, a biological therapeutic, surgery, radiation, diet modulation, environmental modulation, a physical activity modulation, and combinations thereof.

[0305] In some embodiments, the co-therapy or combination therapy can additionally include one or more active agents, including but not limited to, non-biologic compounds and compositions (e.g.,

chemical compounds also referred to in the art as small molecule agents or small molecule therapeutic agents, minerals, vitamins, etc.), biologic agents or molecules (e.g., including, but not limited to, polynucleotides, amino acids, peptides, polypeptides, antibodies, aptamers, ribozymes, hormones). In some embodiments, the secondary active agent is an immunomodulator, antipyretic, anxiolytic, antipsychotic, analgesic, antispasmodic, anti-inflammatory agent, anti-histamine, anti-infective, chemotherapeutic, probiotic, prebiotic, non-virulent microbial cell, or any combinations thereof.

Administration of the Pharmaceutical Formulations

[0306] The pharmaceutical formulations or dosage forms thereof described herein can be administered one or more times hourly, daily, monthly, or yearly (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more times hourly, daily, monthly, or yearly). In some embodiments, the pharmaceutical formulations or dosage forms thereof described herein can be administered continuously over a period of time ranging from minutes to hours to days. Devices and dosages forms are known in the art and described herein that are effective to provide continuous administration of the pharmaceutical formulations described herein. In some embodiments, the first one or a few initial amount(s) administered can be a higher dose than subsequent doses. This is typically referred to in the art as a loading dose or doses and a maintenance dose, respectively. In some embodiments, the pharmaceutical formulations can be administered such that the doses over time are tapered (increased or decreased) overtime so as to wean a subject gradually off of a pharmaceutical formulation or gradually introduce a subject to the pharmaceutical formulation.

[0307] As previously discussed, the pharmaceutical formulation can contain a predetermined amount of a primary active agent, secondary active agent, and/or pharmaceutically acceptable salt thereof where appropriate. In some of these embodiments, the predetermined amount can be an appropriate fraction of the effective amount of the active ingredient. Such unit doses may therefore be administered once or more than once a day, month, or year (e.g., 1, 2, 3, 4, 5, 6, or more times per day, month, or year). Such pharmaceutical formulations may be prepared by any of the methods well known in the art.

[0308] Where co-therapies or multiple pharmaceutical formulations are to be delivered to a subject, the different therapies or formulations can be administered sequentially or simultaneously. Sequential administration is administration where an appreciable amount of time occurs between administrations, such as more than about 15, 20, 30, 45, 60 minutes or more. The time between administrations in sequential administration can be on the order of hours, days, months, or even years, depending on the active agent present in each administration. Simultaneous administration refers to administration of two or more formulations at the same time or substantially at the same time (e.g., within seconds or just a few minutes apart), where the intent is that the formulations be administered together at the same time.

Vaccines

[0309] In some embodiments, the engineered phage or pharmaceutical formulation thereof is a vaccine. In some embodiments, the gene product (e.g., protein or nucleic acid) delivered by the engineered phage is an antigen or antigenic component. Antigenic components include components that specifically trigger the immune response against the antigen or antigens from which the antigenic

components are derived. In some embodiments, delivery of an engineered phage, such as one where the protein delivered by the engineered phage is an antigen, is capable of stimulating an immune response in a subject to which the engineered phage is delivered. In some embodiments, the immune response is an adaptive immune response. In some embodiments, the immune response is a B-cell and/or T-cell immune response. In some embodiments, the vaccine comprises or is co-delivered with one or more adjuvants. The vaccine can be in any appropriate dosage form as previously described.

Methods

[0310] The engineered phages described herein can be used for delivering proteins to a subject in need thereof. In some embodiments, the engineered phages can infect one or more microbes present in or the subject. Without being bound by theory, infection of the host cell results in the host cell producing the gene product(s) (e.g., nucleic acid(s) and/or protein(s)) encoded by the heterologous gene(s) in the engineered phage. In some embodiments, such as when the phage is a lytic (or virulent) phage, the protein produced is released into the subject when the host cell is lysed thereby delivering the protein to the subject. In other embodiments, other secretory pathways of the host cell can result in secretion of the produced protein into the subject, thereby delivering the protein to the subject.

[0311] Described in certain example embodiments herein are methods of delivering a protein to a subject in need thereof and/or treating a disease or a symptom thereof in the subject in need thereof, the method comprising administering an amount of the engineered phage, pharmaceutical formulation, or vaccine as in any one of the preceding claims to the subject in need thereof, whereby after administration the phage infects a host cell of the subject in need thereof, stimulates production of the one or more heterologous gene products (e.g., nucleic acids and/or proteins) by the host cell, and lyses the host cell so as to release the one or more heterologous gene products produced by the host cell thereby delivering the one or more heterologous gene products to the subject in need thereof.

[0312] In certain example embodiments, the host cell is a bacterial cell or a yeast cell. In certain example embodiments, the host cell is part of a microbiome of the subject in need thereof. In certain example embodiments, the microbiome is a gastrointestinal microbiome, a skin microbiome, an airway microbiome (e.g., airway microbiome, tracheal microbiome, sinus microbiome, lung microbiome, or any combination thereof), a vaginal microbiome, or any combination thereof. In certain example embodiments, the gastrointestinal microbiome is an oral cavity microbiome, a stomach microbiome, a small intestine microbiome, a large intestine microbiome, an appendix microbiome, a cecum microbiome, or any combination thereof. In certain example embodiments, administering is oral administration. In some embodiments, administering is topical administration.

[0313] In some embodiments, the host cell is a bacterial cell. In some embodiments, the bacterial cell is an *Escherichia coli* (*E. coli*). In some embodiments, the *E. coli* is *E. coli* B, *E. coli* K-12, *E. coli* C, *E. coli* W3110, *E. coli* BL21. In some embodiments, the bacterial cell is pathogenic to the subject. In some embodiments, the bacterial cell is a commensal bacteria cell. In some embodiments, the bacterial cell is a *Salmonella*, *Shigella*, *Campylobacter*, *E. coli*, or *Yersinia* cell. In some embodiments, the bacterial cell is an *E. coli* or a *Bifidobacterium*. In some embodiments, the bacteria cell is *C. difficile*.

Exemplary microorganisms present in various microbiomes are generally known and will be appreciated by those of skill in the art. Any such cells can be host cells in accordance with the present disclosure.

[0314] In some embodiments, the host cell is a yeast cell. In some embodiments, the yeast cell is pathogenic to the subject. In some embodiments, the yeast cell is a commensal yeast cell. In some embodiments, the yeast cell is a yeast cell of the genus *Candida* or *Saccharomyces*. In some embodiments, the *Candida* is *Candida albicans*. In some embodiments, the *Saccharomyces* is *S. boulardii* or *S. cerevisiae*.

[0315] In some embodiments, administration is one or more, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more times per day, week, month, or year. In some embodiments, the engineered phage or formulation thereof (e.g., a pharmaceutical formulation and/or vaccine) is co-administered with one or more additional agents, such as one or more therapeutic agents and/or agents that modify (e.g., enhance, regulate, etc.) the activity of the engineered phage, one or more proteins delivered by the phage, activity/production of the host cell, etc. The one or more additional active agents can be administered simultaneously with the engineered phage, before or after administration with the engineered phage, or both.

[0316] The methods, such as where the engineered phage is designed to deliver an antigen or antigenic component) can include administering a booster agent in addition to the administration of the composition therein. A booster agent can be an extra administration of the engineered phage that is designed to deliver an antigen or antigenic component) herein or a different agent. A booster (or booster vaccine) may be given after an earlier administration of the engineered phage or formulation thereof. The time of administration between the initial administration of the composition and the booster may be at least 1 minute, at least 5 minutes, at least 10 minutes, at least 20 minutes, at least 30 minutes, at least 40 minutes, at least 50 minutes, at least 1 hour, at least 2 hours, at least 4 hours, at least 8 hours, at least 12 hours, at least 1 day, at least 1 week, at least 2 week, at least 3 week, at least 1 month, at least 2 months, at least 3 months, at least 4 months, at least 5 months, at least 6 months, at least 1 year, at least 5 years, at least 10 years, and any time period in-between.

[0317] Described herein are methods of engineering a microbiome of a subject, that includes delivering an engineered phage of the present disclosure, whereby the engineered phage and/or gene product(s) delivered by the engineered phage is effective to kill or modulate one or more activities (e.g., growth, reproduction, cell product production, etc.) and/or characteristics of one or more species, subspecies, and/or strains of a microbiome of a subject. In certain example embodiments, the microbiome is a gastrointestinal microbiome, a skin microbiome, an airway microbiome, a vaginal microbiome, or any combination thereof. In certain example embodiments, the gastrointestinal microbiome is an oral cavity microbiome, a stomach microbiome, a small intestine microbiome, a large intestine microbiome, an appendix microbiome, a cecum microbiome, or any combination thereof. In certain example embodiments, the engineered microbiome has an increased amount of one or more species, subspecies, or strains of a desired species, strain, or subspecies of microbes as compared to the microbiome prior to administration of the engineered phage(s). In some embodiments, the engineered microbiome has one or more altered functions or effects (e.g., reduction in disease or

modulation of one or more biological responses in the subject) on the subject in need thereof, such as satiety, inflammation, altered mood and/or cognitive function, etc.

[0318] Further embodiments are illustrated in the following Examples which are given for illustrative purposes only and are not intended to limit the scope of the disclosure.

EXAMPLES

[0319] Now having described the embodiments of the present disclosure, in general, the following Examples describe some additional embodiments of the present disclosure. While embodiments of the present disclosure are described in connection with the following examples and the corresponding text and figures, there is no intent to limit embodiments of the present disclosure to this description. On the contrary, the intent is to cover all alternatives, modifications, and equivalents included within the spirit and scope of embodiments of the present disclosure. The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to perform the methods and use the probes disclosed and claimed herein. Efforts have been made to ensure accuracy with respect to numbers (e.g., amounts, temperature, etc.), but some errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, temperature is in °C, and pressure is at or near atmospheric. Standard temperature and pressure are defined as 20 °C and 1 atmosphere.

Example 1

Introduction

[0320] The gastrointestinal tract (GIT) performs important functions in the human body. It facilitates the digestion of food and absorption of nutrients, supports the colonization of commensal microbes, and acts as an immunological interface to the external environment. Unsurprisingly, drugs given orally is an ideal route for administration because of its non-invasive nature. Unfortunately, oral administration remains infeasible for many types of compounds. If the target is in the intestine, the drug must be able to survive the acidic and proteolytic conditions of the upper GIT as well as potential degradation by microbial proteases. For biological therapeutics that are sensitive to denaturation and digestion, the upper GIT is an especially challenging obstacle. If the drug's destination is systemic circulation, then it must also be absorbed at biologically significant quantities, which requires specific physicochemical properties allowing it to traverse the intestinal epithelia. Even if a drug and/or its formulation is capable of overcoming these challenges, frequent dosing (e.g., daily) is associated with noncompliance, and has been shown in half of chronically ill patients not adhering to the prescribed dosing regimen.

[0321] One strategy to overcoming the degradative conditions of the upper GIT is to orally administer biologic-producing microbes. These microbes have greater resilience to the degradative conditions of the upper GI tract and can produce the biologic in situ. Examples of engineered bacteria include *Lactococcus lactis* producing IL-10 to treat colitis and *E. coli* strains producing antimicrobial peptides to clear vancomycin resistant *Enterococcus*, dispersin B to degrade *P. aeruginosa* biofilms, anorexigenic lipid precursors to reduce weight gain, or phenylalanine-degrading enzymes to treat

phenylketonuria. Often, the desired biologic does not occur naturally in the bacteria, so a suitable commensal and genetically tractable bacterium must be engineered to express and release this biologic at physiologically relevant levels. These new features can be metabolically taxing for the bacterium because the cellular resources syphoned towards biologic expression and associated secretory systems are not beneficial for fitness. In the highly competitive ecosystem of the mammalian gut, even slight reductions in metabolic efficiency can sabotage strain fitness and colonization. While one benefit to poor gut colonization is a predictable clearance by these engineered bacteria, this requires frequent administration with high bacterial concentrations for efficacy.

[0322] Obesity is a chronic disease with several co-morbidities including diabetes, cardiovascular diseases, and cancers, among others. By 2030, 51% of the population is projected to be obese. In addition to the complex etiology of obesity, high-fat diets in both humans and animal models lead to the weight gain and low-grade, chronic inflammation that is characteristic of immunologic dysregulation, including overexpression of TNF- α and numerous proinflammatory cytokines. While the resultant systemic inflammation is challenging to mitigate due to the systemic nature of the disease, a durable treatment localized directly at the source of inflammation (i.e., the GIT) could be an effective approach towards disrupting disease pathophysiology.

[0323] Bacteriophages (phages) are prokaryotic viruses that are well-known for their ability to redirect host machinery towards the expression of phage-encoded genes during the infection process. In addition to active research focused on using virulent phages as anti-bacterials ("phage therapy"), emergent work has shown that gut viruses, ~90% of which are phages, are numerically dominant elements of the gut microbiome, can outnumber their bacterial counterparts – especially within the mucosal lining of the GIT, and can persist for long periods of time. Gnotobiotic mouse models colonized with human gut commensals have shown that virulent phages can continuously propagate in the mammalian gut to reach co-existence with their host bacteria. Without being bound by theory, it could be understood that a therapeutic protein may be encoded into a virulent phage so that it is co-expressed with phage genes during the infection of a resident gut bacterium and then released into the extracellular space by phage-induced cell lysis. Further, phage-bacterial coexistence could lead to a sustained production of phage-encoded genes.

[0324] Applicant describes and demonstrates in at least this Example an alternative use of virulent phages. Instead of their conventional application to eradicate a targeted bacterial species, Applicant also envisions their use them as vectors for gene delivery in resident gut bacteria to produce biologic products that are released during the natural process of cell lysis during phage propagation. Furthermore, it may be understood that phage-bacterial coexistence may lead to a sustained production of phage-encoded genes. As provided within this Example, demonstration that genetically engineered virulent phages can be used to reprogram bacteria to express and release proteins, converting resident gut bacteria into personalized microbial factories, is further assessed. Herein, Applicant demonstrates that the virulent phage, T4, releases substantial quantities of the fluorescent marker protein, GFP, into the phage lysate when it is encoded in the host bacterium and when it is encoded in phage genome. The analysis used a mouse model to show that a single dose of phage encoding *gfp* led to significant *in situ* production of GFP in the murine mucosa. Finally, to demonstrate the feasibility of the proposed

approach to generate biologically significant amounts of protein directly in the mammalian gut, Applicant encoded expression of ClpB, which has a discontinuous 5 amino acid motif that is homologous to the anorexigenic neuropeptide α -melanocyte-stimulating hormone, and has shown to induce satiety to reduce weight gain during obesity. In a mouse model of diet-induced obesity (DIO), Applicant shows T4 phage engineered to express ClpB can coexist with the host *E. coli* in the gut for multiple weeks, and that this engineered phage reduces food consumption, increases activity, and reduces weight gain compared to mice receiving no phage or wildtype T4 phage. Collectively, the data demonstrates that virulent phages can be leveraged to express and release heterologous proteins with significant physiological effect on the mammalian host.

Materials and Methods

Molecular Cloning

[0325] Genetic constructs were designed using Benchling and assembled using Golden Gate Assembly (NEB, Catalog # E1601L) with assistance from the NEB Golden Gate Assembly Tool. Plasmids were subsequently transformed into DH5A cells (NEB, Catalog # C2987H) and verified for accuracy via Sanger sequencing (Virginia Tech Genomics Sequencing Center). For challenging plasmids, NEB Stable cells (NEB, Catalog # C3040H) were used. T4 phage promoters were identified from previous screens, and the constitutive *E. coli* promoter J23119 from past work. The fluorescent marker, *sfgfp*, was cloned from a fluorescent *E. coli* K-12 strain (EC001), and *clpB* was cloned from *E. coli* Nissle 1917 using PCR to include an N-terminal His₆-tag. DNA oligos were custom synthesized by Thermo Fisher®.

[0326] In general, genetic constructs in a pGGASelect vector contained a promoter, a ribosome-binding site, and heterologous gene (*e.g.*, *sfgfp* or *clpB*) with flanking regions of ~200 bp of homology to the *ac* gene in the T4 genome. In examples, the overall construct may comprise the following:

5' - Homology Arm 1 (HA1) --- Promoter --- RBS --- Gene --- Homology Arm 2 (HA2) - 3'

DNA sequences are provided in **Table 1**. Recombinant T4 phages were generated by spontaneous recombination, as described previously. Plasmid-containing *E. coli* was grown to log phase and then cultured with wildtype T4 phage (MOI = 0.01) overnight at 37°C. Lysates were treated with chloroform, centrifuged at 6000 rpm for 10 min at 4°C to remove the bacterial debris and sterile-filtered with a 0.45 μ m syringe filter. Disruption of the *ac* gene by insertion of the genetic construct, leads to T4 phage resistance to acriflavine. To select for recombinant phages from the crude phage lysate, 100 μ L phage lysate and 300 μ L log phase *E. coli* DH5A were mixed with 3 mL of molten LB top agar (containing 0.3% agar) at 50°C and 15 μ L of 17 mg/mL acriflavine hydrochloride, then overlaid onto LB agar plates. After overnight culture at 37°C, plaques were purified by suspension into phage buffer and plaque assay against *E. coli* K-12. Recombinant phages were identified by Sanger sequencing from purified plaques.

Table 1

Components (5' to 3')	Sequence*
-----------------------	-----------

HA1	ctgataactttaattgtttacaggcatattaatgagtaattccataacgaacacgtgatgcagttctatcgta acaatcttaaaactaaaggcgtcttcggacgccagtgaggaaaatgaatattgcaaaattataggagtt atttcattattgttgtagtagcatgtgtttaactatctgtatcgat (SEQ ID NO:1)
Promoter	
<i>J23119:</i>	atcattcaagagagtgcaattcaggggtgtgaattgacagctagctcagtcctaggataatagatcttcc gacaccgca (SEQ ID NO:2)
<i>ndd:</i>	atcattcaggatctttccatttaaaaaattcacaaaacagttacataccacaaggaccgtggtacta tacaactatccgacaccgca (SEQ ID NO:3)
<i>modB/A:</i>	atcattcactgggcaaaataattcaaaaagtgtttactttccttctaacgatgatgatagcttctga agtatacgtccgacaccgca (SEQ ID NO:4)
<i>motB:</i>	atcattcacctctatctctccaacaaaacgataaaaagtgtttacttctcgggttagtgggtactat aacaccatccgacaccgca (SEQ ID NO:5)
<i>rIIA:</i>	atcattcaacgcttttcgatatgggagaagccgaagaattaaaaccgctttaagatattagaga taaaattcaatccgacaccgca (SEQ ID NO:6)
<i>gp34i:</i>	atcattcagaactctctgggaccattatacacttaattcttgaagcaaatgagacacaacgtggta cacttcgtgtagctacgcaggctgaagctccgacaccgca (SEQ ID NO:7)
<i>gp18:</i>	atcattcagctgagtatgttccagtatctcatggtatgtcaatggtataaatatattaagcatattaag atccgacaccgca (SEQ ID NO:8)
<i>gp22:</i>	atcattcagcctatatggctctcgcagaggccatgaaaaagcgttataaatattattatctaaacaac atccgacaccgca (SEQ ID NO:9)
<i>gp23:</i>	atcattcaagaaatctaataaagatgaaagcactattactgagagtataaatactcctgatactgaag catccgacaccgca (SEQ ID NO:10)
<i>soc:</i>	atcattcacttaagaatttaattacttattcagaaaagaagtgataaataatcatgtaatttaataaa ggccgacaccgca (SEQ ID NO:11)
RBS	gaattcattaaaggaggaaaaggtacc (SEQ ID NO:12)
Gene	
<i>sfgfp:</i>	atgagcaaaggagaagaactttcactggagttgtccaattctgtgaattagatggtgatgtaaatgggca caaattttctgtccgtggagaggggaagggtgatgtacaacggaaaactcaccttaaatatttgcact actggaaaactactgttccgtggccaacactgtcactactctgacatggtgttaagtctttccggtatc cggatcacatgaaaaggcatgactttcaagagtgccatgccgaagggtatgtacaggaacgcactatat cttcaaagatgacgggacctacaagacgctgctgaagtcaagttgaagggtataccctgttaacgtat cgagttaaagggtattgtttaaagaagatggaacattctggacacaaactcgagtacaacttaactca cacaatgtatacatcacggcagacaaaagaatggaatcaaagctaactcaaaattgccacaac gttgaagatggtccgttaactagcagaccattatcaacaaaatactccaattggcgatggccctgtccttt accagacaaccattacctgtcgacacaatctgtcctttcgaaagatccaacgaaaagcgtgaccacatg gtcctcttgagttgtaactgctgctgggattacacatggcatggatgagctctacaataa (SEQ ID NO:13)
<i>his₆-c/pB:</i>	atgcatcaccatcaccatcacatgctgtggtatgcttactaataaaattccagcttgccttgccgatgccca atcactgcactcgggcacgacaaccaattcatgaaccacttcattaatgagcgccctgctgaatcagga

	aggggggttcggttagtccttattaacatccgctggcataaatgctggccagttgcgcacagatataaatcag gcattaaatcgtttacgcaggtgaaggtactggtggtgatgtccagccatcacaggatctggtgcgcgttct taatctttgcgacaagctggcgcaaaaacgtggtgataactttatctcgtcagaactgttctgttgcggcac ttgagtctcgcggcacgctggccgacatcctgaaagcagcaggggogaccaccgccaacattactcaag cgattgaacaaatcgctggagggtgaagcgatgaacgatcaagggtgtaagaccaacgtcaggctttga aaaaatataccatcgaccttacgaacgagccgaacaaggcaaacctgatccggtgattggtcgtgatga agaaattcgccgtaccattcagggtgctgcaacgctgactaaaaataaccgggtattgattggtgaacccgg cgtcggtaaaactgccatcggtgaaggctggtgcgcagcgtattatcaacggcgaagtgcccgaaaggggtga aaggccgcgcgctactggtggtgatatgggcgcgctggtggtggtggcgaaatatcgcggtgagttga agaacgtttaaaaggcgtacttaacgatcttgctaacaggaaggcaacgtcatcctgtttatcgacgaatt acataccatggtcggcgcggttaaagccgatggcgcaatggacgcgggaaacatgctgaaacctgcac tggcgctggtgaattgcactgcgtagggtccacgacgcttgacgaatatgccagtcacatcgaaaaagat gctgcgctggaacgtcgtttccagaaagtgtgttgccgagcctctgttgaagataccattgctattctgcgtg gcctgaaagaacgttacgagctgcaccaccatgtgcaaactactgacccggcaattgtgcagcagcgac gttgtctcatcgctacattgctgacgctcagctgcgggataaagccatcgacctgatcgatgaagcagcatc cagcattcgatgcagattgactcaaaaccagaagaactcgaccgactcgatcgctgatcatccagctca aactggaacaacaggcggttaataagagctgatgaagccagtaaaaaacgcctggatatgctcaacg aagaactgagcgacaagaacgtcagtactccgagttagaagaagagtggaagcagagaaagcctc gctttctggtacgcagaccattaaagcggaactggaacaggcgaaaatcgctattgaacaggctcgcgt gtgggtgacctggcgcggtatgtgaactgcaatacggcaaaatccgggaactggaaaagcaactgga agccgcaacgcagctcgaaggcaaaactatgcgtctgttgctgtaataaagtaccgatgccgaaattgct gaagtgtggtgcgttggaacgggaattccggtttctgcgatggaagcgagcgcgaaaaactgtgc gtatggagcaagaactgcaccatcgcttattggtcagcacgaagcgggtgatgcggtatctaacgctattc gctgtagccgtgcggggtggtggtaccaaactgcgggattggttcattctgttccggcccaacgggtgt ggggaaaaacagagcttgaaggcgctggcgaaactttatgttgatagcagcagggcgatggtcgtatcg atatgtccgagttatggagaacactcgggtctcgtctggtggtgcgcctccgggatattggttatgaag aagggtgctatctgaccgaagcgggtgcgtcgtcgcgcgtatccggtatccgtggtatgaagtggaaaaa gcgcacccggatgtctcaacattctgttcaggtactggatgatgggcgtctgactgacgggcaggggag aacggtcgacttcgtaatacggctcgtcattatgaacctaacctcgggttcgatctgattcaggaaacgtttgg tgaactggattatgcgcacatgaaagagctggtgctcgtggtgaagccataactccgtccggaattcatt aaccgtatcgatgaagtgcgtcttccatccgctgggtgaacaacacattgcgtcgattgcgcagattcagt tgaacgtctgtacaaacgtctggaagaacgtggttatgaaatccacattctgacgagcgctgaaactg ctgagcgagaacgggtacgatccggtctatggtgcacgtcctctgaaacgtgctatccagcagcagatcga aaaccgctggcacagcaatactgtcgtgtaattggtccgggtaaagtattccctggaagtaatga agaccggatcgtcgcgtccagtaa (SEQ ID NO:14)
HA2	gctagcagtggttttcaagcttttagccagggtatgtgtcatatttaacatttgtgtgttatctaattgattaa gaaaatcttggtctattcattagcccttgcgtcttattggtagcactatattacggagtaattgtcggaattc aagtcgtgctttcattctgatgttatggcactacattcac (SEQ ID NO:15)
	*Non-bolded = Untranslated regions *Bolded = Promoter sequence

Fluorescence Assays

[0327] To measure intracellular protein release by phage infection, *E. coli* K-12 strain EC001 was grown at 37°C to log phase (~0.3-0.5) in TNT media and diluted to an OD_{600nm} of 0.05 before infection with wildtype T4 phage (MOI = 1). To separate supernatant from bacterial cells, samples were centrifuged at 4000 rpm for 5 min. After the supernatant was removed, bacterial pellets were resuspended in an equivalent volume. OD_{600nm} and fluorescence (ex. 485 nm/em. 528 nm) were measured from supernatant and resuspended bacterial pellets.

[0328] To measure GFP production when *sfgfp* is encoded in the T4 phage genome, *E. coli* K-12 was grown to log phase in M9 minimal media supplemented with 0.4% glucose and 1 µg/mL thiamine, diluted to OD_{600nm} of 0.05. The bacterial culture was then mixed with wildtype T4 phage, engineered T4 phages, or the buffer vehicle (MOI = 0.001) and shaken at 37°C for 12 hours. Bacterial cells were removed by centrifugation at 13,000 rpm for 30 sec, and then the supernatant fluorescence was measured (ex. 485 nm/em. 528 nm). Phage concentrations in the supernatant were quantified by spot assay.

In vitro production of the ClpB via T4 phage infection

[0329] *E. coli* K12 cells were grown to log phase (OD₆₀₀ = ~0.3-0.5) in a 50 mL LB broth culture, before infecting with wildtype T4 phage or T4::clpB overnight. Post incubation, a small aliquot of culture was taken and plated for cfu and pfu measurements. The remaining culture was centrifuged at high speed (16,000 xg/1 hr/4°C) to pellet residual bacteria and cellular debris. The resulting supernatant was filtered with a 0.2 µm filter to ensure sterility, storing aliquots at -20°C.

[0330] *In vitro* ClpB protein concentration was accessed via a His6-tag ELISA (GenScript, Catalog #GSCRPT-L00436), adhering to manufacturer's instructions. Interpolated values were normalized to total protein concentration in the lysate, obtained using a Bradford assay (Fisher Scientific®, Catalog #AAJ61522AP).

Housing and husbandry

[0331] All mice were obtained from Jackson Laboratories and allowed to acclimate for one week. Mice were housed in rooms with controlled air, temperature, and humidity with a 12h:12h light:dark cycle. Food and water were provided *ad libitum*, unless otherwise noted. All animal studies were performed in compliance with Institutional Animal Care and Use Committee guidelines at Virginia Tech under protocol # 20-097.

Mucosal expression of sfGFP via engineered phage

[0332] C57Bl/6 mice (Stock 000664, female, 7-8 wks old) were randomized and allowed to acclimate and then received 5 g/L of streptomycin provided in the drinking water (Day -4), which was replaced every 3 days. After 1 day of streptomycin treatment (Day -3), mice were fasted for 4 hours and then received 100 µL of bacteria suspended in PBS. Bacterial inoculum was prepared from streptomycin-resistant *E. coli* K-12 cultures grown overnight in LB broth with 100 µg/mL of streptomycin, washing twice and diluting 10-fold in PBS. After 3 days, on Day 0, mice received 100 µL wildtype T4 phage or T4::sfgfp by oral gavage. Phage solutions were prepared from 10⁸ pfu/mL phage stocks by

diluting 1:10 into 0.1 M sodium bicarbonate immediately before gavage. Food was returned immediately after oral gavage. Stool was regularly collected for bacteria and phage quantification.

[0333] To quantify fecal bacteria and phage, fecal pellets from each mouse were suspended in phage buffer at 50 mg/mL and homogenized manually with sterile wooden sticks and vortexed briefly. Large debris were allowed to settle for 1 min prior to selective plating on MacConkey agar plates containing 100 µg/mL of streptomycin to quantify fecal *E. coli*. The remaining fecal sample was treated with 10 µL of chloroform and centrifuged at 6000 rpm for 10 minutes at 4°C. Phage was quantified by spot assay from this supernatant. Measurements were collected longitudinally for each mouse and experimental condition.

[0334] After euthanasia, colonic tissue was preserved in an OCT frozen block (Tissue-Tek, Sakura Finetek #4583; Tissue-Tek Cryomold, Sakura Finetek #4566) and maintained at -80°C prior to tissue processing.

Engineered phage treatment of diet-induced obesity

[0335] Murine models of DIO mirrored those above with minor alterations. For the DIO model, C57Bl/6 mice (Stock 000664, male, 6 wks old) weighing 19 to 22 grams⁴⁹ were randomly assigned to three experimental groups. On Day 0, mice received 100 µL of a phage buffer vehicle, 100 µL of wildtype T4 phage or 100 µL of T4::*clpB* phage by oral gavage. On Day 3, food was switched from a standard rodent chow (Teklad 2918) to a high-fat diet (Research Diets D12451i). Food was replaced weekly during which stool was collected and body weights measured. Fecal *E. coli* colonies isolated from mice receiving wildtype T4 or T4::*clpB* phages were tested for phage resistance by cross-streaking against 10⁸ pfu/mL of the same purified phage. At least 20 colonies per mouse per timepoint were examined. Measurements were collected longitudinally for each mouse and experimental condition. Prior to sacrifice, a subset of mice was individually housed – cage dimensions mirrored those of the home cage (5.5"W, 6.5"H, 9"L) – and monitored via the Oxymax Comprehensive Lab Animal Monitoring System (CLAMS) (Columbus Instruments). Food (Research Diets D12451i) and water, supplemented with streptomycin at 5 g/L, were provided *ad libitum*. All mice were allowed to acclimate overnight, before collecting CO₂ production, oxygen consumption, animal activity, and energy expenditure every 2 minutes for a complete day-night cycle (7 AM – 7AM).

[0336] Immediately after euthanasia, blood was collected by intracardiac bleeding, followed by colonic and liver tissue preservation in an OCT frozen block (Tissue-Tek, Sakura Finetek #4583; Tissue-Tek Cryomold, Sakura Finetek #4566) or in 10% formalin (Fisherbrand, Catalog # 316-156). Intestinal mucosa was collected from select colon sections by gentle scraping with a metal spatula and suspension at 50 mg/mL in phage buffer. Serum was extracted from the supernatant of blood that was allowed to coagulate at room temperature for 30 minutes and centrifuged at 13,000 rpm for 10 minutes at 4°C. Serum samples (n=5) were assessed for cytokine concentrations with the LEGENDplex (BioLegend, Catalog # 740446) Mouse Inflammation Panel (13-plex). All samples were analyzed using a FACSCanto II (BD Biosciences, Reference # 338960) in technical duplicate and processed according to the manufacturer's instructions.

Histology

[0337] All tissue preparation, slicing, and H&E staining was performed by Virginia Tech Animal Laboratory Services at the Virginia-Maryland College of Veterinary Medicine. Mucin and gastric epithelial cell nuclei staining was accomplished via sectioning of OCT preserved murine colonic tissue. Select tissue slices were fixed and dehydrated as previously described.⁵⁰ Post fixation, all slides were incubated with blocking buffer (2% Fetal Bovine Serum, 1% Bovine Serum Albumin, 0.2% Triton X-100, 0.05% Tween 20) for 1 hour at room temperature and washed twice with PBS in preparation for staining. Slides were incubated with 40 µg/mL WGA (Vector Labs, Catalog # RL-1022), 10 µg/mL UEA I (Vector Labs, Catalog # RL-1062-2), and 1 µg/mL DAPI (Thermo Scientific, Catalog # 62248), diluted in 1X PBS, for 15 minutes at room temperature. All slides were washed briefly with DI water and allowed to air dry before mounting with Prolong Glass Antifade Mountant (ThermoFisher Scientific, Catalog # P36980). Mounted slides were cured for 24 hours in the dark prior to imaging.

Fluorescence Microscopy

[0338] Fluorescence microscopy for unfixed slides was performed with a Nikon Eclipse Ti confocal microscope, utilizing a 20x/0.5 CFI Plan Fluor objective (MRH00201). Mean Fluorescence Intensity of unfixed slides was determined by measuring the intensity of ten mucosal regions per image across two regions per mouse and then normalized to the background tissue intensity. Imaging for WGA, UEA I, and DAPI stained slides was performed with a Nikon CSU-W1 SoRa spinning disk confocal microscope, utilizing a 20x/0.8 CFI Plan Apochromat Lambda D objective (MRD70270) and a 60x/1.49 CFI Apochromat TIRF Oil Immersion objective (MRD01691) under 2.8x magnification. Images were captured with an ORCA-Fusion BT sCMOS camera, with all processing performed on Nikon NIS-Elements software.

Results

Phage propagation releases intracellular protein

[0339] As depicted in **FIG. 1A** and without being bound by theory, Applicant hypothesized that bacterial infection by a virulent phage would enable the release of intracellular protein to the extracellular environment. To demonstrate proof of concept of encode heterologous genes into the phage genome as a protein delivery system, Applicant first determined whether phage-induced lysis of bacteria would lead to the release of significant quantities of intracellular protein. To test this, Applicant incubated the virulent T4 phage with an *E. coli* K-12 strain that constitutively-expresses super-folding green fluorescent protein (sfGFP) from the bacterial chromosome (strain EC001). Applicant found that EC001 is readily lysed by T4 phage (**FIG. 5**) and that the phage-induced lysis of bacteria led to substantially more sfGFP in the culture supernatant. As shown in **FIG. 1B**, EC001 cultures with T4 phage had greater fluorescence in the supernatant ($\sim 0.5\text{--}1.0 \times 10^6$ AU/OD) compared to EC001 cultured without phage ($\sim 7 \times 10^3$ AU/OD). In contrast, fluorescence of the cell pellets was comparatively similar ($\sim 10^5$ AU/OD). Together, this data indicates that T4 phage facilitates the release of intracellular proteins into the extracellular environment.

T4 phage can be engineered to produce and release recombinant proteins

[0340] Next, Applicant aimed to quantitate whether significant amounts of phage-encoded, heterologous protein could be produced and released during infection of bacterial cells. Phage infection

generally leads to a disruption of host cell processes in favor of those that are phage-specific. For T4 phage, the preferential transcription of early T4 promoters over host promoters is due in part to a greater promoter strength of the former and the modification of host RNA polymerases to preferentially transcribe T4 promoters. This ultimately leads to the transcription of phage promoters in sequential stages (FIG. 2A) to optimize the assembly of viable viral particles. Because this process is carefully orchestrated, Applicant determined the promoter and activation (*i.e.*, early, middle, or late) for maximal expression with impact on phage propagation. Applicant generated a small library of recombinant T4 phages containing select promoters driving *sfgfp* expression (FIG. 2B) and inserted this construct into a non-essential hypothetical membrane protein (*ac* gene) that has been previously used for T4 recombination. After 12 hours of their culture with *E. coli*, Applicant measured the sfGFP protein released into the supernatant by fluorescence. As shown in FIG. 2C, the absence of a promoter ("No promoter") or use of a strong constitutive *E. coli* promoter ("J23119") did not lead to significantly more sfGFP fluorescence than a non-engineered T4 phage that lacks the sfGFP gene ("Wildtype"). However, one early promoter (*motB*, $p=0.024$) and two late promoters (*gp18*, $p=0.0035$; and *gp22*, $p=0.01$) led to significantly greater sfGFP levels in the supernatant. Titers of phages encoding these promoters showed about 2-fold reductions compared to wildtype T4 phage, though none achieved statistical significance ($p=0.0807$, 0.0513 , and 0.1556 , respectively) (FIG. 2D). Based on our criteria of maximal sfGFP levels and minimal impact on phage titer, our subsequent experiments used the *gp22* promoter, which regulates transcription of the major prohead scaffolding core protein.

Engineered phage expresses proteins in the mucosa

[0341] To determine the feasibility of the above-mentioned approach for the *in vivo* production of a protein in the gut microbiota, a streptomycin-treated mouse model of *E. coli* colonization was used. Human commensal *E. coli* strains, including the laboratory-adapted strain *E. coli* K-12 originally isolated from human feces, do not reliably colonize the murine gastrointestinal tract. Previous studies have shown that streptomycin administration is effective at maintaining streptomycin-resistant *E. coli* colonization, allowing for the study of *E. coli* function in mouse models of the mammalian gut. As schematically outlined in FIG. 3A, streptomycin was provided in the drinking water, colonized with a single dose of *E. coli* and then administered wildtype T4 phage or T4::*sfgfp* phage ($n=6$ per group). After 4 days of co-colonization between bacteria and phage, it was found that both were largely present (FIG. 3B). Quantitation of fluorescence in the mucosa of unfixed intestinal sections by fluorescence microscopy showed significant increases in GFP ($p<0.0001$) in mice colonized with T4::*sfgfp* phage over those colonized with wildtype T4 phage (FIG. 3C). Representative images show that the wildtype T4 phage condition had low levels of fluorescence due to autofluorescence whereas the T4::*sfgfp* phage condition had substantial mucosal fluorescence (FIG. 3D and 3E, respectively).

[0342] For additional validation that GFP fluorescence was localized in the mucosa, intestinal samples and labeled nuclei were fixed with DAPI and mucins with rhodamine labeled wheat germ agglutinin and *Ulex Europaeus* agglutinin I. Although GFP fluorescence was greatly diminished, it remained sufficiently fluorescent for localization. As shown in FIG. 3F, the FITC-channel fluorescence in the mucosa of mice colonized with wildtype T4 phage was insubstantial when compared to the autofluorescence of the tissue. This was clearly shown at higher magnification (FIG. 3G) with the

channels separated (**FIG. 3H, 3I, and 3J**). In contrast, mice colonized with T4::sfgfp phage had substantial mucosal GFP as evidenced by the orange color due to overlapping green and red (**FIG. 3K**). Higher magnification (**FIG. 3L**) confirmed an increase in GFP (**FIG. 3M**) that colocalized to the mucosa (**FIG. 3N**) on the luminal side of the tissue, away from the nuclei (**FIG. 3O**). Colonic sections were fixed and fluorescently stained with nuclear (blue) and mucin (red) stains (**FIG. 3F-3O**). Macroscopic (**FIG. 3F, 3K**) and higher magnification views (**FIG. 3G, 3L**) show overlays of DAPI, FITC, and TRITC channels whereas individual high magnification channels for FITC (**FIG. 3H, 3M**), TRITC (**FIG. 3I, 3N**), and DAPI (**FIG. 3J, 3O**) are shown separately. Collectively, these results indicated that heterologous genes, driven by phage-specific promoters, could be expressed to significant levels during phage propagation. Furthermore, this production in the mucosa is in close proximity to the intestinal epithelium where these gene products exert effect.

T4::clpB significantly reduces host weight in a diet induced obesity mouse model

[0343] To demonstrate that the infection process by engineered phages can be leveraged towards the *in situ* production of a protein with a sustained impact on the mammalian host, Applicant aimed to produce ClpB, a chaperone protein that, when administered to the GIT, reduces host weight gain by mimicking α -melanocyte-stimulating hormone and inducing satiety. Using the same strategy described above, Applicant inserted the *clpB* gene containing an N-terminal His6-tag into the T4 genome, driven by a *gp22* promoter (T4::clpB phage). Quantitation by ELISA of His6-tagged ClpB in cocultures of T4::clpB with *E. coli* K-12 showed significant levels in the supernatant (**FIG. 6A-6C**) corresponding to $0.88 \pm 0.24 \mu\text{g}$ per g of supernatant protein or $37 \pm 21 \mu\text{g}$ per L of culture (mean $\pm$ s.d., n=3).

[0344] To determine if *in situ* protein production by our phage-based approach can have a physiologically-relevant impact on the host, Applicant used a mouse model of diet-induced obesity (DIO) as outlined in **FIG. 4A**. After colonization with *E. coli* K-12, three experimental groups of mice (n=8 or 9 per group) received a single oral dose of either buffer, wildtype T4 phage, or T4::clpB phage (Day 0). On Day 3, mice were transitioned from a standard mouse chow to a high fat diet (45%) to induce obesity. Fecal phage titers, show that both wildtype T4 and T4::clpB phages maintained colonization of the murine gut, though there was large inter-individual variation, especially for wildtype T4 phage (**FIG. 4B**). Fecal *E. coli* concentrations also showed sustained colonization despite the introduction of wildtype T4 or T4::clpB phages (**FIG. 4C**), highlighting a coexistence between phage and bacteria that is consistent with previous observations. At sacrifice, analysis of the intestinal mucosa showed substantial phage and *E. coli* concentrations, confirming that the mucosa, in addition to the lumen, was colonized (**FIG. 4B and 4C**). For insight into the potential mechanism of observed phage and bacterial coexistence, Applicant cross-streaked fecal *E. coli* colonies against phage to determine the frequency of phage resistance. Of the 1,920 colonies tested across phage-colonized mice and timepoints before and after phage administration (Days 0, 3, and 18), all colonies were phage susceptible. This suggests that resistance development may either be below our detection limit, *E. coli* persists in spatial refuges inaccessible to phage, and/or that phage resistance leads to poor bacterial fitness.

[0345] During the course of this Example, murine weights showed that mice colonized with T4::clpB phage had significantly reduced weight compared to mice colonized with no phage or wildtype T4 phage (**FIG. 4D**). Past work has shown that mice gavaged daily with bacteria encoding *clpB* (e.g.,

E. coli or *Hafnia alvei*) had reduced weight gain that was associated with reduced daily food intake. Applicant similarly observed significantly reduced food consumption in mice colonized with T4::*clpB* phage (FIG. 4E), which was associated with significantly increased night activity (FIG. 4F).

[0346] DIO leads to inflammation in murine models, which mimics the low-grade chronic inflammation in humans. As further support for a physiological effect of *in situ* ClpB production, it was found that several pro-inflammatory cytokines were reduced in mice colonized with T4::*clpB* phage compared to wildtype T4 phage (e.g. IL-1 α and IL-1 β), or both wildtype T4 and bacteria alone (IL-23). Other cytokines were also reduced, including MCP-1, IL-17A, and GM-CSF, but did not reach statistical significance (FIG. 4G). IL-23 is linked to an array of autoimmune inflammatory diseases, however, its role in obesity is less clear. An upregulation in IL-23 stimulates T-cell differentiation, leading to an increase in Th17 cells, which leads to over production of IL-17 and several other proinflammatory cytokines including IL-1. Reductions in IL-23, IL-1 α/β , and IL-17A likely points to a decrease in IL-23/IL-17 axis formation and an overall decrease in systemic inflammation. MCP-1 levels further support this claim, with high concentrations of this cytokine being found in obese individuals and linked to both macrophage invasion of adipose tissue and heightened insulin resistance. As a result, these cumulative results indicate a protective effect conferred by T4::*clpB* against diet-induced inflammation.

Discussion

[0347] The results demonstrate that an engineered virulent phage, T4, can orchestrate the production and release of heterologous proteins in the mammalian gut during the normal phage infection process, leveraging its coexistence with its bacterial host for a sustained *in situ* effect. By engineering T4 phage to express *sfgfp* under a late T4 phage promoter, *gp22*, this reporter protein is shown to be significantly expressed *in vitro* and *in vivo*. The results further demonstrate expression of a protein, ClpB, can significantly reduce weight gain in a DIO murine model.

[0348] Similar to the *in vivo* coexistence of gut phages with their bacterial hosts, a coexistence between T4 phage and *E. coli* was observed. Virulent phages are abundant and prevalent components of the human gut that can exhibit a long-term stability. Exemplar is the prevalence of crAssphage, a *Bacteroides* phage found in healthy humans across diets and geography. Thus, our approach to leverage this coexistence is not an unusual expectation from a microecological perspective. Additionally, using a phage to direct a resident bacteria to express proteins circumvents the challenges of directly using bacteria where extensive engineering to both produce and secrete the biologic to the extracellular environment can sap metabolic resources and compromise its ability to colonize the gut. Notably, utilizing resident bacteria also enables a localization of protein release based on bacterial host. In the present Example, production of a GFP reporter revealed production directly in the mucosa.

[0349] The Example confirmed that heterologous gene expression during T4 phage infection required phage-specific promoters, which ensured that gene expression was contingent on phage infection. Expression from the T4 phage genome progresses through early, middle, and late stages with the transcription of late T4 promoters requiring T4-encoded RNAP-binding proteins Gp33 and Gp55 among other factors. A major concern with any genetically-engineered microbe is potential dissemination to the environment. For bacteria with targeted effect in the gut, a number of

biocontainment strategies have been investigated including nutrient auxotrophy, temperature-sensitive toxin-antitoxin systems, and chemically-dependent CRISPR systems. Similar strategies that conditionally allow gene expression or phage propagation could be integrated into future strategies. Additionally, phages are unable to independently proliferate and require a bacterial host, which provides an additional layer of biocontainment.

[0350] The results demonstrated that the virulent T4 phage, which targets resident non-pathogenic gut *E. coli*, can be engineered to express heterologous proteins that are released during cell lysis. Using a fluorescent protein reporter, T4-specific promoters were identified with maximal protein expression and minimal impact on T4 phage titers. It was further identified that this protein was produced in the intestinal mucosa of murine models of the gut microbiota. Using murine models of diet-induced obesity, the results demonstrated that wildtype and engineered T4 phages persist in the gut with the latter leading to reduced weight gain and an altered immunological profile associated with reduced inflammation. Therefore, the results demonstrate an alternative use for virulent phages in the mammalian gut as engineerable vectors to reprogram resident gut bacteria into personalized pharmacies.

[0351] In some aspects, the results further demonstrate the methods of the disclosure may be beneficial for the synthesis of proteins aimed to combat diverse gastrointestinal diseases, including inflammatory disease or those related to pathogenic microorganisms. This is amplified by the use of diverse phage-bacterial pairings, allowing for unlimited avenues to optimize this system. In some aspects, engineered lytic phage offer a unique opportunity to sustainably produce biologics *in situ*, improving upon previous mechanisms to treat gastrointestinal illness.

Example 2

[0352] As previously described, the engineered T4 phage platform can be used to deliver a variety of cargos. In addition to producing sfGFP and ClpB as described in Example 1 above, Applicant engineered T4 phage to produce a serine protease inhibitor SERPIN B1a. Applicant tested the functional activity of this protein, when produced during T4:serpin propagation, and found that its IC₅₀ value against the pro-inflammatory enzyme, neutrophil elastase, was similar to that of a small molecule drug, sivelestat, which was used as a control. Results are shown in **FIG. 7**.

[0353] Various modifications and variations of the described methods, pharmaceutical compositions, and kits of the disclosure will be apparent to those skilled in the art without departing from the scope and spirit of the disclosure. Although the disclosure has been described in connection with specific embodiments, it will be understood that it is capable of further modifications and that the disclosure as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the disclosure that are obvious to those skilled in the art are intended to be within the scope of the disclosure. This application is intended to cover any variations, uses, or adaptations of the disclosure following, in general, the principles of the disclosure and including such departures from the present disclosure come within known customary practice within

the art to which the disclosure pertains and may be applied to the essential features herein before set forth.

CLAIMS

What is claimed is:

1. An engineered phage comprising:
one or more heterologous cargo genes, wherein the one or more heterologous cargo genes encode one or more heterologous gene products; and
one or more regulatory elements, wherein each of the one or more heterologous cargo genes is operatively coupled to at least one of the one or more regulatory elements.
2. The engineered phage of claim 1, wherein the phage is a lytic phage.
3. The engineered phage of any one of claims 1-2, wherein the phage is a T4 phage, T2 phage, MS2 phage, M13 phage, T3 phage, T5 phage, T7 phage, Lambda phage, P1 phage, PhiX174 phage, Mu phage, P22 phage, *Pseudomonas* phage phiKZ, *Bacillus* phage phi29, *Salmonella* phage P22, *Mycobacterium* phage D29, *Enterobacteria* phage T6, *Mycobacterium smegmatis* phage L5, *Escherichia* phage T5, *listeria* phage A511, *Vibrio* phage KVP40, *Staphylococcus* phage K, or *Staphylococcus* phage PhiETA.
4. The engineered phage of any one of claims 1-3, wherein the phage is capable of infecting a bacterial cell or a yeast cell.
5. The engineered phage of any one of claims 1-4, wherein the phage is capable of infecting a cell of a microbiome.
6. The engineered phage of claim 5, wherein the microbiome is a gastrointestinal microbiome, a skin microbiome, an airway microbiome, a vaginal microbiome, or any combination thereof.
7. The engineered phage of claim 6, wherein the gastrointestinal microbiome is an oral cavity microbiome, a stomach microbiome, a small intestine microbiome, a large intestine microbiome, an appendix microbiome, a cecum microbiome, or any combination thereof.
8. The engineered phage of any one of claims 1-7, wherein the one or more heterologous gene products are therapeutic gene products.
9. The engineered phage of claim 8, wherein the one or more heterologous gene products are capable of (a) stimulating an immune response in a subject in which it is delivered; (b) treating a disease or a symptom thereof; (c) modulating a microbiome; or any combination of (a)-(c).
10. The engineered phage of claim 9, wherein stimulating an immune response comprises stimulating a B-cell and/or T-cell mediated response.

11. The engineered phage of any one of claims 1-10, wherein the one or more regulatory elements comprises a phage promoter, optionally wherein the phage promoter is an endogenous phage promoter.
12. The engineered phage of any one of claims 1-11, wherein the one or more regulatory elements comprises an early phage promoter, a late phage promoter, or both.
13. The engineered phage of any one of claims 1-12, wherein the engineered phage is a T4 phage and the one or more regulatory elements comprise *motB*, *gp18*, *gp22*, or any combination thereof.
14. The engineered phage of any one of claims 1-13, wherein the one or more heterologous gene products comprises one or more signaling domains and/or translocation domains.
15. The engineered phage of claim 14, wherein the one or more signaling domains and/or translocation domains are capable of translocating the one or more heterologous gene products, when expressed in a host cell, to a periplasm.
16. The engineered phage of claim 15, wherein the one or more heterologous gene products comprises an N-terminal signal peptide for periplasmic secretion, optionally wherein the N-terminal signal peptide comprises (a) a signal peptide selected from *Tat*, *appA*, *dsbD*, *degQ*, *cueO*, *pIdA*, *ushA*, *pagP*, *fhuA*, *ptrA*, *lptA*, *bepA*, *dacA*, *agp*, *malS*, *spy*, *amiC*, *tolB*, *rseB*, *amtB*, *dsbC*, *amiA*, *torZ*, *FhuD*, *zinT*, *mgIB*, *eco*, *nikA*, *mepH*, *fadL*, *ompL*, *ompN*, *thiB*, *MalE*, *PhoA*, or any combination thereof; (b) a leader peptide from *spA*, *phoA*, *ribose binding protein*, *pelB*, *ompA*, *ompT*, *dsbA*, *torA*, *torT*, *tolT* or any combination thereof; (c) or any combination thereof.
17. A pharmaceutical formulation comprising:
an engineered phage according to any one of claims 1-16; and
a pharmaceutically acceptable carrier.
18. A vaccine comprising an engineered phage according to any one of claims 1-16.

19. A method of delivering a protein to a subject in need thereof and/or treating a disease or symptom thereof in the subject in need thereof, the method comprising:

administering an amount of the engineered phage, pharmaceutical formulation, or vaccine according to any one of the preceding claims to the subject in need thereof, whereby after administration the phage infects a host cell of the subject in need thereof, stimulates production of the one or more heterologous gene products by the host cell, and lyses the host cell so as to release the one or more heterologous proteins gene products by the host cell thereby delivering the one or more heterologous gene products to the subject in need thereof.

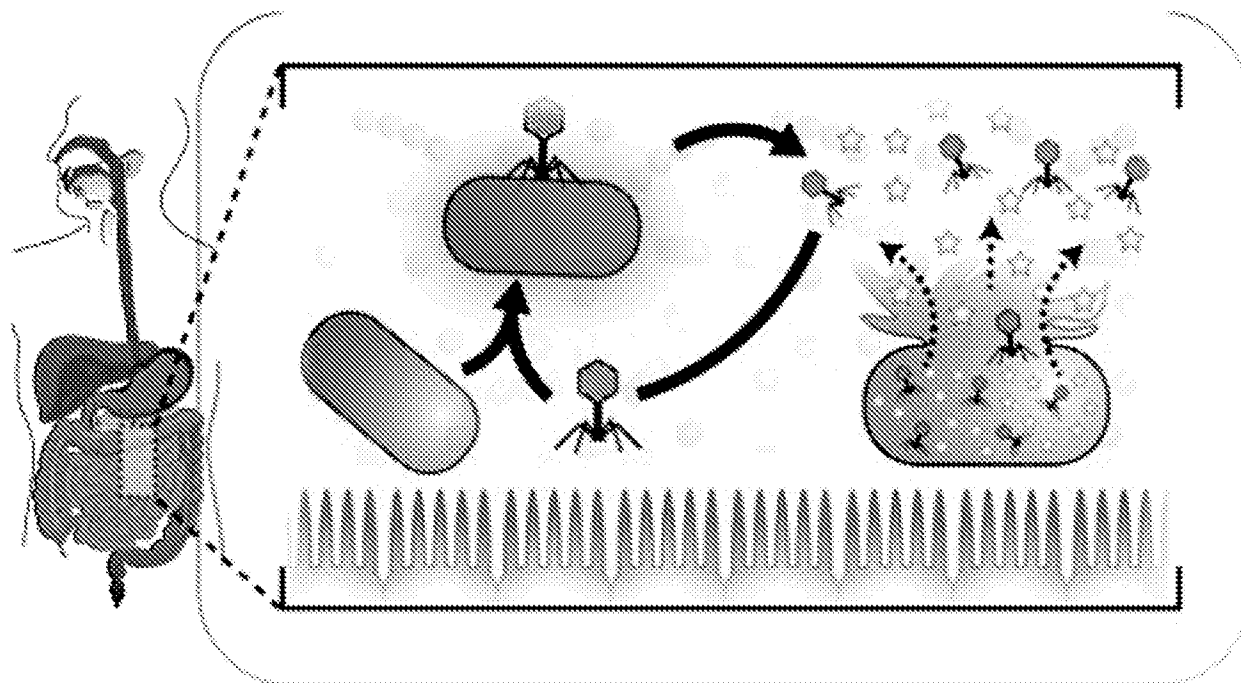
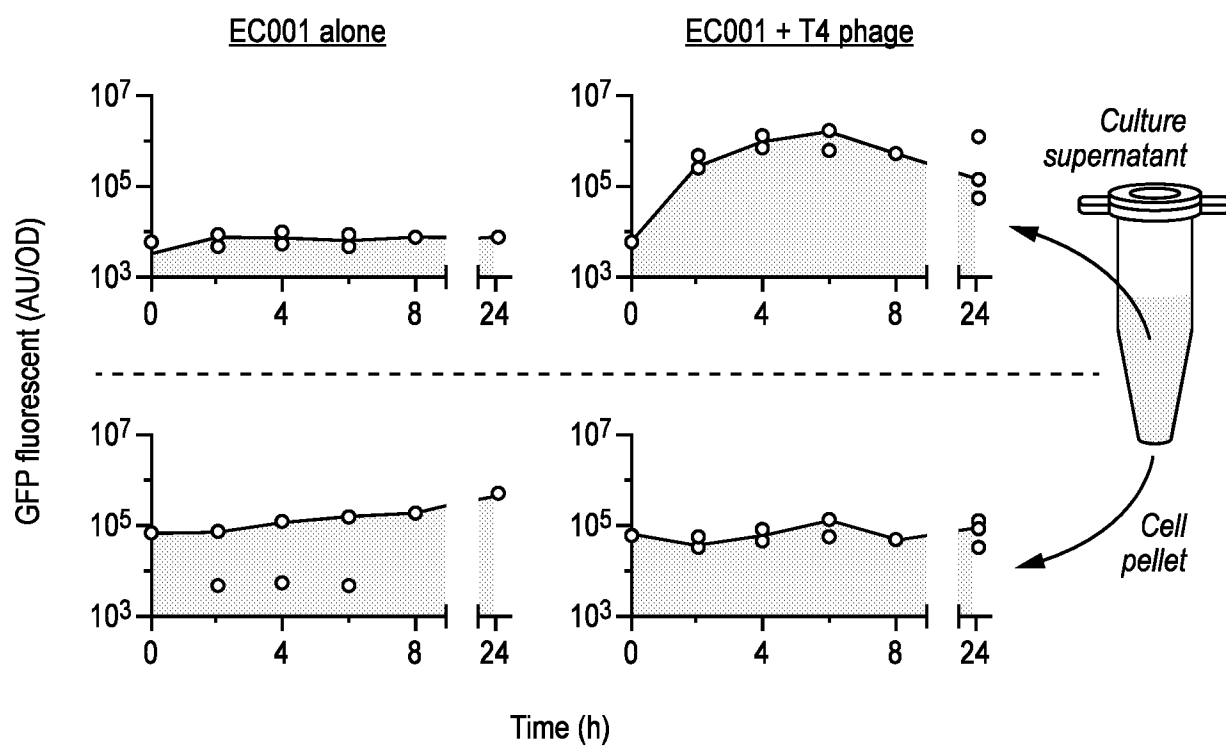
20. The method of claim 19, wherein administering comprises oral administration.

21. The method of any one of claims 19-20, wherein the host cell is a bacterial cell or a yeast cell.

22. The method of any one of claims 19-21, wherein the host cell is part of a microbiome of the subject in need thereof.

23. The method of claim 22, wherein the microbiome is a gastrointestinal microbiome, a skin microbiome, an airway microbiome, a vaginal microbiome, or any combination thereof.

24. The method of claim 23, wherein the gastrointestinal microbiome is an oral cavity microbiome, a stomach microbiome, a small intestine microbiome, a large intestine microbiome, an appendix microbiome, a cecum microbiome, or any combination thereof.

**FIG. 1A****FIG. 1B**

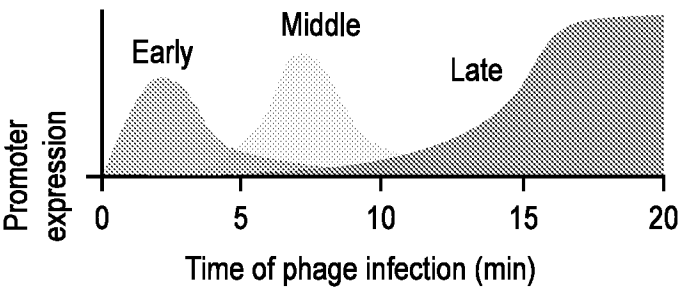


FIG. 2A

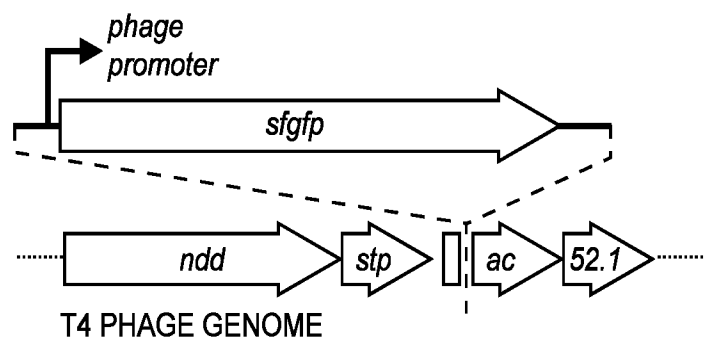


FIG. 2B

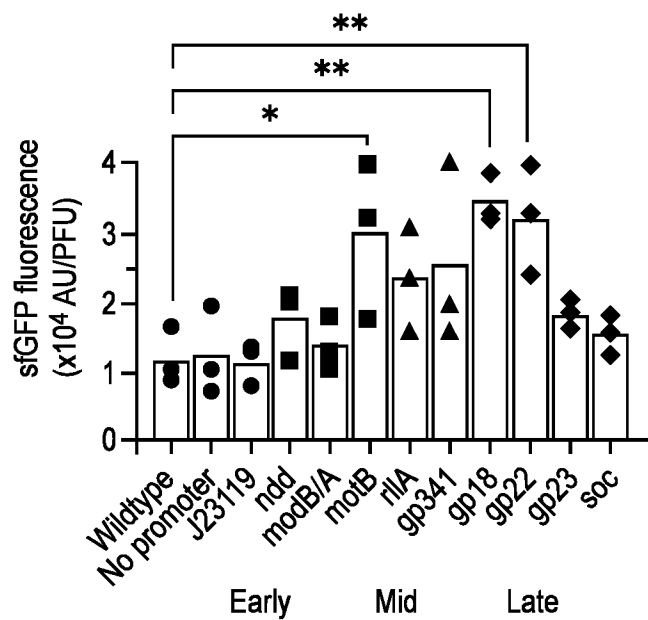


FIG. 2C

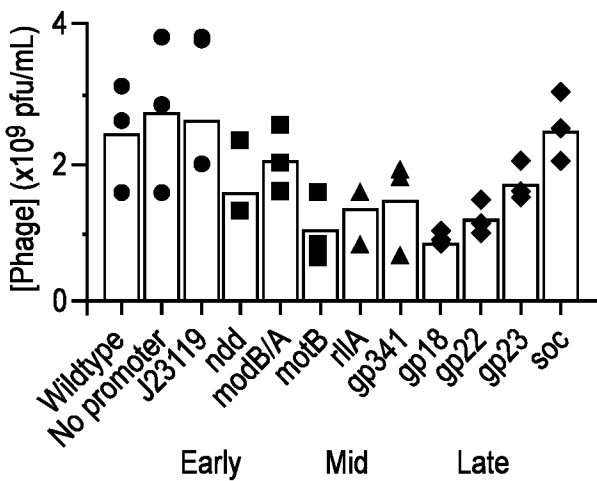


FIG. 2D

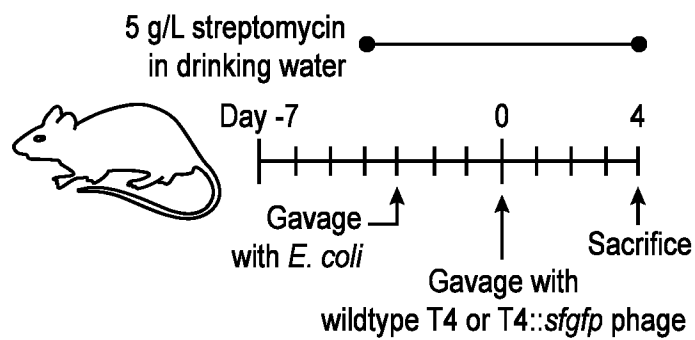


FIG. 3A

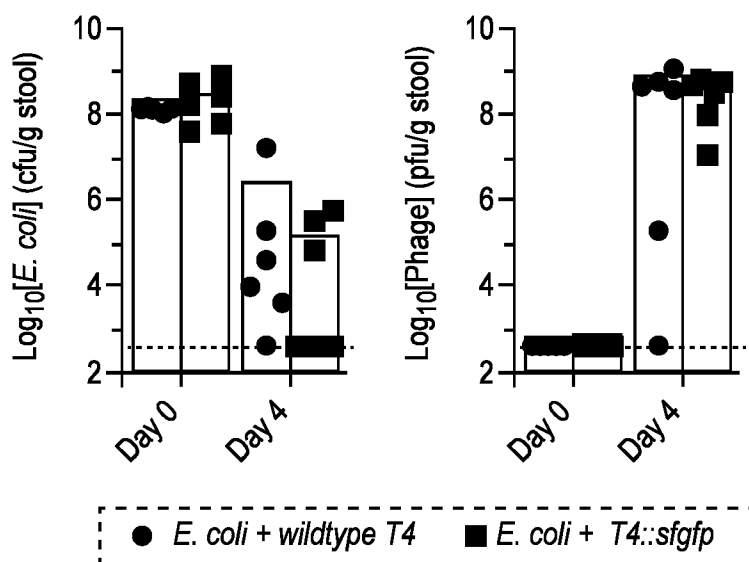


FIG. 3B

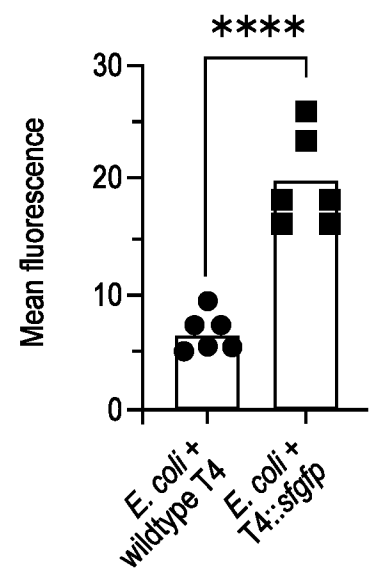
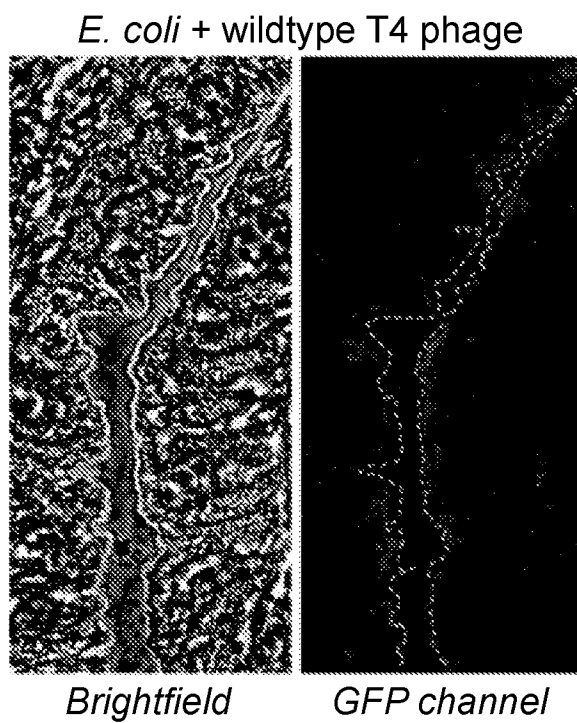
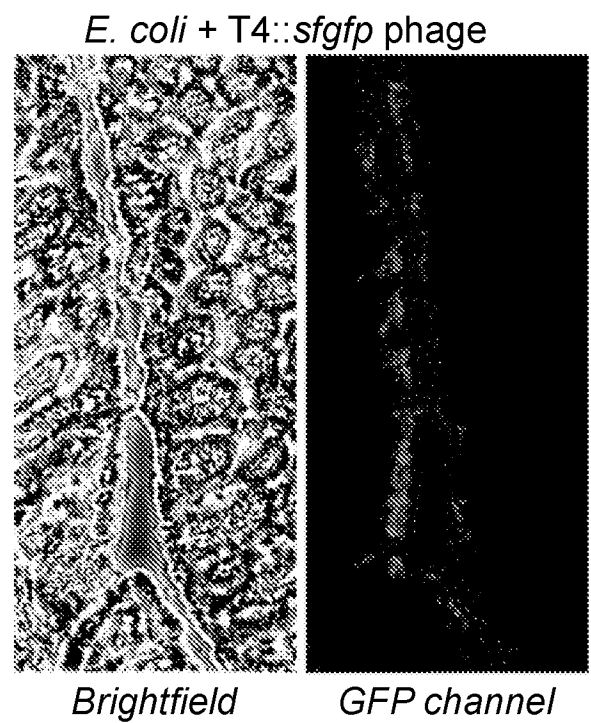
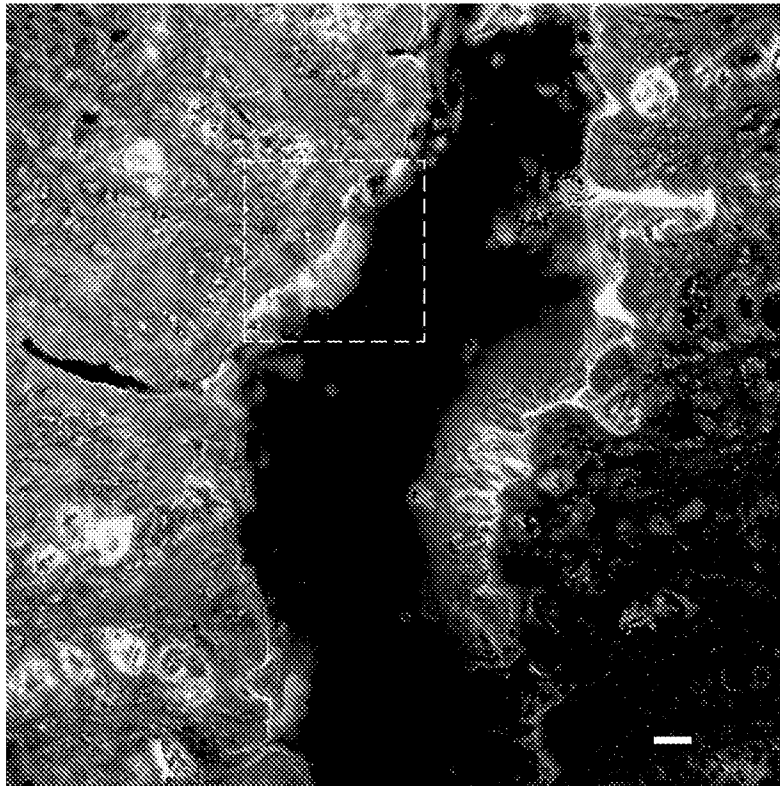
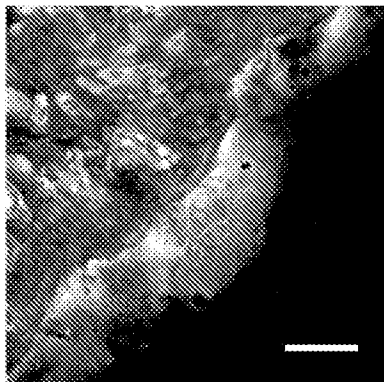
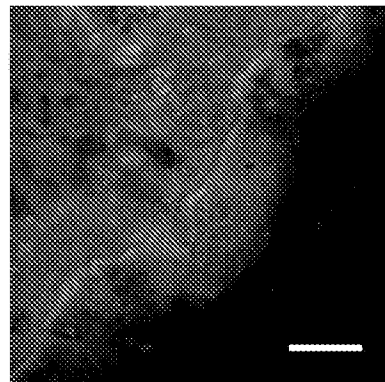
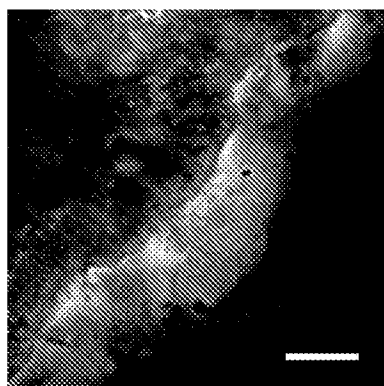
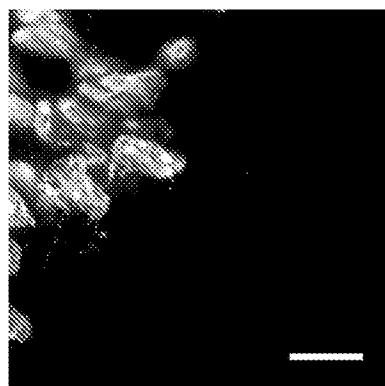


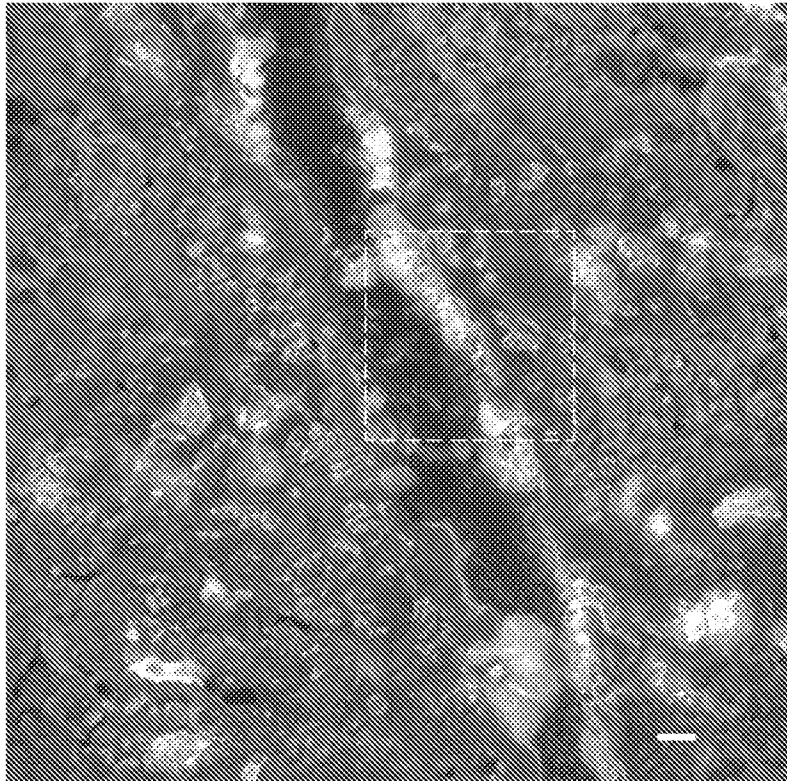
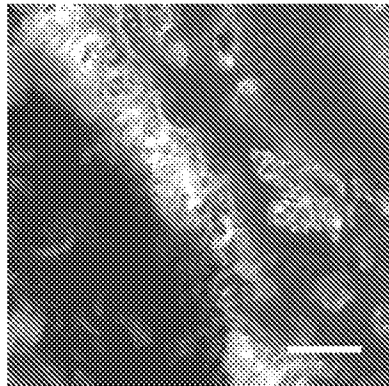
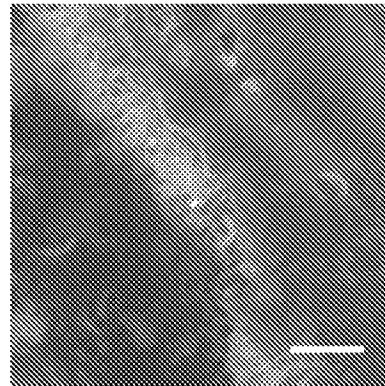
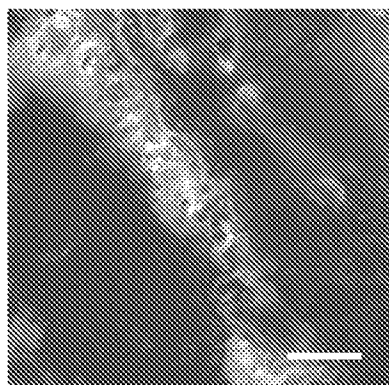
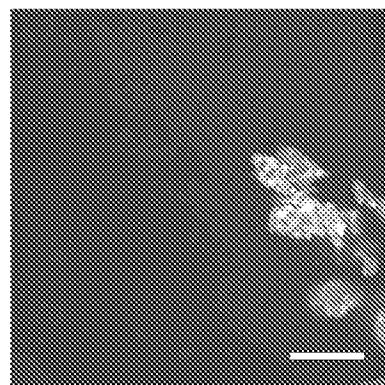
FIG. 3C

**FIG. 3D****FIG. 3E**

5 / 11

E. coli + wildtype T4 phage**FIG. 3F****FIG. 3G****FIG. 3H****FIG. 3I****FIG. 3J**

6 / 11

E. coli + T4::sfgfp phage**FIG. 3K****FIG. 3L****FIG. 3M****FIG. 3N****FIG. 3O**

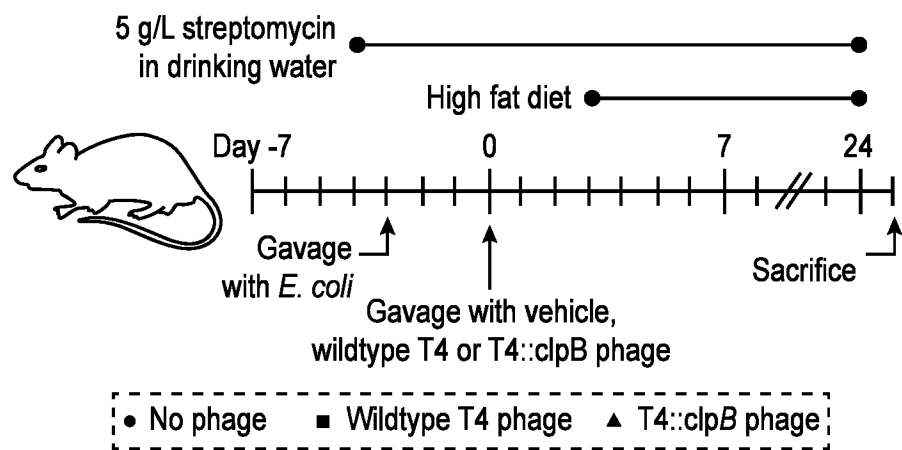


FIG. 4A

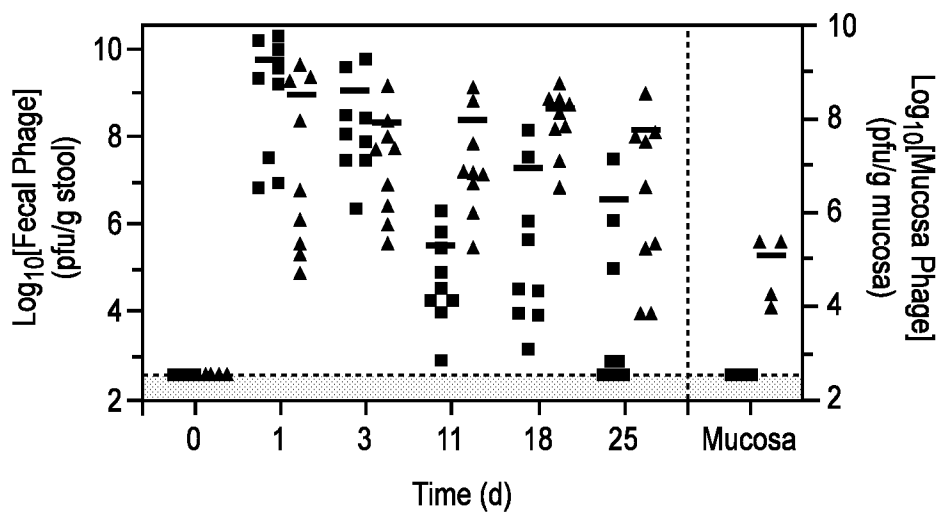


FIG. 4B

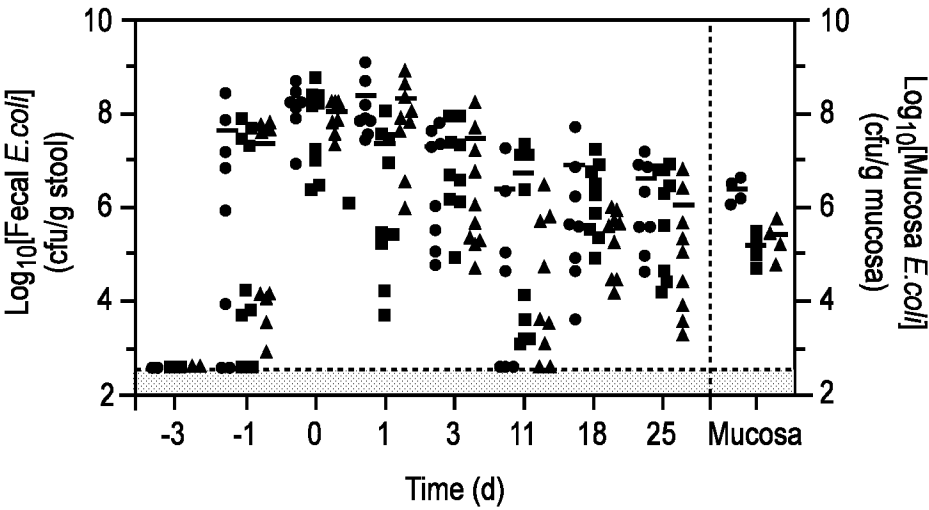


FIG. 4C

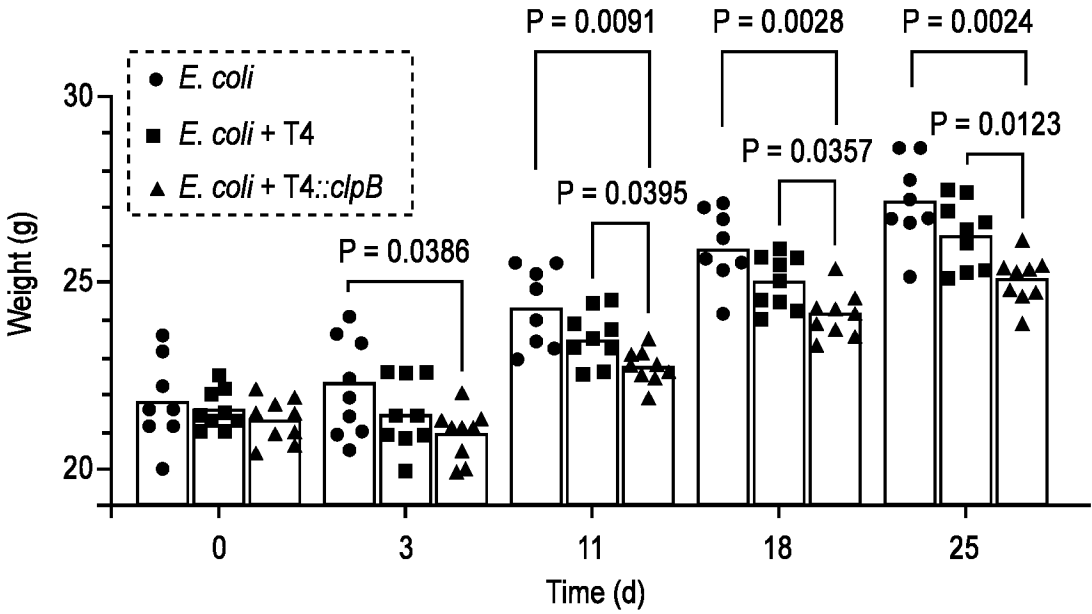


FIG. 4D

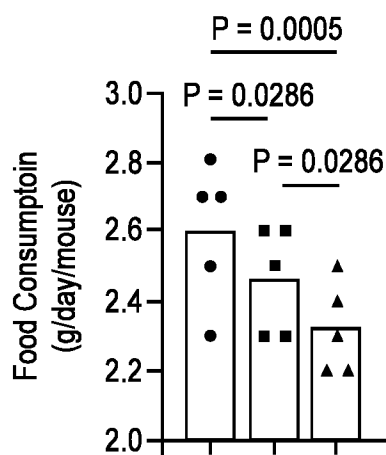


FIG. 4E

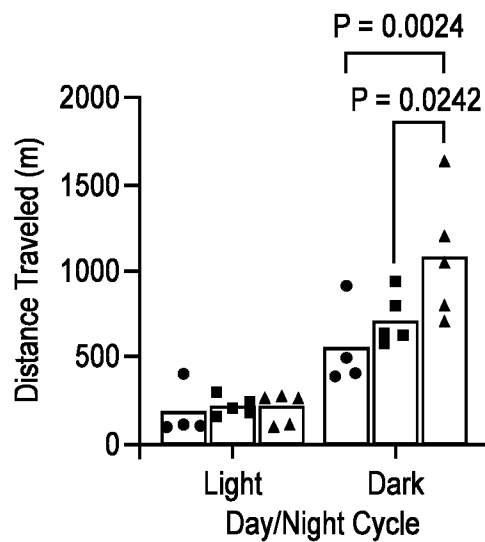


FIG. 4F

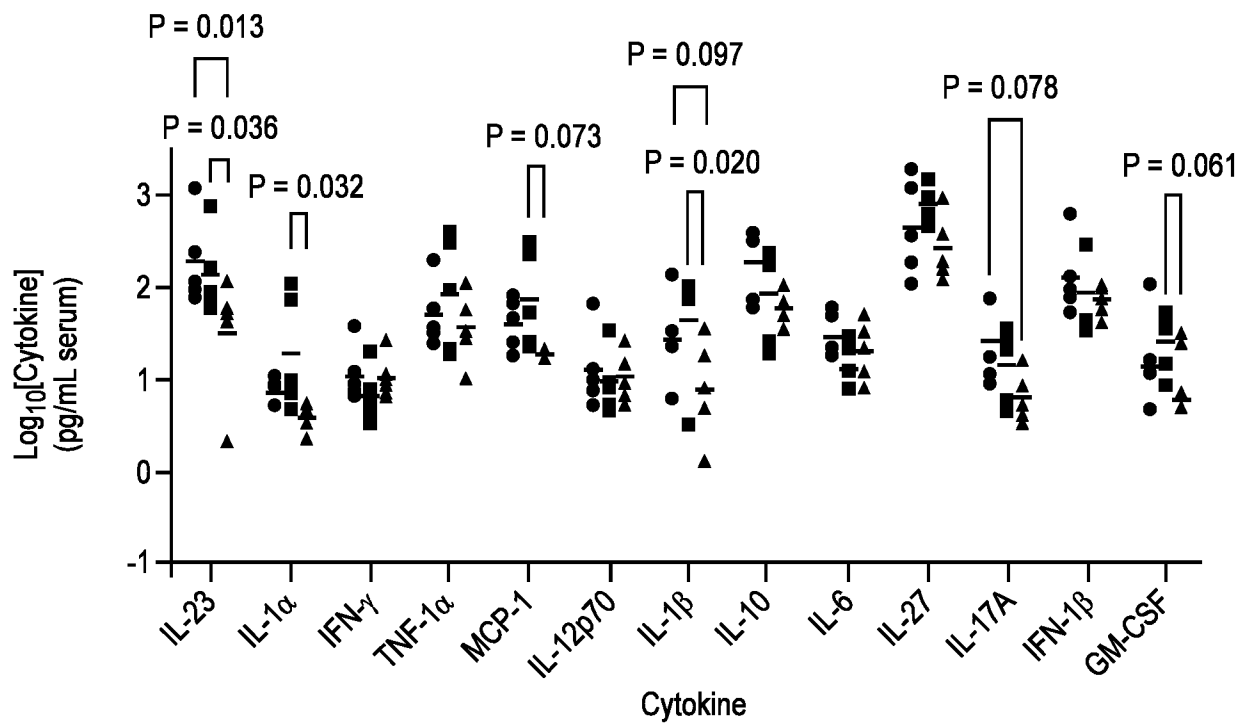


FIG. 4G

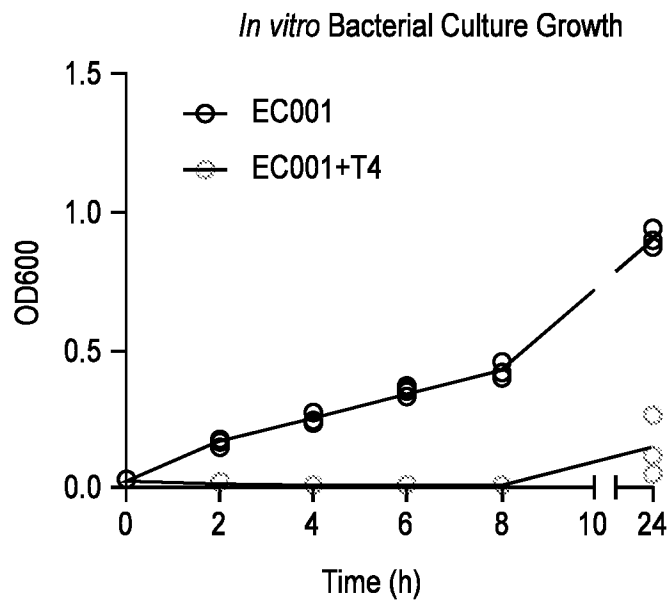


FIG. 5

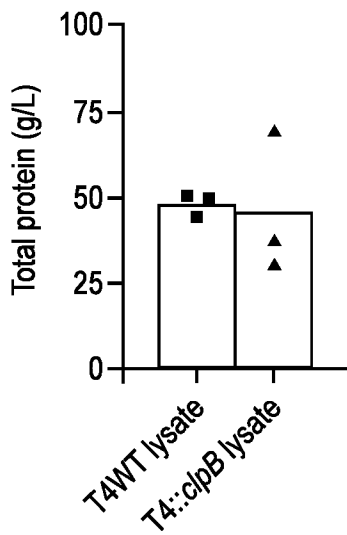


FIG. 6A

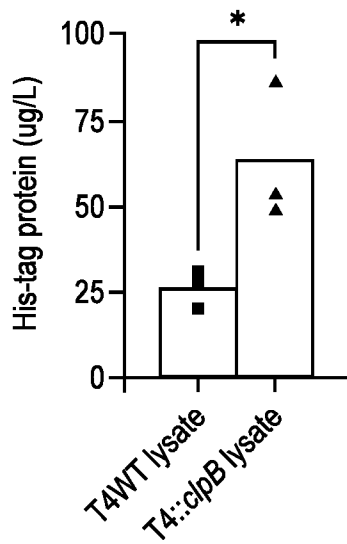


FIG. 6B

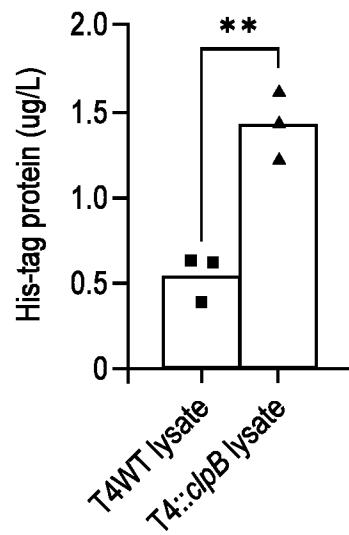


FIG. 6C

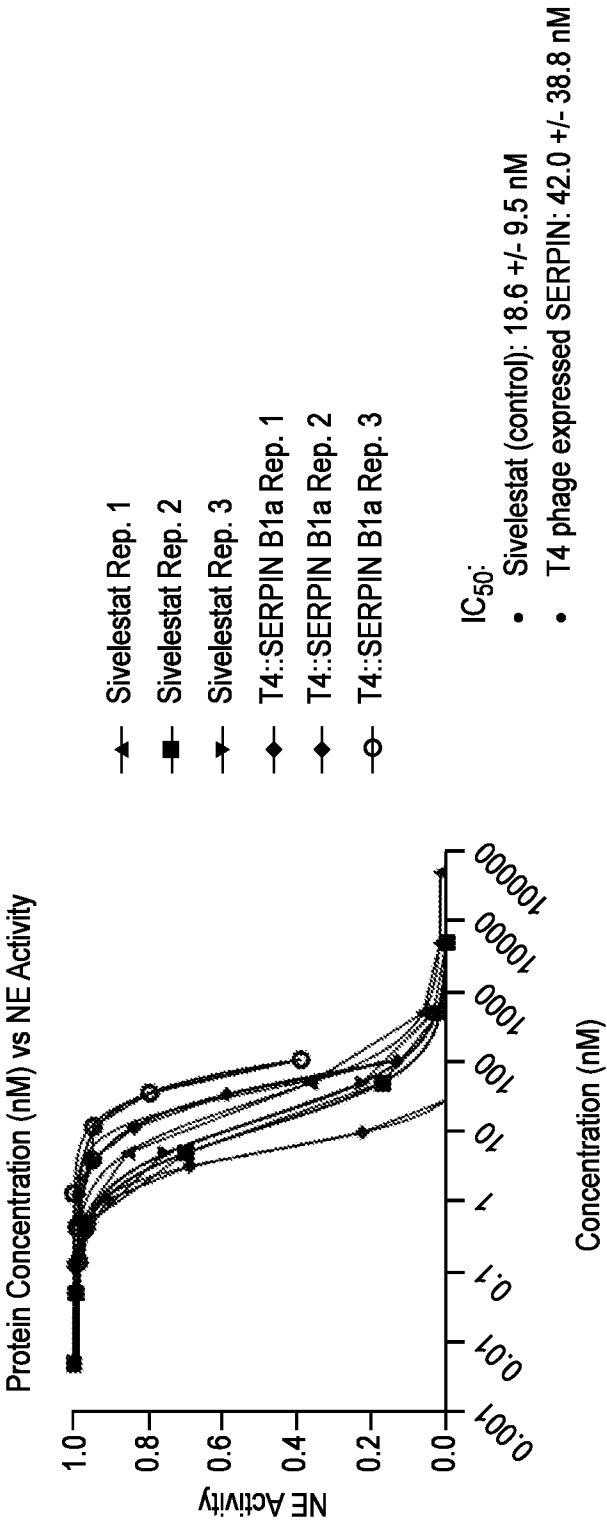


FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2024/033677

A. CLASSIFICATION OF SUBJECT MATTERIPC: **A61K 35/76** (2024.01); **A61K 38/04** (2024.01); **A61K 39/00** (2024.01); **C12N 7/00** (2024.01)CPC: **A61K 35/76**; **C12N 7/00**; **A61K 38/04**; **A61K 39/00**; **C12N 2795/00034**

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)

See Search History Document

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

See Search History Document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History Document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2015/0050717 A1 (COLLINS et al.) 19 February 2015 (19.02.2015) entire document	1-3
A	WO 2021/231689 A1 (CHAN ZUCKERBERG BIOHUB INC. et al.) 18 November 2021 (18.11.2021) entire document	1-3
A	WO 2017/174810 A1 (PHICO THERAPEUTICS LTD.) 12 October 2017 (12.10.2017) entire document	1-3

☐ Further documents are listed in the continuation of Box C.☐ 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

“D” document cited by the applicant in the international application

“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

05 August 2024 (05.08.2024)

Date of mailing of the international search report

19 August 2024 (19.08.2024)

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450, Alexandria, VA 22313-1450

Facsimile No. **571-273-8300**

Authorized officer

MATOS
TAINA

Telephone No. **571-272-4300**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2024/033677

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.
 - b. ☐ furnished subsequent to the international filing date for the purposes of international search (Rule 13ter.1(a)),
☐ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2024/033677

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: **4-24**
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).