

COMPOSITIONS AND METHODS FOR EXPRESSION OF EGFR

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of priority to U.S. Provisional Patent Application Serial No. 63/593,800, filed on October 27, 2023. The disclosure of the above-referenced application is herein expressly incorporated by reference in its entirety, including any drawings.

FIELD

[0002] The present disclosure relates to the field of immunology, and particularly relates to modified viral genomes or self-replicating RNAs (srRNAs) and pharmaceutical compositions containing the same, as well as the use of such nucleic acid molecules and compositions for production of desired products in cell cultures or in a living body. Also provided are methods for eliciting an immune response in a subject in need thereof, as well as methods for preventing and/or treating various health conditions and diseases.

BACKGROUND

[0003] Generation of resistance to cancer therapeutic or prevention agents is a common problem in the treatment of cancer or precancer, and, in several cases, the mechanism of resistance to the therapeutic agent is known. Resistance is often the result of changes in gene expression (over-expression or blocked expression of a protein), change in the gene by mutation, altered sequences by altered splicing or translocation, or altered activation of a protein in the cells (over-activation or blocked activation of a protein).

[0004] One method for dealing with those cancers in which such changes in gene expression, alteration, and mutation occur, has been the development of cancer vaccines. Cancer vaccines target antigens expressed by tumors, but application of these vaccines has not been as effective as once hoped due to induction of immune tolerance by chronic overexpression of the targeted protein in the absence of co-stimulatory molecules and the induction of an immunomodulatory environment. Preventative cancer vaccines may be more promising, but cancers are highly variable, with multiple genetic changes, but few truly universal changes. Thus, it is difficult to predict what antigens will be overexpressed on any specific cancer or whether an individual should be vaccinated and if so, with which antigens.

[0005] The disclosure provided here provides solutions to the problems existing with previous attempts to generate cancer vaccines and potentially offers improved methods for cancer treatment and prevention.

SUMMARY

[0006] The present disclosure relates generally to the development of immuno-therapeutics, such as recombinant nucleic acids constructs and pharmaceutical compositions including the same for use in the prevention and management of various health conditions such as cancer. In particular, as described in greater detail below, some embodiments of the disclosure provide nucleic acid constructs containing sequences that encode a modified viral genome or self-replicating RNA (srRNA), *e.g.*, replicons or self-amplifying RNA, in which at least a portion of the nucleic acid sequence encoding the viral structural proteins of the modified alphavirus genome or srRNA has been replaced by a coding sequence for a polypeptide construct comprising a coding sequence for one or more portions of epidermal growth factor receptor (EGFR) comprising one or more acquired resistance mutations. Also disclosed are recombinant cells that have been engineered to include one or more of the nucleic acid constructs disclosed herein, and pharmaceutical compositions including one or more of the following (a) a nucleic acid construct of the disclosure and/or (b) a recombinant cell of the disclosure. Further provided in particular aspects of the disclosure are compositions and methods for modulating at least one pharmacodynamic effect in a subject, and methods for preventing and/or treating various health conditions, including cancer.

[0007] In one aspect of the disclosure, provided herein are nucleic acid constructs comprising a nucleic acid sequence encoding a modified alphavirus genome or self-replicating RNA (srRNA), wherein at least a portion of the nucleic acid sequence encoding the viral structural proteins of the modified alphavirus genome or srRNA has been replaced by a coding sequence for a polypeptide construct comprising a coding sequence for epidermal growth factor receptor (EGFR) comprising one or more acquired resistance mutations.

[0008] In some embodiments, the modified alphavirus genome or srRNA comprises no nucleic acid sequence encoding viral structural proteins.

[0009] In some embodiments, the nucleic acid sequence encoding the polypeptide construct is operably linked to a promoter sequence. In some embodiments, the promoter sequence is a 26S subgenomic (sg) promoter.

[0010] In some embodiments, the modified alphavirus genome or srRNA is of an alphavirus belonging to the VEEV/EEEV group, or the SFV group, or the SINV group. In some embodiments, the alphavirus is Venezuelan equine encephalitis virus (VEEV), Eastern equine encephalitis virus (EEEV), Madariaga virus (MADV), Chikungunya virus (CHIKV), Western equine encephalitis virus (WEEV), or Sindbis virus (SINV).

[0011] In some embodiments, the srRNA is a capped srRNA comprising a 5'-cap. In some embodiments, the capped srRNA is a co-transcriptionally capped srRNA. In some embodiments, the capped srRNA is an enzymatically capped srRNA.

[0012] In some embodiments, the one or more acquired resistance mutations are configured into a plurality of alteration cassettes arranged in tandem along the length of the coding sequence. In some embodiments, the plurality of alteration cassettes are operably linked to one another by one or more linkers.

[0013] In some embodiments, the one or more acquired resistance mutations comprise one or more (i) activating mutations, (ii) mutations that enhance binding affinity of EGFR to adenosine triphosphate (ATP), and/or (iii) mutations that block binding of EGFR or variant thereof to an inhibitor. In some embodiments, the one or more activating mutations comprises an in-frame insertion in exon 20. In some embodiments, the exon 20 in-frame insertion comprises a mutation selected from the group consisting of A763_Y764insFQEA, S768_D770dup, S768_V769ins, A767_V769dup, D770_N771insX, V769_D770insX, H773_V774insX, H773dup, N771_H773dup, P772_H773insX, and N771_P772insX. In some embodiments, the exon 20 in-frame insertion comprises S768_D770dup, A767_V769dup, and/or H773dup.

[0014] In some embodiments, the one or more activating mutations comprises an L858R substitution.

[0015] In some embodiments, the one or more acquired resistance mutations enhances the binding affinity of EGFR to ATP. In some embodiments, the one or more acquired resistance mutations is selected from the group consisting of T790M, G719X, L858R, L718Q, G724S, L861X, S768I/V, E709X, L747S, D761Y, and T854A.

[0016] In some embodiments, the one or more acquired resistance mutations blocks binding of EGFR to an inhibitor. In some embodiments, the EGFR inhibitor is selected from the group consisting of osimertinib, lazertinib, erlotinib and gefitinib, CO-1686, HM61713, EGF816, ASP8273 and Avitinib, afatinib, mobocertinib, Icotinib, Dacomitinib, Pozotinib Cetuximab, Amivantamab, Mobocertinib, Furmonertinib, DZD9008, CLN-081, STX-721, YK-029A, HS-10376, Zipalertinib, TAK-788, JMT101, ABT-101, BEBT-109, DZG9008, PLB1004, EMB-01, HS-20117, MCLA-129, BLU-945; BDTX-1535; NX-019; JIN-A02; BBT-207; BLU-525; THE-349; STX-241; ABK3376; BI-732; BLU-701; BBT-176, BLU-451, Almonertinib, and Tarloxotinib. In some embodiments, the one or more acquired resistance mutations is selected from the group consisting of C797S, C797G, T790M, L858R, L858M, L718V, L718Q, C796S, L798I, L792X, SV768IL, L692V, G719A, G719S, G719C, G719D, S786I, L861Q, L861R, V834X, V843X, G724S, E709K, E709H, E709A, E709G, E709V, D761Y, D761N, R776C, R776H, and T854A

[0017] In some embodiments, the nucleic acid sequence encoding for EGFR has at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 13-32.

[0018] In some embodiments, the coding sequence for the polypeptide construct comprises a coding sequence for EGFR comprising one or more acquired resistance mutations selected from S768_D770dup, A767_V769dup, H773dup, T790M, L858R, and C797S. In some embodiments, the coding sequence is in a 5' - to 3' -direction.

[0019] Also provided herein is a recombinant cell comprising a nucleic acid construct according to the present disclosure.

[0020] In some embodiments, the recombinant cell is a mammalian cell or an insect cell.

[0021] Also provided herein is a pharmaceutical composition comprising a pharmaceutically acceptable excipient and a nucleic acid construct of the present disclosure.

[0022] In some embodiments, the composition is formulated with a delivery vehicle into a delivery system, wherein the delivery system comprises a liposome, a viral replicon particle (VRP), a lipid-based nanoparticle (LNP), a polymer nanoparticle, a physiologic buffer, a microsphere, an immune stimulating complex (ISCOM), a conjugate of bioactive ligand, or a

combination of any thereof. In some embodiments, the LNP delivery system comprises a cationic lipid, an ionizable cationic lipid, an anionic lipid, or a neutral lipid. In some embodiments, the lipid is present in mass ratio of lipid to RNA from about 100:1 to about 4:1. In some embodiments, the lipid-based nanoparticles have an average diameter of about 25 nm to about 1000 nm.

[0023] In some embodiments, the composition is formulated as a vaccine or immunotherapeutic.

[0024] Also provided herein is a method for inducing an immune response or treating a health condition in a subject in need thereof. The method includes administering to the subject a composition comprising a nucleic acid construct of the present disclosure.

[0025] In some embodiments, the method is a method for inducing an immune response.

[0026] In some embodiments, the method is a method for treating cancer. In some embodiments, the cancer is non-small cell lung cancer (NSCLC).

[0027] In some embodiments, the composition is administered to the subject individually as a single therapy (monotherapy) or as a first therapy in combination with at least one additional therapies.

[0028] Also provided herein is a method for inducing at least one pharmacodynamic effect in a subject. The method includes administering to the subject a composition comprising a nucleic acid construct of the present disclosure.

[0029] In some embodiments, the administered composition results in induced production of one or more of the following: immune responses and mediators selected from TNF, IL-1b, IL-12, IL-2, IFN α , IFN β , IL-6, and IFN γ .

[0030] In some embodiments, the at least one pharmacodynamic effects comprises one

[0031] or more of the following: immunogenicity effect, a biomarker response, a therapeutic effect, a prophylactic effect, a desired effect, an undesired effect, an adverse effect, and effect in a disease model.

[0032] In some embodiments, the administered composition enhances antitumor immunity in a tumor microenvironment. In some embodiments, the subject has a cancer. In some embodiments, the cancer is a lung cancer. In some embodiments, the lung cancer is a NSCLC.

[0033] Each of the aspects and embodiments described herein are capable of being used

together, unless excluded either explicitly or clearly from the context of the embodiment or aspect.

[0034] The foregoing summary is illustrative only and is not intended to be in any way limiting. In addition to the illustrative embodiments and features described herein, further aspects, embodiments, objects and features of the disclosure will become fully apparent from the drawings and the detailed description and the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0035] **FIG. 1** is a bar chart that shows relative expression of EGFR transgene cassettes from VEEV-derived srRNA vectors. ARCA-capped srRNA encoding EGFR comprising one or more acquired resistance mutations was transfected at 500 ng into 500,000 BHK-21 cells by nucleofection and 15 hours later cells were collected and stained with rabbit anti-EGFR antibody EPR15348 and secondary AF647-conjugated goat anti-rabbit IgG H&L. The mean fluorescence intensity (MFI) of transfected cells was determined by flow cytometry (FC).

[0036] **FIG. 2** is a bar chart that shows relative expression of EGFR transgene cassettes from VEEV-derived srRNA vectors. Enzymatically capped srRNA encoding EGFR comprising one or more acquired resistance mutations was transfected at 500 ng into 500,000 BHK-21 cells by nucleofection and 15 hours later cells were collected and stained with rabbit anti-EGFR antibody EPR15348 and secondary AF647-conjugated goat anti-rabbit IgG H&L. The mean fluorescence intensity (MFI) of transfected cells was determined by flow cytometry (FC).

[0037] **FIG. 3** shows determination of optimal cassette design for EGFR monogenic constructs by ELISpot. Mice were injected with a single 10-ug injection I.M. Fourteen days later spleens were collected, and ELISpot was performed to determine the number of IFN γ -expressing T cells present following restimulation with C797S mutant peptides. The graph shows individual mouse values with geomean and 95% CI intervals. Statistics one-way ANOVA.

[0038] **FIG. 4** is a bar chart that shows *in vitro* evaluation of an EGFR transgene cassette expression in different srRNA vectors. srRNA encoding EGFR comprising one or more acquired resistance mutations was transfected at 500 ng into 500,000 BHK-21 cells by nucleofection and 15 hours later cells were collected and stained with rabbit anti-EGFR antibody EPR15348 and secondary AF647-conjugated goat anti-rabbit IgG H&L. The mean fluorescence intensity (MFI) of transfected cells was determined by flow cytometry (FC).

[0039] FIG. 5 is a bar chart that shows *in vitro* evaluation of an EGFR transgene cassette expression in different srRNA vectors. srRNA encoding EGFR comprising one or more acquired resistance mutations was transfected at 50 ng into 500,000 BHK-21 cells by nucleofection and 15 hours later cells were collected and stained with rabbit anti-EGFR antibody EPR15348 and secondary AF488-conjugated goat anti-rabbit IgG H&L. The mean fluorescence intensity (MFI) of transfected cells was determined by flow cytometry (FC).

[0040] FIGs. 6A-6B show T cell responses as measured by IFN γ ELISpot 7 days after injecting mice with 2x10⁶ doses of srRNA 3 weeks apart in HLA-A2 (FIG. 6A) or HLA-A11 (FIG. 6B) transgenic mice. Graph shows individual mouse values with mean and standard deviation. Statistics one-way ANOVA.

DETAILED DESCRIPTION OF THE DISCLOSURE

[0041] Provided herein are, *inter alia*, viral expression systems including self-replicating RNAs (srRNAs) based on RNA viruses (*e.g.*, alphaviruses) with superior expression potential which are suitable for expressing heterologous molecules such as, for example, therapeutic polypeptides, in recombinant cells. For example, some embodiments of the disclosure relate generally to nucleic acid constructs expressing epidermal growth factor receptor (EGFR) comprising one or more acquired resistance mutations, for the purposes of therapeutic treatment of human health conditions or diseases, such as, for example, cancer. These constructs address the problem with treatment modalities, such as tyrosine kinase inhibitor administration, due to the development of acquired resistance mutations that have been previously demonstrated. In some embodiments, provided herein are gene expression systems with superior expression potential which are suitable for expressing a coding sequence for an EGFR comprising one or more acquired resistance mutations, in recombinant cells. For example, some embodiments of the disclosure relate to nucleic acid constructs such as, *e.g.* expression constructs and vectors, containing a modified genome or srRNA of an alphavirus in which at least a portion of the nucleic acid sequence encoding the viral structural proteins of the modified alphavirus genome or srRNA has been replaced by a coding sequence for a polypeptide construct comprising (a) a coding sequence for EGFR comprising one or more acquired resistance mutations. In some embodiments, the polypeptide construct does not comprise a dimerization domain. Further provided are recombinant cells that are genetically engineered to include one or more of the

nucleic acid constructs disclosed herein. Biomaterials and recombinant products derived from such recombinant cells are also within the scope of the application. Also provided are compositions and methods useful for (i) modulating a pharmacodynamic effect and for preventing and/or treating a health condition in a subject in need thereof.

[0042] As described in greater detail below, EGFR inhibitors have been approved or tested for treatment of a variety of cancers, including non-small cell lung cancer (NSCLC), head and neck cancer, colorectal carcinoma, and Her2-positive breast cancer, and are increasingly being added to standard therapy. EGFR inhibitors, which may target either the intracellular tyrosine kinase domain or the extracellular domain of the EGFR target, are generally plagued by low population response rates, leading to ineffective or non-optimal chemotherapy in many instances, as well as unnecessary drug toxicity and expense. For example, a reported clinical response rate for treatment of colorectal carcinoma with cetuximab (a chimeric monoclonal antibody targeting the extracellular domain of EGFR) is about 11% (Cunningham et al, N Engl J Med 2004; 351: 337-45), and a reported clinical response rate for treatment of NSCLC with erlotinib is about 8.9% (Shepherd F A, et al, N Engl J Med 2005; 353:123-132). Since the identification of activating mutations of the (EGFR in NSCLC (Lynch et al., N Engl J Med 2004; 2004; 350: 2129-2139), patients whose tumor displays such mutations are treated with EGFR TKI (tyrosine kinase inhibitors), resulting in a dramatic increase in the clinical response rate (Chong and Janne, Nat Med 2013; 19: 1389-1400). However, a resistance mechanism almost invariably occurs when EGFR inhibitors are used in the treatment of this type of patients.

[0043] The disclosure provided here provides, *inter alia*, solutions to the development of acquired resistance mutations in cancer and thereby offers improved methods for treatment of health conditions including cancer. As illustrated in Examples below, srRNA vectors capable of expressing various EGFR acquired resistance mutations can be generated by using the compositions and methods disclosed herein.

DEFINITIONS

[0044] Unless otherwise defined, all terms of art, notations, and other scientific terms or terminology used herein are intended to have the meanings commonly understood by those of skill in the art to which this application pertains. In some cases, terms with commonly understood meanings are defined herein for clarity and/or for ready reference, and the inclusion

of such definitions herein should not necessarily be construed to represent a substantial difference over what is generally understood in the art. Many of the techniques and procedures described or referenced herein are well understood and commonly employed using conventional methodology by those skilled in the art.

[0045] The singular form “a”, “an”, and “the” include plural references unless the context clearly dictates otherwise. For example, the term “a cell” includes one or more cells, comprising mixtures thereof. “A and/or B” is used herein to include all of the following alternatives: “A”, “B”, “A or B”, and “A and B”.

[0046] The terms “administration” and “administering”, as used herein, refer to the delivery of a bioactive composition or formulation by an administration route comprising, but not limited to, intranasal, transdermal, intravenous, intra-arterial, intramuscular, intranodal, intraperitoneal, subcutaneous, intramuscular, oral, intravaginal, and topical administration, or combinations thereof. The term includes, but is not limited to, administering by a medical professional and self-administering.

[0047] The terms “cell”, “cell culture”, and “cell line” refer not only to the particular subject cell, cell culture, or cell line but also to the progeny or potential progeny of such a cell, cell culture, or cell line, without regard to the number of transfers or passages in culture. It should be understood that not all progeny are exactly identical to the parental cell. This is because certain modifications may occur in succeeding generations due to either mutation (*e.g.*, deliberate or inadvertent mutations) or environmental influences (*e.g.*, methylation or other epigenetic modifications), such that progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein, so long as the progeny retain the same functionality as that of the original cell, cell culture, or cell line.

[0048] The term “construct” refers to a recombinant molecule, *e.g.*, recombinant nucleic acid or polypeptide, including one or more isolated nucleic acid sequences or amino acid sequences from heterologous sources. For example, polypeptide constructs can be chimeric polypeptide molecules in which two or more amino acid sequences of different origin are operably linked to one another in a single polypeptide construct. Similarly, nucleic acid constructs can be chimeric nucleic acid molecules in which two or more nucleic acid sequences of different origin are assembled into a single nucleic acid molecule. Thus, representative nucleic

acid constructs can include any recombinant nucleic acid molecules, linear or circular, single stranded or double stranded DNA or RNA nucleic acid molecules, derived from any source, such as a plasmid, cosmid, virus, autonomously replicating polynucleotide molecule, phage, capable of genomic integration or autonomous replication, comprising a nucleic acid molecule where one or more nucleic acid sequences have been operably linked. Two or more nucleic acid constructs can be contained within a single nucleic acid molecule, such as a single vector, or can be containing within two or more separate nucleic acid molecules, such as two or more separate vectors.

[0049] In some embodiments of the disclosure, the nucleic acid construct may be incorporated within a vector. The term “vector” is used herein to refer to a nucleic acid molecule or sequence capable of transferring or transporting another nucleic acid molecule. Thus, the term “vector” encompasses both DNA-based vectors and RNA-based vectors. The term “vector” includes cloning vectors and expression vectors, as well as viral vectors and integrating vectors. An “expression vector” is a vector that includes a regulatory region, thereby capable of expressing DNA sequences and fragments *in vitro*, *ex vivo*, and/or *in vivo*. In some embodiments, a vector may include sequences that direct autonomous replication in a cell such as, for example a plasmid (DNA-based vector) or a self-replicating RNA vector. In some embodiments, a vector may include sequences sufficient to allow integration into host cell DNA. Useful vectors include, for example, plasmids (*e.g.*, DNA plasmids or RNA plasmids), transposons, cosmids, bacterial artificial chromosomes, and viral vectors. In some embodiments, the vector of the disclosure can be single-stranded vector (*e.g.*, ssDNA or ssRNA). In some embodiments, the vector of the disclosure can be double-stranded vector (*e.g.*, dsDNA or dsRNA). In some embodiments, a vector is a gene delivery vector. In some embodiments, a vector is used as a gene delivery vehicle to transfer a gene into a cell.

[0050] In addition to the components of the construct, the vector may include, for example, one or more selectable markers, one or more origins of replication, such as prokaryotic and eukaryotic origins, at least one multiple cloning site, and/or elements to facilitate stable integration of the construct into the genome of a cell. As described above, two or more constructs can be incorporated within a single nucleic acid molecule, such as a single vector, or can be incorporated within two or more separate nucleic acid molecules, such as two or more separate

vectors. An “expression construct” generally includes at least a control sequence operably linked to a nucleotide sequence of interest. In this manner, for example, promoters in operable connection with the nucleotide sequences to be expressed are provided in expression constructs for expression in a cell. For the practice of the present disclosure, compositions and methods for preparing and using constructs and cells are known to one skilled in the art.

[0051] The term “effective amount”, “therapeutically effective amount”, or “pharmaceutically effective amount” of a composition of the disclosure, *e.g.*, nucleic acid constructs (*e.g.*, srRNA constructs), recombinant cells, and/or pharmaceutical compositions, generally refers to an amount sufficient for the composition to accomplish a stated purpose relative to the absence of the composition (*e.g.*, achieve the effect for which it is administered, stimulate an immune response, prevent or treat a disease, or reduce one or more symptoms of a disease, disorder, infection, or health condition). An example of an “effective amount” is an amount sufficient to contribute to the treatment, prevention, or reduction of a symptom or symptoms of a disease, which could also be referred to as a “therapeutically effective amount.” A “reduction” of a symptom means decreasing of the severity or frequency of the symptom(s), or elimination of the symptom(s). The exact amount of a composition including a “therapeutically effective amount” will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using known techniques (see, *e.g.*, Lieberman, *Pharmaceutical Dosage Forms* (vols. 1-3, 1992); Lloyd, *The Art, Science and Technology of Pharmaceutical Compounding* (1999); Pickar, *Dosage Calculations* (1999); and Remington: *The Science and Practice of Pharmacy*, 20th Edition, 2003, Gennaro, Ed., Lippincott, Williams & Wilkins).

[0052] The term “naked” as used herein when referencing nucleic acids that are substantially free of other macromolecules, such as lipids, polymers, and proteins. A “naked” nucleic acid, such as a self-replicating RNA, is not formulated with other macromolecules to improve cellular uptake. Accordingly, a naked nucleic acid is not encapsulated in, absorbed on, or bound to a liposome, a microparticle, a nanoparticle, a cationic emulsion, and the like.

[0053] The term “operably linked”, as used herein, denotes a physical or functional linkage between two or more elements, *e.g.*, polypeptide sequences or polynucleotide sequences, which permits them to operate in their intended fashion. For example, the term “operably linked” when used in context of the nucleic acid molecules described herein or the coding sequences and

promoter sequences in a nucleic acid molecule means that the coding sequences and promoter sequences are in-frame and in proper spatial and distance away to permit the effects of the respective binding by transcription factors or RNA polymerase on transcription. It should be understood that operably linked elements may be contiguous or non-contiguous (*e.g.*, linked to one another through a linker). In the context of polypeptide constructs, “operably linked” refers to a physical linkage (*e.g.*, directly or indirectly linked) between amino acid sequences (*e.g.*, different segments, portions, regions, or domains) to provide for a described activity of the constructs. Operably linked segments, portions, regions, and domains of the polypeptides or nucleic acid molecules disclosed herein may be contiguous or non-contiguous (*e.g.*, linked to one another through a linker).

[0054] The term “portion” as used herein refers to a fraction. With respect to a particular structure such as a polynucleotide sequence or an amino acid sequence or protein the term “portion” thereof may designate a continuous or a discontinuous fraction of said structure. For example, a portion of an amino acid sequence comprises at least 1%, at least 5%, at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, and at least 90% of the amino acids of said amino acid sequence. In addition or alternatively, if the portion is a discontinuous fraction, said discontinuous fraction is composed of 2, 3, 4, 5, 6, 7, 8, or more parts of a structure (*e.g.*, domains of a protein), each part being a continuous element of the structure. For example, a discontinuous fraction of an amino acid sequence may be composed of 2, 3, 4, 5, 6, 7, 8, or more, for example not more than 4 parts of said amino acid sequence, wherein each part comprises at least 1, at least 2, at least 3, at least 4, at least 5 continuous amino acids, at least 10 continuous amino acids, at least 20 continuous amino acids, or at least 30 continuous amino acids of the amino acid sequence.

[0055] The term “percent identity,” as used herein in the context of two or more nucleic acids or proteins, refers to two or more sequences or subsequences that are the same or have a specified percentage of nucleotides or amino acids that are the same (*e.g.*, about 60% sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region, when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by

manual alignment and visual inspection. See *e.g.*, the NCBI website at ncbi.nlm.nih.gov/BLAST. Such sequences are then said to be “substantially identical.” This definition also refers to, or may be applied to, the complement of a query sequence. This definition includes sequence comparison performed by a BLAST algorithm wherein the parameters of the algorithm are selected to give the largest match between the respective sequences over the entire length of the respective reference sequences. This definition also includes sequences that have deletions and/or additions, as well as those that have substitutions. Sequence identity can be calculated over a region that is at least about 20 amino acids or nucleotides in length, or over a region that is 10-100 amino acids or nucleotides in length, or over the entire length of a given sequence. Sequence identity can be calculated using published techniques and widely available computer programs, such as the GCS program package (Devereux *et al.*, *Nucleic Acids Res* (1984) 12:387), BLASTP, BLASTN, FASTA (Atschul *et al.*, *J Mol Biol* (1990) 215:403). Sequence identity can be measured using sequence analysis software such as the Sequence Analysis Software Package of the Genetics Computer Group at the University of Wisconsin Biotechnology Center (1710 University Avenue, Madison, Wis. 53705), with the default parameters thereof. Additional methodologies that can suitably be utilized to determine similarity or identity amino acid sequences include those relying on position-specific structure-scoring matrix (P3SM) that incorporates structure-prediction scores from Rosetta, as well as those based on a length-normalized edit distance as described previously in, *e.g.*, Setcliff *et al.*, *Cell Host & Microbe* 23(6), May 2018.

[0056] The term “pharmaceutically acceptable excipient” as used herein refers to any suitable substance that provides a pharmaceutically acceptable carrier, additive, or diluent for administration of a compound(s) of interest to a subject. As such, “pharmaceutically acceptable excipient” can encompass substances referred to as pharmaceutically acceptable diluents, pharmaceutically acceptable additives, and pharmaceutically acceptable carriers. As used herein, the term “pharmaceutically acceptable carrier” includes, but is not limited to, saline, solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Supplementary active compounds (*e.g.*, antibiotics and additional therapeutic agents) can also be incorporated into the compositions.

[0057] The term “recombinant” when used with reference to a cell, a nucleic acid, a protein, or a vector, indicates that the cell, nucleic acid, protein or vector has been altered or produced through human intervention such as, for example, has been modified by or is the result of laboratory methods. Thus, for example, recombinant proteins and nucleic acids include proteins and nucleic acids produced by laboratory methods. Recombinant proteins can include amino acid residues not found within the native (non-recombinant or wild-type) form of the protein or can include amino acid residues that have been modified, *e.g.*, labeled. The term can include any modifications to the peptide, protein, or nucleic acid sequence. Such modifications may include the following: any chemical modifications of the peptide, protein or nucleic acid sequence, including of one or more amino acids, deoxyribonucleotides, or ribonucleotides; addition, deletion, and/or substitution of one or more of amino acids in the peptide or protein; creation of a fusion protein, *e.g.*, a fusion protein comprising an antibody fragment; and addition, deletion, and/or substitution of one or more of nucleic acids in the nucleic acid sequence. The term “recombinant” when used in reference to a cell is not intended to include naturally-occurring cells but encompass cells that have been engineered/modified to include or express a polypeptide or nucleic acid that would not be present in the cell if it was not engineered/modified.

[0058] As used herein, a “subject” or an “individual” includes animals, such as human (*e.g.*, human individuals) and non-human animals. In some embodiments, a “subject” or “individual” is a patient under the care of a physician. Thus, the subject can be a human patient or an individual who has, is at risk of having, or is suspected of having a health condition of interest (*e.g.*, autoimmune disease, inflammatory disease, or cardiovascular disease) and/or one or more symptoms of the health condition. The subject can also be an individual who is diagnosed with a risk of the health condition of interest at the time of diagnosis or later. The term “non-human animals” includes all vertebrates, *e.g.*, mammals, *e.g.*, rodents, *e.g.*, mice, non-human primates, and other mammals, such as *e.g.*, sheep, dogs, cats, cows, chickens, and non-mammals, such as amphibians, reptiles, *etc.*

[0059] It is understood that aspects and embodiments of the disclosure described herein include “comprising”, “consisting”, and “consisting essentially of” aspects and embodiments. As used herein, “comprising” is synonymous with “including”, “containing”, or “characterized by”,

and is inclusive or open-ended and does not exclude additional, unrecited elements or method steps. As used herein, "consisting of" excludes any elements, steps, or ingredients not specified in the claimed composition or method. As used herein, "consisting essentially of" does not exclude materials or steps that do not materially affect the basic and novel characteristics of the claimed composition or method. Any recitation herein of the term "comprising", particularly in a description of components of a composition or in a description of steps of a method, is understood to encompass those compositions and methods consisting essentially of and consisting of the recited components or steps.

[0060] 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.

[0061] Certain ranges are presented herein with numerical values being preceded by the term "about" which, as used herein, has its ordinary meaning of approximate. The term "about" is used herein to provide literal support for the exact number that it precedes, as well as a number that is near to or approximately the number that the term precedes. In determining whether a number is near to or approximately a specifically recited number, the near or approximating unrecited number may be a number which, in the context in which it is presented, provides the substantial equivalent of the specifically recited number. If the degree of approximation is not otherwise clear from the context, "about" means either within plus or minus 10% of the provided value, or rounded to the nearest significant figure, in all cases inclusive of the provided value. In some embodiments, the term "about" indicates the designated value $\pm$ up to 10%, up to $\pm$ 5%, or up to $\pm$ 1%.

[0062] Where a range of values is provided, it is understood by one having ordinary skill in the art that all ranges disclosed herein encompass any and all possible sub-ranges and combinations of sub-ranges thereof. Any listed range can be easily recognized as sufficiently

describing and enabling the same range being broken down into at least equal halves, thirds, quarters, fifths, tenths, *etc.* As a non-limiting example, each range discussed herein can be readily broken down into a lower third, middle third and upper third, *etc.* As will also be understood by one skilled in the art all language such as “up to”, “at least”, “greater than”, “less than”, and the like include the number recited and refer to ranges which can be subsequently broken down into sub-ranges as discussed above. Finally, as will be understood by one skilled in the art, a range includes each individual member. Thus, for example, a group having 1-3 articles refers to groups having 1, 2, or 3 articles. Similarly, a group having 1-5 articles refers to groups having 1, 2, 3, 4, or 5 articles, and so forth.

[0063] All genes, gene names, and gene products disclosed herein are intended to correspond to homologs from any species for which the compositions and methods disclosed herein are applicable. Thus, the terms include, but are not limited to genes and gene products from humans and mice. It is understood that when a gene or gene product from a particular species is disclosed, this disclosure is intended to be exemplary only, and is not to be interpreted as a limitation unless the context in which it appears clearly indicates. Thus, for example, for the genes or gene products disclosed herein, which in some embodiments relate to mammalian nucleic acid and amino acid sequences, are intended to encompass homologous and/or orthologous genes and gene products from other animals including, but not limited to other mammals, fish, amphibians, reptiles, and birds. In some embodiments, the genes, nucleic acid sequences, amino acid sequences, peptides, polypeptides and proteins are human. The term “gene” is also intended to include variants thereof.

[0064] Headings, *e.g.*, (a), (b), (i) *etc.*, are presented merely for ease of reading the specification and claims. The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described. The use of headings in the specification or claims does not require the steps or elements be performed in alphabetical or numerical order or the order in which they are presented.

[0065] It is appreciated that certain features of the disclosure, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. Conversely, various features of the disclosure, which are, for brevity, described in the context of a single embodiment, may also be provided separately or in any

suitable sub-combination. All combinations of the embodiments pertaining to the disclosure are specifically embraced by the present disclosure and are disclosed herein just as if each and every combination was individually and explicitly disclosed. In addition, all sub-combinations of the various embodiments and elements thereof are also specifically embraced by the present disclosure and are disclosed herein just as if each and every such sub-combination was individually and explicitly disclosed herein.

EGFR

[0066] Epidermal growth factor receptor, “EGFR”, “ErbB” or “HER” is a receptor protein tyrosine kinase which belongs to the ErbB receptor family and includes ErbB1 (or HER1 or EGFR), ErbB2 (or HER2), ErbB3 (or HER3) and ErbB4 (or HER4) receptors (Ullrich, 1984). The ErbB receptor generally comprises an extracellular domain, which may bind an EGFR ligand; a lipophilic transmembrane domain; a conserved intracellular tyrosine kinase domain; and a carboxyl-terminal signaling domain harboring several tyrosine residues which may be phosphorylated. Being activated by their six structurally related agonists-EGF, tumor growth factor α (TGF α), heparin-binding EGF-like growth factor (HB-EGF), amphiregulin, betacellulin and epiregulin- the receptors promote pathways entailing proliferation and transformation. Activated EGFRs homo- or heterodimerize and subsequently autophosphorylation of cytoplasmic tyrosine residues is initiated. These phosphorylated amino acids represent docking sites for a variety of different proteins (Prenzel 2001). Tyrosine phosphorylation of the EGFR leads to the recruitment of diverse signaling proteins, including the Adaptor proteins GRB2 (Growth Factor Receptor-Bound Protein-2) and Nck (Nck Adaptor Protein), PLC-Gamma (Phospholipase-C-Gamma), SHC (Src Homology-2 Domain Containing Transforming Protein), and STATS (Signal Transducer and Activator of Transcription 5).

[0067] The epidermal growth factor receptor (EGFR) has been identified as a relevant target for treatment of solid tumors, as it is involved in regulating cellular functions important in the proliferation and survival of cancer cells. EGFR is commonly expressed in a range of tumors, and high expression is often related to poor prognosis. A new class of targeted therapies directed at inhibiting the EGFR, tyrosine kinase inhibitors, have appeared. Two known examples are gefitinib (Iressa) or erlotinib (Tarceva). Despite initial responses of some patients to these therapies, patients eventually progress by mechanisms of "acquired" resistance.

Self-replicating RNA

[0068] As will be appreciated by the skilled artisan, the term “self-replicating RNA” (srRNA) refers to RNA molecule that contains all of the genetic information required for directing its own amplification or self-replication within a permissive cell. Therefore, srRNA is sometimes also referred to as “self-amplifying RNA” (saRNA). In some embodiments, the srRNA is a “replicon,” which can be a linear or circular section of DNA or RNA which replicates sequentially as a unit. Non-limiting examples of replicons include “replicon RNA” or “RNA replicon.” To direct its own replication, the srRNA generally (1) encodes polymerase, replicase, or other proteins which may interact with viral or host cell-derived proteins, nucleic acids or ribonucleoproteins to catalyze the RNA amplification process; and (2) contain *cis*-acting RNA sequences required for replication and transcription of the subgenomic RNA. These sequences may be bound during the process of replication to its self-encoded proteins, or non-self-encoded cell-derived proteins, nucleic acids or ribonucleoproteins, or complexes between any of these components. In some embodiments of the disclosure, the replicon, *e.g.*, srRNA, is derived from a Venezuelan Equine Encephalitis Virus (VEEV). In some embodiments of the disclosure, a VEEV srRNA construct (*e.g.*, srRNA, saRNA, or RNA replicon molecule) generally contains the following elements: 5' viral or defective-interfering RNA sequence(s) required in *cis* for replication, sequences coding for biologically active alphavirus non-structural proteins (*e.g.*, nsP1, nsP2, nsP3, and nsP4), a subgenomic promoter (sg) for the subgenomic RNA (sgRNA), 3' viral sequences required in *cis* for replication, and optionally a polyadenylate tract (poly(A)). In some instances, a subgenomic promoter (sg) that directs expression of a heterologous sequence can be included in the srRNA construct of the disclosure.

[0069] Further, the term srRNA molecule (*e.g.*, srRNA, saRNA, or RNA replicon molecule) generally refers to a molecule of positive polarity, or “message” sense, and the srRNA may be of length different from that of any known, naturally-occurring alphavirus. In some embodiments of the present disclosure, the srRNA does not contain at least a portion of the coding sequence for one or more of the alphavirus structural proteins; and/or sequences encoding structural genes can be substituted with heterologous sequences. In those instances, where the srRNA is to be packaged into a recombinant alphavirus particle, it can contain one or more sequences, so-called packaging signals, which serve to initiate interactions with alphavirus

structural proteins that lead to particle formation.

[0070] The srRNA constructs of the disclosure generally have a length of at least about 2 kb. For example, the srRNA can have a length of at least about 2 kb, at least about 3 kb, at least about 4 kb, at least about 5 kb, at least about 6 kb, at least about 7 kb, at least about 8 kb, at least about 9 kb, at least about 10 kb, at least about 11 kb, at least about 12 kb or more than 12 kb. In some embodiments, the srRNA can have a length of about 4 kb to about 20 kb, about 4 kb to about 18 kb, about 5 kb to about 16 kb, about 6 kb to about 14 kb, about 7 kb to about 12 kb, about 8 kb to about 16 kb, about 9 kb to about 14 kb, about 10 kb to about 18 kb, about 11 kb to about 16 kb, about 5 kb to about 18 kb, about 6 kb to about 20 kb, about 5 kb to about 10 kb, about 5 kb to about 8 kb, about 5 kb to about 7 kb, about 5 kb to about 6 kb, about 6 kb to about 12 kb, about 6 kb to about 11 kb, about 6 kb to about 10 kb, about 6 kb to about 9 kb, about 6 kb to about 8 kb, about 6 kb to about 7 kb, about 7 kb to about 11 kb, about 7 kb to about 10 kb, about 7 kb to about 9 kb, about 7 kb to about 8 kb, about 8 kb to about 11 kb, about 8 kb to about 10 kb, about 8 kb to about 9 kb, about 9 kb to about 11 kb, about 9 kb to about 10 kb, or about 10 kb to about 11 kb. In some embodiments, the srRNA can have a length of about 6 kb to about 14 kb. In some embodiments, the srRNA can have a length of about 6 kb to about 16 kb.

[0071] In some embodiments, the srRNA constructs of the disclosure can include a 5' cap. The term "5' cap" as used herein refers to a structure found on the 5' end of some eukaryotic RNAs, e.g., RNA transcripts, and generally includes a dinucleotide or a guanosine nucleotide connected to an RNA, e.g., mRNA, via a 5'- to 5'-triphosphate linkage (also referred to as Gppp or G(5')ppp(5')). In some embodiments, a guanosine nucleoside included in a 5' cap may be modified, for example, by methylation at one or more positions (e.g., at the 7-position) on a base (guanine), and/or by methylation at one or more positions of a ribose. In some embodiments, a guanosine nucleoside included in a 5' cap includes a 2'O methylation at a ribose (2'OMeG). In some embodiments, a guanosine nucleoside included in a 5' cap includes methylation at the 7-position of guanine (m7G). In some embodiments, a guanosine nucleoside included in a 5' cap includes methylation at the 7-position of guanine and a 2' O methylation at a ribose (m7(2'OMeG)).

[0072] Multiple distinct cap structures can be used to generate the 5' cap of *in vitro* transcribed synthetic srRNA. In some embodiments, providing an srRNA with a 5' cap disclosed

herein or a 5' cap analog may be achieved by *in vitro* transcription, in which a 5' cap is co-transcriptionally expressed into an RNA strand, or may be attached to an RNA post-transcriptionally using capping enzymes. Accordingly, in some embodiments, 5' capping of synthetic srRNA can be performed co-transcriptionally with chemical cap analogs (i.e., capping during *in vitro* transcription). For example, CleanCap® technology provides high efficiency capping (90%+) in a co-transcriptional reaction using commercially available reagents with an AG initiator to provide a natural Cap 1 structure with a 2'-O-methyl group and N7 methyl on separate guanine components. As another example, the Anti-Reverse Cap Analog (ARCA) cap contains a 5'-5'-triphosphate guanine-guanine linkage where one guanine contains an N7 methyl group as well as a 2'-O-methyl group. Alternatively, in some embodiments, synthetic srRNA molecules may also be enzymatically capped post-transcriptionally. These may generate a more authentic 5' cap structure that more closely mimics, either structurally or functionally, the endogenous 5' cap which have enhanced binding of cap binding proteins, increased half-life, reduced susceptibility to 5' endonucleases and/or reduced 5' decapping. Numerous synthetic 5' cap analogs have been developed and are known in the art to enhance mRNA stability and translatability (see, e.g., Grudzien-Nogalska E. et al., Synthetic mRNAs with superior translation and stability properties in Synthetic Messenger RNA and Cell Metabolism Modulation in Methods in Molecular Biology v.969 (Rabinovich, P. H. Ed), 2013).

[0073] Thus, in some embodiments, the srRNA of the disclosure is co-transcriptionally capped srRNA. Exemplary co-transcriptionally capped srRNAs include, without limitation, anti-reverse cap analogs (ARCAs) and CleanCap® srRNA.

[0074] In other embodiments, the srRNA of the disclosure is enzymatically capped srRNA. Exemplary capping enzymes include, without limitation, Vaccinia virus capping enzyme (VCE) and Faustovirus capping enzyme (FCE).

COMPOSITIONS OF THE DISCLOSURE

[0075] As described in greater detail below, one aspect of the present disclosure relates to nucleic acid constructs sequences that encode a modified alphavirus genome or srRNA where at least a portion of the nucleic acid sequence encoding the viral structural proteins of the modified alphavirus genome or srRNA has been replaced by a coding sequence for a polypeptide construct comprising a coding sequence for EGFR comprising one or more acquired resistance mutations.

Also provided are recombinant cells and cell cultures that have been engineered to include a nucleic acid construct as disclosed herein.

A. Nucleic acid constructs

[0076] As described in greater detail below, one aspect of the present disclosure relates to nucleic acid constructs including a nucleic acid sequence encoding a modified alphavirus genome or srRNA where at least a portion of the nucleic acid sequence encoding the viral structural proteins of the modified alphavirus genome or srRNA has been replaced by an coding sequence for a polypeptide construct comprising a coding sequence for an epidermal growth factor receptor (EGFR) comprising one or more acquired resistance mutations. In some embodiments, the sequence encoding a srRNA construct can be operably linked, *e.g.*, placed under the control of elements required for expression (*e.g.*, promoter sequences), which allow expression of the srRNA construct in a host cell, in a subject, or in an *ex-vivo* cell-free expression system.

[0077] The terms "nucleic acid molecule" and "polynucleotide" are used interchangeably herein, and refer to both RNA and DNA molecules, including nucleic acid molecules comprising cDNA, genomic DNA, synthetic DNA, and DNA or RNA molecules containing nucleic acid analogs. A nucleic acid molecule can be double-stranded or single-stranded (*e.g.*, a sense strand or an antisense strand). A nucleic acid molecule may contain unconventional or modified nucleotides. The terms "polynucleotide sequence" and "nucleic acid sequence" as used herein interchangeably refer to the sequence of a polynucleotide molecule. The nomenclature for nucleotide bases as set forth in 37 CFR §1.822 is used herein.

[0078] Nucleic acid molecules of the present disclosure can be of any length, including for example, between about 1.5 Kb and about 50 Kb, between about 5 Kb and about 40 Kb, between about 5 Kb and about 30 Kb, between about 5 Kb and about 20 Kb, or between about 10 Kb and about 50 Kb, for example between about 15 Kb to 30 Kb, between about 20 Kb and about 50 Kb, between about 20 Kb and about 40 Kb, about 5 Kb and about 25 Kb, or about 30 Kb and about 50 Kb.

[0079] Non-limiting exemplary embodiments of the nucleic acid constructs (*e.g.*, srRNA constructs) of the disclosure can include one or more of the following features. In some embodiments of the disclosure, the coding sequence for EGFR is optimized for one of more

desired characteristics. In some embodiments, the coding sequence for EGFR is optimized for one or more of the following: (a) enhancing RNA stability, (b) enhancing expression level, (c) minimizing rare codon usage, (d) minimizing secondary structures, (e) facilitating better srRNA replication, and (f) facilitating better RNA manufacturing process.

[0080] In some embodiments, the modified alphavirus genome or srRNA vector is devoid of at least a portion of the nucleic acid sequence encoding one or more of the viral structural proteins CP, E1, E2, E3, and 6K of the alphavirus genome or srRNA vector. In some embodiments, the modified alphavirus genome or srRNA vector is devoid of a portion of or the entire sequence encoding CP. In some embodiments, the modified alphavirus genome or srRNA vector is devoid of a portion of or the entire sequence encoding E1. In some embodiments, the modified alphavirus genome or srRNA vector is devoid of a portion of or the entire sequence encoding E2. In some embodiments, the modified alphavirus genome or srRNA vector is devoid of a portion of or the entire sequence encoding E3. In some embodiments, the modified alphavirus genome or srRNA vector is devoid of a portion of or the entire sequence encoding 6K. In some embodiments, the modified alphavirus genome or srRNA vector is devoid of a portion of or the entire sequence encoding a combination of CP, E1, E2, E3, and 6K. In some embodiments of the disclosure, the coding sequence for nonstructural proteins nsP1, nsP2, nsP3, and nsP4 of the modified alphavirus genome or srRNA vector is present, however at least a portion of or the entire sequence encoding one or more structural proteins (*e.g.*, CP, E1, E2, E3, and 6K) of the modified alphavirus genome or srRNA vector is absent.

[0081] In some embodiments, the modified alphavirus genome or srRNA vector is devoid of a substantial portion of the nucleic acid sequence encoding one or more viral structural proteins. The skilled artisan will understand that a substantial portion of a nucleic acid sequence encoding a viral structural polypeptide can include enough of the nucleic acid sequence encoding the viral structural polypeptide to afford putative identification of that polypeptide, either by manual evaluation of the sequence by one skilled in the art, or by computer-automated sequence comparison and identification using algorithms such as BLAST (see, for example, in “Basic Local Alignment Search Tool”; Altschul SF *et al.*, J. Mol. Biol. 215:403-410, 1993). Accordingly, a substantial portion of a nucleotide sequence comprises enough of the sequence to afford specific identification and/or isolation of a nucleic acid fragment comprising the sequence.

For example, a substantial portion of a nucleic acid sequence can include at least about 20%, for example, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 95% of the full-length nucleic acid sequence.

[0082] In some embodiments, the modified alphavirus genome or srRNA vector of the present disclosure is devoid of the entire sequence encoding viral structural proteins, *e.g.*, the modified alphavirus genome or srRNA vector includes no nucleic acid sequence encoding the viral structural proteins.

[0083] The nucleic acid constructs of the disclosure further include a coding sequence for a polypeptide construct that replaces at least a portion of the nucleic acid sequence encoding the viral structural proteins of the modified alphavirus genome or srRNA. A coding sequence for a polypeptide construct can be a construct of genetic material that contains coding sequences and enough regulatory information to direct proper transcription and/or translation of the coding sequences in a cell, *in vivo* and/or *ex vivo*. The coding sequence for a polypeptide construct can be inserted into a vector for targeting to a desired host cell and/or into a subject. Accordingly, in some embodiments, the term “coding sequence for a polypeptide construct” can be used interchangeably with the term “expression construct.” In some embodiments, a coding sequence for a polypeptide construct can be a nucleic acid construct that includes a gene encoding a protein or functional RNA operably linked to regulatory elements such as, for example, a promoter and/or a termination signal, and optionally, any or a combination of other nucleic acid sequences that affect the transcription or translation of the gene.

[0084] As described above, the nucleic acid constructs described herein include coding sequences for EGFR. In some embodiments, the nucleic acid constructs encode polypeptides containing peptide/epitopes or combinations thereof of EGFR that are able to elicit an immune response. The variants of EGFR can encompass coding sequences for a polypeptide having an amino acid sequence that is the same or essentially the same as that of the reference protein (*e.g.*, EGFR) except having at least one amino acid modified, for example, deleted, inserted, or replaced, respectively. The amino acid replacement may be a conservative amino acid substitution, preferably at a non-essential amino acid residue in the protein. A “conservative amino acid substitution” is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains

are known in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), non-polar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). A variant of a protein may have an amino acid sequence at least about 80%, 90%, 95%, or 99%, preferably at least about 90%, more preferably at least about 95%, identical to the amino acid sequence of the protein. Preferably, a variant is a functional variant of a protein that retains the same function as the protein. The terms “variant”, when used in reference to a nucleic acid sequence, refer to a nucleic acid sequence that differs by one or more nucleotides from another, usually related nucleotide acid sequence. As such, the term “variant” can refer to a change of one or more nucleotides of a reference nucleic acid which includes the insertion of one or more new nucleotides, deletion of one or more nucleotides, and substitution of one or more existing nucleotides. A variant can also include a point mutation, multiple mutation, single nucleotide polymorphism (SNP), deletion, insertion, and translocation. Thus, variants of the coding sequences described herein include nucleic acids that encode polypeptides that can be, for example, full length, mutated, truncated, inactivated, peptide/epitopes or combinations thereof of EGFR.

[0085] The full-length amino acid sequence of the human EGFR is set forth in SEQ ID NO: 1 (1210 aa) as follows:

```
MRPSGTAGAA LLALLAALCP ASRALEEKV CQGTSNKLTQ LGTFEDHFLS LQRMFNCE
VVLGNLEITYV QRNYDLSFLK TIQEVAGYVL IALNTVERIP LENLQIIRGN MYYENSYALA
VLSNYDANKT GLKELPMRNL QEILHGAVRF SNNPALCNVE SIQWRDIVSS DFSLNMSMDF
QNHLSGSCQKC DPSCPNGSCW GAGEENCQKL TKIICAQQCS GRCRGKSPSD CCHNQCAAGC
TGPRESDECLV CRKFRDEATC KDTCPPLMLY NPTTYQMDVN PEGKYSFGAT CVKKCPRNYV
VTDHGSCVRA CGADSYEMEE DGVKCKKCE GPCRKVCNGI GIGEFKDSLS INATNIKHFK
NCTSIGDLH ILPVAFRGDS FTHTPPLDPQ ELDILKTVE ITGFLLIQAW PENRTDLHAF
ENLEIIRGRT KQHGQFSLAV VSLNITSLGL RSLKEISDGD VIISGNKNLC YANTINWKKL
FGTSGQKTKI ISNRGENSCK ATGQVCHALC SPEGCWGPEP RDCVSCRNVS RGRECVDK
CNLLEGEPREFV ENSECIQCHP ECLPQAMNIT CTGRGPDNCI QCAHYIDGPH CVKTCFAGVM
GENNTLVWKY ADAGHVCHLC HPNCTYGCTG PGLEGCPNG PKIPSIATGM VGALLLLLV
VALGIGLFMR RRHIVRKRTL RRLQERELV EPLTPSGEAP NQALLRILKE TEFKKIKVLGSG
AFGTVYKGL WIPEGEKVKI PVAIKELREA TSPKANKEIL DEAYVMASVD NPHVCRLGIC
```

LTSTVQLIT QLMFPGCLLD YVREHKDNIG SQYLLNWCVQ IAKGMNYLED RRLVHRDLAAR
 NVLVKTPQH VKITDFGLAK LLGAEEKEYH AEGGKVPIKW MALESILHRI YTHQSDVWSY
 GVTWELMTF GSKPYDGIPA SEISSILEKG ERLPQPPICT IDVYMIMVKC WMIDADSRPK
 FRELIIEFSK MARDPQRYLV IQGDERMHLP SPTDSNFYRA LMDEEDMDDV VDADEYLIPQ
 QGFFSSPSTS RTPLLSLSA TSNNSTVACI DRNGLQSCPI KEDSFLQRY S DPTGALTED
 SDDTFLPVP EYINQSVPKR PAGSVQNPVY HNQLNPAPS RDPHYQDPHS TAVGNPEYLN
 TVQPTCVNST FDSPAHWQK GSHQISLDNP DYQQDFFPKE AKPNGIFKGS TAENAEYLRV
 APQSSEFIGA

[0086] As described above, the coding sequence for EGFR in the nucleic acid constructs of the disclosure may include one or more molecular alterations, e.g., mutations. In some embodiments, the coding sequence for EGFR in the nucleic acid constructs includes one or more acquired resistance mutations. As used herein, the term “acquired resistance mutation” generally refers to any mutation that results in an EGFR becoming resistant to the effects of a therapeutic drug and/or becoming substantially unresponsive after the EGFR has been exposed to the drug for a period of time. Exemplary types of molecular alterations, e.g., mutations in the coding sequences described herein can be one or more of deletions, substitutions, insertion, duplications, mutations, frameshift variants, splice variants, and combinations of any thereof.

[0087] In some embodiments, the therapeutic drug is a tyrosine kinase inhibitor (TKI). This is in contrast to neoantigens as well as intrinsic or primary resistance which occurs when an EGFR mutant has little or no objective response to a first line of treatment (e.g., a TKI).

[0088] Amino acid sequence numbering for EGFR and EGFR acquired resistance mutations as used herein is generally with reference to the amino acid sequence and numbering of wild type human EGFR as presented SEQ ID NO: 1. In some embodiments, the coding sequence for the EGFR protein in the nucleic acid constructs described herein encodes the amino acid sequence of SEQ ID NO: 1 comprising one or more acquired resistance mutations. In some embodiments, the one or more acquired resistance mutations is selected from A763_Y764insFQEA, S768_D770dup, S768_V769ins, A767_V769dup, D770_N771insX, V769_D770insX, H773_V774insX, H773dup, N771_H773dup, P772_H773insX, and N771_P772insX, T790M, G719X, L858R, L718Q, G724S, S768I, G719X, L861X, S768I/V, E709X, L747S, D761Y, T854A, C797S, C797G, T790M, L858R, L858M, L718V, L718Q, C796S, L798I, L792X, SV768IL, L692V, G719A, G719S, G719C, G719D, S786I, L861Q, L861R, V834X, V843X, G724S, E709K, E709H, E709A, E709G, E709V, D761Y, D761N,

R776C, and R776H. In some embodiments, the one or more acquired resistance mutations is selected from S768_D770dup, A767_V769dup, H773dup, T790M, L858R, and C797S.

[0089] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR protein having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 1. In some embodiments, the coding sequence for the EGFR protein encodes smaller portions of the amino acid sequence of SEQ ID NO: 1. These smaller portions can include at least 8, 10, 12, 14, 16, 18, 20, 30 or more amino acids of SEQ ID NO: 1. Exemplary portions of EGFR that are useful in the constructs disclosed herein include those in Table 1 below:

Table 1.

Amino acid sequence	SEQ ID NO
LITQLMPFGSLLDYVREHKDNIGS	2
DEAYVMASVDSVDNPHVCRLLGIC	3
DEAYVMASVASVDNPHVCRLLGIC	4
DEAYVMASVDNPHHVCRLLGICLTS	5
VKTPQHVKITDFGRAKLLGAEEK	6
GICLTSTVQLIMQLMPFGCL	7

[0090] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding a portion of EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NOs: 2-7.

[0091] In some embodiments, the one or more mutations are configured into a plurality of alteration cassettes. In some embodiments, the plurality of alteration cassettes are arranged in tandem along the length of the coding sequence. In some embodiments, the length and amino acid composition of the alteration cassettes can be optimized to achieve a desired activity or property of the coding sequence or variant thereof. In some embodiments, an alteration cassette of the plurality of alteration cassettes includes about 2 to about 50 amino acid residues, such as about 5 to about 45, about 10 to about 40, about 15 to about 30, about 20 to about 50, about 2 to

about 30, about 3 to about 25, about 4 to about 20, about 5 to about 15, about 6 to about 10, about 3 to about 15, about 4 to about 10, about 5 to about 30, about 2 to about 5, about 3 to about 5, about 4 to about 8 amino acid residues. In some embodiments, an alteration cassette of the plurality of alteration cassettes includes one, two, three, four, five, or more mutations.

[0092] In some embodiments, the EGFR described herein comprises one or more acquired resistance mutations that are activating mutations. EGFR activating mutations are generally found in exons 18 to 21 of the EGFR gene, which is part of the gene coding for the tyrosine kinase domain of the EGFR protein. These activating mutations lead to a ligand-independent activation of tyrosine kinase activity in EGFR and can render the EGFR insensitive to one or more tyrosine kinase inhibitors.

[0093] In some embodiments, the one or more activating mutations comprises an in-frame insertion in exon 20. The one or more EGFR exon 20 mutations may be located at one or more residues selected from the group consisting of A763, A767, S768, V769, D770, N771, P772, and H773. Exemplary EGFR exon 20 insertions may include H773_V774insH, A767_v769ASV, N771_P772insH, D770_N771insG, H779_V774insH, N771delinsHH, S768_D770dupDVD, A767_V769dupASV, A767_V769dupASV, P772_H773dup, N771_H773dupNPH, S768_D770dupSVD, N771delinsGY, S768_D770delinsSVD, D770_D770delinsGY, A767_V769dupASV, and/or H773dup. In particular aspects, the exon 20 mutations are A763insFQEA, A767insASV, S768dupSVD, V769insASV, D770insSVD, D770insNPG, H773insNPH, N771del insGY, N771del insFH, and/or N771dupNPH.

[0094] In some embodiments, the exon 20 in-frame insertion comprises a mutation selected from the group consisting of A763_Y764insFQEA, S768_D770dup, S768_V769ins, A767_V769dup, D770_N771insX, V769_D770insX, H773_V774insX, H773dup, N771_H773dup, P772_H773insX, and N771_P772insX. In some embodiments, the exon 20 mutation is an in-frame insertion is selected from S768_D770dup, A767_V769dup, and/or H773dup.

[0095] In some embodiments, the exon 20 in-frame insertion comprises S768_D770dup, A767_V769dup, and/or H773dup.

[0096] In some embodiments, the activating mutation is located in exon 21 of EGFR. An exemplary acquired resistance mutation in exon 21 of EGFR includes, without limitation,

L858R.

[0097] The one or more EGFR acquired resistance mutations can also include mutations that enhance the binding affinity of EGFR to ATP. Exemplary acquired resistance mutations of this type include those located at the entrance to a hydrophobic pocket in the back of the ATP binding cleft of EGFR. Substitution of one or more of these residues in EGFR can cause resistance by interference with binding of TKIs. The one or more EGFR acquired resistance mutations may be located at one or more residues selected from the group consisting of T790, G719, L858, C797, L718, G724, S768, L861, E709, L747, D761, and T854. Exemplary acquired resistance mutations that enhance the binding affinity of EGFR to ATP include, without limitation, T790M, G719X, L858R, C797S, L718Q, G724S, L861X, S768I/V, E709X, L747S, D761Y, and T854A.

[0098] The one or more EGFR acquired resistance mutations can also include one or more acquired resistance mutations blocks binding of EGFR to an inhibitor. An EGFR inhibitor can refer to a molecule having the ability to inhibit a biological function of a native EGFR. While preferred inhibitors herein specifically interact with (e.g. bind to) an EGFR, molecules that inhibit an EGFR biological activity by interacting with other members of the EGFR signal transduction pathway are also specifically included. A preferred EGFR biological activity inhibited by an EGFR inhibitor is associated with the development, growth, or spread of a tumor. EGFR inhibitors, without limitation, include peptides, non-peptide small molecules, antibodies, antibody fragments, antisense molecules, and oligonucleotide decoys.

[0099] In some embodiments, the EGFR inhibitor is a tyrosine kinase inhibitor. As used herein the term "tyrosine kinase inhibitors (TKIs)" can refer to a small molecule capable of inhibiting an ErbB signaling pathway. Typically, TKIs as contemplated herein may be categorized to four groups: (1) ATP-competitive inhibitors, which bind predominantly to the ATP-binding site of the kinase when this site is in the active conformation; (2) inhibitors that recognize and bind to the non-active conformation of the ATP-binding site of the kinase, thus making activation energetically unfavorable; (3) allosteric inhibitors, that bind outside of the ATP-binding site, modifying the tridimensional structure of the receptor and disrupting the interaction between the ATP and the kinase pocket; and (4) covalent inhibitors, that bind irreversibly by covalently bonding to the ATP-binding site of the target kinase. Exemplary

EGFR tyrosine kinase inhibitors include, without limitation, Osimertinib, Lazertinib, Erlotinib and Gefitinib, CO-1686, HM61713, EGF816, ASP8273 and Avitinib, Afatinib, Mobocertinib, Icotinib, Dacomitinib, Poziotinib, Cetuximab, Amivantamab, Mobocertinib, Furmonertinib, DZD9008, CLN-081, STX-721, YK-029A, HS-10376, Zipalertinib, TAK-788, JMT101, ABT-101, BEBT-109, DZG9008, PLB1004, EMB-01, HS-20117, MCLA-129, BLU-945; BDTX-1535; NX-019; JIN-A02; BBT-207; BLU-525; THE-349; STX-241; ABK3376; BI-732; BLU-701; BBT-176, BLU-451, Almonertinib, and Tarloxotinib.

[0100] In some embodiments, the one or more acquired resistance mutations that block binding of EGFR to an inhibitor include, without limitation, C797S, C797G, T790M, L858R, L858M, L718V, L718Q, C796S, L798I, L792X, SV768IL, L692V, G719A, G719S, G719C, G719D, S786I, L861Q, L861R, V834X, V843X, G724S, E709K, E709H, E709A, E709G, E709V, D761Y, D761N, R776C, R776H, and T854A.

[0101] In some embodiments, the one or more acquired resistance mutations are operably linked to one another by a linker. The linker can be a peptide linker, which joins together two adjacent alteration acquired resistance mutation cassettes, as described herein. In some embodiments, the length and amino acid composition of the peptide linker sequence can be optimized to vary the orientation, flexibility, and/or proximity of the alteration cassettes relative to one another to achieve a desired activity or property of the EGFR.

[0102] In some embodiments, a polypeptide linker includes a single-chain polypeptide sequence comprising about 1 to about 30 amino acid residues (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, etc. amino acid residues). In some embodiments, a linker sequence includes about 2 to 30, about 3 to 25, about 4 to 20, about 5 to 15, about 6 to 10, about 3 to 15, about 4 to 10, about 5 to 30, about 2 to 5, about 3 to 5, about 4 to 8 amino acid residues.

[0103] In some embodiments, the length and amino acid composition of the linker polypeptide sequence can be optimized to vary the orientation, flexibility, and/or proximity of the acquired resistance mutation cassettes relative to one another to achieve a desired activity or property of the encoded polypeptide. In some embodiments, the orientation, flexibility, and/or proximity of the acquired resistance mutation cassettes relative to one another can be varied as a “tuning” tool to achieve a tuning effect that would enhance or reduce the activity of the encoded polypeptide or encoded polypeptide variant. In certain embodiments, the linker contains only

glycine and/or serine residues (e.g., glycine-serine linker). Examples of such polypeptide linkers include: Gly, Ser; Gly Ser; Gly Gly Ser; Ser Gly Gly; Gly Gly Gly Ser; Ser Gly Gly Gly; Gly Gly Gly Gly Ser; Ser Gly Gly Gly Gly; Gly Gly Gly Gly Gly Ser; Ser Gly Gly Gly Gly Gly; Gly Gly Gly Gly Gly Gly Ser; Ser Gly Gly Gly Gly Gly Gly; (Gly Gly Gly Gly Ser)_n, wherein n is an integer of one or more; and (Ser Gly Gly Gly Gly)_n, wherein n is an integer of one or more. In some embodiments, the polypeptide linkers are modified such that the amino acid sequence Gly Ser Gly (GSG) (that occurs at the junction of traditional Gly/Ser linker polypeptide repeats) is not present. In some embodiments, the peptide linker includes an amino acid sequence selected from the group consisting of SEQ ID NOs: 8-12.

[0104] In some embodiments, the coding sequence for a polypeptide construct of the nucleic acid construct described herein encodes an EGFR comprising portions of the EGFR amino acid sequence operably linked with GGGGS linkers (underlined). An exemplary amino acid sequence comprises that of SEQ ID NO: 13 as follows:

GICLTSTVQLIMQLMPFGCLGGGGSVKTPQHVKITDFGRAKLLGAEKGGGGSLITQ
LMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVDSVDNPHVCRLLGICGGGGSDE
AYVMASVASVDNPHVCRLLGICGGGGSDEAYVMASVDNPHHVCRLLGICLTSGGG
GS

[0105] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 13.

[0106] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 14 as follows:

DEAYVMASVASVDNPHVCRLLGICGGGGSDEAYVMASVDSVDNPHVCRLLGICGG
GSDEAYVMASVDNPHHVCRLLGICLTSGGGGSVKTPQHVKITDFGRAKLLGAEK
GGGGSLITQLMPFGSLLDYVREHKDNIGSGGGGSGICLTSTVQLIMQLMPFGCLGGG
GS

[0107] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an

amino acid sequence of SEQ ID NO: 14.

[0108] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 15 as follows:

GICLTSTVQLIMQLMPFGCLGGGGSLITQLMPFGSLLDYVREHKDNIGSGGGGSVKT
PQHVKITDFGRAKLLGAEEKGGGGSEAYVMASVDNPHHVCRLLGICLTSGGGGSE
AYVMASVDSVDNPHVCRLLGICGGGGSEAYVMASVASVDNPHVCRLLGICGGGG
S

[0109] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 15.

[0110] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 16 as follows:

GICLTSTVQLIMQLMPFGCLGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGSLITQ
LMPFGSLLDYVREHKDNIGSGGGGSEAYVMASVASVDNPHVCRLLGICGGGGSE
AYVMASVDSVDNPHVCRLLGICGGGGSEAYVMASVDNPHHVCRLLGICLTSGGG
GS

[0111] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 16.

[0112] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 17 as follows:

DEAYVMASVDSVDNPHVCRLLGICGGGGSEAYVMASVDNPHHVCRLLGICLTSG
GGGSVKTPQHVKITDFGRAKLLGAEEKGGGGSLITQLMPFGSLLDYVREHKDNIGSG
GGSGICLTSTVQLIMQLMPFGCLGGGGSEAYVMASVASVDNPHVCRLLGICGGG
GS

[0113] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an

amino acid sequence of SEQ ID NO: 17.

[0114] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 18 as follows:

LITQLMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVASVDNPHVCRLLGICGGG
GSDEAYVMASVDSVDNPHVCRLLGICGGGGSGICLTSTVQLIMQLMPFGCLGGGGS
DEAYVMASVDNPHHVCRLLGICLTSGGGGSVKTPQHVKITDFGRAKLLGAEEKGGG
GS

[0115] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 18.

[0116] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 19 as follows:

DEAYVMASVDNPHHVCRLLGICLTSGGGGSDEAYVMASVASVDNPHVCRLLGICG
GGGSDEAYVMASVDSVDNPHVCRLLGICGGGGLITQLMPFGSLLDYVREHKDNIG
SGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGSGICLTSTVQLIMQLMPFGCLGG
GGG

[0117] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 19.

[0118] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 20 as follows:

GICLTSTVQLIMQLMPFGCLGGGGSDEAYVMASVDSVDNPHVCRLLGICGGGGGSDE
AYVMASVASVDNPHVCRLLGICGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGG
DEAYVMASVDNPHHVCRLLGICLTSGGGGSLITQLMPFGSLLDYVREHKDNIGSGG
GGG

[0119] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an

amino acid sequence of SEQ ID NO: 20.

[0120] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 21 as follows:

DEAYVMASVDSVDNPHVCRLLGICGGGSLITQLMPFGSLLDYVREHKDNIGSGGG
GSDEAYVMASVASVDNPHVCRLLGICGGGGSDEAYVMASVDNPHHVCRLLGICLTS
GGGSGICLTSTVQLIMQLMPFGCLGGGGSVKTPQHVKITDFGRAKLLGAEEKGGG
GS

[0121] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 21.

[0122] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 22 as follows:

LITQLMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVDSVDNPHVCRLLGICGGG
GSGICLTSTVQLIMQLMPFGCLGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGSDE
AYVMASVASVDNPHVCRLLGICGGGGSDEAYVMASVDNPHHVCRLLGICLTS
GGGGS

[0123] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 22.

[0124] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 23 as follows:

DEAYVMASVDNPHHVCRLLGICLTSGGGSGICLTSTVQLIMQLMPFGCLGGGGS
EAYVMASVASVDNPHVCRLLGICGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGS
DEAYVMASVDSVDNPHVCRLLGICGGGSLITQLMPFGSLLDYVREHKDNIGSGGG
GS

[0125] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an

amino acid sequence of SEQ ID NO: 23.

[0126] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 24 as follows:

VKTPQHVKITDFGRAKLLGAEEKGGGGSDEAYVMASVASVDNPHVCRLLGICGGG
GSGICLTSTVQLIMQLMPFGCLGGGSLITQLMPFGSLLDYVREHKDNIGSGGGGSD
 EAYVMASVDNPHHVCRLLGICLTSGGGGSDEAYVMASVDSVDNPHVCRLLGICGG
GS

[0127] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 24.

[0128] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 25 as follows:

DEAYVMASVDNPHHVCRLLGICLTSGGGGSGICLTSTVQLIMQLMPFGCLGGGGSV
 KTPQHVKITDFGRAKLLGAEEKGGGSLITQLMPFGSLLDYVREHKDNIGSGGGGS
 D EAYVMASVASVDNPHVCRLLGICGGGGSDEAYVMASVDSVDNPHVCRLLGICGGG
GS

[0129] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 25.

[0130] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 26 as follows:

LITQLMPFGSLLDYVREHKDNIGSGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGS
 DEAYVMASVASVDNPHVCRLLGICGGGSGICLTSTVQLIMQLMPFGCLGGGGSDE
 AYVMASVDSVDNPHVCRLLGICGGGGSDEAYVMASVDNPHHVCRLLGICLTSGGG
GS

[0131] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an

amino acid sequence of SEQ ID NO: 26.

[0132] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 27 as follows:

GICLTSTVQLIMQLMPFGCLGGGGSDEAYVMASVASVDNPHVCRLLGICGGGGSVK
TPQHVKITDFGRAKLLGAEEKGGGGSDEAYVMASVDSVDNPHVCRLLGICGGGSLI
TQLMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVDNPHHVCRLLGICLTSGGGGS

[0133] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 27.

[0134] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 28 as follows:

DEAYVMASVDNPHHVCRLLGICLTSGGGGSICLTSTVQLIMQLMPFGCLGGGGS
TQLMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVDSVDNPHVCRLLGICGGGGS
VKTPQHVKITDFGRAKLLGAEEKGGGGSDEAYVMASVASVDNPHVCRLLGICGGG
GS

[0135] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 28.

[0136] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 29 as follows:

VKTPQHVKITDFGRAKLLGAEEKGGGGSDEAYVMASVDSVDNPHVCRLLGICGGG
GSLITQLMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVASVDNPHVCRLLGICG
GGSGICLTSTVQLIMQLMPFGCLGGGGSDEAYVMASVDNPHHVCRLLGICLTSGG
GS

[0137] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 29.

[0138] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 30 as follows:

DEAYVMASVASVDNPHVCRLLGICGGGGSGICLTSTVQLIMQLMPFGCLGGGGSLIT
QLMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVDNPHHVCRLLGICLTSGGGGS
DEAYVMASVDSVDNPHVCRLLGICGGGGSVKTPQHVKITDFGRAKLLGAEEKGGG
GS

[0139] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 30.

[0140] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 31 as follows:

GICLTSTVQLIMQLMPFGCLGGGGSGDEAYVMASVASVDNPHVCRLLGICGGGGSGDE
AYVMASVDNPHHVCRLLGICLTSGGGGSLITQLMPFGSLLDYVREHKDNIGSGGGG
SDEAYVMASVDSVDNPHVCRLLGICGGGGSVKTPQHVKITDFGRAKLLGAEEKGGG
GS

[0141] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 31.

[0142] In some embodiments, the amino acid sequence comprises that of SEQ ID NO: 32 as follows:

DEAYVMASVDSVDNPHVCRLLGICGGGGSGDEAYVMASVASVDNPHVCRLLGICGG
GGSVKTPQHVKITDFGRAKLLGAEEKGGGGSLITQLMPFGSLLDYVREHKDNIGSGG
GGSGDEAYVMASVDNPHHVCRLLGICLTSGGGGSGICLTSTVQLIMQLMPFGCLGGG
GS

[0143] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding an EGFR having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence of SEQ ID NO: 32.

[0144] In some embodiments, the nucleic acid sequence encoding the polypeptide construct is operably linked to a promoter sequence. In some embodiments, the promoter sequence is a subgenomic (*sg*) promoter. In some embodiments, the *sg* promoter sequence is a 26S subgenomic promoter. In some embodiments, the subgenomic promoter is heterologous relative to the remainder of the modified viral genome or srRNA. In some embodiments, the subgenomic promoter is an alphavirus subgenomic promoter.

[0145] In some embodiments of the disclosure, at least one nonstructural protein (nsP), or a portion thereof, of the modified viral genome or srRNA is heterologous relative to the remainder of the modified viral genome or srRNA. In some embodiments, the nucleic acid constructs disclosed herein further include a nucleic acid sequence encoding a heterologous nsP or a portion thereof. In some embodiments, the nucleic acid constructs disclosed herein further include one or more untranslated regions (UTRs). In some embodiments, at least one of the UTRs is a heterologous UTR.

[0146] In some embodiments of the methods described herein, the recombinant alphavirus srRNA is of a virus belonging to the *Alphavirus* genus of the *Togaviridae* family. In some embodiments of the disclosure, the modified alphavirus genome or srRNA is of an alphavirus belonging to the Venezuelan equine encephalitis virus/Eastern Equine Encephalitis virus (VEEV/EEEV) group, or the Semliki Forest virus (SFV) group, or the Sindbis virus (SINV) group. In some embodiments, the modified alphavirus genome or srRNA is of an alphavirus belonging to the BFV complex, EEEV complex, MIDV complex, NDUV, complex, SFV complex, VEEV complex, WEEV complex. In some embodiments, the alphavirus is Eastern equine encephalitis virus (EEEV), Venezuelan equine encephalitis virus (VEEV), Everglades virus (EVEV), Mucambo virus (MUCV), Pixuna virus (PIXV), Middleburg virus (MIDV), Chikungunya virus (CHIKV), O'Nyong-Nyong virus (ONNV), Ross River virus (RRV), Barmah Forest virus (BF), Getah virus (GET), Sagiyama virus (SAGV), Bebaru virus (BEBV), Mayaro virus (MAYV), Una virus (UNAV), Sindbis virus (SINV), Aura virus (AURAV), Whataroa virus (WHAV), Babanki virus (BABV), Kyzylagach virus (KYZV), Western equine encephalitis virus (WEEV), Highland J virus (HJV), Fort Morgan virus (FMV), Ndumu virus (NDUV), Madariaga virus (MADV), or Buggy Creek virus. In some embodiments, the alphavirus is VEEV, EEEV, CHIKV, or SINV. In some embodiments, the alphavirus is VEEV. In some

embodiments, the alphavirus is EEEV. In some embodiments, the alphavirus is Western Equine Encephalitis virus (WEEV). In some embodiments, the alphavirus is CHIKV. In some embodiments, the alphavirus is SINV.

[0147] In some embodiments, the alphavirus is Chikungunya virus (CHIKV). Non-limiting examples of CHIKV strains suitable for the compositions and methods of the disclosure include CHIKV S27, CHIKV LR2006-OPY-1, CHIKV YO123223, CHIKV DRDE, CHIKV 37997, CHIKV 99653, CHIKV Ag41855, and Nagpur (India) 653496 strain. Virulent and avirulent CHIKV strains are both suitable. Additional examples of CHIKV strains suitable for the compositions and methods of the disclosure include but are not limited to those described in Afreen *et al. Microbiol. Immunol.* 2014, 58:688-696, Lanciotti and Lambert *ASTMH* 2016, 94(4):800-803 and Langsjoen *et al. mBio.* 2018, 9(2):e02449-17. In some embodiments, the modified CHIKV genome or replicon RNA (*e.g.*, self-replicating RNA) is derived from CHIKV strain S27. In some embodiments, the modified CHIKV genome or replicon RNA is derived from CHIKV strain DRDE. In some embodiments, the modified CHIKV genome or replicon RNA is derived from CHIKV strain DRDE-06. In some embodiments, the modified CHIKV genome or replicon RNA is derived from CHIKV strain DRDE-07. In some embodiments, the modified CHIKV genome or replicon RNA is derived from CHIKV strain S27.

[0148] In some embodiments, the alphavirus is Eastern Equine Encephalitis virus (EEEV). Non-limiting examples of EEEV strains suitable for the compositions and methods of the disclosure include EEEV 792138, 783372, BeAn5122, BeAr300851, BeAr436087, C-49, FL91-4679, FL93-939, GML903836, MP-9, PE6, and V105-00210. Virulent and avirulent EEEV strains are both suitable. Additional suitable EEEV strains include, but are not limited to those described in the Virus Pathogen Resource website (ViPR; which is publicly available at www.viprbrc.org/brc/vipr_genome_search.spg?method=SubmitForm&blockId=868&decorator=toga). In some embodiments, the modified EEEV genome or replicon RNA (*e.g.*, self-replicating RNA) is derived from EEEV strain FL93-939.

[0149] In some embodiments, the alphavirus is Sindbis virus (SINV). In some embodiments, the modified genome or RNA replicon (*e.g.*, self-replicating RNA) is of a SINV strain. Non-limiting examples of SINV strains suitable for the compositions and methods of the disclosure include SINV strain AR339, AR86, and Girdwood. Examples of SINV strains suitable

for the compositions and methods of the disclosure include, but are not limited to those described in Sammels *et al. J. Gen. Virol.* 1999, 80(3):739-748, Lundström and Pfeffer *Vector Borne Zoonotic Dis.* 2010, 10(9):889-907, Sigei *et al. Arch. of Virol.* 2018, 163:2465-2469 and Ling *et al. J. Virol.* 2019, 93:e00620-19. Additional suitable SINV strains include, but are not limited to those described in the Virus Pathogen Resource website (ViPR; which is publicly available at www.viprbrc.org/brc/vipr_genome_search.spg?method=SubmitForm&blockId=868&decorator=toga). Virulent and avirulent SINV strains are both suitable. In some embodiments, the modified genome or RNA replicon is of a SINV strain Girdwood. In some embodiments, the modified genome or RNA replicon is of a SINV strain AR86. In some embodiments, the modified SINV genome or replicon RNA is derived from SINV strain Girdwood. In some embodiments, the modified SINV genome or replicon RNA is derived from SINV strain AR86. In some embodiments, the at least one heterologous nsP or portion thereof of the modified genome or RNA replicon is derived from a SINV strain AR86. In some embodiments, the at least one heterologous nsP or portion thereof is nsP1, nsP3, nsP4, or a portion of any thereof, or a combination of any of the foregoing. In some embodiments, the modified genome or RNA replicon is of a SINV strain AR86.

[0150] In some embodiments, the alphavirus is Western Equine Encephalitis virus (WEEV). Non-limiting examples of WEEV strains suitable for the compositions and methods of the disclosure include WEEV California, McMillan, IMP181, Imperial, Imperial181, IMPR441, 71V-1658, AG80-646, BFS932, COA592, EP-6, E1416, BFS1703, BFS2005, BSF3060, BSF09997, CHLV53, KERN5547, 85452NM, Montana-64, S8-122, and TBT-235. Additional examples of WEEV strains suitable for the compositions and methods of the disclosure include 5614, 93A27, 93A30, 93A38, 93A79, B628(CI 15), CBA87, CNTR34, CO921356, Fleming, Lake43, PV012357A, PV02808A, PV72102, R02PV001807A, R02PV002957B, R02PV003422B, R05PV003422B, R0PV003814A and R0PV00384A. Virulent and avirulent WEEV strains are both suitable. Additional suitable WEEV strains include, but are not limited to those described in Bergren NA *et al., J. Virol.* 88(16): 9260–9267, Aug 2014, and in the Virus Pathogen Resource website (ViPR; which is publicly available at https://www.viprbrc.org/brc/vipr_genome_search.spg?method=SubmitForm&blockId=57240&decorator=toga). In some embodiments, the modified WEEV genome or srRNA is derived from

WEEV strain Imperial. In some embodiments, the modified WEEV genome or srRNA is derived from WEEV strain McMillan.

[0151] In some embodiments, the alphavirus is Madariaga virus (MADV), formerly referred to as South American Eastern Equine Encephalitis virus (SA EEEV). Non-limiting examples of MADV strains suitable for the compositions and methods of the disclosure include ArgLL, ArgB, BeAn-5122, ArgM, 24443 (TR59), 25714 (BG60), BeAr 18205, 900188 (PA62), BeAr 81828, BeAr 126650, 68U231, 77U1104 (PE70), 75V1496, BeAr 300851, 75U40, and El Delirio (Arrigo NC *et al.*, supra 2010). Additional examples of MADV strains suitable for the compositions and methods of the disclosure include 76V25343, 77U1 (BR77), BeAr348998, IVICPan57151, BeAn416361, 903836 (PA84), BeAr436087, 435731 (PA86), C49 (CO92), PE-0.0155-96 (0.0155), PE-3.0815-96 (3.0815), PE-16.0050-98 (16.0050), PE-18.0140-99 (18.0140), and PE-18.0172-99 (18.0172) (Arrigo NC *et al.*, supra 2010). Additional suitable MADV strains include, but are not limited to those described in Arrigo NC *et al.*, supra 2010, and in the Virus Pathogen Resource website (ViPR; which is publicly available at www.viprbrc.org/brc/vipr_genome_search.spg?method=SubmitForm&blockId=868&decorator=toga). In some embodiments, the modified MADV genome or srRNA is derived from MADV strain BeAr300851.

[0152] In some embodiments of the disclosure, the coding sequence for the polypeptide construct includes, in 5' to 3' direction (*i.e.*, in N-terminus to C-terminus direction of the polypeptide sequence): a coding sequence for epidermal growth factor receptor (EGFR) comprising one or more acquired resistance mutations selected from S768_D770dup, A767_V769dup, H773dup, T790M, L858R, and C797S.

[0153] In some embodiments, the coding sequence of the EGFR protein is redesigned and/or optimized for a desired property, such as increased stability, potency, and expression (*e.g.*, translation efficiency), which in turns can maximize the impact of producing, delivering, and administering the biotherapeutic EGFR. For example, in some embodiments, the coding sequence of the EGFR is optimized for expression at a level higher than the expression level of a reference coding sequence. In some embodiments, the coding sequence of the EGFR is optimized for one or more of the following: (a) enhancing RNA stability, (b) enhancing expression level, (c) minimizing rare codon usage, (c) minimizing secondary structures, (d)

facilitating better srRNA replication, and (e) facilitating better RNA manufacturing process.

[0154] With respect to sequence-optimization of nucleotide sequences, degeneracy of the genetic code provides the possibility to substitute at least one base of the protein encoding sequence of a gene with a different base without causing the amino acid sequence of the polypeptide produced from the gene to be changed. Hence, the nucleic acid constructs of the present disclosure may also have any base sequence that has been changed from any polynucleotide sequence disclosed herein by substitution in accordance with degeneracy of the genetic code. References describing codon usage are readily publicly available. In some embodiments, polynucleotide sequence variants can be produced for a variety of reasons, *e.g.*, to optimize expression for a particular host (*e.g.*, changing codon usage in the alphavirus mRNA to those preferred by other organisms such as human, non-human primates, hamster, mice, or monkey). Accordingly, in some embodiments, the coding sequence is optimized for expression in a target host cell through the use of codons optimized for expression. The techniques for the construction of synthetic nucleic acid sequences encoding genes using preferred codons optimal for host cell expression may be determined by computational methods analyzing the commonality of codon usage for encoding native proteins of the host cell genome and their relative abundance by techniques well known in the art. The codon usage database (<http://www.kazusa.or.jp/codon>) may be used for generation of codon optimized sequences in mammalian cell environments. Furthermore, a variety of software tools are available to convert sequences from one organism to the optimal codon usage for a different host organism such as the JCat Codon Optimization Tool (www.jcat.de), Integrated DNA Technologies (IDT) Codon Optimization Tool (<https://www.idtdna.com/CodonOpt>) or the Optimizer online codon optimization tool (<http://genomes.urv.es/OPTIMIZER>). Such synthetic sequences may be constructed by techniques known in the art for the construction of synthetic nucleic acid molecules and may be obtained from a variety of commercial vendors.

[0155] Accordingly, in some embodiments, the coding sequence of the EGFR is optimized for expression at a level higher than the expression level of a reference coding sequence, such as, for example, a coding sequence that has not been codon-optimized. In some embodiments, the codon-optimized sequence of the EGFR results in an increased expression level by at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or

at least 100% compared to a reference coding sequence that has not been codon-optimized. In some embodiments, the codon-optimized sequence of the EGFR results in an increased expression level by at least 2-fold, at least 3-fold, at least 4-fold, or at least 5-fold compared to a reference coding sequence that has not been codon-optimized.

[0156] In some embodiments, the coding sequence of the EGFR is optimized for enhanced RNA stability and/or expression. The stability of RNA generally relates to the “half-life” of RNA. “Half-life” relates to the period of time which is needed to eliminate half of the activity, amount, or number of molecules. In the context of the present disclosure, the half-life of an RNA is indicative for the stability of said RNA. The half-life of RNA may influence the “duration of expression” of the RNA. Additional information regarding principles, strategies, and methods for use in enhancing RNA stability can be found at, for example, Leppek K. *et al.*, Nature Communications, 22 Mar 2022, 13(1):1536 .

[0157] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding a polypeptide construct of the disclosure, wherein the nucleic acid sequence has at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence of SEQ ID NO: 13-32.

[0158] In some embodiments, the nucleic acid constructs of the disclosure include a nucleic acid sequence encoding a polypeptide construct having 100% sequence identity to the amino acid sequence of SEQ ID NO: 13-32.

[0159] Nucleic acid sequences having a high degree of sequence identity (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%) to a sequence of a modified alphavirus genome or srRNA of interest can be identified and/or isolated by using the sequences identified herein (*e.g.*, SEQ ID NO: 1) or any others as they are known in the art, by genome sequence analysis, hybridization, and/or PCR with degenerate primers or gene-specific primers from sequences identified in the respective alphavirus genome or srRNA.

[0160] The molecular techniques and methods by which these new nucleic acid constructs were assembled and characterized are described more fully in the Examples of the present application. In some embodiments, the nucleic acid molecules are recombinant nucleic acid

molecules. As used herein, the term recombinant means any molecule (*e.g.* DNA, RNA, polypeptide), that is, or results, however indirect, from human manipulation. As non-limiting examples, a cDNA is a recombinant DNA molecule, as is any nucleic acid molecule that has been generated by *in vitro* polymerase reaction(s), or to which linkers have been attached, or that has been integrated into a vector, such as a cloning vector or expression vector. As non-limiting examples, a recombinant nucleic acid molecule: 1) has been synthesized or modified *in vitro*, for example, using chemical or enzymatic techniques (for example, by use of chemical nucleic acid synthesis, or by use of enzymes for the replication, polymerization, exonucleolytic digestion, endonucleolytic digestion, ligation, reverse transcription, transcription, base modification (including, *e.g.*, methylation), or recombination (including homologous and site-specific recombination) of nucleic acid molecules; 2) includes conjoined nucleotide sequences that are not conjoined in nature; 3) has been engineered using molecular cloning techniques such that it lacks one or more nucleotides with respect to the naturally occurring nucleotide sequence; and/or 4) has been manipulated using molecular cloning techniques such that it has one or more sequence changes or rearrangements with respect to the naturally occurring nucleotide sequence.

[0161] In some embodiments, the nucleic acid molecules disclosed herein are produced using recombinant DNA technology (*e.g.*, polymerase chain reaction (PCR) amplification, cloning, *etc.*) or chemical synthesis. Nucleic acid molecules as disclosed herein include natural nucleic acid molecules and homologs thereof, including, but not limited to, natural allelic variants and modified nucleic acid molecules in which one or more nucleotide residues have been inserted, deleted, and/or substituted, in such a manner that such modifications provide the desired property in effecting a biological activity as described herein.

[0162] A nucleic acid molecule, including a variant of a naturally-occurring nucleic acid sequence, can be produced using a number of methods known to those skilled in the art (see, for example, Sambrook *et al.*, In: Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (1989)). The sequence of a nucleic acid molecule can be modified with respect to a naturally-occurring sequence from which it is derived using a variety of techniques including, but not limited to, classic mutagenesis techniques and recombinant DNA techniques, such as but not limited to site-directed mutagenesis, chemical treatment of a nucleic acid molecule to induce mutations, restriction enzyme cleavage of a

nucleic acid fragment, ligation of nucleic acid fragments, PCR amplification and/or mutagenesis of selected regions of a nucleic acid sequence, recombinational cloning, and chemical synthesis, including chemical synthesis of oligonucleotide mixtures and ligation of mixture groups to "build" a mixture of nucleic acid molecules, and combinations thereof. Nucleic acid molecule homologs can be selected from a mixture of modified nucleic acid molecules by screening for the function of the protein or the srRNA encoded by the nucleic acid molecule and/or by hybridization with a wild-type gene or fragment thereof, or by PCR using primers having homology to a target or wild-type nucleic acid molecule or sequence.

B. Recombinant cells

[0163] The nucleic acid constructs of the present disclosure can be introduced into a host cell to produce a recombinant cell containing the nucleic acid molecule. Accordingly, prokaryotic or eukaryotic cells that contain a nucleic acid construct encoding a modified alphavirus genome or srRNA as described herein are also features of the disclosure. In a related aspect, some embodiments disclosed herein relate to methods of transforming a cell which includes introducing into a host cell, such as an animal cell, a nucleic acid construct as provided herein, and then selecting or screening for a transformed cell. Introduction of the nucleic acid constructs of the disclosure into cells can be achieved by methods known to those skilled in the art such as, for example, viral infection, transfection, conjugation, protoplast fusion, lipofection, electroporation, nucleofection, calcium phosphate precipitation, polyethyleneimine (PEI)-mediated transfection, DEAE-dextran mediated transfection, liposome-mediated transfection, particle gun technology, direct micro-injection, nanoparticle-mediated nucleic acid delivery, and the like.

[0164] In one aspect, some embodiments of the disclosure relate to recombinant cells, for example, recombinant animal cells that include a nucleic acid construct described herein. The nucleic acid construct can be stably integrated in the host genome, or can be episomally replicating, or present in the recombinant host cell as a mini-circle expression vector for a stable or transient expression. Accordingly, in some embodiments of the disclosure, the nucleic acid construct is maintained and replicated in the recombinant host cell as an episomal unit. In some embodiments, the nucleic acid construct is stably integrated into the genome of the recombinant cell. Stable integration can be completed using classical random genomic recombination

techniques or with more precise genome editing techniques such as using guide RNA directed CRISPR/Cas9 or TALEN genome editing. In some embodiments, the nucleic acid construct present in the recombinant host cell as a mini-circle expression vector for a stable or transient expression.

[0165] In some embodiments, the recombinant cell is a prokaryotic cell, such as the bacterium *E. coli*, or a eukaryotic cell, such as an insect cell (*e.g.*, a mosquito cell or a Sf21 cell), or mammalian cells (*e.g.*, COS cells, NIH 3T3 cells, or HeLa cells). In some embodiments, the cell is *in vivo*. In some embodiments, the cell is *ex vivo*. In some embodiments, the cell is *in vitro*. In some embodiments, the recombinant cell is a eukaryotic cell. In some embodiments, the recombinant cell is an animal cell. In some embodiments, the animal cell is a vertebrate animal cell or an invertebrate animal cell. In some embodiments, the recombinant cell is a mammalian cell. Non-limiting examples of recombinant cells suitable for the methods and compositions of the disclosure include monkey kidney CV1 cells transformed by SV40 (*e.g.*, COS-7 cells), human embryonic kidney cells (*e.g.*, HEK 293 or HEK 293 cells) or derivative cells thereof (*e.g.*, BHK-21 or BHK-570 cells), baby hamster kidney cells (BHK), mouse sertoli cells (*e.g.*, TM4 cells), monkey kidney cells (*e.g.*, CV1 cells), human cervical carcinoma cells (*e.g.*, HeLa cells), canine kidney cells (MDCK cells), buffalo rat liver cells (*e.g.*, BRL 3A cells), human lung cell (*e.g.*, W138 cells), human liver cell (*e.g.*, Hep G2 cells), mouse mammary tumor (*e.g.*, MMT 060562 cells), TRI cells, FS4 cells, Chinese hamster ovary cells (CHO cells), African green monkey kidney cell (*e.g.*, Vero cells), human A549 cells, human cervix cells, human CHME5 cells, human PER.C6 cells, NS0 murine myeloma cells, human epidermoid larynx cells, human fibroblast cells, human HUH-7 cells, human MRC-5 cells, human muscle cells, human endothelial cells, human astrocyte cells, human macrophage cells, human RAW 264.7 cells, mouse 3T3 cells, mouse L929 cells, mouse connective tissue cells, mouse muscle cells, and rabbit kidney cells.

[0166] In some embodiments, the recombinant cell is an immune cell. In some embodiments, the immune cell is a B cell, a monocyte, a natural killer (NK) cell, a natural killer T (NKT) cell, a basophil, an eosinophil, a neutrophil, a dendritic cell (DC), a macrophage, a regulatory T cell, a helper T cell (T_H), a cytotoxic T cell (T_{CTL}), a memory T cell, a gamma delta ($\gamma\delta$) T cell, a hematopoietic stem cell, or a hematopoietic stem cell progenitor. In some

embodiments, the immune cell is a B cell, a T cell, a macrophage, or a dendritic cell (DC). In some embodiments, the immune cell is a B cell. In some embodiments, the immune cell is a T cell.

[0167] In some embodiments, the recombinant cell is a cell derived from a cell described above (*i.e.*, a derivative cell of an original cell described herein) such as, for example, a cell that is either expanded from a clone of the original cell, an engineered version of the original cell, or a reclassification of the original cell after it has undergone extensive passaging, or has been passaged through another host.

[0168] In some embodiments, the recombinant cell is an insect cell, *e.g.*, cell of an insect cell line. In some embodiments, the recombinant cell is a Sf21 cell. Additional suitable insect cell lines include, but are not limited to, cell lines established from insect orders Diptera, Lepidoptera and Hemiptera, and can be derived from different tissue sources. In some embodiments, the recombinant cell is a cell of a lepidopteran insect cell line. In the past few decades, the availability of lepidopteran insect cell lines has increased at about 50 lines per decade. More information regarding available lepidopteran insect cell lines can be found in, *e.g.*, Lynn D.E., *Available lepidopteran insect cell lines*. Methods Mol Biol. 2007;388:117-38, which is herein incorporated by reference. In some embodiments, the recombinant cell is a mosquito cell, *e.g.*, a cell of mosquito species within *Anopheles* (*An.*), *Culex* (*Cx.*) and *Aedes* (*Stegomyia*) (*Ae.*) genera. Exemplary mosquito cell lines suitable for the compositions and methods described herein include cell lines from the following mosquito species: *Aedes aegypti*, *Aedes albopictus*, *Aedes pseudoscutellaris*, *Aedes triseriatus*, *Aedes vexans*, *Anopheles gambiae*, *Anopheles stephensi*, *Anopheles albimanus*, *Culex quinquefasciatus*, *Culex theileri*, *Culex tritaeniorhynchus*, *Culex bitaeniorhynchus*, and *Toxorhynchites amboinensis*. Suitable mosquito cell lines include, but are not limited to, CCL-125, Aag-2, RML-12, C6/26, C6/36, C7-10, AP-61, A.t. GRIP-1, A.t. GRIP-2, UM-AVE1, Mos.55, Sua1B, 4a-3B, Mos.43, MSQ43, and LSB-AA695BB. In some embodiments, the mosquito cell is a cell of a C6/26 cell line.

[0169] In another aspect, provided herein are cell cultures including at least one recombinant cell as disclosed herein, and a culture medium. Generally, the culture medium can be any suitable culture medium for culturing the cells described herein. Techniques for transforming a wide variety of the above-mentioned host cells and species are known in the art

and described in the technical and scientific literature. Accordingly, cell cultures including at least one recombinant cell as disclosed herein are also within the scope of this application. Methods and systems suitable for generating and maintaining cell cultures are known in the art.

D. Pharmaceutical compositions

[0170] The nucleic acid constructs and recombinant cells of the disclosure can be incorporated into compositions, including pharmaceutical compositions. Such compositions generally include one or more of the nucleic acid constructs and recombinant cells described and provided herein, and a pharmaceutically acceptable excipient, *e.g.*, carrier. In some embodiments, the compositions of the disclosure are formulated for the prevention, treatment, or management of a health condition such as an autoimmune disease, an inflammatory disease, or a cardiovascular disease. For example, the compositions of the disclosure can be formulated as a prophylactic composition, a therapeutic composition, or a pharmaceutical composition comprising a pharmaceutically acceptable excipient, or a mixture thereof. In some embodiments, the compositions of the present disclosure are formulated for use as a vaccine or an immunotherapeutic. In some embodiments, the compositions of the present application are formulated for use as an adjuvant.

[0171] Accordingly, in one aspect, provided herein are pharmaceutical compositions including a pharmaceutically acceptable excipient and: (a) a nucleic acid construct of the disclosure; and/or (b) a recombinant cell of the disclosure.

[0172] Non-limiting exemplary embodiments of the pharmaceutical compositions of the disclosure can include one or more of the following features. In some embodiments, provided herein are compositions including a nucleic acid construct as disclosed herein and a pharmaceutically acceptable excipient. In some embodiments, provided herein are compositions including a recombinant cell as disclosed herein and a pharmaceutically acceptable excipient.

[0173] In some embodiments, the nucleic acid constructs of the disclosure (*e.g.*, a vectors or srRNA molecules) can be used in a naked form or formulated with a delivery vehicle. Exemplary delivery vehicles suitable for the compositions and methods of the disclosure include, but are not limited to liposomes (*e.g.*, neutral or anionic liposomes), microspheres, immune stimulating complexes (ISCOMS), lipid-based nanoparticles (LNP), solid lipid nanoparticles (SLN), polyplexes, polymer nanoparticles, viral replicon particles (VRPs), or conjugated with

bioactive ligands, which can facilitate delivery and/or enhance the immune response. These compounds are readily available to one skilled in the art; for example, see *Liposomes: A Practical Approach*, RCP New Ed, IRL press (1990). Adjuvants other than liposomes and the like are also used and are known in the art. Adjuvants may protect the antigen (*e.g.*, nucleic acid constructs, vectors, srRNA molecules) from rapid dispersal by sequestering it in a local deposit, or they may contain substances that stimulate the host to secrete factors that are chemotactic for macrophages and other components of the immune system. An appropriate selection can be made by those skilled in the art, for example, from those described below.

[0174] The composition of the disclosure can be formulated in a format to be compatible with its intended route of administration, such as liposome, lipid-based nanoparticle (LNP), a polymer nanoparticle, a polyplex, viral replicon particle (VRP), microsphere, immune stimulating complex (ISCOM), conjugate of bioactive ligand, or a combination of any thereof. Accordingly, in some embodiments, the compositions of the disclosure can be formulated in a liposome.

Polymer nanoparticle

[0175] In some embodiments, the compositions of the disclosure can be formulated in a polymer nanoparticle. In some embodiments, the polymeric nanoparticle includes a cationic polymer, a non-cationic polymer, or a combination thereof. In some embodiments, the cationic polymer includes a naturally-derived cationic polymer. In some embodiments, the naturally-derived cationic polymer includes chitosan, gelatin, dextran, cellulose, cyclodextrin, or a combination thereof. In some embodiments, the cationic polymer includes a synthetic cationic polymer. In some embodiments, the synthetic cationic polymer includes a polyethyleneimine (PEI), poly-L-lysine (PLL), a poly(amino acid) (PAA), a poly(amidoamine) (PAMAM), a poly(cystamine bisacrylamide-co-4-amino-1-butanol) (pABOL), a poly(amino-co-ester) (PAE), poly(2-N,N-dimethylaminoethylmethacrylate), a poly(beta-amino ester) (PBAE), an imidazole-containing polymer, a tertiary-amine containing polymer, poly(2-(dimethylamino)ethyl methacrylate), poly-N-(2-hydroxy-propyl) methacrylamide, a polyamidoamine dendrimer, a cationic glycopolymer, or derivatives thereof.

[0176] In some embodiments, the non-cationic polymer is negatively-charged (*i.e.*, anionic) or electronically neutral. In some embodiments, the non-cationic polymer includes a

polyethylene glycol (PEG), a polyester (*e.g.*, polylactic acid (PLA), poly (lactic-co-glycolic acid) (PLGA), poly glycolic acid (PGA), polycaprolactone (PCL)), and polysarcosine (pSar), or derivatives thereof. In some embodiments, the polymer is water-soluble and/or biodegradable.

[0177] In some embodiments of the disclosure, the polymeric nanoparticle includes one or more of the following: poly-(γ -L-glutamylglutamine) (PGGA), poly-(γ -L-aspartylglutamine) (PGAA), poly-L-lactic acid (PLLA), poly-(lactic acid-co-glycolic acid) (PLGA), polyalkylcyanoacrylate (PACA), polyanhydrides, polyhydroxyacids, polypropylfumerate, polyamide, polyacetal, polyether, polyester, poly(orthoester), polycyanoacrylate, [N-(2-hydroxypropyl)methacrylamide] (HPMA) copolymer, polyvinyl alcohol, polyurethane, polyphosphazene, polyacrylate, polyurea, polyamine polyepsilon-caprolactone (PCL), and copolymers thereof.

Lipid-based nanoparticle (LNP)

[0178] In some embodiments, the compositions of the disclosure can be formulated in a lipid-based nanoparticle (LNP). For example, the nucleic acid constructs of the disclosure can be delivered to a cell or a subject by a lipid-based nanoparticle (LNP). LNP are generally less immunogenic than viral particles. While many humans have preexisting immunity to viral particles there is no pre-existing immunity to LNP. In addition, adaptive immune response against LNP is unlikely to occur which enables repeat dosing of LNP.

[0179] The lipids suitable for the compositions and methods described herein can be cationic lipids, ionizable cationic lipids, anionic lipids, or neutral lipids.

[0180] In some embodiments, the LNP of the disclosure can include one or more ionizable lipids. As used herein, the term "ionizable lipid" refers to a lipid that is cationic or becomes ionizable (protonated) as the pH is lowered below the pKa of the ionizable group of the lipid, but is more neutral at higher pH values. At pH values below the pKa, the lipid is then able to associate with negatively charged nucleic acids (*e.g.*, oligonucleotides). As used herein, the term "ionizable lipid" includes lipids that assume a positive charge on pH decrease from physiological pH, and any of a number of lipid species that carry a net positive charge at a selective pH, such as physiological pH. Permanently cationic lipids such as DOTMA have proven too toxic for clinical use. The ionizable lipid can be present in lipid formulations according to other embodiments, preferably in a ratio of about 30 to about 70 Mol%, in some embodiments, about

30 Mol%, in other embodiments, about 40 Mol%, in other embodiments, about 45 Mol% in other embodiments, about 47.5 Mol% in other embodiments, about 50 Mol%, in still other embodiments, and about 60 Mol% in yet others (“Mol%” means the percentage of the total moles that is of a particular component). the LNP of the disclosure can include DODMA, or 1,2-dioleoyloxy-3-dimethylaminopropane, which is an ionizable lipid, as is DLin-MC3-DMA or 0-(Z,Z,Z,Z-heptatriaconta-6,9,26,29-tetraen-19-yl)-4-(N,N- dimethylamino) (“MC3”).

[0181] Exemplary ionizable lipids suitable for the compositions and methods of the disclosure includes those described in PCT publications WO2020252589A1 and WO2021000041A1, U.S. Patent Nos. 8,450,298 and 10,844,028, and Love K.T. *et al.*, *Proc Natl Acad Sci USA*, Feb. 2, 2010 107 (5) 1864-1869, all of which are hereby incorporated by reference in their entirety. Accordingly, in some embodiments, the LNP of the disclosure includes one or more lipid compounds described in Love K.T. *et al.*, 2010 *supra*, such as C16-96, C14-110, and C12-200. In some embodiments, the LNP includes an ionizable cationic lipid selected from the group consisting of ALC-0315, C12-200, LN16, MC3, MD1, SM-102, and a combination of any thereof. In some embodiments, the LNP of the disclosure includes C12-200. The structure of C12-200 lipid is known in the art and described in, *e.g.*, U.S. Patent Nos. 8,450,298 and 10,844,028, which are hereby incorporated by reference in their entirety. In some embodiments the C12-200 is combined with cholesterol, C14-PEG2000, and DOPE. In some embodiments, the C12-200 is combined with DSPC and DMG-PEG2000.

[0182] In some embodiments, the LNP of the disclosure includes one or more cationic lipids. Suitable cationic lipids include, but are not limited to, 98N12-5, C12-200, C14-PEG2000, DLin-KC2-DMA (KC2), DLin-MC3-DMA (MC3), XTC, MD1, and 7C1. In some embodiments, the LNP of the disclosure includes one or more neutral lipids. Non-limiting neutral lipids suitable for the compositions and methods of the disclosure include DPSC, DPPC, POPC, DOPE, and SM. In some embodiments, the LNP of the disclosure includes one or more ionizable lipid compounds described in PCT publications WO2020252589A1 and WO2021000041A1, which are hereby incorporated by reference in their entirety.

[0183] A number of other lipids or combination of lipids that are known in the art can be used to produce a LNP. Non-limiting examples of lipids suitable for use to produce LNPs include DOTMA, DOSPA, DOTAP, DMRIE, DC-cholesterol, DOTAP-cholesterol, GAP-

DMORIE-DPyPE, and GL67A-DOPE-DMPE-polyethylene glycol (PEG). Non-limiting examples of cationic lipids include 98N12-5, C12-200, C14-PEG2000, DLin-KC2-DMA (KC2), DLin-MC3-DMA (MC3), XTC, MD1, 7C1, and a combination of any thereof. Non-limiting examples of neutral lipids include DPSC, DPPC, POPC, DOPE, and SM. Non-limiting examples of PEG-modified lipids include PEG-DMG, PEG-CerC14, and PEG-CerC20.

[0184] In some embodiments, the LNP of the disclosure includes at least one lipid selected from the group consisting of C12-200, C14-PEG2000, DOPE, DMG-PEG2000, DPSC, DOTMA, DOSPA, DOTAP, DMRIE, DC-cholesterol, DOTAP-cholesterol, GAP-DMORIE-DPyPE, and GL67A-DOPE-DMPE-polyethylene glycol (PEG). In some embodiments the C12-200 is combined with cholesterol, C14-PEG2000, and DOPE. In some embodiments, the C12-200 is combined with DPSC and DMG-PEG2000.

[0185] In some embodiments, the mass ratio of lipid to nucleic acid in the LNP delivery system is about 100:1 to about 3:1, about 70:1 to 10:1, or 16:1 to 4:1. In some embodiments, the mass ratio of lipid to nucleic acid in the LNP delivery system is about 16:1 to 4:1. In some embodiments, the mass ratio of lipid to nucleic acid in the LNP delivery system is about 20:1. In some embodiments, the mass ratio of lipid to nucleic acid in the LNP delivery system is about 8:1. In some embodiments, the lipid-based nanoparticles have an average diameter of less than about 1000 nm, about 500 nm, about 250 nm, about 200 nm, about 150 nm, about 100 nm, about 75 nm, about 50 nm, or about 25 nm. In some embodiments, the LNPs have an average diameter ranging from about 70 nm to 100 nm. In some embodiments, the LNPs have an average diameter ranging from about 88 nm to about 92 nm, from 82 nm to about 86 nm, or from about 80 nm to about 95 nm.

[0186] As described above, neural lipids, also known as “structural lipids” or “helper lipids” can also be incorporated into lipid formulations and lipid particles in some embodiments. The lipid formulations and lipid particles can include one or more structural lipids at about 10 to 40 Mol% of the composition. Suitable structural lipids support the formation of particles during manufacture. Structural lipids refer to any one of a number of lipid species that exist in either in an anionic, uncharged or neutral zwitterionic form at physiological pH. Representative structural lipids include diacylphosphatidylcholines, diacylphosphatidylethanolamines, diacylphosphatidylglycerols, ceramides, sphingomyelins, dihydrosphingomyelins, cephalins, and

cerebrosides.

[0187] Exemplary structural lipids include zwitterionic lipids, for example, distearoylphosphatidylcholine (DSPC), dioleoylphosphatidylcholine (DOPC), dipalmitoylphosphatidylcholine (DPPC), dioleoyl-phosphatidylethanolamine (DOPE), palmitoyloleoylphosphatidylcholine (POPC), palmitoyloleoyl-phosphatidylethanolamine (POPE) and dioleoyl-phosphatidylethanolamine 4-(N-maleimidomethyl)-cyclohexane-1-carboxylate (DOPE-mal), dipalmitoyl phosphatidyl ethanolamine (DPPE), dimyristoylphosphoethanolamine (DMPE), distearoyl-phosphatidylethanolamine (DSPE), 16-O-monomethyl PE, 16-O-dimethyl PE, 18-1-trans PE, 1-stearoyl-2-oleoyl- phosphatidyethanol amine (SOPE), and 1,2-dielaidoyl-sn-glycero-3-phosphoethanolamine (trans DOPE).

[0188] In another embodiment, the structural lipid can be any lipid that is negatively charged at physiological pH. These lipids include phosphatidylglycerols such as dioleoylphosphatidylglycerol (DOPG), dipalmitoylphosphatidylglycerol (DPPG), palmitoyloleoylphosphatidylglycerol (POPG), cardiolipin, phosphatidylinositol, diacylphosphatidylserine, diacylphosphatidic acid, and other anionic modifying groups joined to neutral lipids. Other suitable structural lipids include glycolipids (e.g., monosialoganglioside GM1).

[0189] Stabilizing agents can be included in lipid formulations embodiments to ensure integrity of the mixtures. Stabilizing agents are a class of molecules which disrupt or help form the hydrophobic-hydrophilic interactions among molecules. Suitable Stabilizing agents include, but are not limited to, polysorbate 80 (also known as Tween 80, IUPAC name 2-[2-[3,4-bis(2-hydroxyethoxy)oxolan-2-yl]-2-(2-hydroxyethoxy)ethoxy]ethyl octadec-9-enoate), Myrj52 (Polyoxyethylene (40) stearate), and Brij™ S10 (Polyoxyethylene (10) stearyl ether). Polyethylene glycol conjugated lipids may also be used. The stabilizing agents may be used alone or in combinations with each other.

[0190] In some embodiments, the stabilizing agents comprises about 0.1 to 3 Mol% of the overall lipid mixture. In some embodiments, the stabilizing agents comprise about 0.5 to 2.5 Mol% of the overall lipid mixture. In some embodiments, the stabilizing agent is present at greater than 2.5Mol%. In some embodiments the stabilizing agent is present at 5 Mol%. In some embodiments the stabilizing agent is present at 10 Mol%. In some embodiments, the stabilizing

agent is about 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, and so forth. In other embodiments, the stabilizing agent is 2.6-10 Mol% of the lipid mixture. In other embodiments, the stabilizing agents is present at greater than 10 Mol% of the lipid mixture.

[0191] Steroids can also be included in the lipid compositions for certain applications, and lipid particles made therefrom include sterols, such as cholesterol and phytosterol.

[0192] In some embodiments, the immunogenic compositions are substantially non-immunogenic or minimally immunogenic (*e.g.* compositions that minimally stimulate an immune response in a subject. In some embodiments, the non-immunogenic or minimally immunogenic compositions are formulated as a biotherapeutic. In some embodiments, the pharmaceutical compositions are formulated for one or more of intranasal administration, transdermal administration, intraperitoneal administration, intramuscular administration, intranodal administration, intratumoral administration, intraarticular administration, intravenous administration, subcutaneous administration, intravaginal administration, cardiac administration and oral administration.

[0193] Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™. (BASF, Parsippany, N.J.), or phosphate buffered saline (PBS), tris (tromethamine), and HEPES. In these cases, the composition should be sterile and should be fluid to the extent that easy syringability exists. It can be stable under the conditions of manufacture and storage, and can be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants, *e.g.*, sodium dodecyl sulfate. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be generally to include

isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, sucrose, trehalose, and/or sodium chloride in the composition. In some embodiments, the composition comprises tris and sucrose. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

[0194] Sterile injectable solutions can be prepared by incorporating the active compound in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle, which contains a basic dispersion medium and the required other ingredients from those enumerated above.

[0195] In some embodiments, the pharmaceutical compositions are formulated for one or more of intranasal administration, transdermal administration, intrathecal administration, intraperitoneal administration, intramuscular administration, intratracheal administration, intranodal administration, intratumoral administration, intraarticular administration, intravenous administration, subcutaneous administration, intravaginal administration, intraocular administration, rectal administration, intravesicular bladder instillation and oral administration.

[0196] In some embodiments, the pharmaceutical compositions of the disclosure are formulated for inhalation, such as an aerosol, spray, mist, liquid, or powder. Administration by inhalation may be in the form of either dry powders or aerosol formulations, which are inhaled by a subject (*e.g.*, a patient) either through use of an inhalation device, *e.g.*, a microspray, a pressurized metered dose inhaler, or nebulizer.

METHODS OF THE DISCLOSURE

[0197] Administration of any one of the therapeutic compositions described herein, *e.g.*, nucleic acid constructs, recombinant cells, and/or pharmaceutical compositions, can be used for modulating at least one pharmacodynamics effect in a subject, or can be used in the treatment of relevant health conditions, such as cancer.

[0198] Non-limiting examples of cancer suitable for the methods of the disclosure include carcinoma, lymphoma, blastoma, sarcoma, and leukemia. More particular examples of such cancers include, but are not limited to, basal cell carcinoma, biliary tract cancer; bladder cancer; bone cancer; brain and CNS cancer; breast cancer; cancer of the peritoneum; cervical cancer;

choriocarcinoma; colon and rectum cancer; connective tissue cancer; cancer of the digestive system; endometrial cancer; esophageal cancer; eye cancer; cancer of the head and neck; gastric cancer (including gastrointestinal cancer); glioblastoma (GBM); hepatic carcinoma; hepatoma; intra-epithelial neoplasm; kidney or renal cancer; larynx cancer; leukemia; liver cancer; lung cancer (e.g., small-cell lung cancer, non-small cell lung cancer, adenocarcinoma of the lung, and squamous carcinoma of the lung); lymphoma including Hodgkin's and non-Hodgkin's lymphoma; melanoma; myeloma; neuroblastoma; oral cavity cancer (e.g., lip, tongue, mouth, and pharynx); ovarian cancer; pancreatic cancer; prostate cancer; retinoblastoma; rhabdomyosarcoma; rectal cancer; cancer of the respiratory system; salivary gland carcinoma; sarcoma; skin cancer; squamous cell cancer; stomach cancer; testicular cancer; thyroid cancer; uterine or endometrial cancer; cancer of the urinary system; vulval cancer; as well as other carcinomas and sarcomas; as well as B-cell lymphoma (including low grade/follicular non-Hodgkin's lymphoma (NHL); small lymphocytic (SL) NHL; intermediate grade/follicular NHL; intermediate grade diffuse NHL; high grade immunoblastic NHL; high grade lymphoblastic NHL; high grade small non-cleaved cell NHL; bulky disease NHL; mantle cell lymphoma; AIDS-related lymphoma; and Waldenstrom's Macroglobulinemia); chronic lymphocytic leukemia (CLL); acute lymphoblastic leukemia (ALL); Hairy cell leukemia; chronic myeloblastic leukemia; and post-transplant lymphoproliferative disorder (PTLD), as well as abnormal vascular proliferation associated with phakomatoses, edema (such as that associated with brain tumors), and Meigs' syndrome. In some embodiments, the cancer is non-small cell lung cancer.

[0199] In some embodiments, the nucleic acid constructs, recombinant cells, and/or pharmaceutical compositions as described herein can be useful for modulating, *e.g.*, eliciting or suppressing a pharmacodynamic effect in a subject in need thereof. In some embodiments, the pharmacodynamic effect includes eliciting an immune response in the subject. Non-limiting examples of pharmacodynamic effect include immunogenicity effects, biomarker responses, therapeutic effects, prophylactic effects, desired effects, undesired effects, adverse effects, and effects in a disease model.

[0200] Accordingly, one aspect of the disclosure relates to methods for modulating a pharmacodynamic effect in a subject in need thereof, the methods include administering to the

subject a composition including one or more of the following: (a) a nucleic acid construct as described herein; (c) a recombinant cell as described herein; and (c) a pharmaceutical composition as described herein. In some embodiments, the pharmacodynamic effect includes one or more of the following: immunogenicity effect, a biomarker response, a therapeutic effect, a prophylactic effect, a desired effect, an undesired effect, an adverse effect, and effect in a disease model. In some embodiments, the pharmacodynamic effect includes eliciting an immune response in the subject.

[0201] Accordingly, in another aspect, provided herein are methods for preventing or treating a health condition in a subject, the methods include prophylactically or therapeutically administering to the subject a composition including one or more of the following: (a) a replicon, *e.g.*, self-replicating RNA construct (srRNA) as described herein; (b) a nucleic acid as described herein; (c) a recombinant cell as described herein; and (d) a pharmaceutical composition as described herein. In some embodiments, the administered composition elicits an immune response in the subject. In some embodiments, the administered composition induces production of one or more pro-inflammatory molecules in the subject. In some embodiments, the one or more pro-inflammatory molecules includes interleukin-1 alpha (IFN α), interleukin-1 beta (IFN β), interleukin-18 (IL-18), interleukin-6 (IL-6), interleukin-1 alpha (IL-1 α), interleukin-1 beta (IL-1 β), interleukin-12 (IL-12), interleukin-2 (IL-2), IL-23, IL-27, interferon gamma (IFN γ), cytokines, TNF- α , GM-CSF, and MIP1 α , granzyme B, granzyme A, perforin, or a combination of any thereof. In some embodiments, the subject has been previously treated with one or more therapies and has developed at least a partial resistance to said one or more therapies.

[0202] As described above, administration of any one of the therapeutic compositions described herein, *e.g.*, nucleic acid constructs, recombinant cells, and/or pharmaceutical compositions, can be used for inducing at least one pharmacodynamic effect in a subject. In some embodiments, the nucleic acid constructs, recombinant cells, and/or pharmaceutical compositions described herein are analyzed for their capacity to confer at least one pharmacodynamic effect is carried out *in vivo* or *ex vivo*. Examples of pharmacodynamic effects that can be analyzed include: immunogenicity effect (*e.g.*, eliciting an immune response *in vivo*), a biomarker response, a therapeutic effect, a prophylactic effect, a desired effect, an undesired effect, an adverse effect, and effect in a disease model. An effect in a disease model can include

a pharmacodynamic effect that results in control, lessening, or reversal of a disease and/or its attributes, such as prevention or reduction of mortality, or morbidity, or disease severity. In some embodiments, the assessment of pharmacodynamic effects includes assessing induction of an immune response *in vivo*. In some embodiments, the assessment of pharmacodynamic effects includes assessing induction of cytokine pathways that can potentiate an immune response and prevent angiogenesis and metastasis.

[0203] In some embodiments, the disclosed composition is formulated to be compatible with its intended route of administration. For example, the nucleic acid constructs (*e.g.*, srRNA constructs), recombinant cells, and/or pharmaceutical compositions of the disclosure may be given orally or by inhalation, but it is more likely that they will be administered through a parenteral route. Examples of parenteral routes of administration include, for example, intramuscular, intratumoral, intraocular, intravenous, intranodal, intradermal, subcutaneous, transdermal (topical), transmucosal, intravaginal, and rectal administration. In some embodiments, the composition is administered intramuscularly. In some embodiments, the composition is administered intratumorally or *via* cardiac administration. Solutions or suspensions used for parenteral application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid (EDTA); buffers such as acetates, citrates, phosphates, tris, sucrose and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as mono- and/or di-basic sodium phosphate, hydrochloric acid or sodium hydroxide (*e.g.*, to a pH of about 7.2-7.8, *e.g.*, 7.5). The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

[0204] Dosage, toxicity and therapeutic efficacy of such subject nucleic acid constructs, recombinant cells, and/or pharmaceutical compositions of the disclosure can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD₅₀/ED₅₀. Compounds that

exhibit high therapeutic indices are generally suitable. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

[0205] For example, the data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies generally within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the disclosure, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC₅₀ (*e.g.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

[0206] The therapeutic compositions described herein, *e.g.*, nucleic acid constructs, *e.g.*, srRNA constructs, recombinant cells, and/or pharmaceutical compositions, can be administered one from one or more times per day to one or more times per week; including once every other day. The skilled artisan will appreciate that certain factors may influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of the subject multivalent polypeptides and multivalent antibodies of the disclosure can include a single treatment or, can include a series of treatments. In some embodiments, the compositions are administered every 8 hours for five days, followed by a rest period of 2 to 14 days, *e.g.*, 9 days, followed by an additional five days of administration every 8 hours. With regard to nucleic acid constructs (*e.g.*, srRNA constructs), the therapeutically effective amount of a nucleic acid construct of the disclosure (*e.g.*, an effective dosage) depends on the nucleic acid construct selected.

[0207] As discussed *supra*, a therapeutically effective amount includes an amount of a

therapeutic composition that is sufficient to promote a particular effect when administered to a subject, such as one who has, is suspected of having, or is at risk for a health condition, *e.g.*, an autoimmune disease, an inflammatory disease, or a cardiovascular disease. In some embodiments, an effective amount includes an amount sufficient to prevent or delay the development of a symptom of the disease, alter the course of a symptom of the disease (for example but not limited to, slow the progression of a symptom of the disease), or reverse a symptom of the disease.

[0208] The efficacy of a treatment including a disclosed therapeutic composition for the treatment of health condition or disease can be determined by the skilled clinician. However, a treatment is considered effective treatment if at least any one or all of the signs or symptoms of disease are improved or ameliorated. Efficacy can also be measured by failure of an individual to worsen as assessed by hospitalization or need for medical interventions (*e.g.*, progression of the health condition or disease is halted or at least slowed). Methods of measuring these indicators are known to those of skill in the art and/or described herein. Treatment includes any treatment of a health condition or disease in a subject or an animal (some non-limiting examples include a human, or a mammal) and includes: (1) inhibiting the health condition or disease, *e.g.*, arresting, or slowing the progression of symptoms; or (2) relieving the health condition or disease, *e.g.*, causing regression of symptoms; and (3) preventing or reducing the likelihood of the development of symptoms.

[0209] In some embodiments, the nucleic acid constructs (*e.g.*, srRNA constructs), recombinant cells, and/or pharmaceutical compositions of the disclosure can be administered to a subject in a composition having a pharmaceutically acceptable carrier and in an amount effective to stimulate an immune response. Generally, a subject can be immunized through an initial series of injections (or administration through one of the other routes described below) and subsequently given boosters to increase the protection afforded by the original series of administrations. The initial series of injections and the subsequent boosters are administered in such doses and over such a period of time as is necessary to stimulate an immune response in a subject. In some embodiments, the administered composition results in an increased production of interferon in the subject. In some embodiments of the disclosed methods, the subject is a mammal. In some embodiments, the mammal is a human subject.

[0210] When the nucleic acid constructs, recombinant cells, and/or pharmaceutical compositions are suitably protected, as described above, they may be orally administered, for example, with an inert diluent or an assimilable edible carrier. The nucleic acid constructs, recombinant cells, and/or pharmaceutical compositions and other ingredients may also be enclosed in a hard or soft shell gelatin capsule, compressed into tablets, or incorporated directly into the individual's diet. For oral therapeutic administration, the active compound may be incorporated with excipients and used in the form of ingestible tablets, buccal tablets, troches, capsules, elixirs, suspensions, syrups, wafers, and the like.

[0211] In some embodiments, the nucleic acid constructs of the disclosure can be delivered to a cell or a subject by a lipid-based nanoparticle (LNP). LNP are generally less immunogenic than viral particles. While many humans have preexisting immunity to viral particles there is no pre-existing immunity to LNP. In addition, adaptive immune response against LNP is unlikely to occur which enables repeat dosing of LNP.

Additional therapies

[0212] In some embodiments, a composition according to the present disclosure is administered to the subject individually as a single therapy (monotherapy) or as a first therapy in combination with at least one additional therapies (e.g., second therapy). In some embodiments, the second therapy is selected from the group consisting of chemotherapy, radiotherapy, immunotherapy, hormonal therapy, toxin therapy, targeted therapy, and surgery. In some embodiments, the second therapy is selected from the group consisting of chemotherapy, radiotherapy, immunotherapy, hormonal therapy, toxin therapy or surgery. In some embodiments, the first therapy and the second therapy are administered concomitantly. In some embodiments, the first therapy is administered at the same time as the second therapy. In some embodiments, the first therapy and the second therapy are administered sequentially. In some embodiments, the first therapy is administered before the second therapy. In some embodiments, the first therapy is administered after the second therapy. In some embodiments, the first therapy is administered before and/or after the second therapy. In some embodiments, the first therapy and the second therapy are administered in rotation. In some embodiments, the first therapy and the second therapy are administered together in a single formulation.

KITS

[0213] Also provided herein are various kits for the practice of a method described herein as well as written instructions for making and using the same. In particular, some embodiments of the disclosure provide kits for modulating (*e.g.*, inducing, eliciting, or suppressing) a pharmacodynamic effect. Some embodiments of the disclosure provide kits for eliciting an immune response in a subject. Some other embodiments relate to kits for the prevention of a health condition, *e.g.*, an autoimmune disease, in a subject in need thereof. Some other embodiments relate to kits for methods of treating a health condition, *e.g.*, an autoimmune disease, in a subject in need thereof. For example, provided herein, in some embodiments, are kits that include one or more of the nucleic acid constructs, recombinant cells, and/or pharmaceutical compositions as provided and described herein, as well as written instructions for making and using the same.

[0214] In some embodiments, the kits of the disclosure further include one or more means useful for the administration of any one of the provided nucleic acid constructs, recombinant cells, and/or pharmaceutical compositions to a subject. For example, in some embodiments, the kits of the disclosure further include one or more syringes (including pre-filled syringes) and/or catheters (including pre-filled syringes) used to administer any one of the provided nucleic acid constructs, recombinant cells, and/or pharmaceutical compositions to a subject. In some embodiments, a kit can have one or more additional therapeutic agents that can be administered simultaneously or sequentially with the other kit components for a desired purpose, *e.g.*, for diagnosing, preventing, or treating a condition in a subject in need thereof.

[0215] Any of the above-described kits can further include one or more additional reagents, where such additional reagents can be selected from: dilution buffers, reconstitution solutions, wash buffers, control reagents, control expression vectors, negative controls, positive controls, reagents suitable for *in vitro* production of the provided nucleic acid constructs, recombinant cells, and/or pharmaceutical compositions of the disclosure.

[0216] In some embodiments, the components of a kit can be in separate containers. In some other embodiments, the components of a kit can be combined in a single container. Accordingly, in some embodiments of the disclosure, the kit includes one or more of the nucleic acid constructs, recombinant cells, and/or pharmaceutical compositions as provided and described herein in one container (*e.g.*, in a sterile glass or plastic vial) and a further therapeutic

agent in another container (*e.g.*, in a sterile glass or plastic vial).

[0217] In another embodiment, the kit includes a combination of the compositions described herein, including one or more nucleic acid constructs, recombinant cells, and/or pharmaceutical compositions of the disclosure in combination with one or more further therapeutic agents formulated together, optionally, in a pharmaceutical composition, in a single, common container.

[0218] If the kit includes a pharmaceutical composition for parenteral administration to a subject, the kit can include a device (*e.g.*, an injection device or catheter) for performing such administration. For example, the kit can include one or more hypodermic needles or other injection devices as discussed above containing one or more nucleic acid constructs, recombinant cells, and/or pharmaceutical compositions of the disclosure.

[0219] In some embodiments, the components of a kit can be in separate containers. In some other embodiments, the components of a kit can be combined in a single container.

[0220] In some embodiments, a kit can further include instructions for using the components of the kit to practice the methods disclosed herein. For example, the kit can include a package insert including information concerning the pharmaceutical compositions and dosage forms in the kit. Generally, such information aids patients and physicians in using the enclosed pharmaceutical compositions and dosage forms effectively and safely. For example, the following information regarding a combination of the disclosure may be supplied in the insert: pharmacokinetics, pharmacodynamics, clinical studies, efficacy parameters, indications and usage, contraindications, warnings, precautions, adverse reactions, over-dosage, proper dosage and administration, how supplied, proper storage conditions, references, manufacturer/distributor information and intellectual property information.

[0221] The instructions for practicing the methods are generally recorded on a suitable recording medium. For example, the instructions can be printed on a substrate, such as paper or plastic, *etc.* The instructions can be present in the kit as a package insert, in the labeling of the container of the kit or components thereof (*e.g.*, associated with the packaging or sub-packaging), *etc.* The instructions can be present as an electronic storage data file present on a suitable computer readable storage medium, *e.g.* CD-ROM, diskette, flash drive, *etc.* In some instances, the actual instructions are not present in the kit, but means for obtaining the

instructions from a remote source (e.g., via the internet), can be provided. An example of this embodiment is a kit that includes a web address where the instructions can be viewed and/or from which the instructions can be downloaded. As with the instructions, this means for obtaining the instructions can be recorded on a suitable substrate.

[0222] All publications and patent applications mentioned in this disclosure are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

[0223] No admission is made that any reference cited herein constitutes prior art. The discussion of the references states what their authors assert, and the Applicant reserves the right to challenge the accuracy and pertinence of the cited documents. It will be clearly understood that, although a number of information sources, including scientific journal articles, patent documents, and textbooks, are referred to herein; this reference does not constitute an admission that any of these documents forms part of the common general knowledge in the art.

[0224] The discussion of the general methods given herein is intended for illustrative purposes only. Other alternative methods and alternatives will be apparent to those of skill in the art upon review of this disclosure, and are to be included within the spirit and purview of this application.

[0225] Additional embodiments are disclosed in further detail in the following examples, which are provided by way of illustration and are not in any way intended to limit the scope of this disclosure or the claims.

EXAMPLES

[0226] The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, cell biology, biochemistry, nucleic acid chemistry, and immunology, which are well known to those skilled in the art. Such techniques are explained fully in the literature, such as Sambrook, J., & Russell, D. W. (2012). *Molecular Cloning: A Laboratory Manual* (4th ed.). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory and Sambrook, J., & Russel, D. W. (2001). *Molecular Cloning: A Laboratory Manual* (3rd ed.). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory (jointly referred to herein as “Sambrook”); Ausubel, F. M. (1987). *Current Protocols in Molecular Biology*. New York, NY: Wiley (including supplements through 2014); Bollag, D. M. *et al.* (1996). *Protein*

Methods. New York, NY: Wiley-Liss; Huang, L. *et al.* (2005). *Nonviral Vectors for Gene Therapy*. San Diego: Academic Press; Kaplitt, M. G. *et al.* (1995). *Viral Vectors: Gene Therapy and Neuroscience Applications*. San Diego, CA: Academic Press; Lefkovits, I. (1997). *The Immunology Methods Manual: The Comprehensive Sourcebook of Techniques*. San Diego, CA: Academic Press; Doyle, A. *et al.* (1998). *Cell and Tissue Culture: Laboratory Procedures in Biotechnology*. New York, NY: Wiley; Mullis, K. B., Ferré, F. & Gibbs, R. (1994). *PCR: The Polymerase Chain Reaction*. Boston: Birkhauser Publisher; Greenfield, E. A. (2014). *Antibodies: A Laboratory Manual* (2nd ed.). New York, NY: Cold Spring Harbor Laboratory Press; Beaucage, S. L. *et al.* (2000). *Current Protocols in Nucleic Acid Chemistry*. New York, NY: Wiley, (including supplements through 2014); and Makrides, S. C. (2003). *Gene Transfer and Expression in Mammalian Cells*. Amsterdam, NL: Elsevier Sciences B.V., the disclosures of which are incorporated herein by reference.

[0227] Additional embodiments are disclosed in further detail in the following examples, which are provided by way of illustration and are not in any way intended to limit the scope of this disclosure or the claims.

EXAMPLE 1

Construction of modified alphavirus vectors

[0228] This Example describes the experiments performed to construct base alphavirus vectors (*e.g.*, without a heterologous gene) that were subsequently used for construction of vectors that express a gene of interest (*e.g.*, EGFR).

EEEV base vector

[0229] The base EEEV vector (*i.e.* without a heterologous gene of interest) was constructed as follows: The base EEEV vector was synthesized *de novo* in four ~4 kb parts (Twist Bioscience) from a reference sequence (Genbank EF151502) with several modifications. Silent mutations G301A, A3550C, G4516A, G5725A, G7399A were incorporated to eliminate restriction enzyme cut sites. A unique restriction enzyme cut site (SpeI, 5'-A'CTAG,T-3') was incorporated in place of the coding sequence of the native EEEV structural genes (where the 5' A matches the location of the structural polyprotein ATG start codon, and the 3' T matches the location of the structural polyprotein stop codon TAA). A 5' adaptor sequence (5'-CTGGAGACGTGGAGGAGAACCCTGGACCT-3'; SEQ ID NO: 2) was inserted upstream of

the SpeI site, and a 3' adaptor sequence (5'-GACCGCTACGCCCCAATGACCCGACCAGC-3'; SEQ ID NO: 3) was inserted downstream of the SpeI site for subsequent Gibson Assembly® procedures (Gibson et al., Nat. Methods 6, 343–345, 2009). A bacteriophage T7 RNA polymerase promoter (5'-TAATACGACTCACTATAG-3'; SEQ ID NO: 4) was included upstream of the EEEV genome sequence, and downstream contained a poly(A) sequence followed by a SapI site, which cuts upstream of the recognition site. Immediately downstream of the SapI site is a T7 terminator sequence (5'-AACCCCTCTCTAAACGGAGGGGTTTTTTT-3'; SEQ ID NO: 5) followed by a unique restriction enzyme cut site (NotI, 5'-GC'GGCC,GC-3'). The parts were combined in a five-piece Gibson Assembly® reaction: a linearized pYL backbone and the four synthesized fragments to result in the EEEV base vector.

CHIKV base vectors

[0230] The base CHIKV S27 vector was synthesized de novo in four ~4 kb parts (Twist Bioscience, Thermo Fisher GeneArt) from a reference sequence (Genbank AF369024) with a silent A5366G mutation, and with a unique restriction enzyme cut site (SpeI, 5'-A'CTAG,T-3') in place of the coding sequence of the CHIKV structural genes (where the 5' A matches the location of the structural polyprotein's ATG start codon, and the 3' T matches the location of the structural polyprotein's stop codon TAA). A 5' adaptor sequence (5'-CTGGAGACGTGGAGGAGAACCCTGGACCT-3'; SEQ ID NO: 2) was inserted upstream of the SpeI site, and a 3' adaptor sequence (5'-GACCGCTACGCCCCAATGACCCGACCAGC-3'; SEQ ID NO: 3) was inserted downstream of the SpeI site for subsequent Gibson Assembly® procedures. A bacteriophage T7 RNA polymerase promoter (5'-TAATACGACTCACTATAG-3'; SEQ ID NO: 4) was included upstream of the CHIKV genome sequence, and downstream contained a poly(A) sequence followed by a SapI site, which cuts upstream of the recognition site. Immediately downstream of the SapI site is a T7 terminator sequence (5'-AACCCCTCTCTAAACGGAGGGGTTTTTTT-3'; SEQ ID NO: 5) followed by a unique restriction enzyme cut site (NotI, 5'-GC'GGCC,GC-3'). The parts were combined in a five-piece Gibson Assembly® reaction a linearized pYL backbone and the four synthesized fragments to result in the CHIKV S27 base vector.

[0231] The CHIKV DRDE base vector was similarly constructed from a reference sequence (Genbank EF210157), except the S27 3' UTR was used in place of the DRDE 3' UTR.

SINV base vectors

[0232] The base SINV Girdwood vector was synthesized de novo in four ~4 kb parts (Twist Bioscience, Thermo Fisher GeneArt) from a Girdwood strain reference sequence (Genbank MF459683) with a unique restriction enzyme cut site (SpeI, 5'-A'CTAG,T-3') in place of the coding sequence of the SINV structural genes (where the 5' A is the next nucleotide after a P2A sequence following nucleotide 93 of the structural polyprotein gene, and the 3' T matches the location of the structural polyprotein's stop codon TGA). A bacteriophage T7 RNA polymerase promoter (5'-TAATACGACTCACTATAG-3'; SEQ ID NO: 4) was included upstream of the SINV genome sequence, and downstream contained a poly(A) sequence followed by a SapI site, which cuts upstream of the recognition site. Immediately downstream of the SapI site is a T7 terminator sequence (5'-AACCCCTCTCTAAACGGAGGGGTTTTTTT-3'; SEQ ID NO: 5) followed by a unique restriction enzyme cut site (NotI, 5'-GC'GGCC,GC-3'). The parts were combined in a five-piece Gibson Assembly® reaction (e.g., a linearized pYL backbone and the four synthesized fragments) to result in the SINV Girdwood base vector.

[0233] The base SINV AR86 vector was similarly constructed from a reference sequence (Genbank U38305), except the nsP2 coding sequence was derived from the Girdwood reference sequence.

VEE base vector

[0234] The base VEE vector was synthesized de novo in four ~4 kb parts (Twist Bioscience, Thermo Fisher GeneArt) from a TC-83 strain reference sequence (Genbank L01443) with a silent A2087G mutation, and a unique restriction enzyme cut site (SpeI, 5'-A'CTAG,T-3') in place of the coding sequence of the VEE structural genes (where the 5' A is the next nucleotide after a P2A sequence following nucleotide 93 of the structural polyprotein gene, and the 3' T matches the location of the structural polyprotein's stop codon TGA). A 5' adaptor sequence (5'-CTGGAGACGTGGAGGAGAACCCTGGACCT-3'; SEQ ID NO: 2) was inserted upstream of the SpeI site, and a 3' adaptor sequence (5'-GACCGCTACGCCCCAATGACCCGACCAGC-3'; SEQ ID NO: 3) was inserted downstream of the SpeI site for subsequent Gibson Assembly® procedures. A bacteriophage T7 RNA polymerase promoter (5'-TAATACGACTCACTATAG-3'; SEQ ID NO: 4) was included upstream of the VEE genome sequence, and downstream contained a poly(A) sequence followed

by a SapI site, which cuts upstream of the recognition site. Immediately downstream of the SapI site is a T7 terminator sequence (5'-AACCCCTCTCTAAACGGAGGGGTTTTTTT-3'; SEQ ID NO: 5) followed by a unique restriction enzyme cut site (NotI, 5'-GC'GGCC,GC-3'). The parts were combined in a five-piece Gibson Assembly® reaction (e.g., a linearized pYL backbone and the four synthesized fragments) to result in the VEE base vector.

MADV base vector

[0235] The base MADV vector (i.e. without a heterologous gene of interest) was constructed as follows: The base MADV vector was synthesized de novo in three 869 to 4992 bp parts from a reference sequence (Genbank KJ469641) with several modifications. A unique restriction enzyme cut site (SpeI, 5'-A'CTAG,T-3') was incorporated in place of the coding sequence of the native MADV structural genes (where the 5' A matches the location of the structural polyprotein ATG start codon, and the 3' T matches the location of the structural polyprotein stop codon TAA). A 5' adaptor sequence (5'-CTGGAGACGTGGAGGAGAACCCTGGACCT-3'; SEQ ID NO: 2) was inserted upstream of the SpeI site, and a 3' adaptor sequence (5'-GACCGCTACGCCCCAATGACCCGACCAGC-3'; SEQ ID NO: 3) was inserted downstream of the SpeI site for subsequent Gibson Assembly® procedures (Gibson et al., Nat. Methods 6, 343–345, 2009). A bacteriophage T7 RNA polymerase promoter (5'-TAATACGACTCACTATAG-3'; SEQ ID NO: 4) was included upstream of the MADV genome sequence, and downstream contained a poly(A) sequence followed by a SapI site, which cuts upstream of the recognition site. Immediately downstream of the SapI site is a T7 terminator sequence (5'-AACCCCTCTCTAAACGGAGGGGTTTTTTT-3'; SEQ ID NO: 5) followed by a unique restriction enzyme cut site (NotI, 5'-GC'GGCC,GC-3'). The parts were combined in a five-piece Gibson Assembly® reaction: a linearized pYL backbone and the four synthesized fragments to result in the MADV base vector.

WEEV base vectors

[0236] The base WEEV vectors (i.e. without a heterologous gene of interest) were constructed as follows: The base WEEV Imperial181 vector was synthesized de novo in four 200 to 3875 bp parts from a reference sequence (Genbank GQ287641) with several modifications. Ambiguous base calls in the reference sequence were assigned nucleotides: K1655T, W4518C, and M7804A. Silent mutations A6948G and A7242G were incorporated to eliminate restriction

enzyme cut sites. A unique restriction enzyme cut site (SpeI, 5'-A'CTAG,T-3') was incorporated in place of the coding sequence of the native WEEV structural genes (where the 5' A matches the location of the structural polyprotein ATG start codon, and the 3' T matches the location of the structural polyprotein stop codon TAA). A 5' adaptor sequence (5'-CTGGAGACGTGGAGGAGAACCCTGGACCT-3'; SEQ ID NO: 2) was inserted upstream of the SpeI site, and a 3' adaptor sequence (5'-GACCGCTACGCCCCAATGACCCGACCAGC-3'; SEQ ID NO: 3) was inserted downstream of the SpeI site for subsequent Gibson Assembly® procedures (Gibson et al., Nat. Methods 6, 343–345, 2009). A bacteriophage T7 RNA polymerase promoter (5'-TAATACGACTCACTATAG-3'; SEQ ID NO: 4) was included upstream of the WEEV genome sequence, and downstream contained a poly(A) sequence followed by a SapI site, which cuts upstream of the recognition site. Immediately downstream of the SapI site is a T7 terminator sequence (5'-AACCCCTCTCTAAACGGAGGGGTTTTTTT-3'; SEQ ID NO: 5) followed by a unique restriction enzyme cut site (NotI, 5'-GC'GGCC,GC-3'). The parts were combined in a five-piece Gibson Assembly® reaction: a linearized pYL backbone and the four synthesized fragments to result in the WEEV Imperial181 base vector. A similar approach was used to generate the WEEV McMillan base vector.

[0237] The EGFR transgenes were synthesized (IDT) with flanking sequences homologous to the 5' and 3' adaptor sequences, and inserted into the SpeI-linearized base vectors by Gibson Assembly® to result in the final vectors.

EXAMPLE 2

In vitro evaluation of modified alphavirus vectors

[0238] This Example describes the results of *in vitro* experiments performed to evaluate expression levels of the synthetic srRNA constructs described in Example 1 above, and to investigate any differential behavior thereof (e.g., protein expression).

[0239] *In vitro* transcription: RNA was prepared by *in vitro* transcription from a SapI-linearized plasmid template with bacteriophage T7 RNA polymerase with either a 5' ARCA cap (HiScribe™ T7 ARCA mRNA Kit, NEB) or by uncapped transcription (HiScribe™ T7 High Yield RNA Synthesis Kit, NEB) followed by addition of a 5' cap 1 (Vaccinia Capping System, mRNA Cap 2'-O-Methyltransferase, NEB). RNA is then purified using phenol/chloroform

extraction, or column purification (Monarch® RNA Cleanup Kit, NEB). RNA concentration is determined by absorbance at 260 nm (Nanodrop, Thermo Fisher Scientific).

[0240] Replication: RNA was transformed by electroporation into BHK-21 or Vero cells (e.g., 4D-Nucleofector™, Lonza). At 15-22 hours following transformation, the cells were fixed and permeabilized (eBioscience™ Foxp3 / Transcription Factor Staining Buffer Set, Invitrogen) and stained using a PE-conjugated anti-dsRNA mouse monoclonal antibody (J2, Scicons) to quantify the frequency of dsRNA+ cells by fluorescence flow cytometry.

[0241] Protein expression: RNA was transformed by electroporation into BHK-21 cells (e.g., 4D-Nucleofector™, Lonza). At 15-22 hours following transformation, the cells were fixed and permeabilized (eBioscience™ Foxp3 / Transcription Factor Staining Buffer Set, Invitrogen) and stained using a rabbit anti-EGFR antibody EPR15348 followed by AF647-conjugated or or AF488-conjugated goat anti-rabbit IgG H&L. The mean fluorescence intensity (MFI) was used as the readout of EGFR expression. Results are shown in **FIGs. 1, 2, 4, and 5**.

[0242] **FIG. 1** shows that ARCA-capped srRNA vectors containing EGFR transgene cassettes have the property of EGFR expression, and that the identity of the EGFR transgene cassette can influence the relative level of protein expression. This data allows for the selection of EGFR transgene cassettes that have the property of high levels of EGFR expression. **FIG. 2** shows that enzymatically-capped srRNA vectors containing EGFR transgene cassettes have the property of EGFR expression, and that the identity of the EGFR transgene cassette can influence the relative level of protein expression. This data allows for the selection of EGFR transgene cassettes that have the property of high levels of EGFR expression. **FIGs. 4 and 5** show that different srRNA vectors containing the same EGFR transgene cassette retain the property of EGFR expression, and that the identity of the srRNA vector can influence the relative level of protein expression.

EXAMPLE 3

In vivo evaluation of modified alphavirus vectors

[0243] This Example describes the results of *in vivo* experiments performed to evaluate the srRNA constructs described herein (e.g., both unformulated and LNP formulated vectors).

[0244] In these experiments, synthetic srRNA constructs derived from various alphavirus strains were designed and subsequently evaluated.

[0245] Mice and injections:

[0246] HLA-A2 or HLA-A1101 transgenic mice were purchased from Charles River Labs, Envigo, or Jackson Laboratories. On day of dosing, between 10 µg of material was injected intramuscularly either into one or split into both quadricep muscles. Animals were monitored for body weight and other general observations throughout the course of the study. For immunogenicity studies, animals were dosed on Day 0 only or Day 0 and Day 21.

[0247] LNP formulation: srRNA was formulated in lipid nanoparticles using a microfluidics mixer and analyzed for particle size, polydispersity using dynamic light scattering, and encapsulation efficiency using a dye exclusion assay (Ribogreen). Lipids were suspended in ethanol. Each srRNA was suspended in 100 mM NaOAc pH 4.0 at a concentration of 82 µg/ml and is mixed at a flow rate of 3:1 (aqueous:organic).

[0248] ELISpot. To measure the magnitude of EGFR-specific T cell responses, IFN γ ELISpot analysis was performed using Mouse IFN γ ELISpot PLUS Kit (HRP) (MabTech) as per manufacturer's instructions. Results of mouse IFN γ detecting ELISpot assay as measured by spot-forming units corresponding to responder splenic T cells 14 days after 2 intramuscular injections of srRNA encoding EGFR is shown in **FIG. 6**. The total T cell responses (plotted as counted spot-forming units per million of cells) are shown the y-axis.

[0249] These data demonstrate that EGFR mutations are immunogenic in mice expressing human MHC molecules and predict immunogenicity in humans. Results from this study led to the selection of an optimal vector.

[0250] While particular alternatives of the present disclosure have been disclosed, it is to be understood that various modifications and combinations are possible and are contemplated within the true spirit and scope of the appended claims. There is no intention, therefore, of limitations to the exact abstract and disclosure herein presented.

SEQ ID NO	Sequence	Description
1	MRPSGTAGAALLALLAALCPASRALEEKKVCQGTSNKLTQLGT FEDHFLSLQRMFNNCEVVLGNLEITYVQRNYDL SFLKTIQEVAG YVLIALNTVERIPLLENLQIIRGNMYYENSYALAVLSNYDANKTG LKELPMRNLQEILHGAVRFSNNPALCNVESIQWRDIVSSDFLSN MSMDFQNLHGSCQKCDPSCPNGSCWGAGEENCQKLTKIICAQ QCSGRRCRGKSPSDCCHNQCAAGCTGPRES DCLVCRKFRDEATC KDTCPPLMLYNPTTYQMDVNPEGKYSFGATCVKKCPRNYVVT DHGSCVRACGADSYEMEEDGVRKCKKCEGPCRKVCNGIGIGEF KDSL SINATNIKHFKNCT SISGDLHILPVAFRGDSFTHTPPLDPQE LDILKTVKEITGFLLIQAWPENRTDLHAFENLEIIRGRTKQHGQS LAVVSLNITSLGLRSLKEISDGDVIISGNKNLCYANTINWKKLFG TSGQKTKIISNRGENSCKATGQVCHALCSPEGCWGPEPRDCVSC RNVSRGRECVDKCNLLEGEPRFVENSECIQCHPECLPQAMNIT CTGRGPDNCIQCAHYIDGPHCVKTCPAGVMGENNTLVWKYAD AGHVCHLCHPNCTYGCTGPGLEGCP TNGPKIPSIATGMVGALL LLLVALGIGLFMRRRHIVRKRTLRLRLQERELVEPLTPSGEAP NQALLRILKETEFKKIKVLGSGAFGT VYKGLWIPEGEKV KIPVAI KELREATSPKANKEILDEAYVMASVDNPHVCRLLGICLTSTVQL ITQLMPFGCLLDYVREHKDNIGSQYLLNWCVQIAKGMNYLED R RLVHRDLAARNVLVKT PQHVKITDFGLAKLLGAEEKEYHAEG GKVPIKWMALESILHRIYTHQSDVWSYGVTVWELMTFGSKPY DGIPASEISSILEKGERLPQPPICTIDVYMIMVKCWMIDADSRPKF RELIIEF SKMARDPQRYLVIQGDERMHLPSPTDSNFYRALMDEE DMDDVVDAD EYLIPQQGFFSSPSTSRTPLLSLSATSNNSTVACI DRNGLQSCPIKEDSFLQRYSSDPTGALTEDSIDD TFLPVPEYINQ SVPKRPA GSVQNPVYHNQPLNPAPSRDPHYQDPHSTAVGNPEY LNTVQPTCVNSTFD SPAHWAQKGSHQISLDNPDYQQDFFPKEA KPNGIFKGSTAENAEYLRVAPQSSEFIGA	Wildtype EGFR
2	LITQLMPFGSLLDYVREHKDNIGS	EGFR C797S mutation and flanking residues
3	DEAYVMASVDSVDNPHVCRLLGIC	EGFR exon 20 insertion and flanking residues
4	DEAYVMASVASVDNPHVCRLLGIC	EGFR exon 20 insertion and flanking residues
5	DEAYVMASVDNPHHVCRLLGICLTS	EGFR exon

		20 insertion and flanking residues
6	VKTPQHVKITDFGRAKLLGAEEK	EGFR L858R mutation and flanking residues
7	GICLTSTVQLIMQLMPFGCL	EGFR T790M mutation and flanking residues
8	AAY	Peptide linker
9	EAAAK	Peptide linker
10	RVRR	Peptide linker
11	GGGGS	Peptide linker
12	GPGPG	Peptide linker
13	GICLTSTVQLIMQLMPFGCLGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGSGLITQLMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVDSVDNPHVCRLLGICGGGGSDEAYVMASVASVDNPHVCRLLGICGGGGSDEAYVMASVDNPHHVCRLLGICLTSGGGGS	RBI-212
14	DEAYVMASVASVDNPHVCRLLGICGGGGSDEAYVMASVDSVDNPHVCRLLGICGGGGSDEAYVMASVDNPHHVCRLLGICLTSGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGSGLITQLMPFGSLLDYVREHKDNIGSGGGSGICLTSTVQLIMQLMPFGCLGGGGS	RBI-213
15	GICLTSTVQLIMQLMPFGCLGGGGSGLITQLMPFGSLLDYVREHKDNIGSGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGSDEAYVMASVDNPHHVCRLLGICLTSGGGGSEAYVMASVDSVDNPHVCRLLGICGGGGSDEAYVMASVASVDNPHVCRLLGICGGGGS	RBI-214
16	GICLTSTVQLIMQLMPFGCLGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGSGLITQLMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVASVDNPHVCRLLGICGGGGSDEAYVMASVDSVDNPHVCRLLGICGGGGSDEAYVMASVDNPHHVCRLLGICLTSGGGGS	RBI-215
17	DEAYVMASVDSVDNPHVCRLLGICGGGGSDEAYVMASVDNPHHVCRLLGICLTSGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGSGLITQLMPFGSLLDYVREHKDNIGSGGGSGICLTSTVQLIMQLMPFGCLGGGGSDEAYVMASVASVDNPHVCRLLGICGGGGS	RBI-216
18	LITQLMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVASVDNPHVCRLLGICGGGGSDEAYVMASVDSVDNPHVCRLLGICGGGGS	RBI-217

	SGICLTSTVQLIMQLMPFGCLGGGGSDEAYVMASVDNPHHVCRL LLGICLTSGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGS	
19	DEAYVMASVDNPHHVCRLGICLTSGGGGSDEAYVMASVASV DNPHVCRLLGICGGGGSDEAYVMASVDSVDNPHVCRLLGICGG GGSLITQLMPFGSLLDYVREHKDNIGSGGGGSVKTPQHVKITDF GRAKLLGAEEKGGGSGICLTSTVQLIMQLMPFGCLGGGGS	RBI-218
20	GICLTSTVQLIMQLMPFGCLGGGGSDEAYVMASVDSVDNPHVC RLLGICGGGGSDEAYVMASVASVDNPHVCRLLGICGGGGSVKT PQHVKITDFGRAKLLGAEEKGGGGSDEAYVMASVDNPHHVCRL LLGICLTSGGGGSLITQLMPFGSLLDYVREHKDNIGSGGGGS	RBI-219
21	DEAYVMASVDSVDNPHVCRLGICGGGGSLLITQLMPFGSLLDY VREHKDNIGSGGGGSDEAYVMASVASVDNPHVCRLLGICGGG GSDEAYVMASVDNPHHVCRLGICLTSGGGGSGICLTSTVQLIM QLMPFGCLGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGS	RBI-220
22	LITQLMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVDSVDN PHVCRLLGICGGGSGICLTSTVQLIMQLMPFGCLGGGGSVKTP QHVKITDFGRAKLLGAEEKGGGGSDEAYVMASVASVDNPHVC RLLGICGGGGSDEAYVMASVDNPHHVCRLGICLTSGGGGS	RBI-221
23	DEAYVMASVDNPHHVCRLGICLTSGGGGSGICLTSTVQLIMQ LMPFGCLGGGGSDEAYVMASVASVDNPHVCRLLGICGGGGSV KTPQHVKITDFGRAKLLGAEEKGGGGSDEAYVMASVDSVDNP HVCRLLGICGGGGSLLITQLMPFGSLLDYVREHKDNIGSGGGGS	RBI-222
24	VKTPQHVKITDFGRAKLLGAEEKGGGGSDEAYVMASVASVDN PHVCRLLGICGGGSGICLTSTVQLIMQLMPFGCLGGGGSLLITQ LMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVDNPHHVCRL LLGICLTSGGGGSDEAYVMASVDSVDNPHVCRLLGICGGGGS	RBI-223
25	DEAYVMASVDNPHHVCRLGICLTSGGGGSGICLTSTVQLIMQ LMPFGCLGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGSLLIT QLMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVASVDNPH VCRLLGICGGGGSDEAYVMASVDSVDNPHVCRLLGICGGGGS	RBI-224
26	LITQLMPFGSLLDYVREHKDNIGSGGGGSVKTPQHVKITDFGRA KLLGAEEKGGGGSDEAYVMASVASVDNPHVCRLLGICGGGGS GICLTSTVQLIMQLMPFGCLGGGGSDEAYVMASVDSVDNPHVC RLLGICGGGGSDEAYVMASVDNPHHVCRLGICLTSGGGGS	RBI-225
27	GICLTSTVQLIMQLMPFGCLGGGGSDEAYVMASVASVDNPHVC RLLGICGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGSDEAY VMASVDSVDNPHVCRLLGICGGGSLITQLMPFGSLLDYVREHK DNIGSGGGGSDEAYVMASVDNPHHVCRLGICLTSGGGS	RBI-226
28	DEAYVMASVDNPHHVCRLGICLTSGGGGSGICLTSTVQLIMQ LMPFGCLGGGGSLLITQLMPFGSLLDYVREHKDNIGSGGGGSDE AYVMASVDSVDNPHVCRLLGICGGGGSVKTPQHVKITDFGRA KLLGAEEKGGGGSDEAYVMASVASVDNPHVCRLLGICGGGGS	RBI-227
29	VKTPQHVKITDFGRAKLLGAEEKGGGGSDEAYVMASVDSVDN PHVCRLLGICGGGGSLLITQLMPFGSLLDYVREHKDNIGSGGGGS DEAYVMASVASVDNPHVCRLLGICGGGSGICLTSTVQLIMQL	RBI-228

	MPFGCLGGGGSDEAYVMASVDNPHHVCRLLGICLTSGGGS	
30	DEAYVMASVASVDNPHVCRLLGICGGGSGICLTSTVQLIMQL MPFGCLGGGSLITQLMPFGSLLDYVREHKDNIGSGGGGSDEA YVMASVDNPHHVCRLLGICLTSGGGGSDEAYVMASVDSVDNP HVCRLLGICGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGS	RBI-229
31	GICLTSTVQLIMQLMPFGCLGGGGSDEAYVMASVASVDNPHVC RLLGICGGGGSDEAYVMASVDNPHHVCRLLGICLTSGGGGSLIT QLMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVDSVDNPH VCRLLGICGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGGS	RBI-230
32	DEAYVMASVDSVDNPHVCRLLGICGGGGSDEAYVMASVASVD NPHVCRLLGICGGGGSVKTPQHVKITDFGRAKLLGAEEKGGGG SLITQLMPFGSLLDYVREHKDNIGSGGGGSDEAYVMASVDNPH HVCRLLGICLTSGGGGSGICLTSTVQLIMQLMPFGCLGGGGS	RBI-231

CLAIMS

WHAT IS CLAIMED IS:

1. A nucleic acid construct comprising a nucleic acid sequence encoding a modified alphavirus genome or self-replicating RNA (srRNA), wherein at least a portion of the nucleic acid sequence encoding the viral structural proteins of the modified alphavirus genome or srRNA has been replaced by a coding sequence for a polypeptide construct comprising a coding sequence for epidermal growth factor receptor (EGFR) comprising one or more acquired resistance mutations.
2. The nucleic acid construct of claim 1, wherein the modified alphavirus genome or srRNA comprises no nucleic acid sequence encoding viral structural proteins.
3. The nucleic acid construct of claim 1, wherein the nucleic acid sequence encoding the polypeptide construct is operably linked to a promoter sequence.
4. The nucleic acid construct of Claim 3, wherein the promoter sequence is a 26S subgenomic (sg) promoter.
5. The nucleic acid construct of Claim 1, wherein the modified alphavirus genome or srRNA is of an alphavirus belonging to the VEEV/EEEV group, or the SFV group, or the SINV group.
6. The nucleic acid construct of Claim 5, wherein the alphavirus is Venezuelan equine encephalitis virus (VEEV), Eastern equine encephalitis virus (EEEV), Madariaga virus (MADV), Chikungunya virus (CHIKV), Western equine encephalitis virus (WEEV), or Sindbis virus (SINV).
7. The nucleic acid construct of Claim 1, wherein the srRNA is a capped srRNA comprising a 5'-cap.
8. The nucleic acid construct of Claim 7, wherein the capped srRNA is a co-

transcriptionally capped srRNA.

9. The nucleic acid construct of Claim 7, wherein the capped srRNA is an enzymatically capped srRNA.

10. The nucleic acid construct of Claim 1, wherein the one or more acquired resistance mutations are configured into a plurality of alteration cassettes arranged in tandem along the length of the coding sequence.

11. The nucleic acid construct of Claim 10, wherein the plurality of alteration cassettes are operably linked to one another by one or more linkers.

12. The nucleic acid construct of Claim 1, wherein the one or more acquired resistance mutations comprise one or more (i) activating mutations, (ii) mutations that enhance binding affinity of EGFR to adenosine triphosphate (ATP), and/or (iii) mutations that block binding of EGFR or variant thereof to an inhibitor.

13. The nucleic acid construct of Claim 12, wherein the one or more activating mutations comprises an in-frame insertion in exon 20.

14. The nucleic acid construct of Claim 13, wherein the exon 20 in-frame insertion comprises a mutation selected from the group consisting of A763_Y764insFQEA, S768_D770dup, S768_V769ins, A767_V769dup, D770_N771insX, V769_D770insX, H773_V774insX, H773dup, N771_H773dup, P772_H773insX, and N771_P772insX.

15. The nucleic acid construct of Claim 14, wherein the exon 20 in-frame insertion comprises S768_D770dup, A767_V769dup, and/or H773dup.

16. The nucleic acid construct of Claim 12, wherein the one or more activating mutations comprises an L858R substitution.

17. The nucleic acid construct of Claim 12, wherein the one or more acquired resistance mutations enhances the binding affinity of EGFR to ATP.
18. The nucleic acid construct of Claim 17, wherein the one or more acquired resistance mutations is selected from the group consisting of T790M, G719X, L858R, L718Q, G724S, L861X, S768I/V, E709X, L747S, D761Y, and T854A.
19. The nucleic acid construct of Claim 12, wherein the one or more acquired resistance mutations blocks binding of EGFR to an inhibitor.
20. The nucleic acid of claim 19, wherein the EGFR inhibitor is selected from the group consisting of osimertinib, lazertinib, erlotinib and gefitinib, CO-1686, HM61713, EGF816, ASP8273 and Avitinib, afatinib, mobocertinib, Icotinib, Dacomitinib, Pozotinib Cetuximab, Amivantamab, Mobocertinib, Furmonertinib, DZD9008, CLN-081, STX-721, YK-029A, HS-10376, Zipalertinib, TAK-788, JMT101, ABT-101, BEBT-109, DZG9008, PLB1004, EMB-01, HS-20117, MCLA-129, BLU-945; BDTX-1535; NX-019; JIN-A02; BBT-207; BLU-525; THE-349; STX-241; ABK3376; BI-732; BLU-701; BBT-176, BLU-451, Almonertinib, and Tarloxotinib.
21. The nucleic acid construct of Claim 19, wherein the one or more acquired resistance mutations is selected from the group consisting of C797S, C797G, T790M, L858R, L858M, L718V, L718Q, C796S, L798I, L792X, SV768IL, L692V, G719A, G719S, G719C, G719D, S786I, L861Q, L861R, V834X, V843X, G724S, E709K, E709H, E709A, E709G, E709V, D761Y, D761N, R776C, R776H, and T854A.
22. The nucleic acid construct of Claim 1, wherein the nucleic acid sequence encoding for EGFR has at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from the

group consisting of SEQ ID NOS: 13-32.

23. The nucleic acid construct of Claim 1, wherein the coding sequence for the polypeptide construct comprises a coding sequence for EGFR comprising one or more acquired resistance mutations selected from S768_D770dup, A767_V769dup, H773dup, T790M, L858R, and C797S.

24. The nucleic acid construct of claim 23, wherein the coding sequence is in a 5' - to 3' - direction.

25. A recombinant cell comprising a nucleic acid construct according to Claim 1.

26. The recombinant cell of Claim 25, wherein the recombinant cell is a mammalian cell or an insect cell.

27. A pharmaceutical composition comprising a pharmaceutically acceptable excipient and a nucleic acid construct of Claim 1.

28. The pharmaceutical composition of Claim 27, wherein the composition is formulated with a delivery vehicle into a delivery system, wherein the delivery system comprises a liposome, a viral replicon particle (VRP), a lipid-based nanoparticle (LNP), a polymer nanoparticle, a physiologic buffer, a microsphere, an immune stimulating complex (ISCOM), a conjugate of bioactive ligand, or a combination of any thereof.

29. The pharmaceutical composition of Claim 28, wherein the LNP delivery system comprises a cationic lipid, an ionizable cationic lipid, an anionic lipid, or a neutral lipid.

30. The pharmaceutical composition of Claim 29, wherein the lipid is present in mass ratio of lipid to RNA from about 100:1 to about 4:1.

31. The pharmaceutical composition of Claim 27, wherein the lipid-based nanoparticles have

an average diameter of about 25 nm to about 1000 nm.

32. The pharmaceutical composition of Claim 27, wherein the composition is formulated as a vaccine or immunotherapeutic.

33. A method for inducing an immune response or treating a health condition in a subject in need thereof, the method comprises administering to the subject a composition comprising a nucleic acid construct of Claim 1.

34. The method of Claim 33, wherein the method is a method for inducing an immune response.

35. The method of Claim 33, wherein the method is a method for treating cancer.

36. The method of Claim 33, wherein the cancer is non-small cell lung cancer (NSCLC).

37. The method of Claim 33, wherein the composition is administered to the subject individually as a single therapy (monotherapy) or as a first therapy in combination with at least one additional therapies.

38. A method for inducing at least one pharmacodynamic effect in a subject, the method comprises administering to the subject a composition comprising a nucleic acid construct of Claim 1.

39. The method of Claim 38, wherein the administered composition results in induced production of one or more of the following: immune responses and mediators selected from TNF, IL-1b, IL-12, IL-2, IFN α , IFN β , IL-6, and IFN γ .

40. The method of Claim 38, wherein the at least one pharmacodynamic effects comprises one or more of the following: immunogenicity effect, a biomarker response, a therapeutic effect, a prophylactic effect, a desired effect, an undesired effect, an adverse effect, and effect in a disease

model.

41. The method of Claim 38, wherein the administered composition enhances antitumor immunity in a tumor microenvironment.

42. The method of Claim 41, wherein the subject has a cancer.

43. The method of claim 42, wherein the cancer is a lung cancer.

44. The method of claim 43, wherein the lung cancer is a NSCLC.